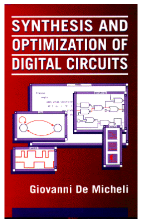
Design Technologies for Integrated Systems

Giovanni De Micheli

Integrated Systems Laboratory



Course objectives

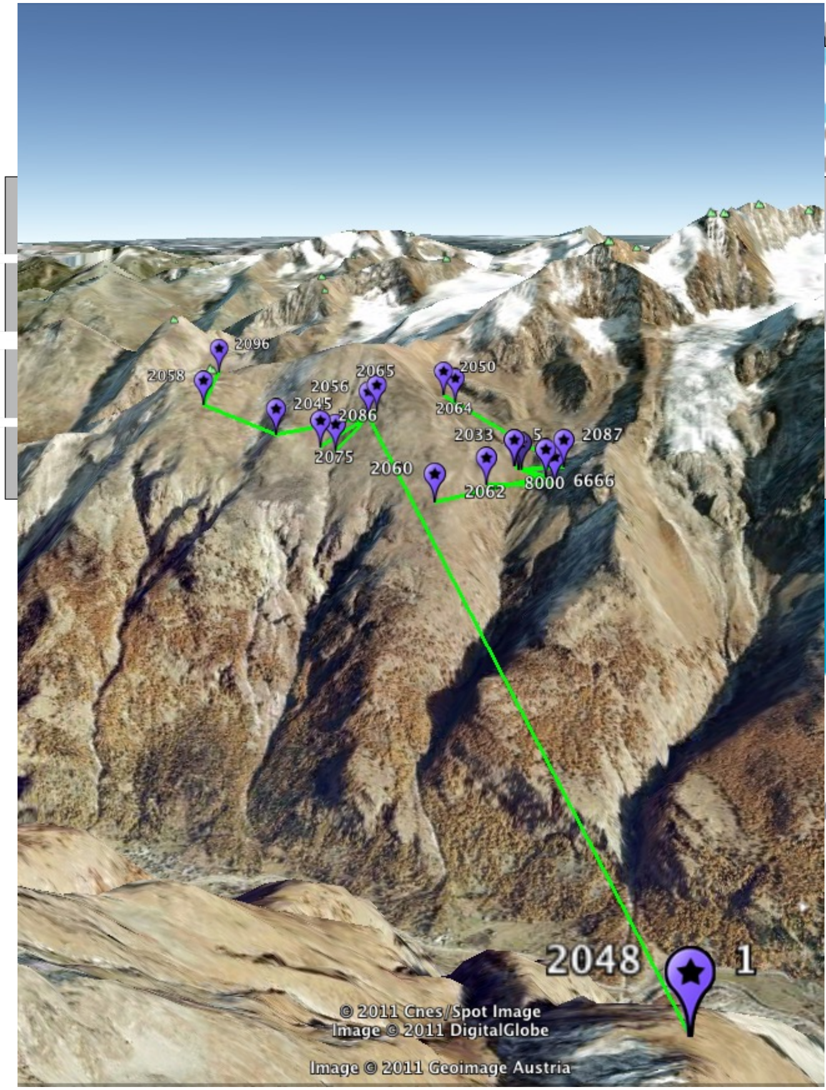
- A state-of-the art “*chip*” (integrated circuit) consists of up to trillions of transistors
- These are designed using *Electronic Design Automation* (EDA) tools
- In this course you will:
  - Understand the inner workings of (some) EDA tools
    - Models and algorithms
  - How to use them effectively to design digital hardware

Module 1

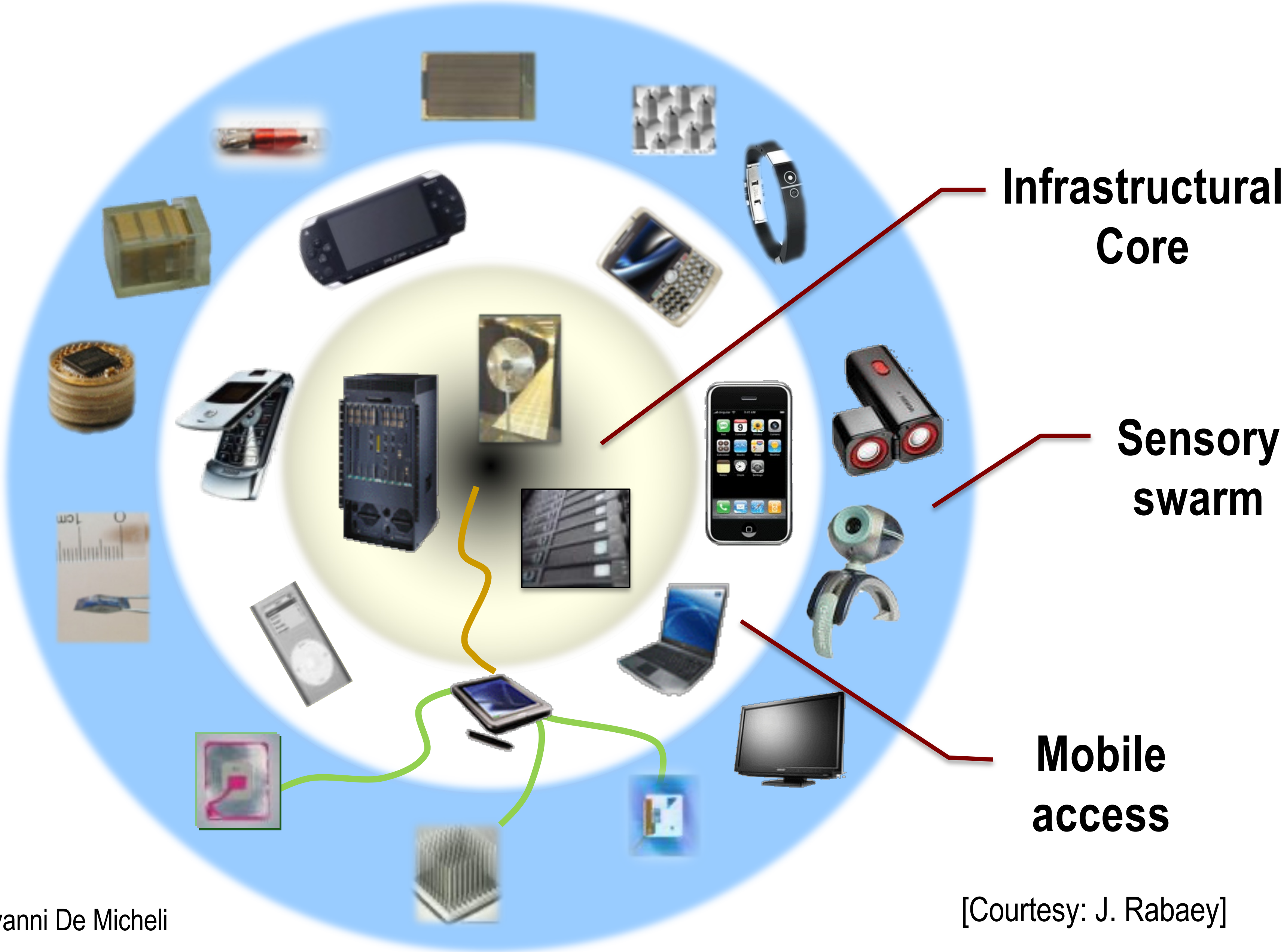
- **Objective**
  - **Electronic systems and their requirements**
  - **Integrated circuits**
  - **Design styles**

