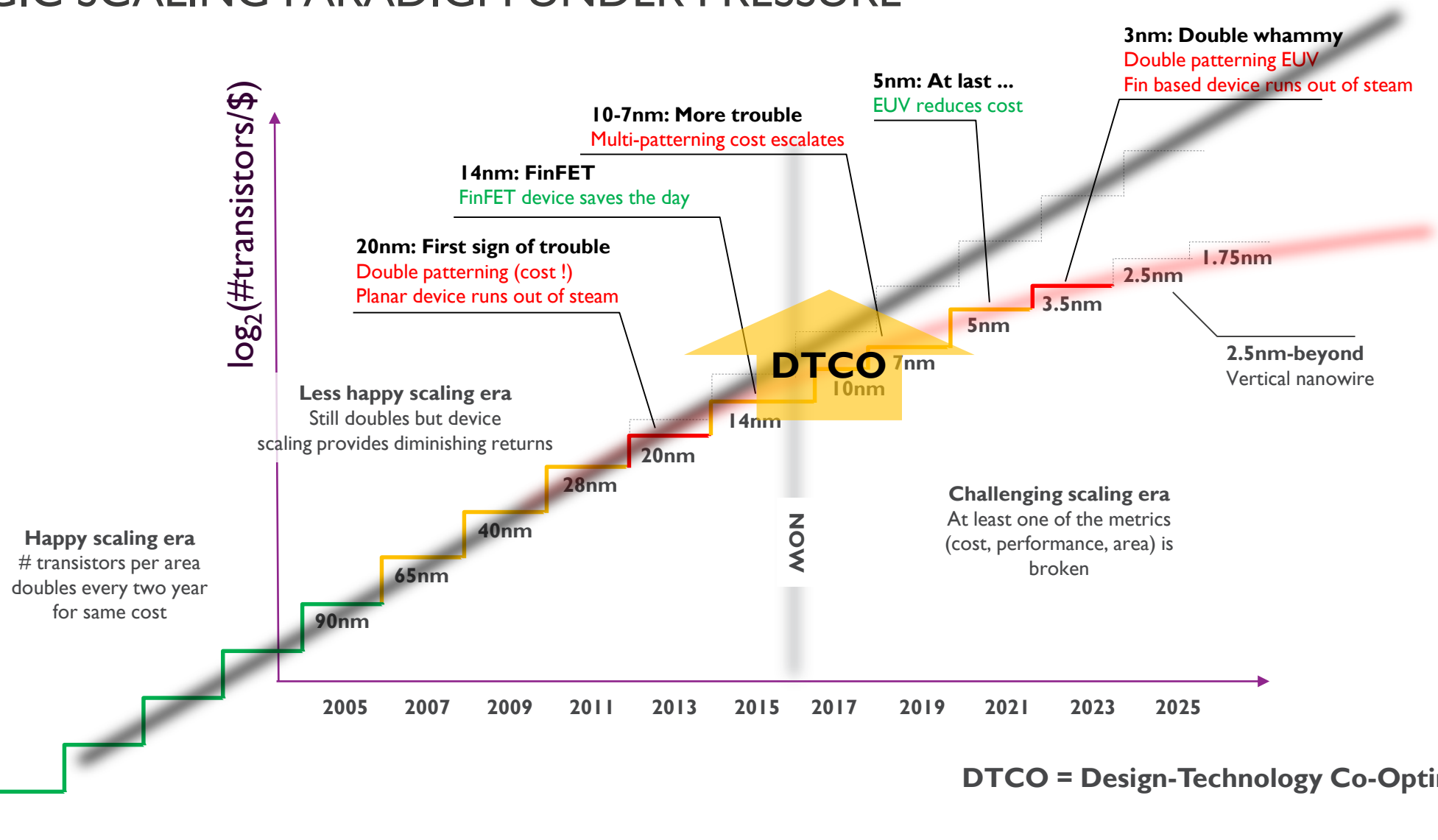
SILICON SCALING BY EXPLOITING THE 3RD DIMENSION

JULIEN RYCKAERT

TECHNOLOGY TRENDS
LOGIC SCALING PARADIGM UNDER PRESSURE

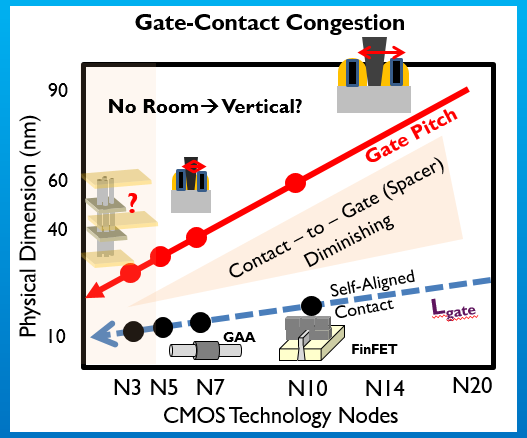


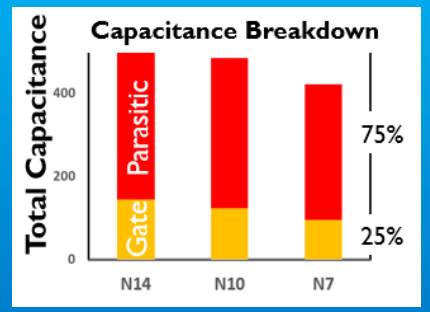
DTCO = Design-Technology Co-Optimization