Computing today

Computing today

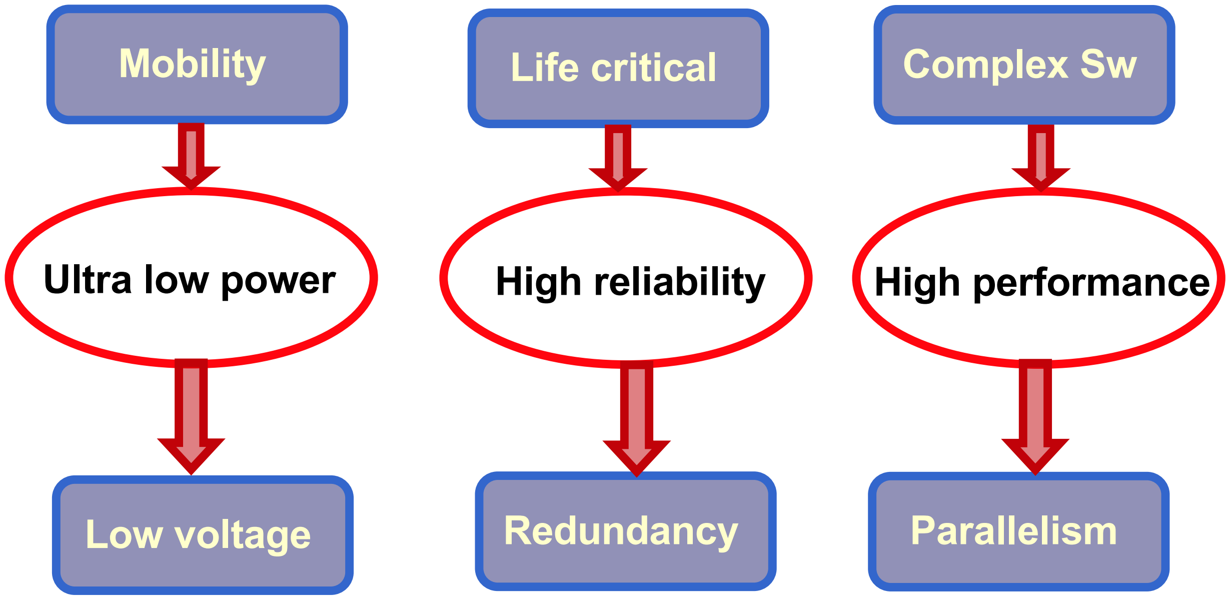


The Internet of Things

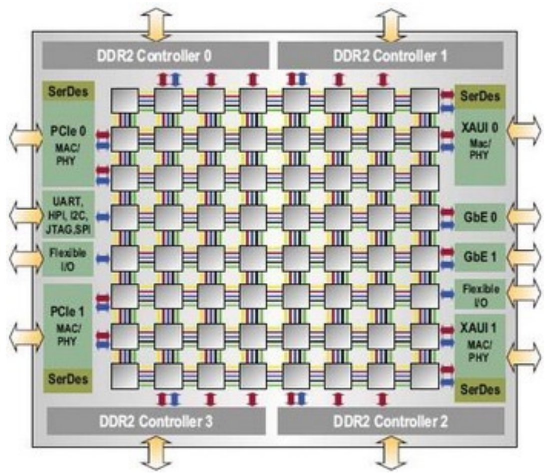


[Courtesy: J. Rabaey]

Requirements for electronic chip design



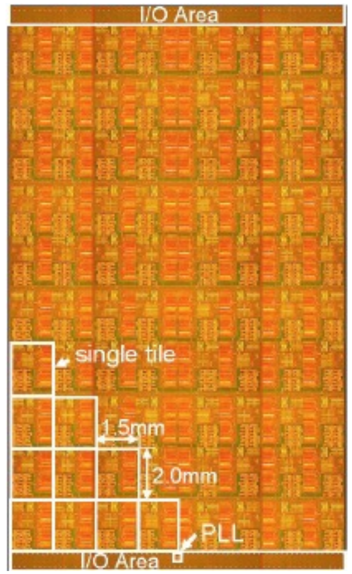
- From processors to multi-processors
  - Clock speed limitaions
  - Performance gain from parallelism



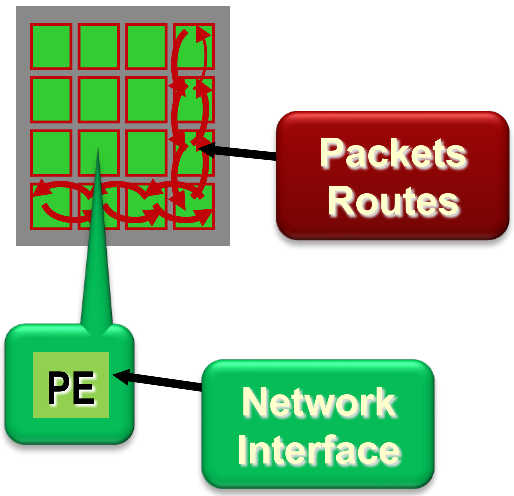
[Tilera]

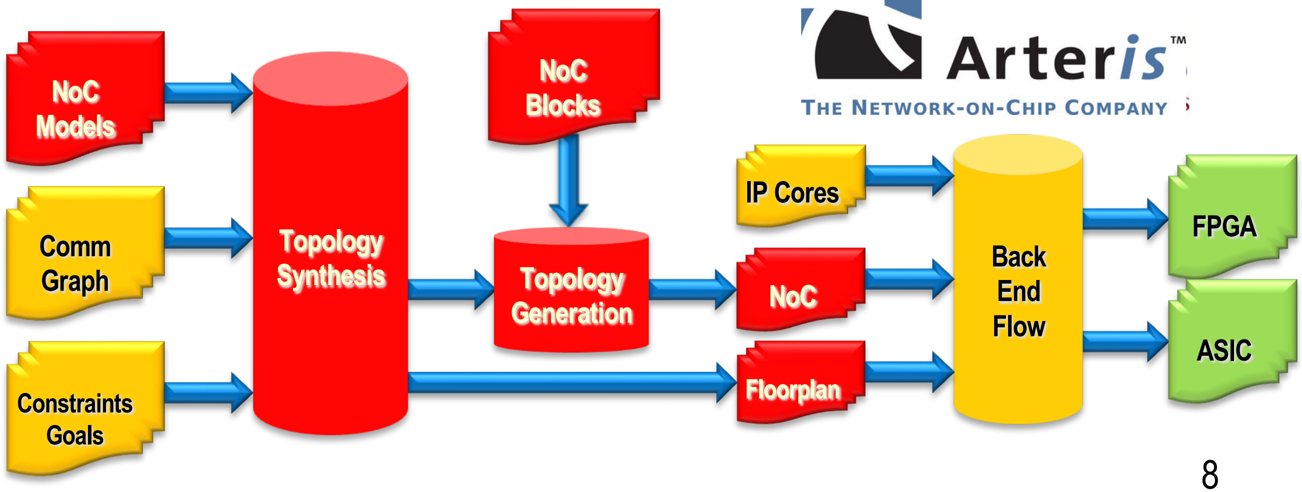
New communication structures

- **Design requirements:**
  - Predictable design
  - Fast design closure
- **Network on Chip communication**
  - Modular and flexible interconnect
  - Reliable on-chip communication
  - Structured design with synthesis and optimization support



[Vangal, Intel07]

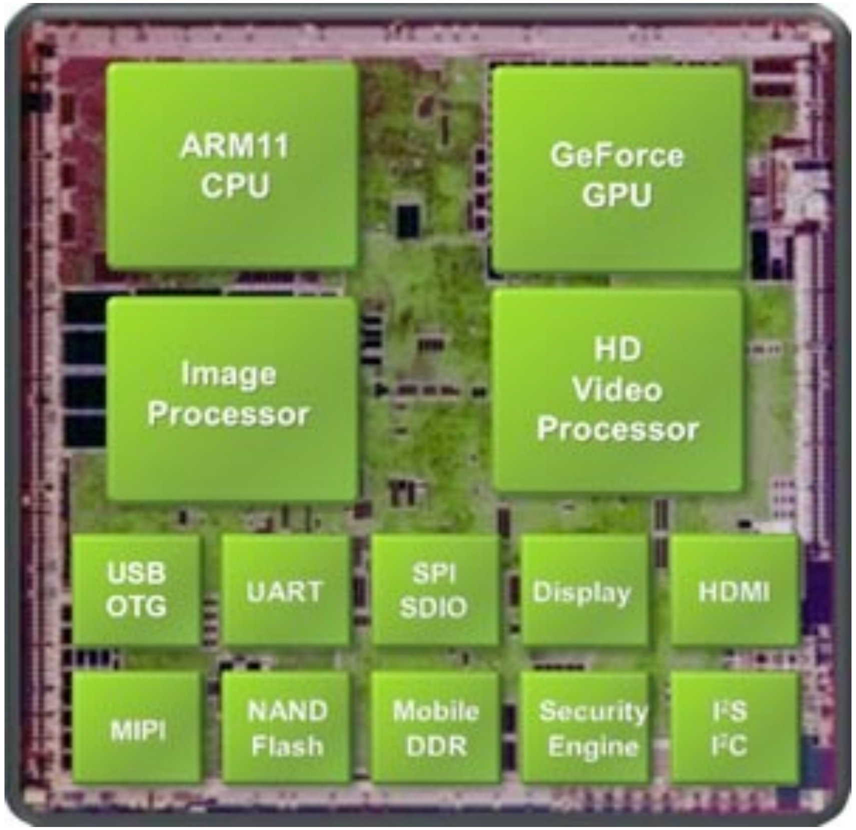




Arteris™
THE NETWORK-ON-CHIP COMPANY

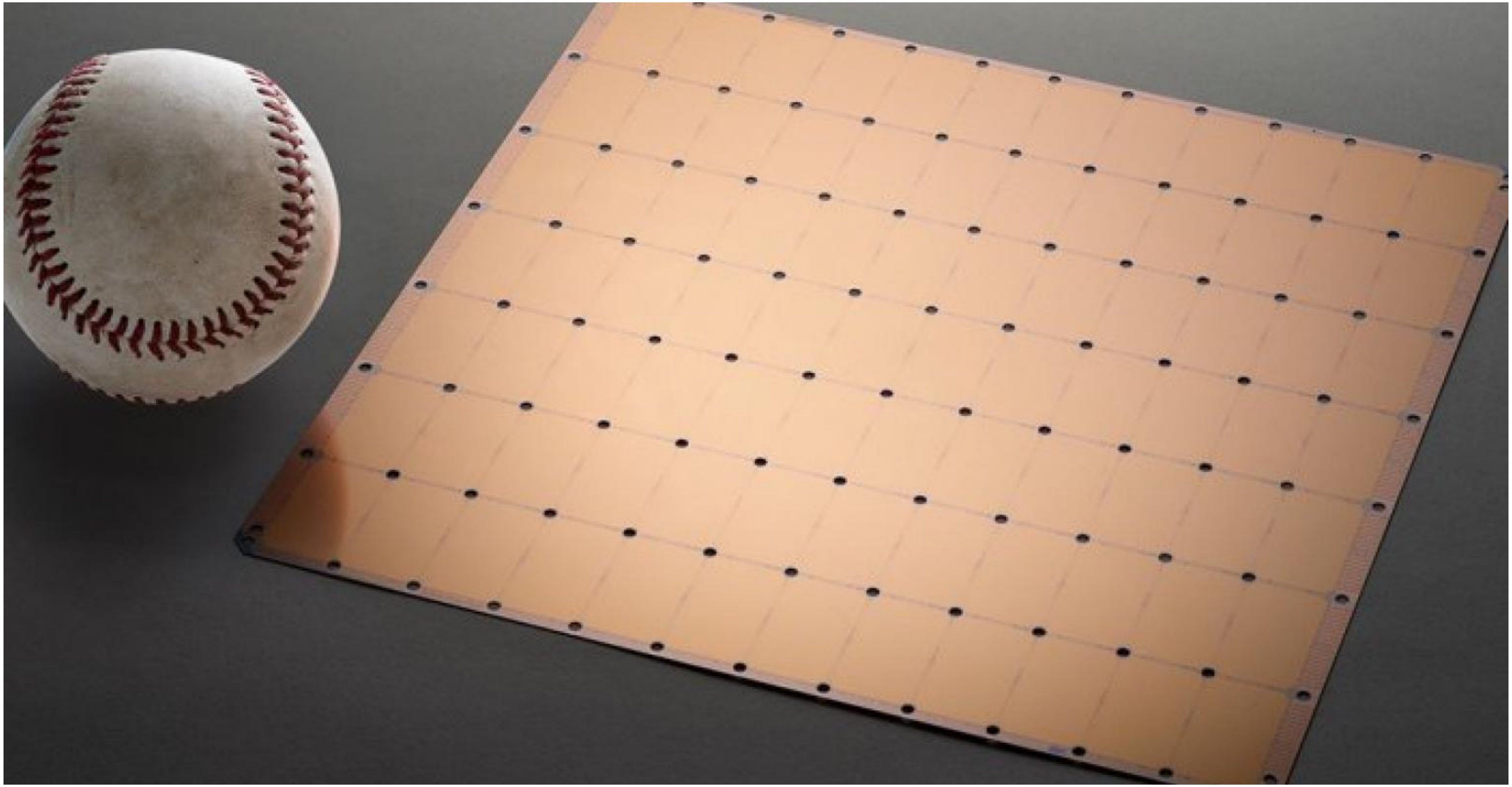
Systems on Chip

- Large-scale
  - Billion transistor chips
  - Multi-cores, multi-threaded SW
- Power-consumption limited
  - Dark silicon
- Very expensive to design
  - *Non recurring engineering* (NRE) costs
  - Migration toward software



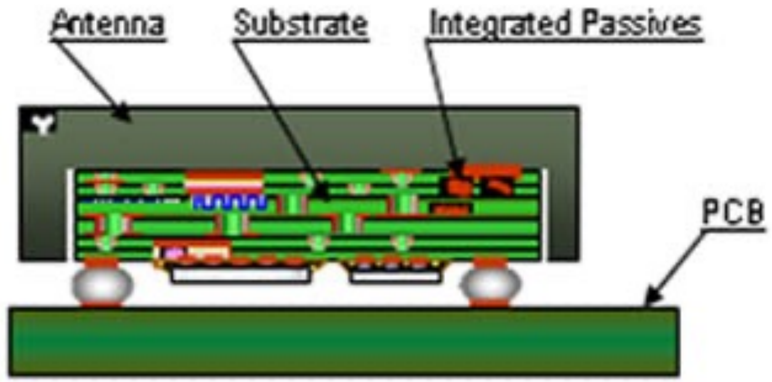
NVIDIA TEGRA Processor

Wafer scale integration (Cerebras)



Systems in Package

- There is a diminishing return in integrating everything on Silicon
  - Heterogeneous technologies
  - Multiple voltages
  - Thermal issues
- Packaging technology
- Chiplets

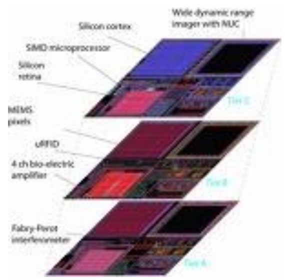


Schematic diagram of RF-SIP

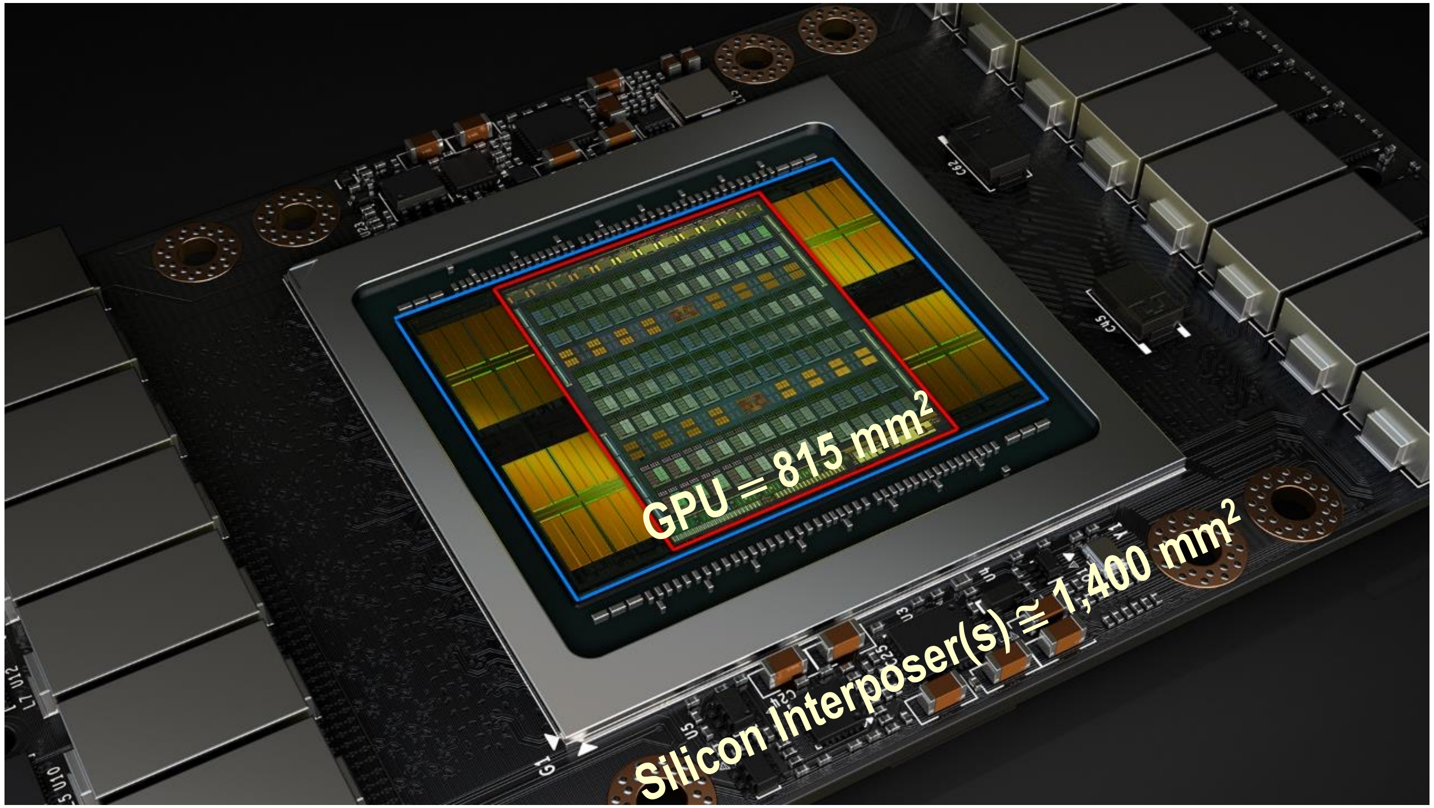
New packaging technology

- **From planar to 3D integration**
  - **Chips have limited wiring resources**
  - **Electrical and manufacturing constraints limit heterogeneous planar integration**
- ***Through silicon vias* allow designer to stack:**
  - **Computing arrays**
  - **Memory arrays**
  - **Analog and RF circuitry**

[Fraunhofer]



NVidia 2.5D-IC High BW Memory Application



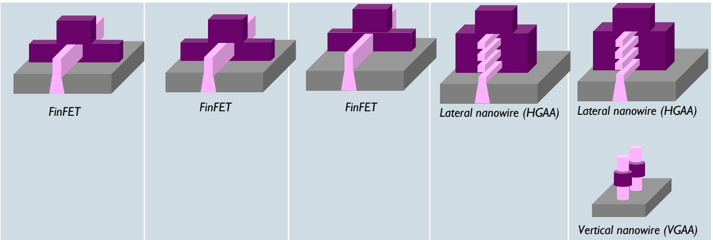
GPU Die (21B Transistors @ 12 Nanometers TSMC, 300W) + 4 HBM2 Stacks (4 × 4GB) Onto Several Silicon Interposer Die Stitched Together, price at launch $10,000

"Volta" 5.5D-IC (3D + 2.5D) Integration; Source: J.-H. Huang, et al., NVIDIA, GPU Technology Conference 2017