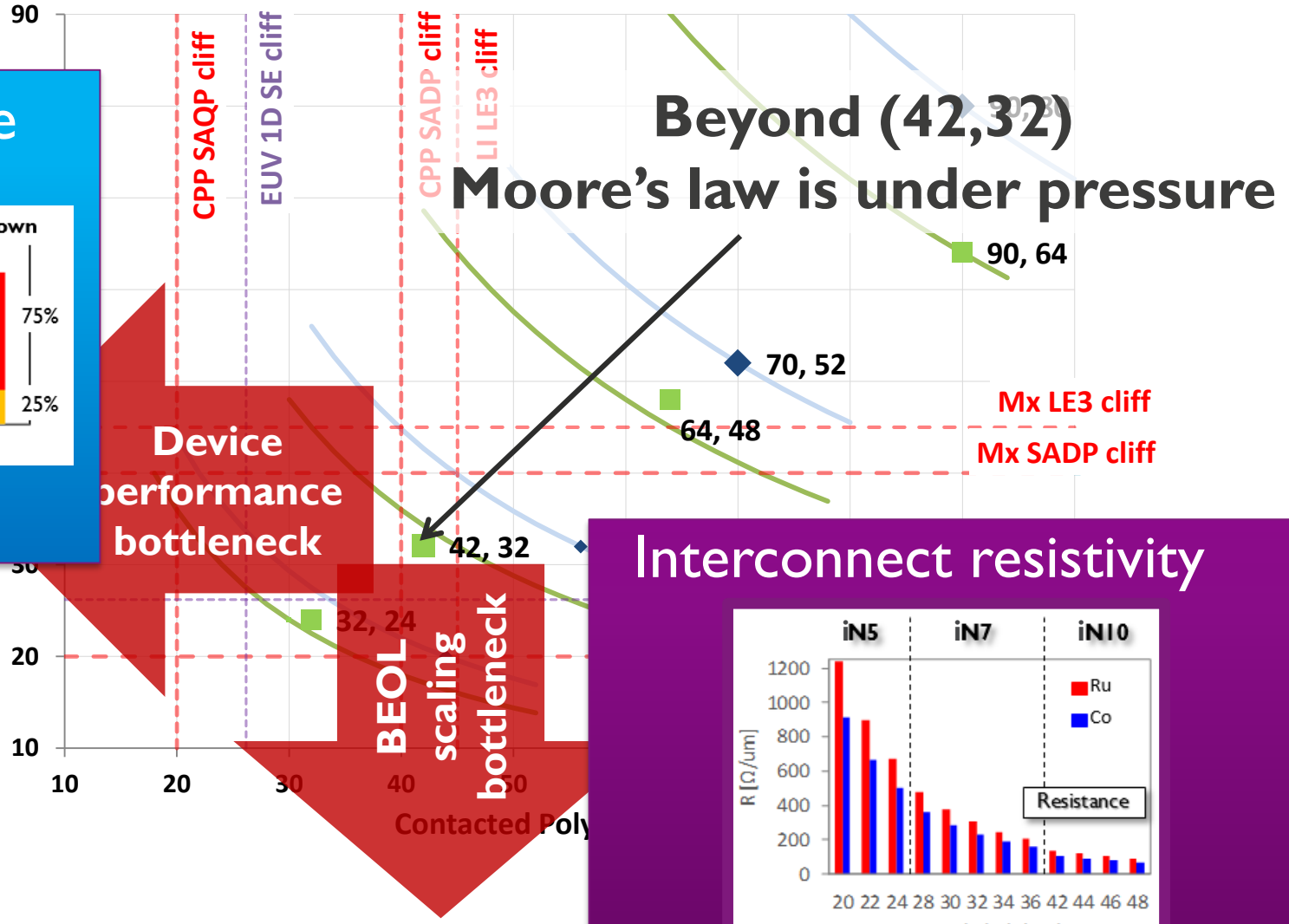
DIMENSIONAL SCALING LOOSES STEAM

NEED TO FIND OTHER MEANS FOR SCALING

Device power & performance





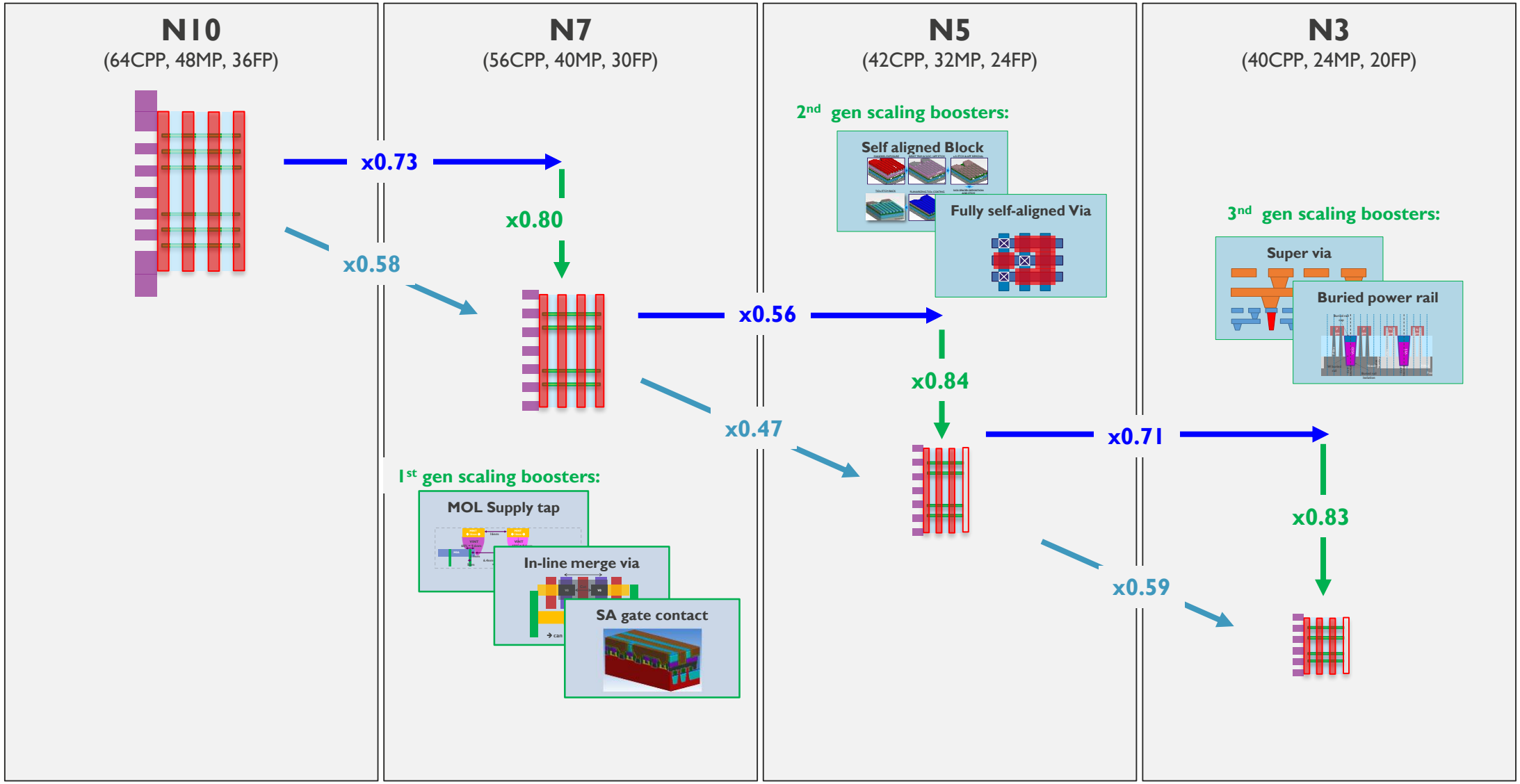


Interconnect resistivity

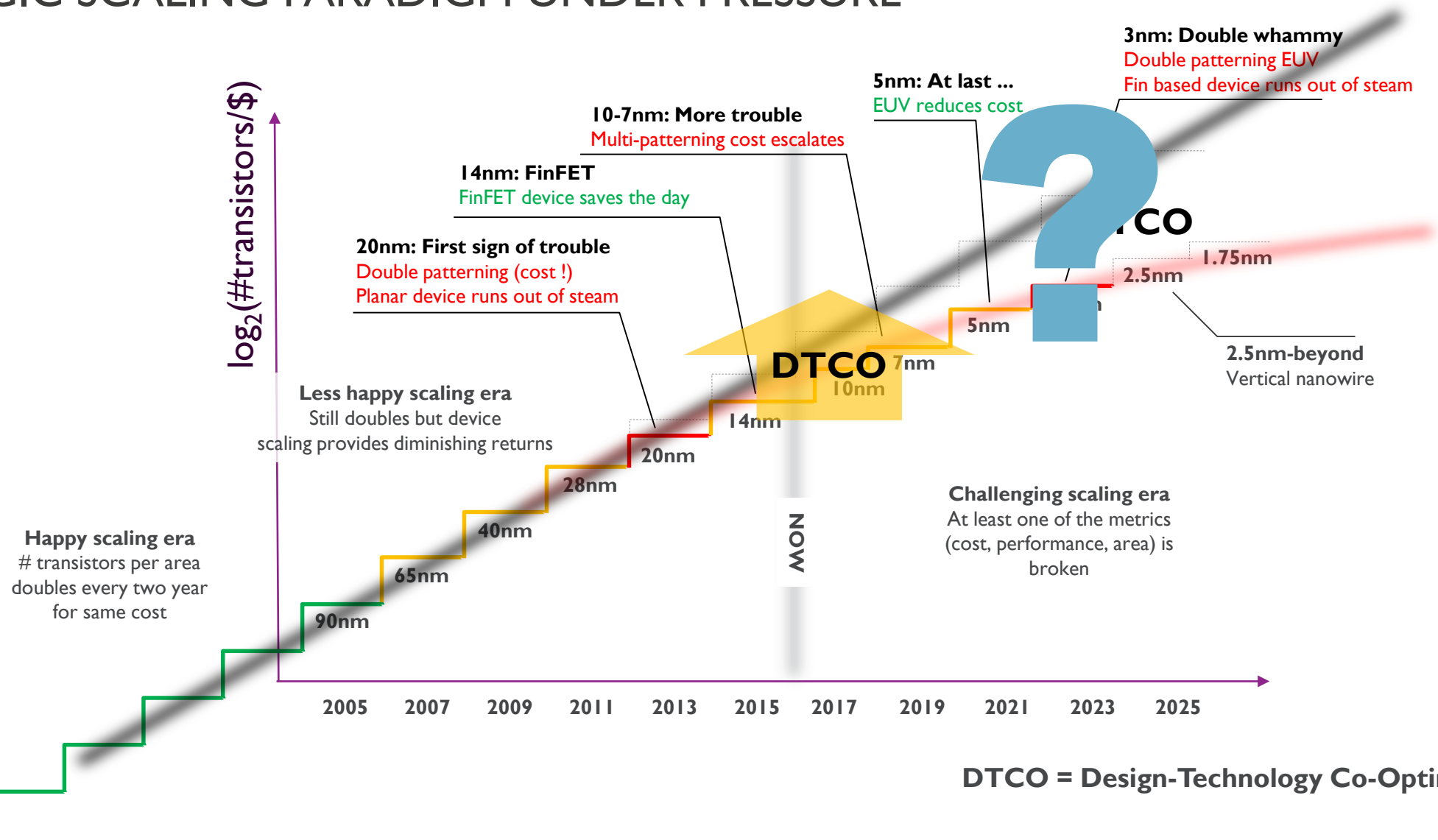
imec

SCALING BOOSTERS TO ENABLE FURTHER SCALING

DIFFERENT GENERATIONS

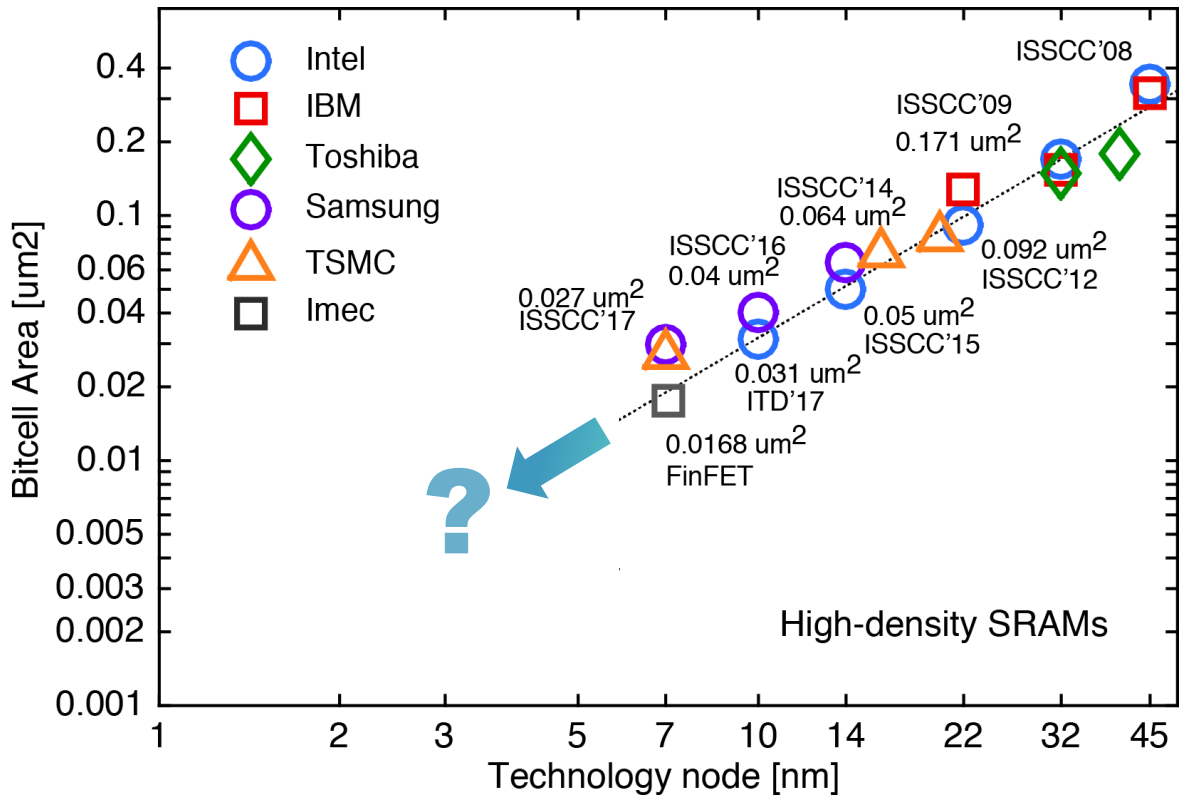


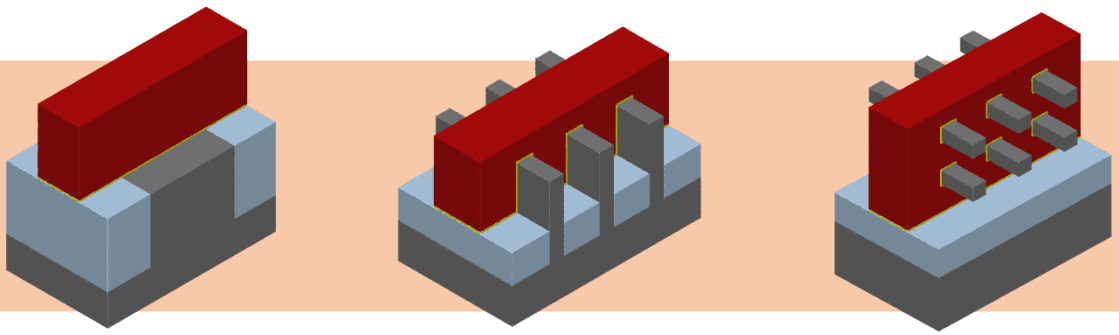