Credit: Stanford EE292A Lecture 1

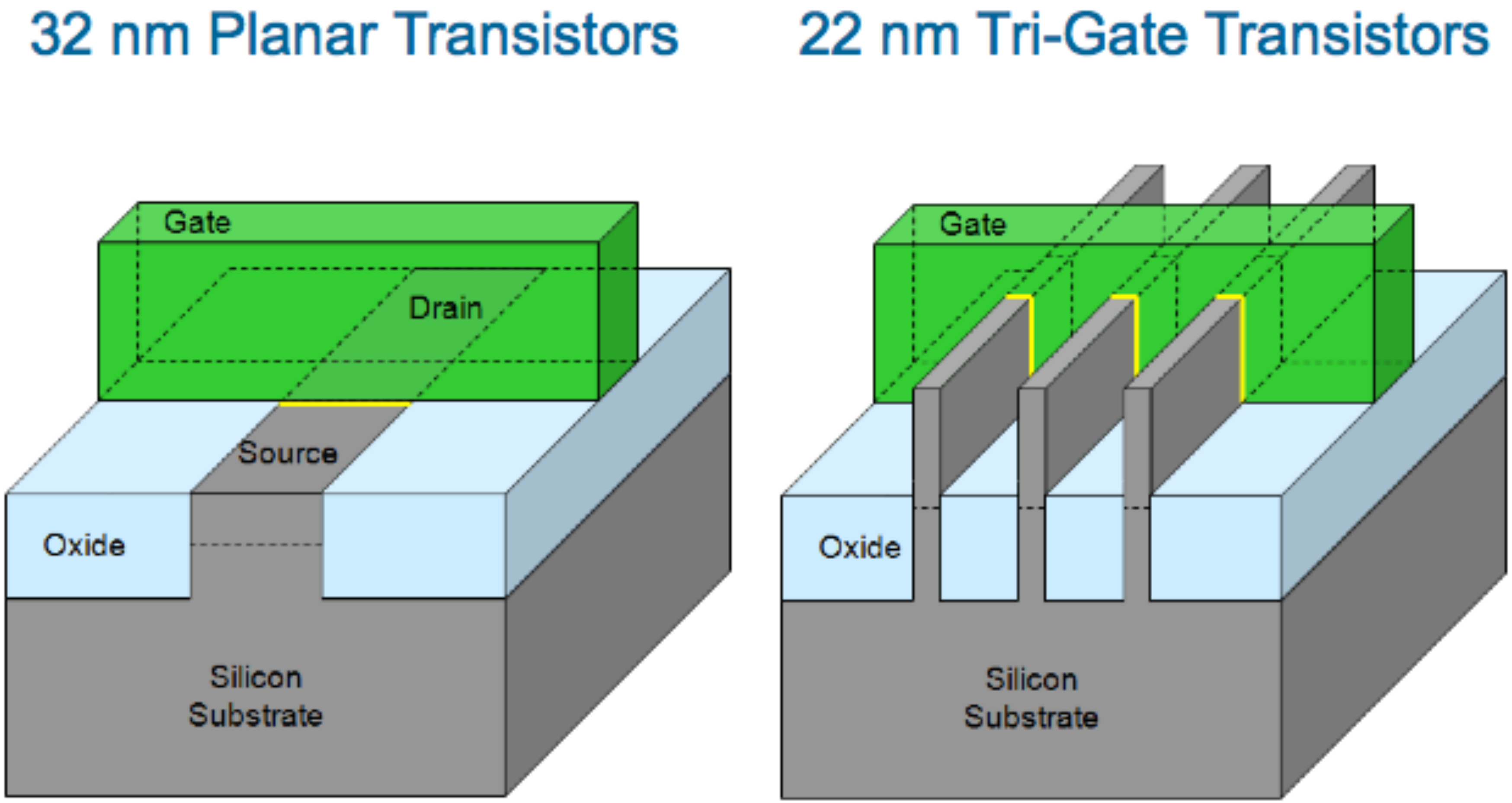
The emerging nano-technologies

- *Enhanced* silicon CMOS is likely to remain the main manufacturing process in the medium term
  - The 3nm technology node is here (Since 2023)
- What are the candidate techs for the 3nm node and beyond?
  - Silicon Nanowires/sheets (SiNW))
  - Carbon Nanotubes (CNT)
  - 2D devices (flatronics)
- What is in common from a design standpoint?



Early production	2014 iN14	2016 iN10	2017-2018 iN7	2018-2019 iN5	>2020 iN3
Vdd (V)	0.8	0.8-0.7	0.7-0.6	0.7-0.5	0.6-0.5
Gate Pitch (nm)	70-90, 193i	52-64, 193i	36-46, 193i	26-36, EUV, 193i	18-28, EUV, 193i
Device	FinFET	FinFET	FinFET {HGAA}	HGAA	HGAA {VGAA}
Channel nfet/pfet	Si / Si	Si / Si {SiGe}	Si / SiGe	Si/ SiGe	High mobility

22 nm Tri-Gate Transistors



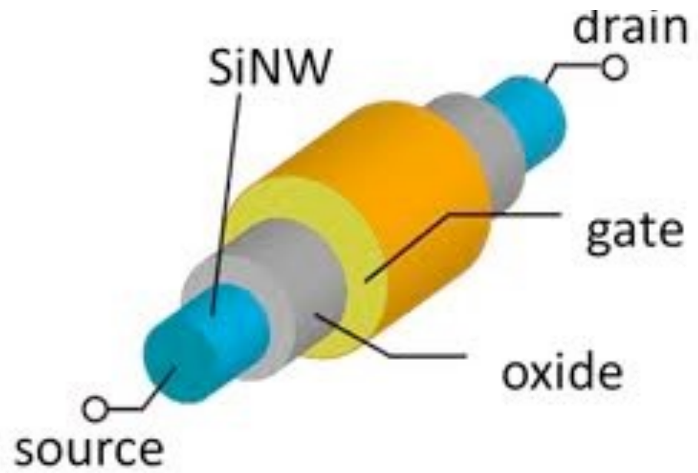
[Courtesy: M. Bohr]

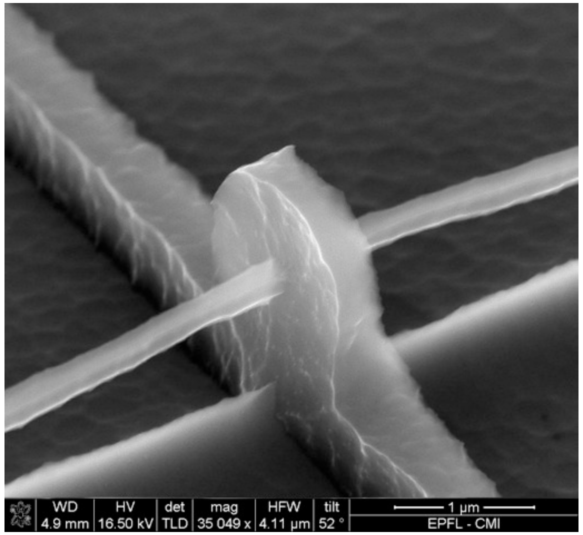
Fully Depleted SoI Transistors

[Courtesy: STMicroelectronics]

- Transistor is built on top of buried oxide (BOX)
- Thin, undoped channel (fully depleted)
- Fine power-consumption control through body bias

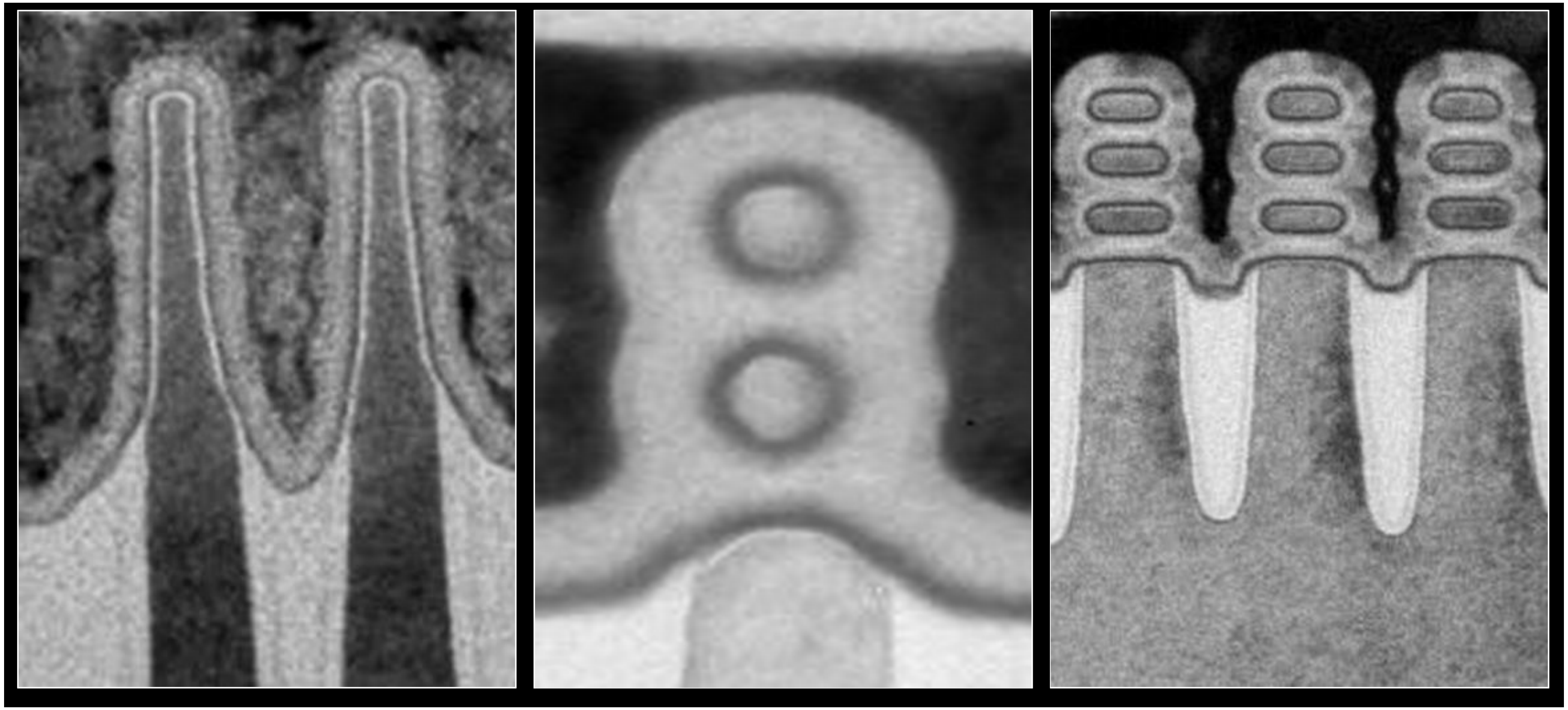
Silicon Nanowire Transistor





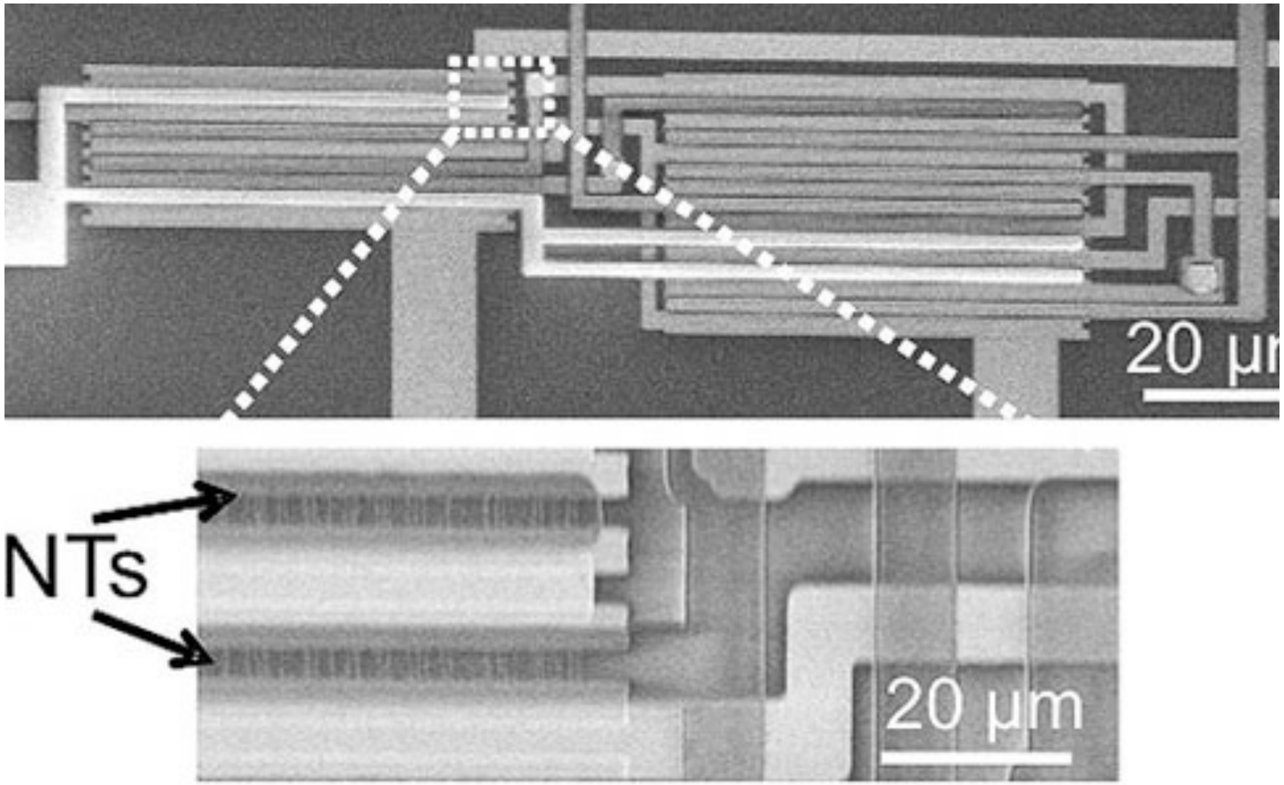
- Fully compatible with CMOS process
- Gate all around
- High I_{on} / I_{off} ratio

FINFETs, NanoWires and NanoSheets



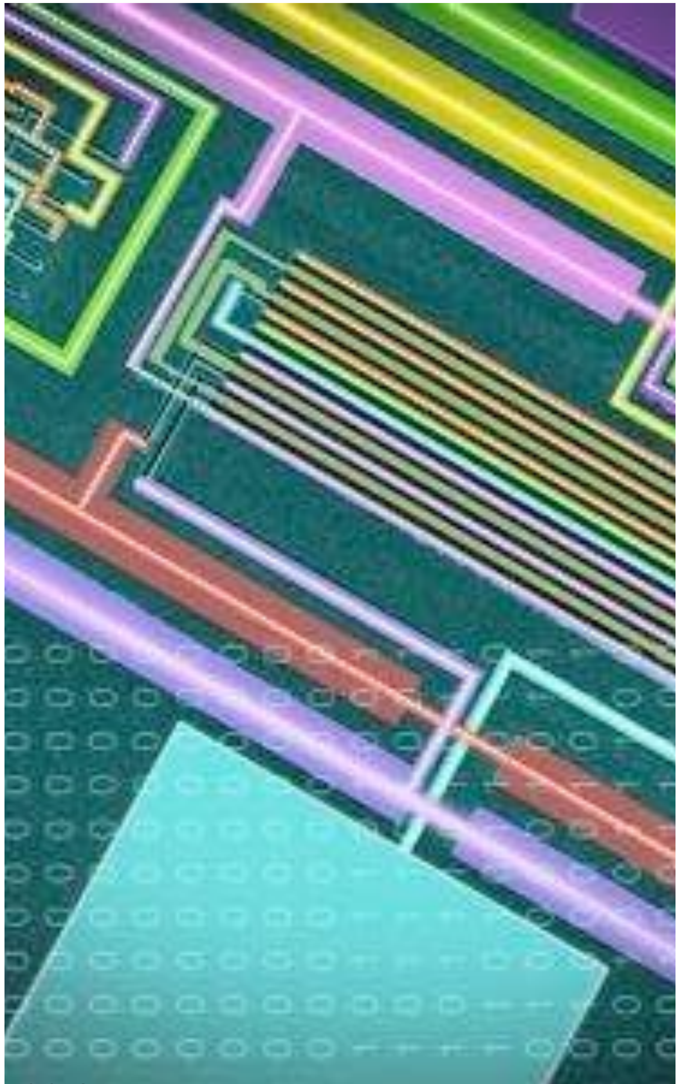
[INTEL, 2017]

Carbon Nanotube Transistors

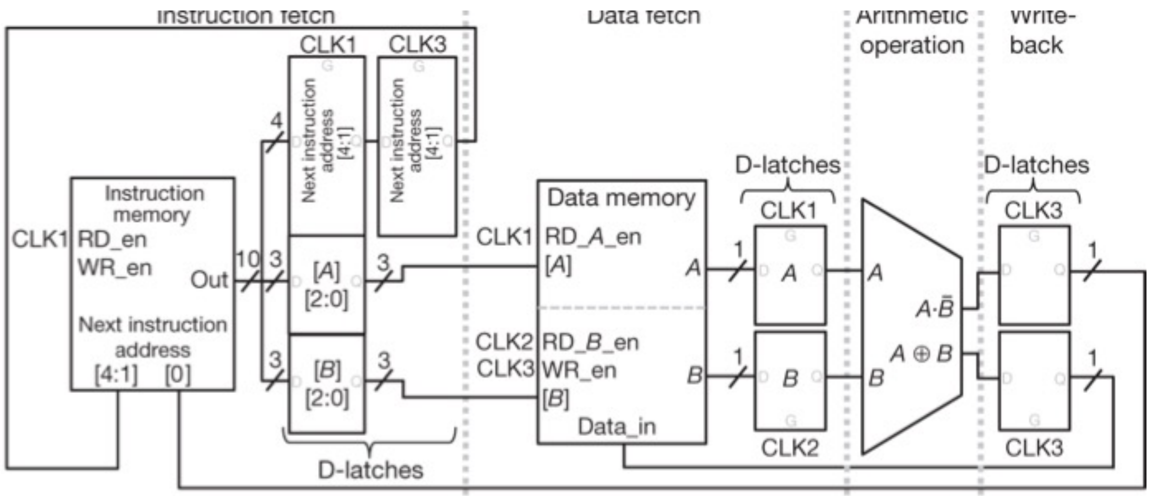


- CNTs benefit from higher mobility and thus higher currents
- CNTs grown separately but can be ported to Si wafers
- Handling CNT imperfections is major design and fabrication issue