TECHNOLOGY TRENDS
LOGIC SCALING PARADIGM UNDER PRESSURE



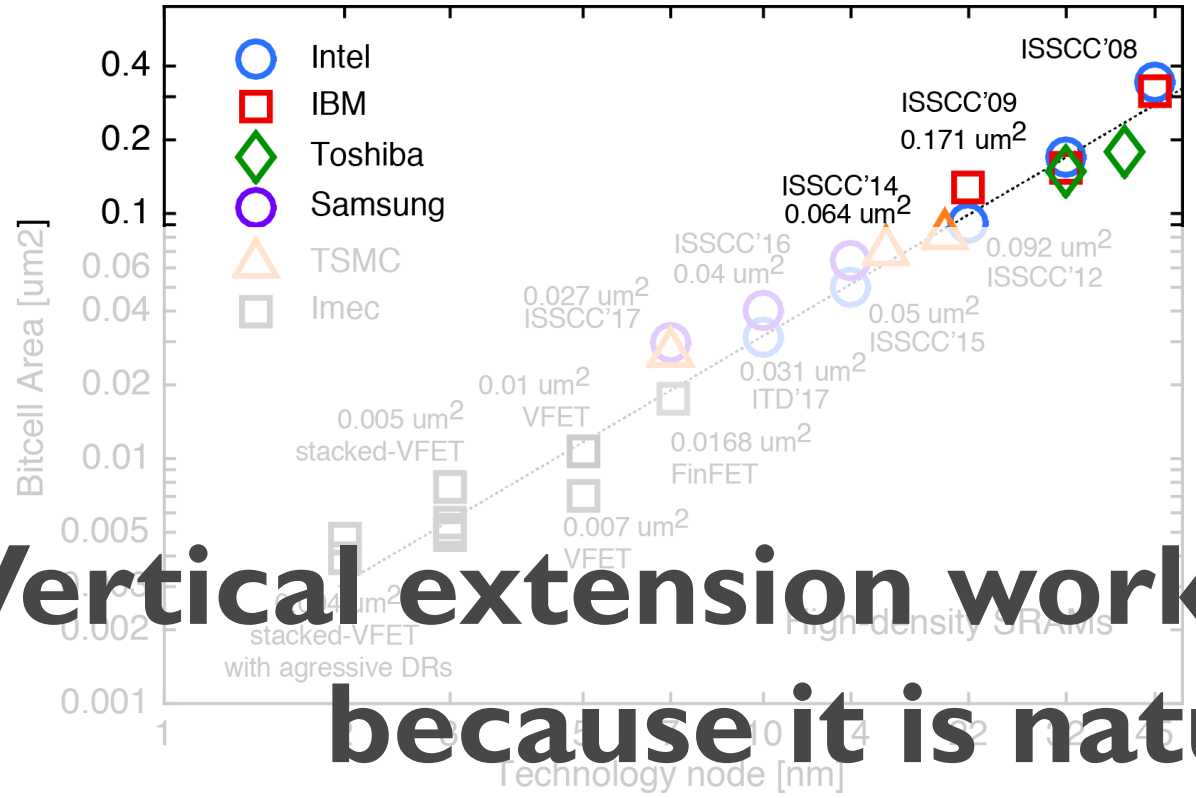
DTCO = Design-Technology Co-Optimization

NEW DEVICE ARCHITECTURES MAINTAIN SRAM SCALING



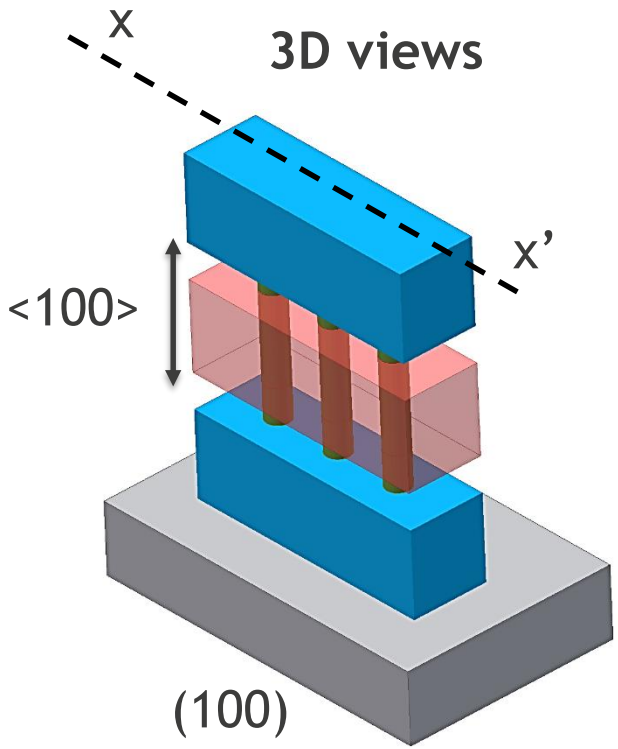


VERTICAL INTEGRATION INCREASES SRAM DENSITY



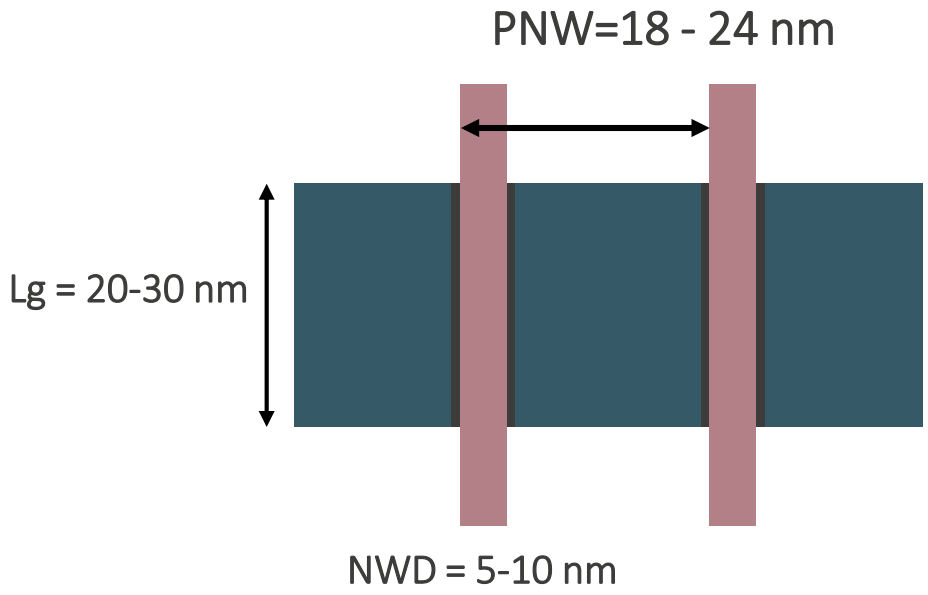
Vertical extension works well for the SRAM because it is naturally regular