CNT nanocomputer

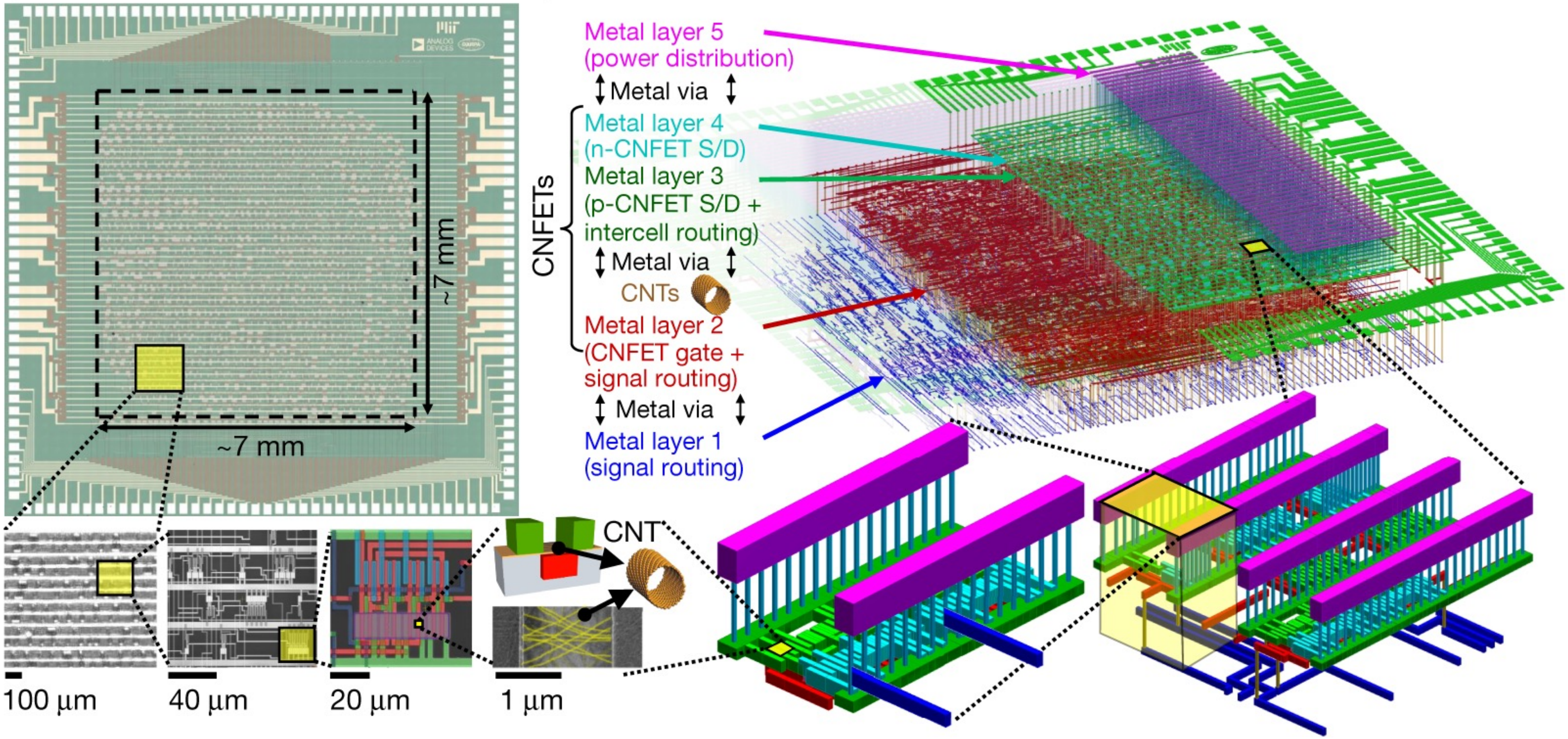


- First CNT computing engine
- Runs 20 MIPS instructions
- Multitasking



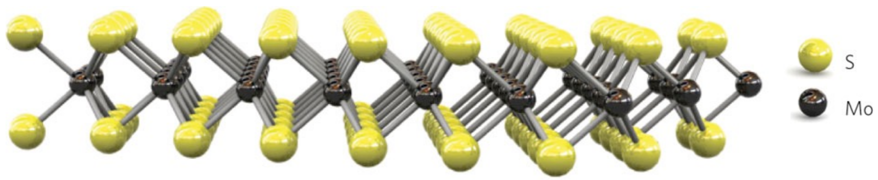
[Shulaker, Wong, Mitra et al, NatureNano 13]

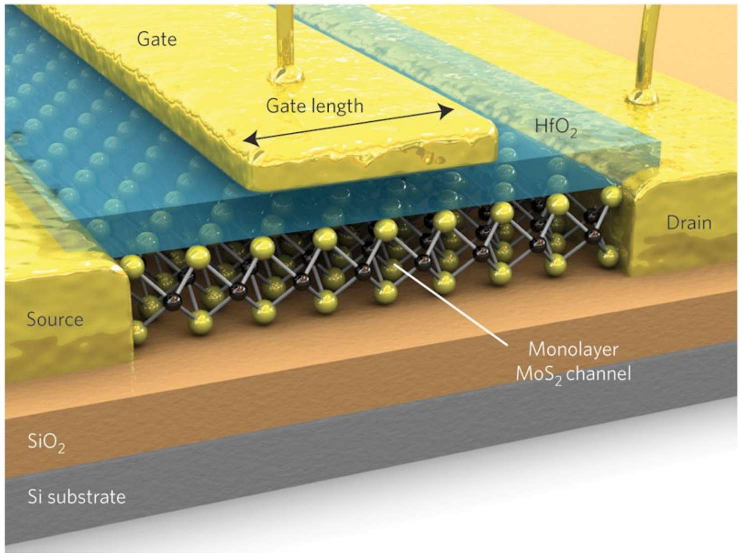
16B RISC-V 14'000 CNFET transistor chip



[Shulaker et al., Nature 19]

2D electronic technologies



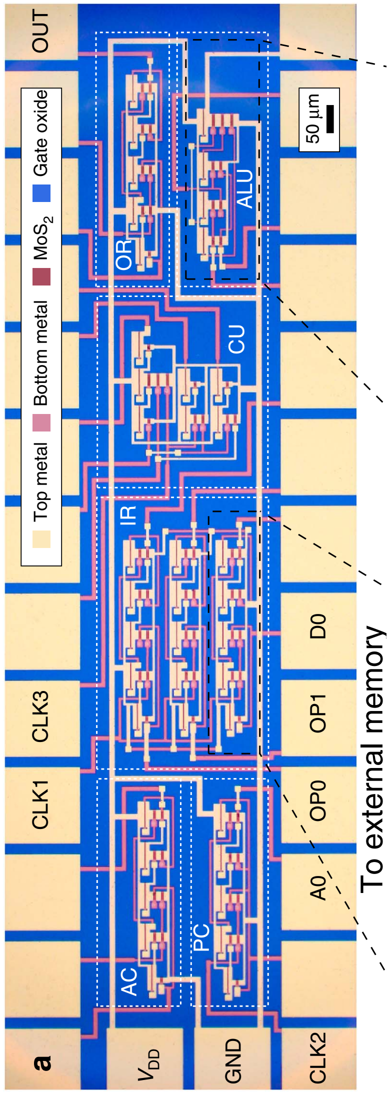


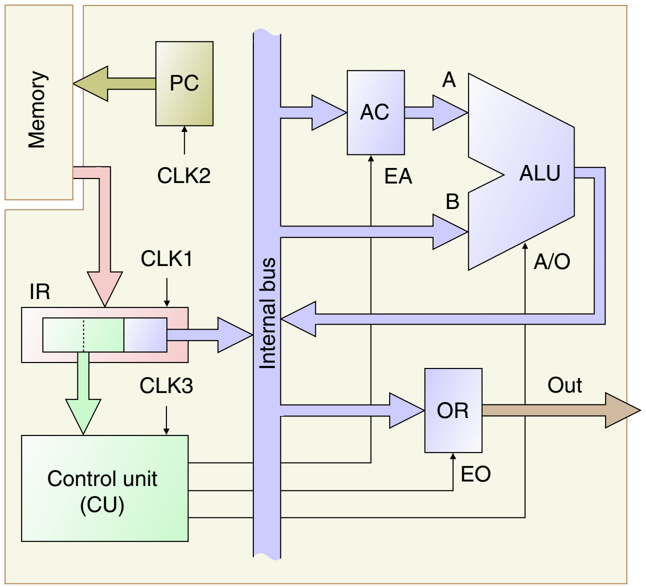
[Kis, Nature Nano 2011]