VERTICAL GAA NW TRANSISTORS
DIMENSIONS



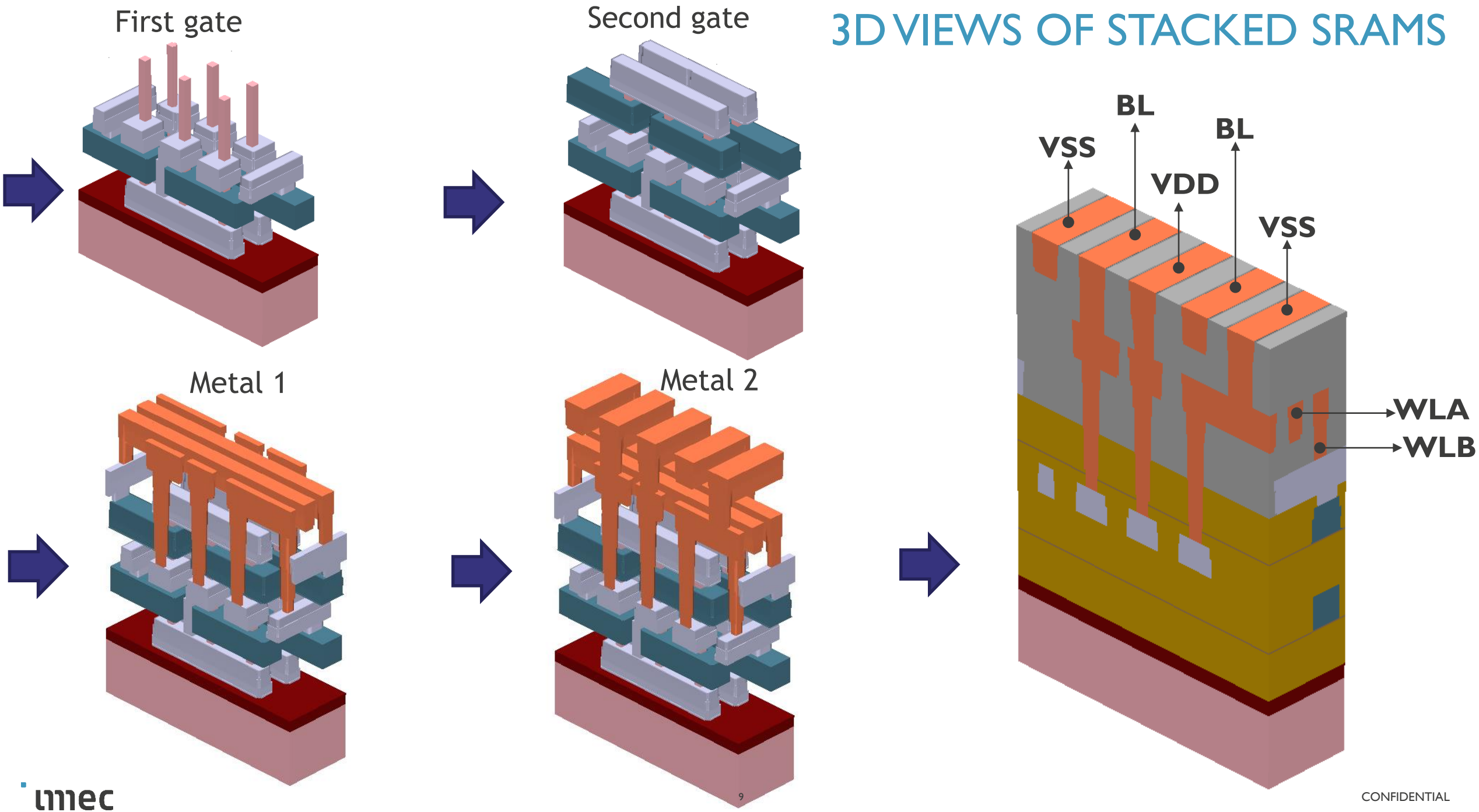
3D views

Cross-section views



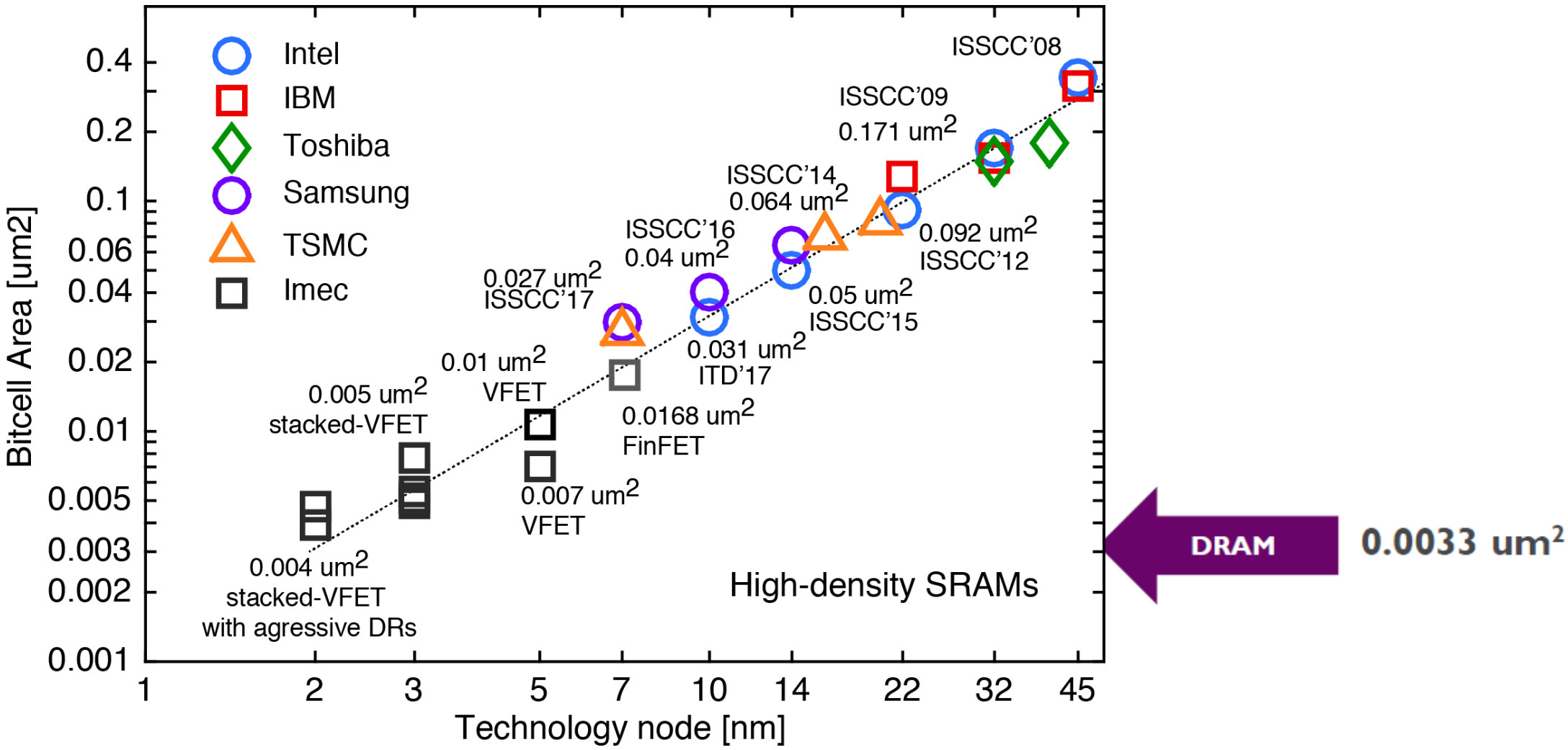
Need to control the alignement of the gate and the S/D junctions

3D VIEWS OF STACKED SRAMS



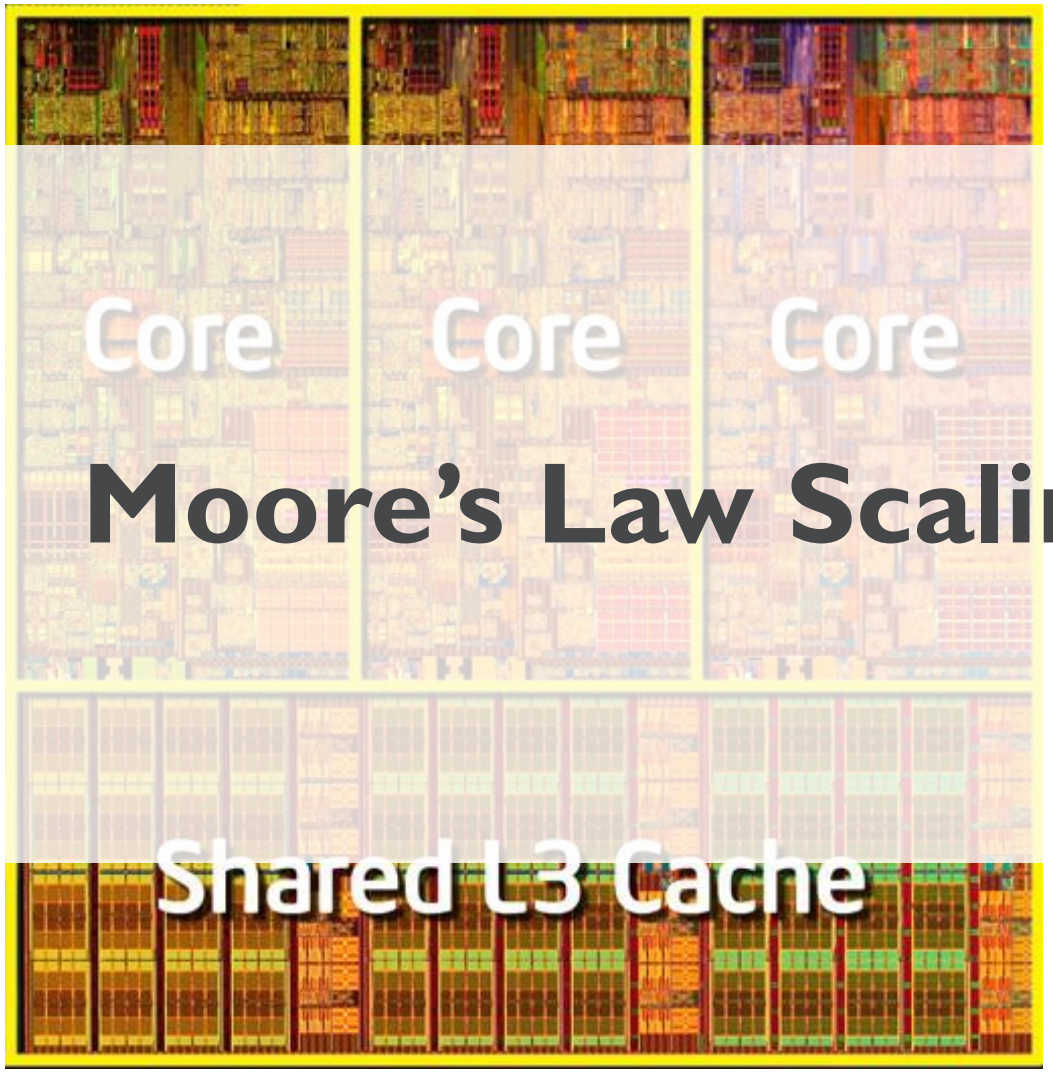
SRAM CAN BECOME COMPETITIVE TO DRAM IN SIZE

THANKS TO VERTICAL INTEGRATION



THE CMOS TECHNOLOGY ANTAGONISM

LOGIC AND MEMORY NEED TO SCALE TOGETHER



LOGIC
→ STANDARD CELLS

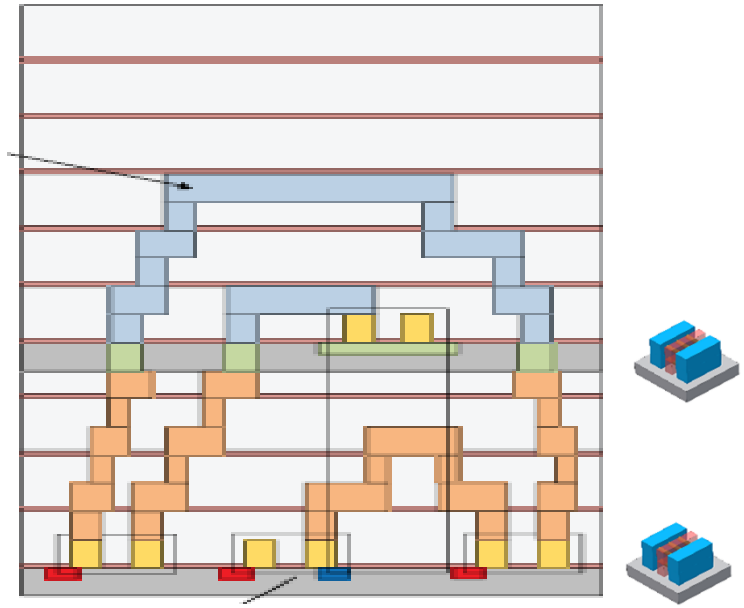
Moore's Law Scaling requires both to scale

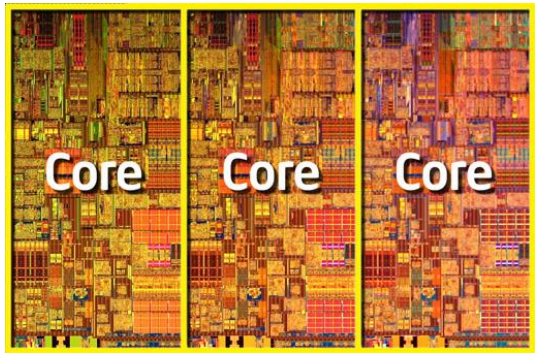
MEMORY (on-chip)
→ **SRAM**

VERTICAL LOGIC… REALLY?

POTENTIAL GAINS IN LOGIC ARE NOT EVIDENT…

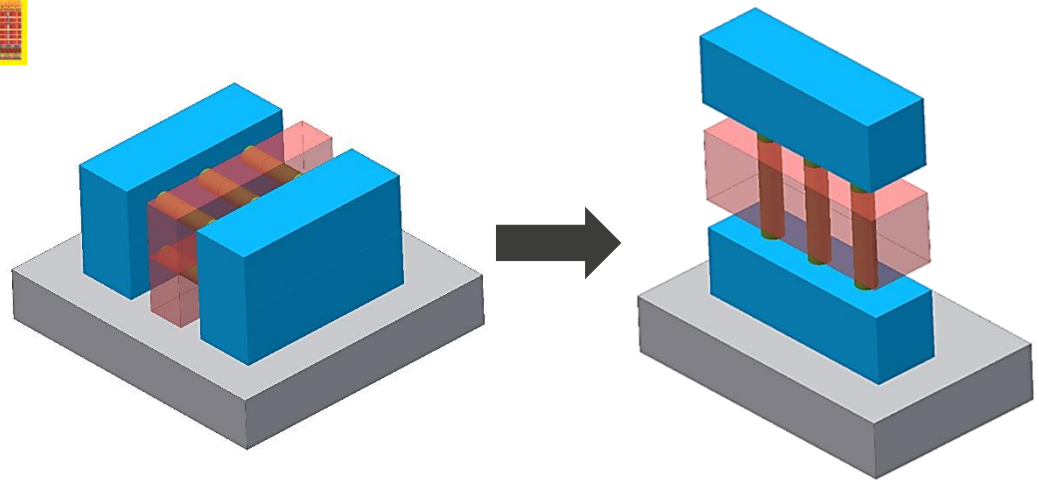
Monolithic 3D





Requires heavy interconnect
→ 2x$

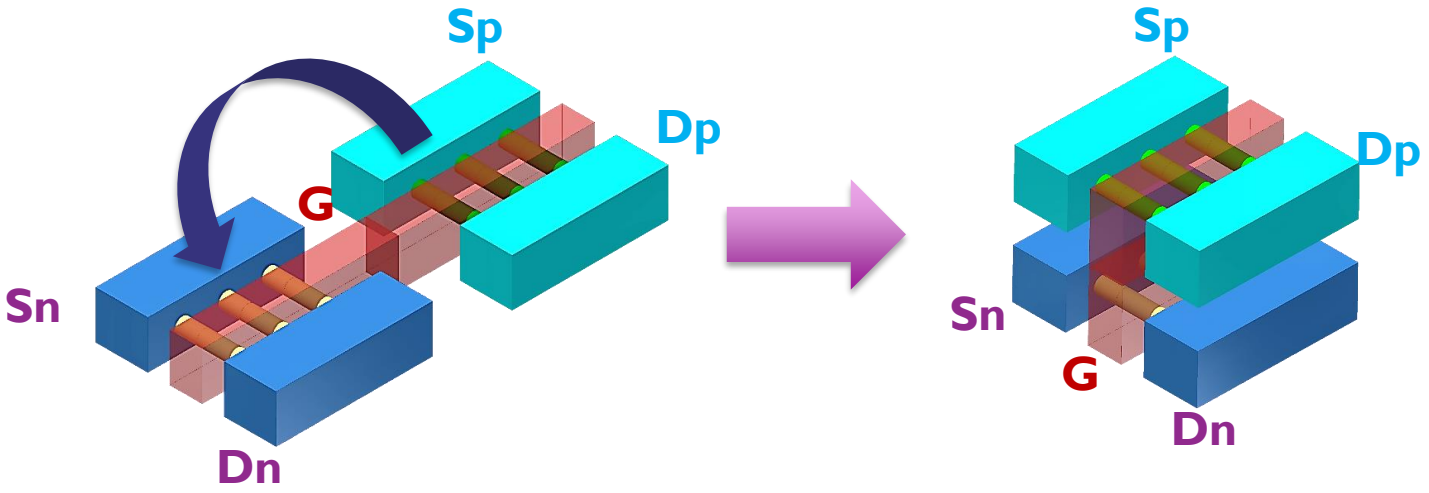
Vertical Device



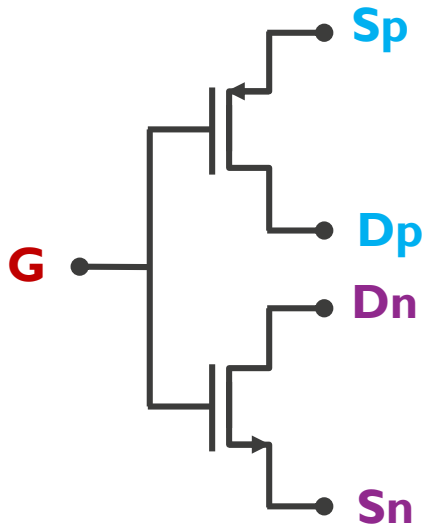
Bottom electrode "hidden"
→ Standard cell won't scale

COMPLEMENTARY FET SCALING

THE FIRST STEP TOWARDS FUNCTIONAL SCALING



Stacked p-n LNW to fold a CMOS structure into a 5 terminal Complementary FET device (CFET)

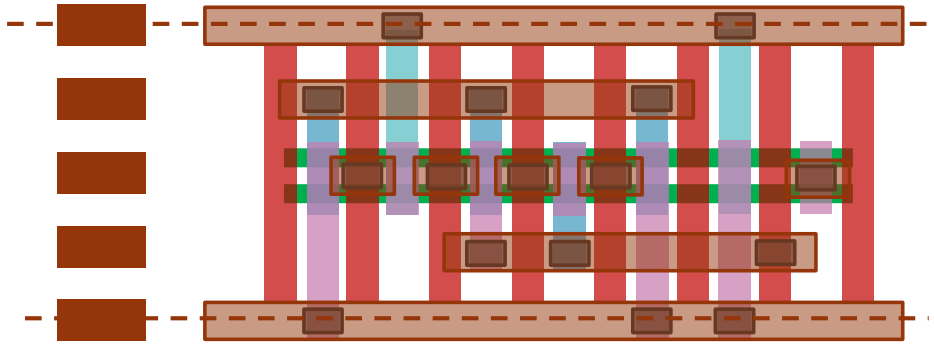


→Requires complex 2 level MOL integration

→Common gate for p and n devices

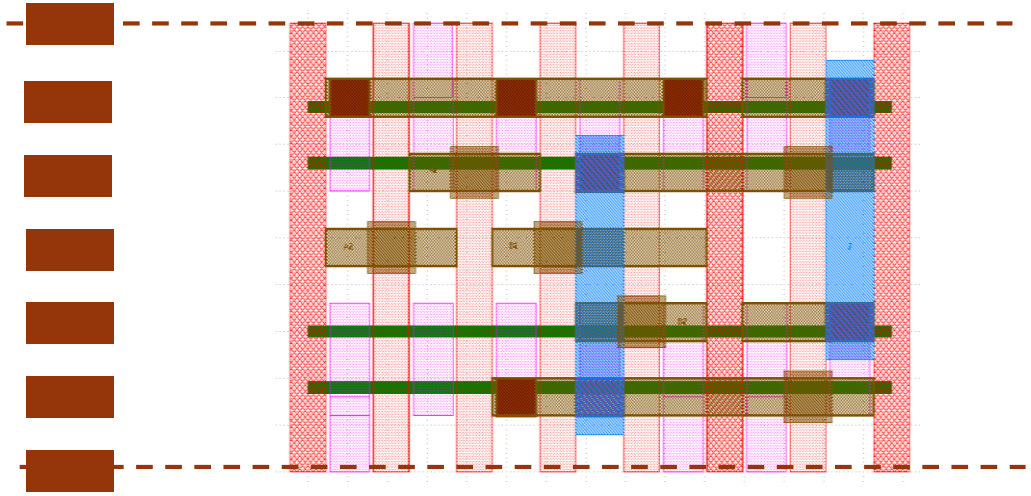
COMPARISON WITH CONVENTIONAL 6T CELL

CFET ENABLES REDUCTION TO 4T CELLS



4T P-N stacked

(SDB, SAGC, out-bound PR)



6T

(SDB, SAGC, out-bound PR)