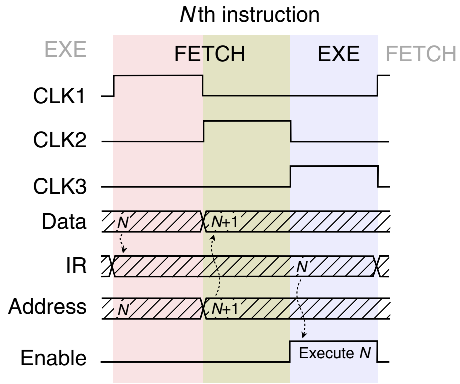
- Graphene, MoS_2 and other materials
- Single or few atomic layers
- High I_{on} / I_{off} ratio for MoS_2 (10^8) but n-type mainly

MoS_2 nanocomputer

- First MoS_2 computing engine
- Runs 4instructions
- 115 N-xtors (enhancement load)
- 2 micron feature size







c

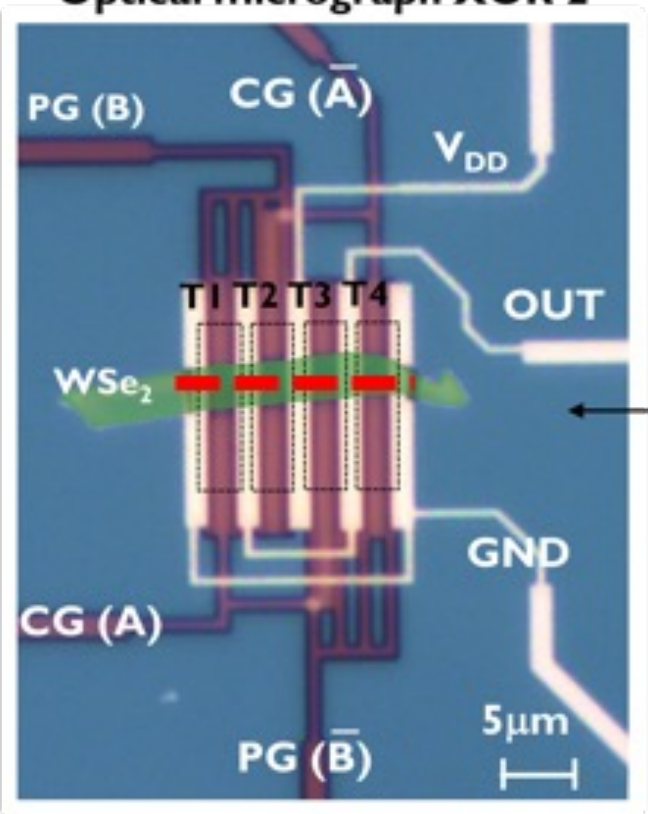
Command	Code	Example
NOP	00	NOP
LDA	01	LDA 0
AND	10	AND 1
OR	11	OR 0

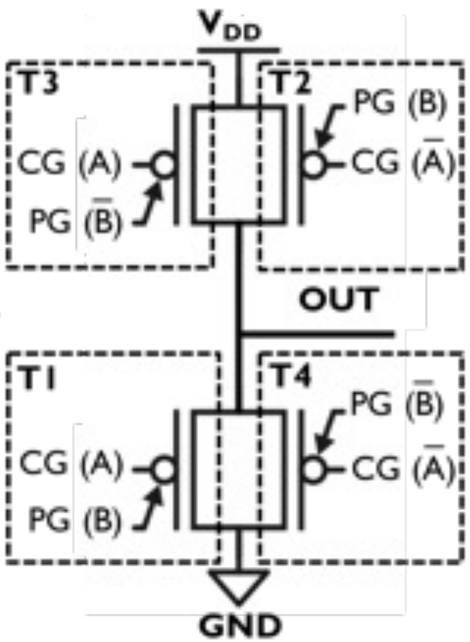
[Wacter et al. Nature Comm, 2017]

Controllable polarity transistors in 2D

2D Controllable-polarity EXOX in WSe_2

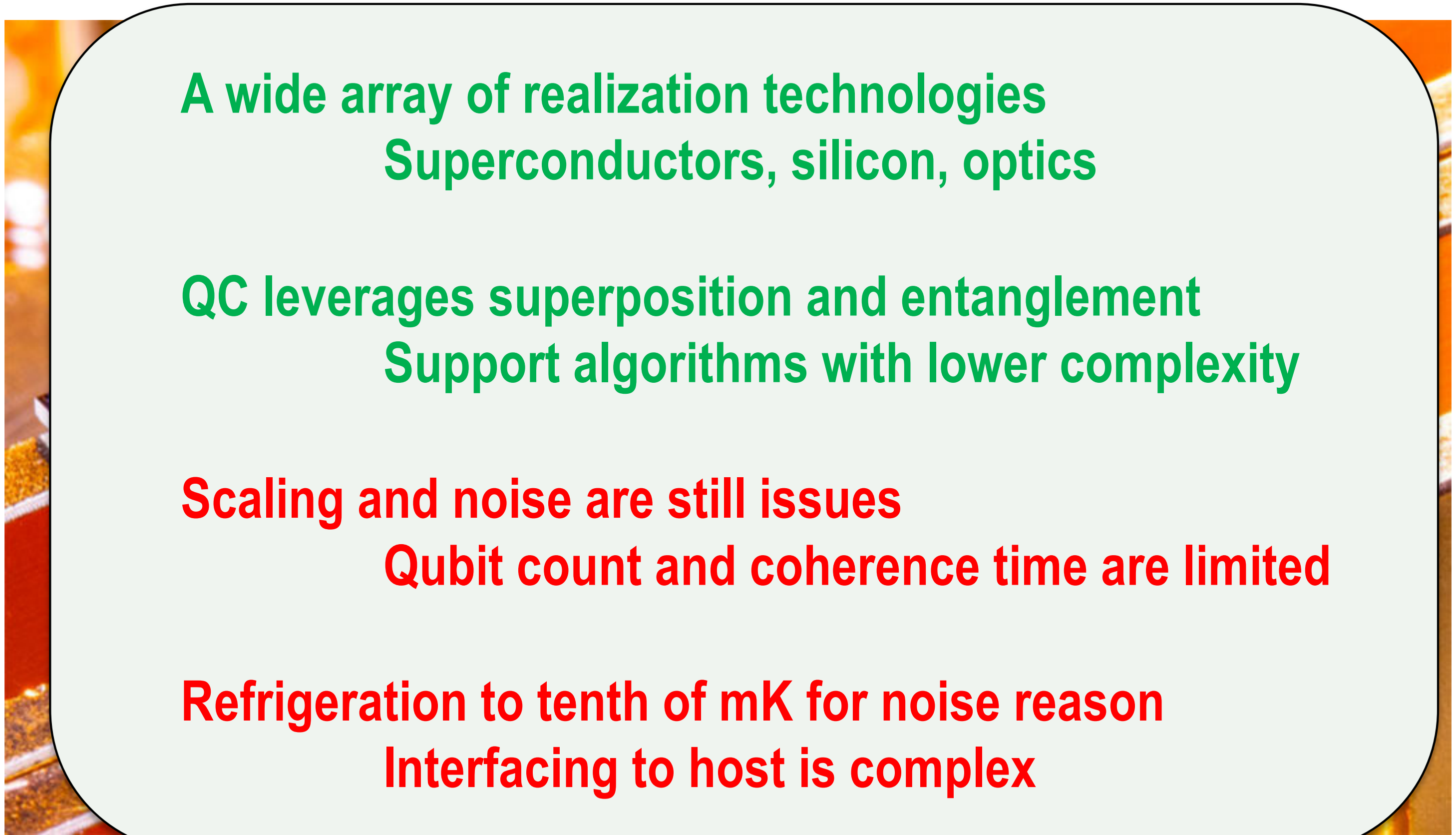
Optical micrograph XOR 2





[Resta, ACS Nano 2017]

Quantum computing technologies



A wide array of realization technologies

Superconductors, silicon, optics

QC leverages superposition and entanglement

Support algorithms with lower complexity

Scaling and noise are still issues

Qubit count and coherence time are limited

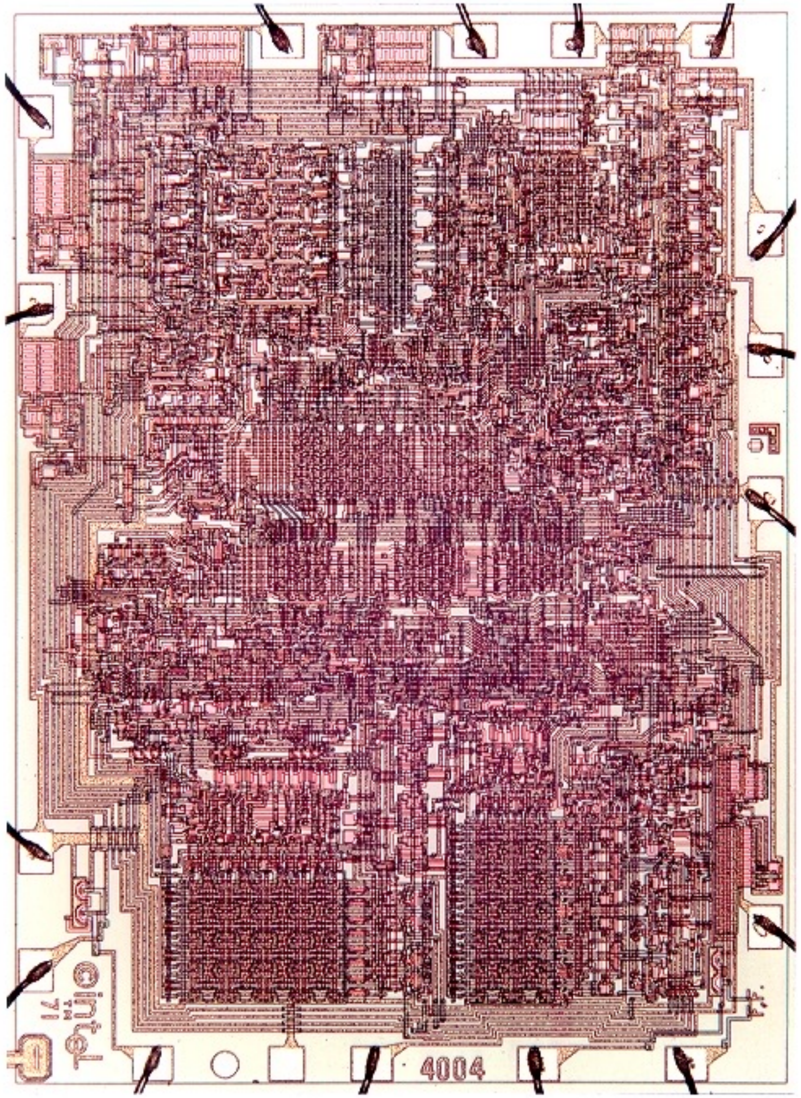
Refrigeration to tenth of mK for noise reason