→ 33% area reduction from 6T

→ Less M1 usage in the cell

KEY STEPS IN THE CFET FLOW

INNER SPACER AND S/D MOL STACK AS KEY CHALLENGES

Fin Etch with Cut

Dummy Gate w spacer

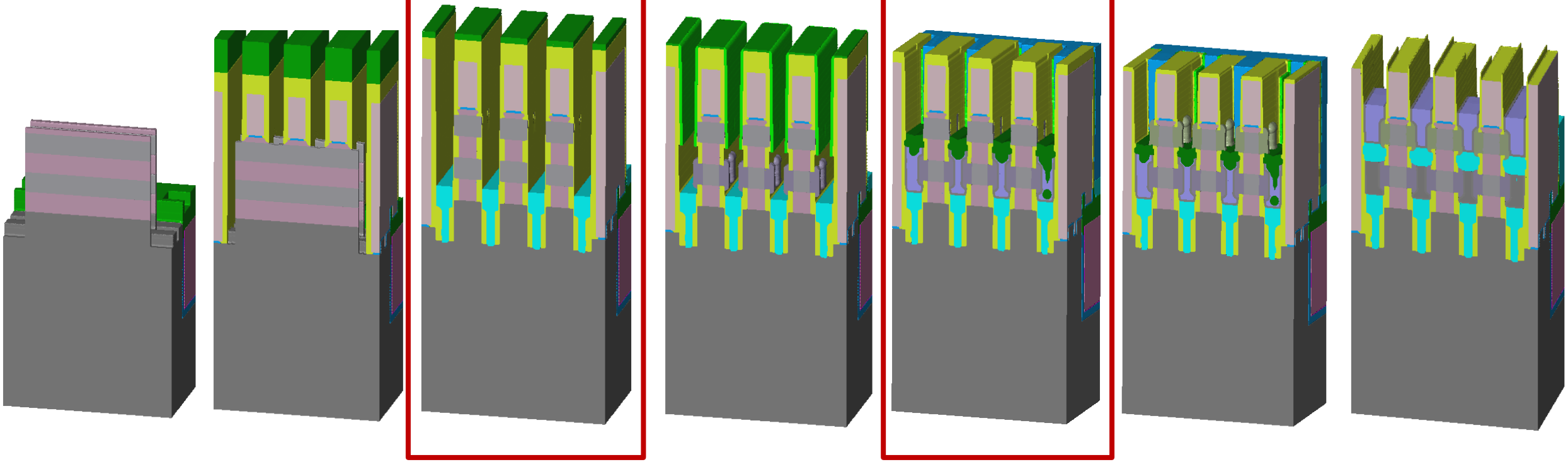
SiGe Recess and Inner Spacer

n-Source-Drain Grow

Metal Fill and recess

p-Source-Drain Grow

Upper Trench Contact



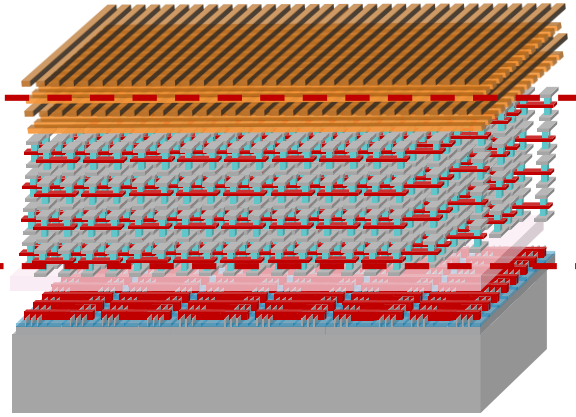
Inner spacer between largely spaced NWs

Dielectric growth between 2 metal electrodes

MOORE'S LAW ON THE VERGE OF MORPHING

NEW 3D LOGIC PARADIGM FOR FUNCTIONAL SCALING

NANOFABRIC BASED LOGIC

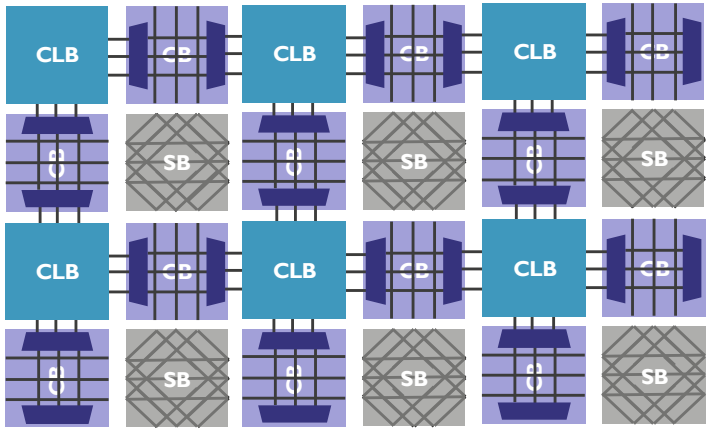


Global BEOL → Design Specific

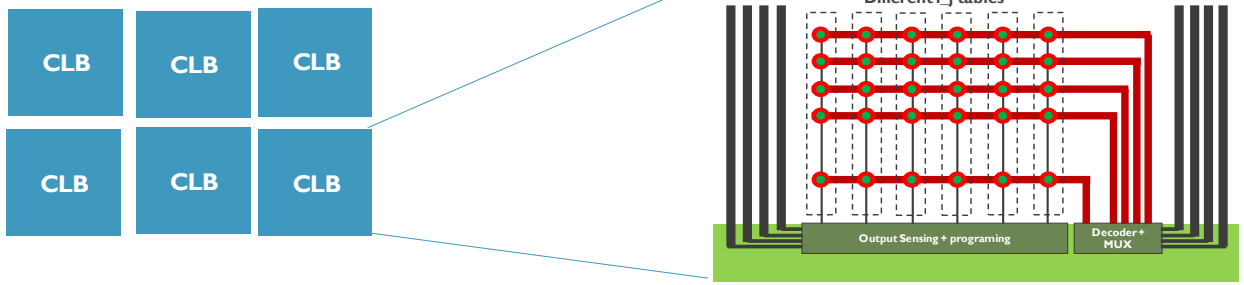
Nanofabric → stacked scalable process

Baseline technology → "intelligence" in the fabric utilization

FPGA



Primitive functions in a nanofabric structure

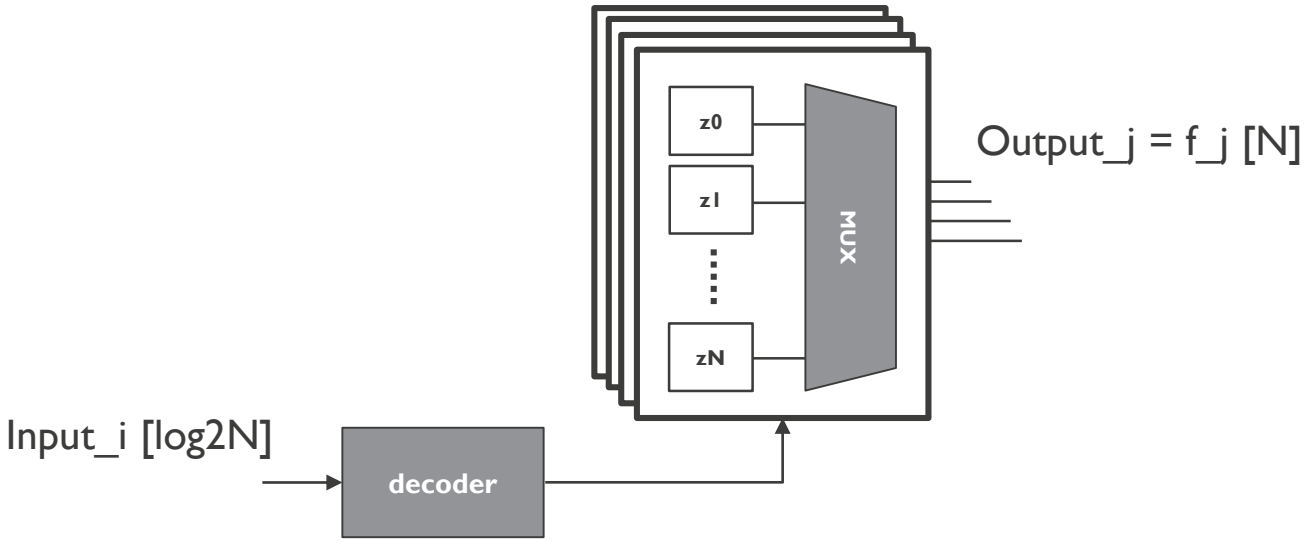


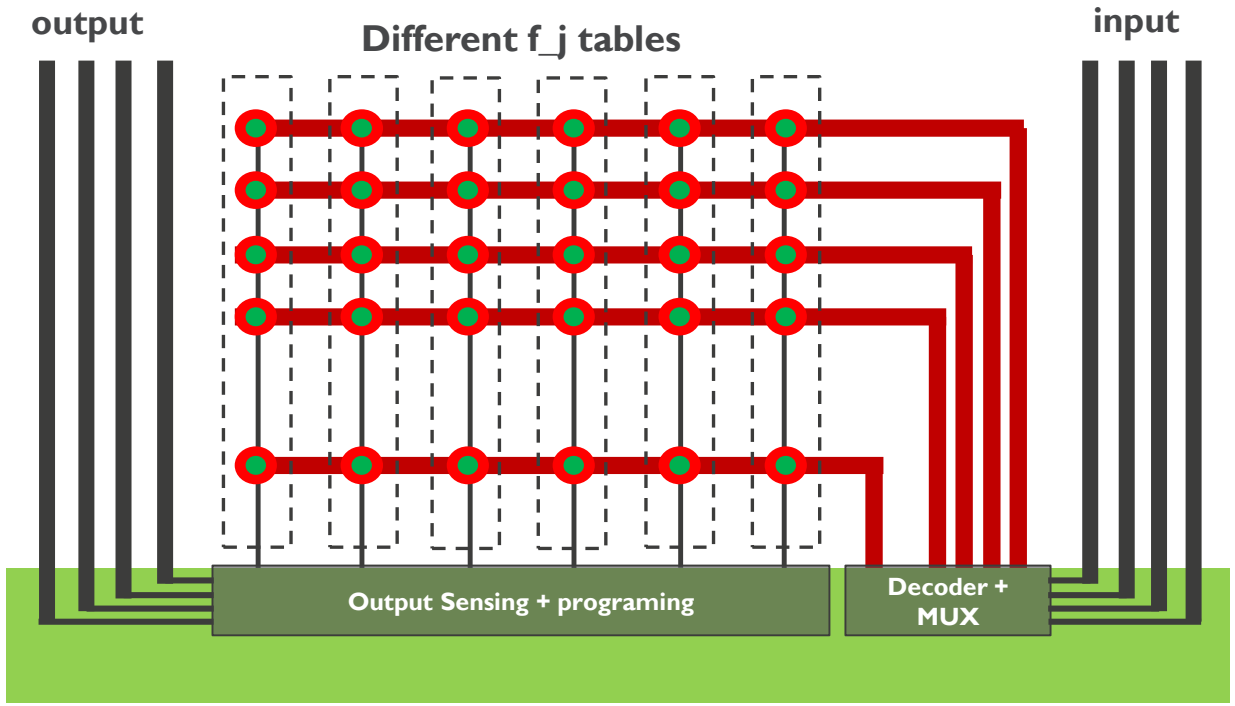
Concept builds on the FPGA as a tile of functional primitives

Key is creating highest expressivity in a small area but no need for interconnect reconfigurability!

LUT BASED CLB AS THE MOST COMPLETE SOLUTION

ALLOWS EXPLORING THE REQUIRED FUNCTIONAL GRANULARITY





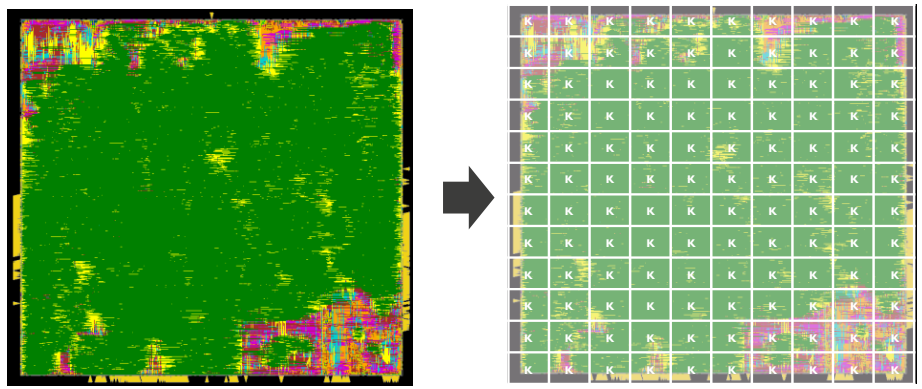
- **Expressivity driven by input/output characteristic of the primitive**
- **LUT can emulate all possible logic truth tables**

➔ **LUT based Synthesis with standard BEOL PnR**

FUNCTIONAL SCALING AS A SCALING VECTOR

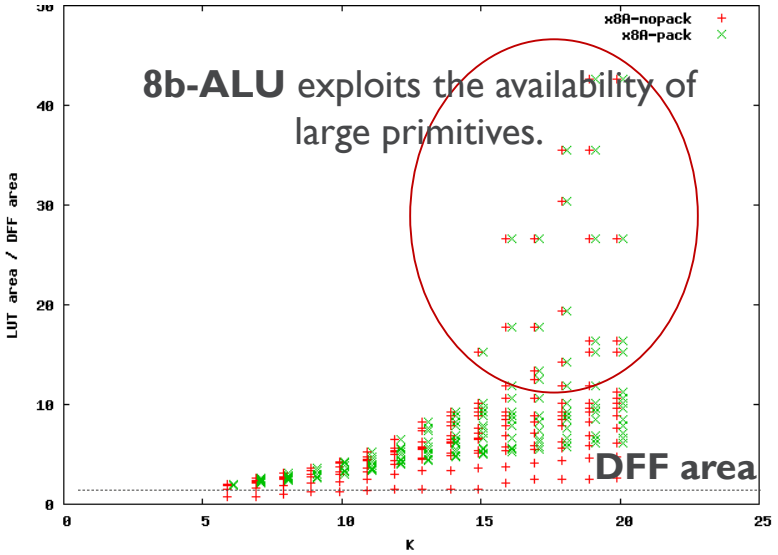
PARTITION INTO LARGER PRIMITIVES

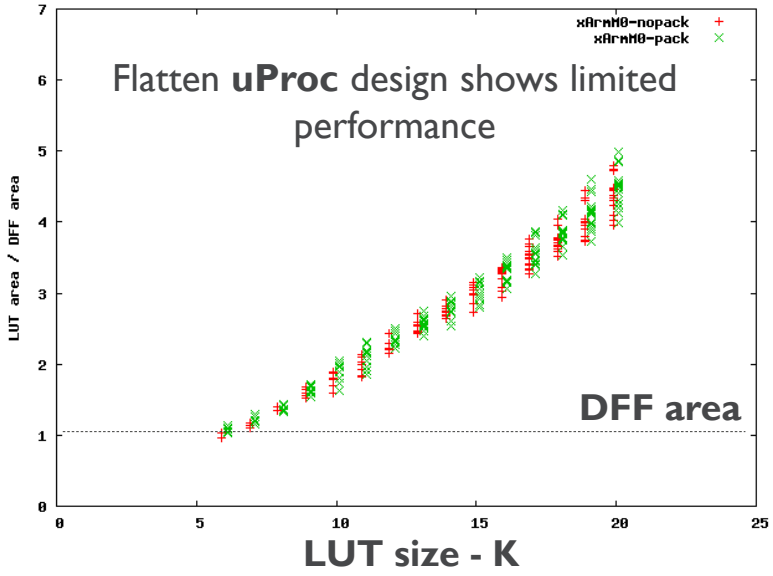
Comparison of LUT to standard cell implementation



Standard cell implementation

LUT-K implementation



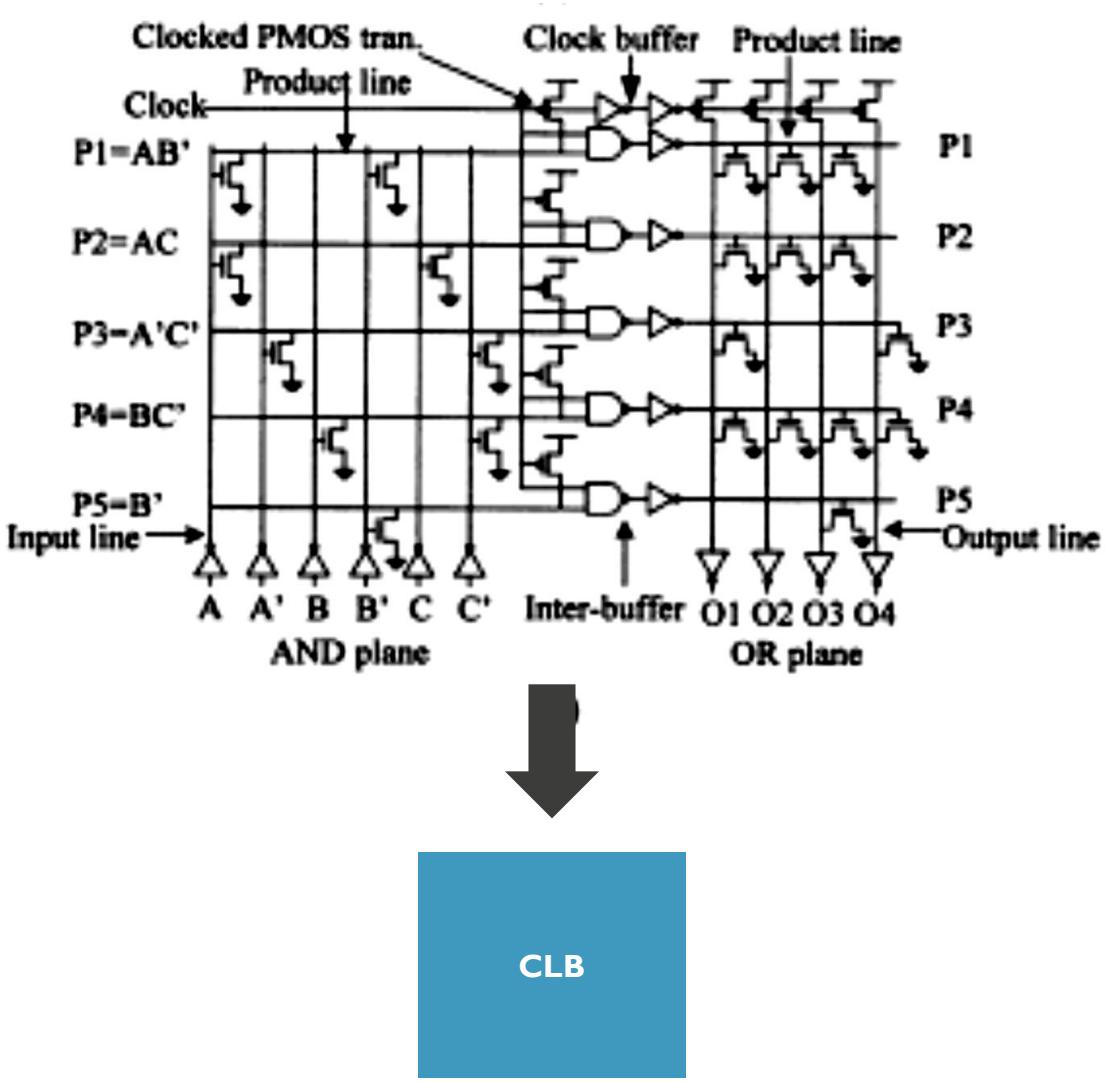


➔ **Explore functional scaling as a potential universal scaling vector for logic**

➔ **Develop new optimization for synthesis (collaboration with EPFL and Univ Utah)**

OTHE CLB IMPLEMENTATIONS

GATE ARRAYS (PLA) CAN EMULATE A LOGIC FUNCTION



Only sum of products can be emulated

→ High entropy (NAND/NOR)

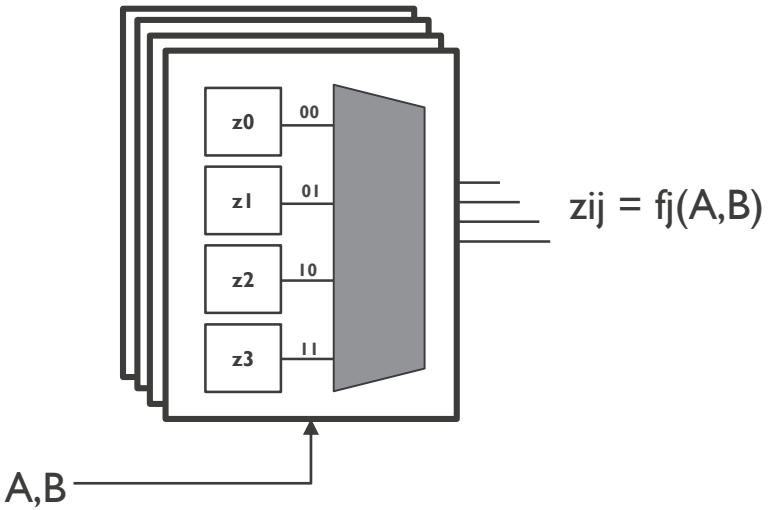
→ Lots of "waste" in the physical implementation

→ Mask-programmed (requires seq process)

RESTRICTION ON THE OUTPUT SIZE

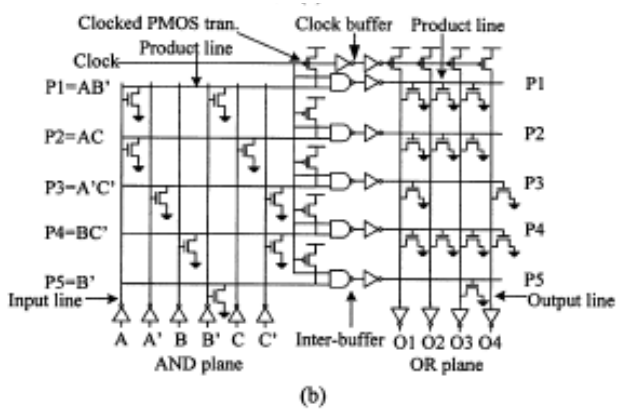
DIFFERENT IMPLEMENTATIONS LEAD TO DIFFERENT I/O CONFIG

LUT approach



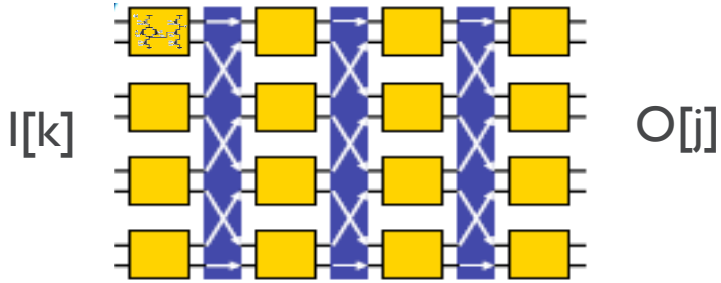
Multiple functions are created from the same inputs by duplicating the LUT structures with different truth tables

PLA approach



The PLA offers by construction multiple output SoP (O[j])

BBD Matrices



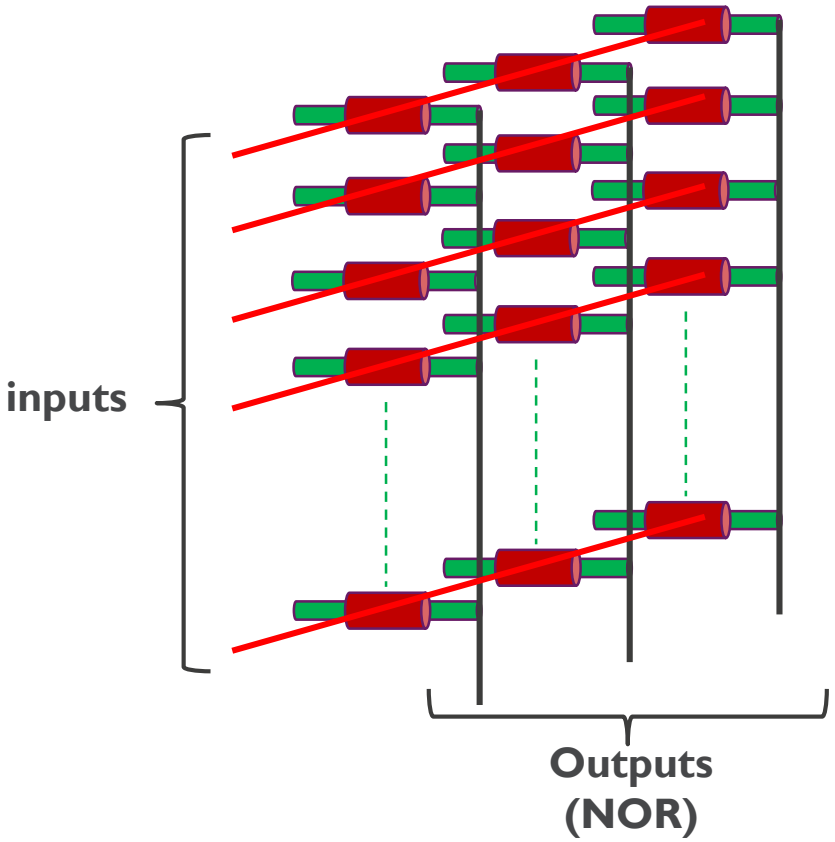
Network of LUTs can implement multiple O[j] even in a fixed network

... Others?

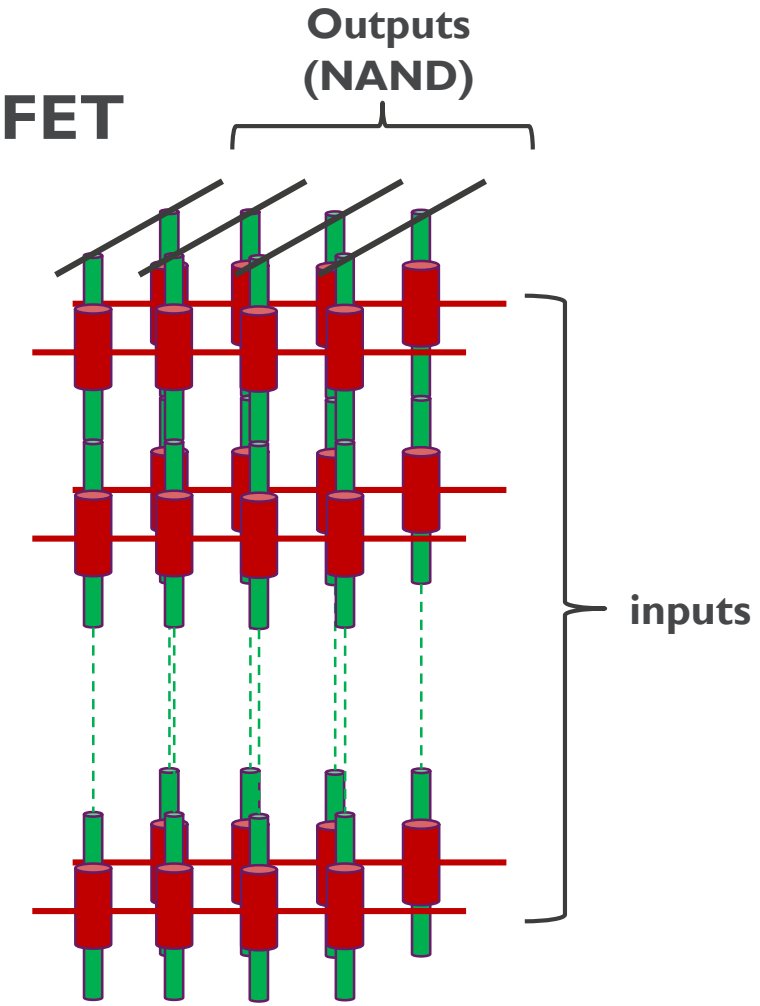
3D IMPLEMENTATION OF A FABRIC

ORIENTATION OF THE CHANNEL IMPACTS PRIMITIVE

Stacked LFET



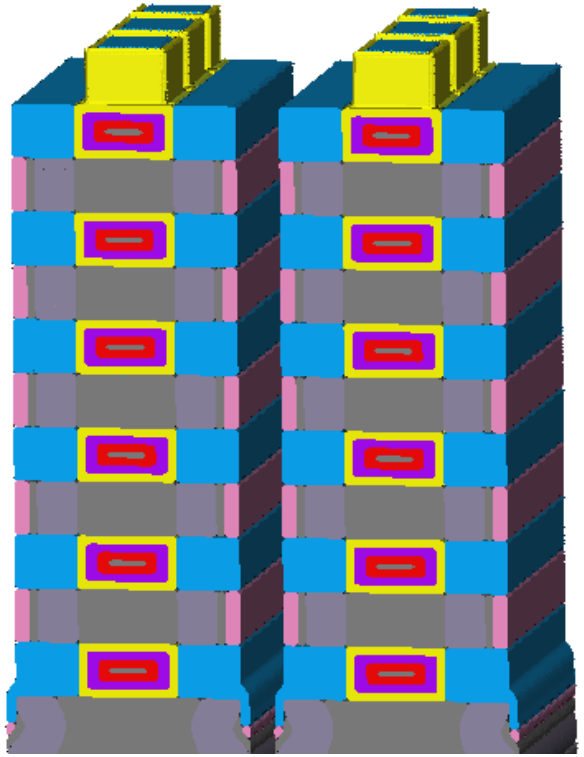
Stacked VFET

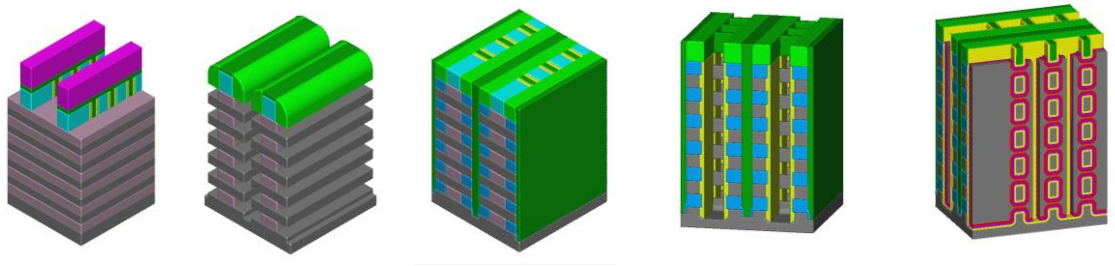


NOR has a preference for Lateral channels and are most likely more efficient

FIRST ATTEMPTS TO EXPLORE MANUFACTURABILITY

BUILD ON WHAT WE CAN DO TODAY





Build an array of FETs with "classical" process steps

...but interestingly we can deviate from typical standard cell layout paradigms and foresee relaxation of some dimensions to best exploit the 3D

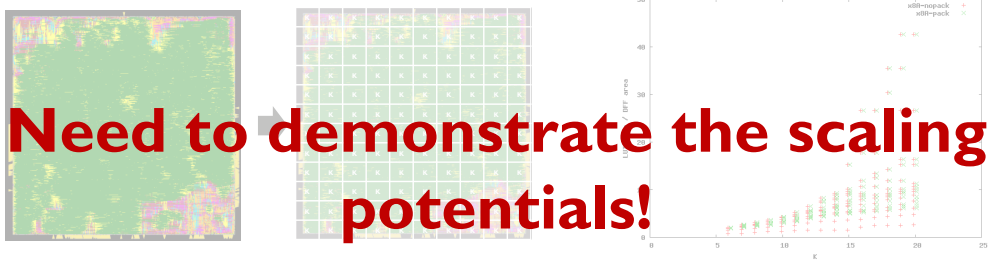
NANOFABRIC ARCHITECTURE

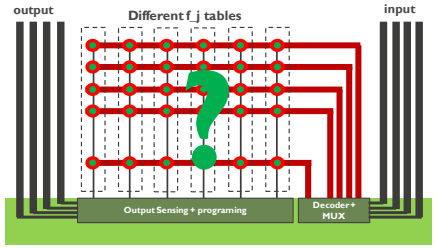
TOP-DOWN AND BOTTOM-UP APPROACH

System architecture

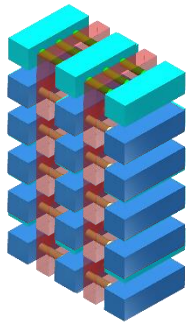
Circuits

Manufacturing





3D circuit primitive design: LUT, PLA, ...



Efficient and manufacturable device stacking for logic

THANKS

Besides the many colleagues at imec...

- Pierre-Emmanuel Gaillardon
- Mathias Soeken
- Eleonora Testa
- Edouard Giacomin
- Giovanni De Micheli