Interfacing to host is complex

Integrated Circuit Design Styles

- **Custom:**
  - All geometries are designed *ad hoc*
  - Mainly abandoned, but useful for high-performance components
- **Semicustom:**
  - Optimality is traded-off for regularity
  - Performance penalty is small, cost is lower

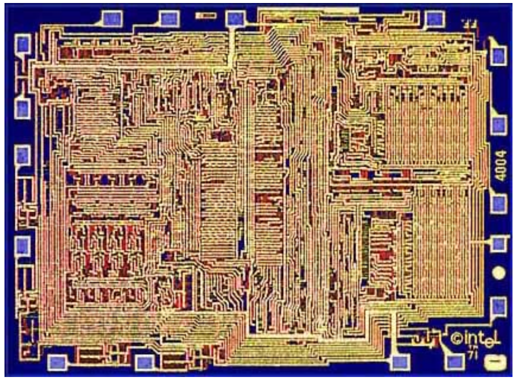
The Custom Approach



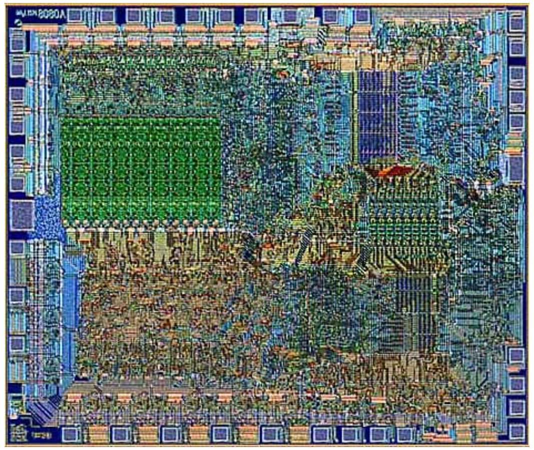
Intel 4004

Courtesy Intel

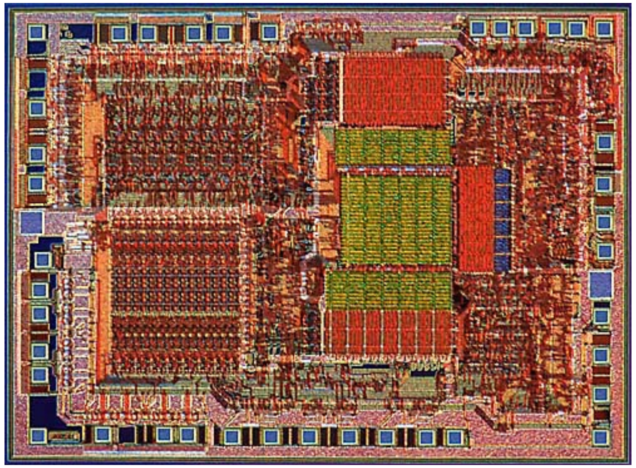
Transition to Automation and Regular Structures



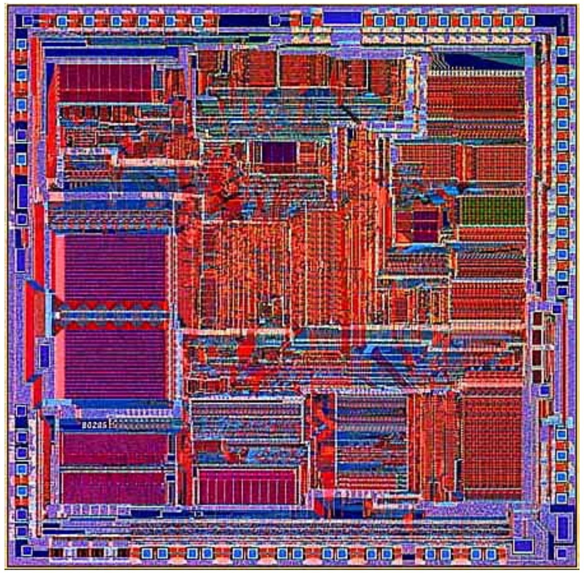
Intel 4004 (‘71)



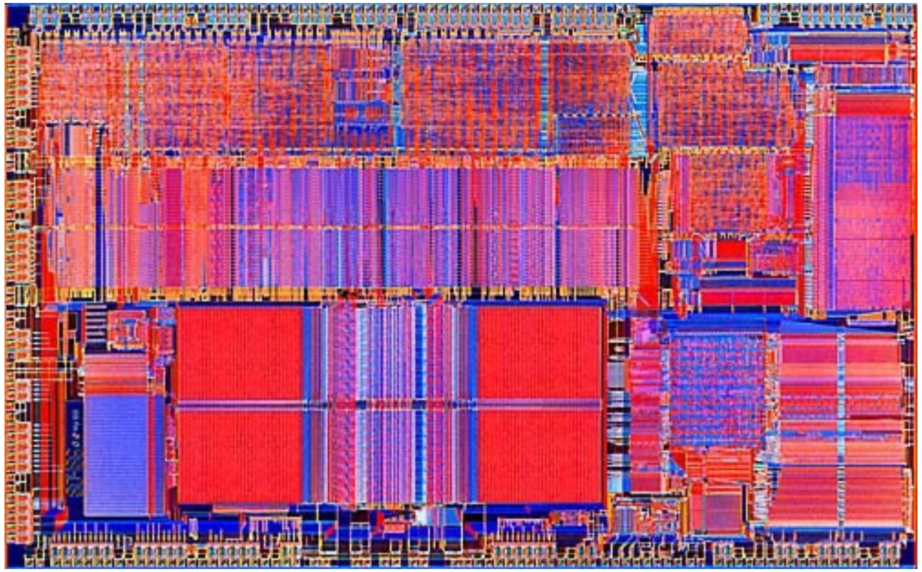
Intel 8080



Intel 8085



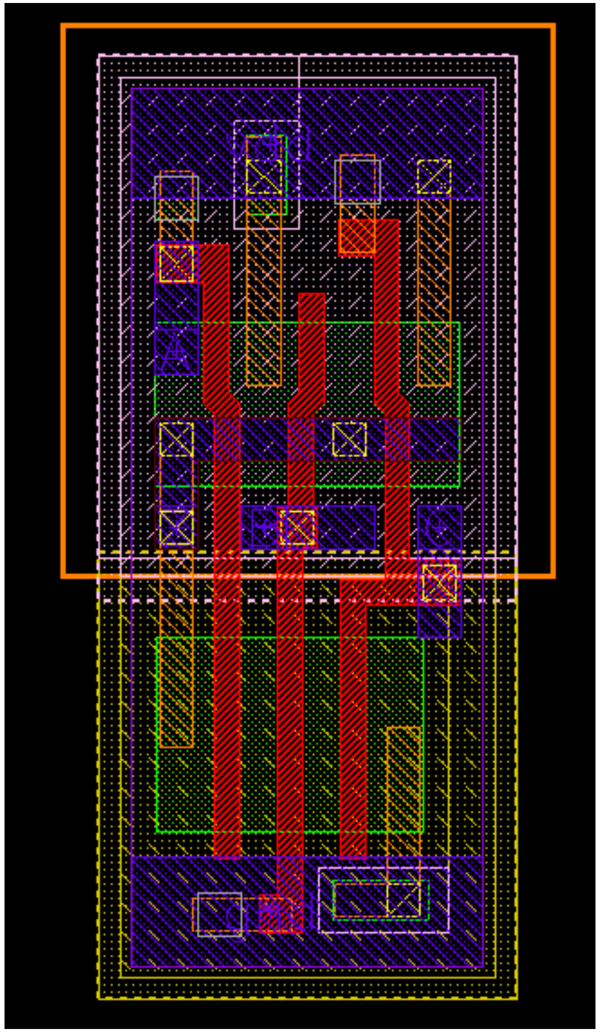
Intel 8286



Intel 8486

Courtesy Intel

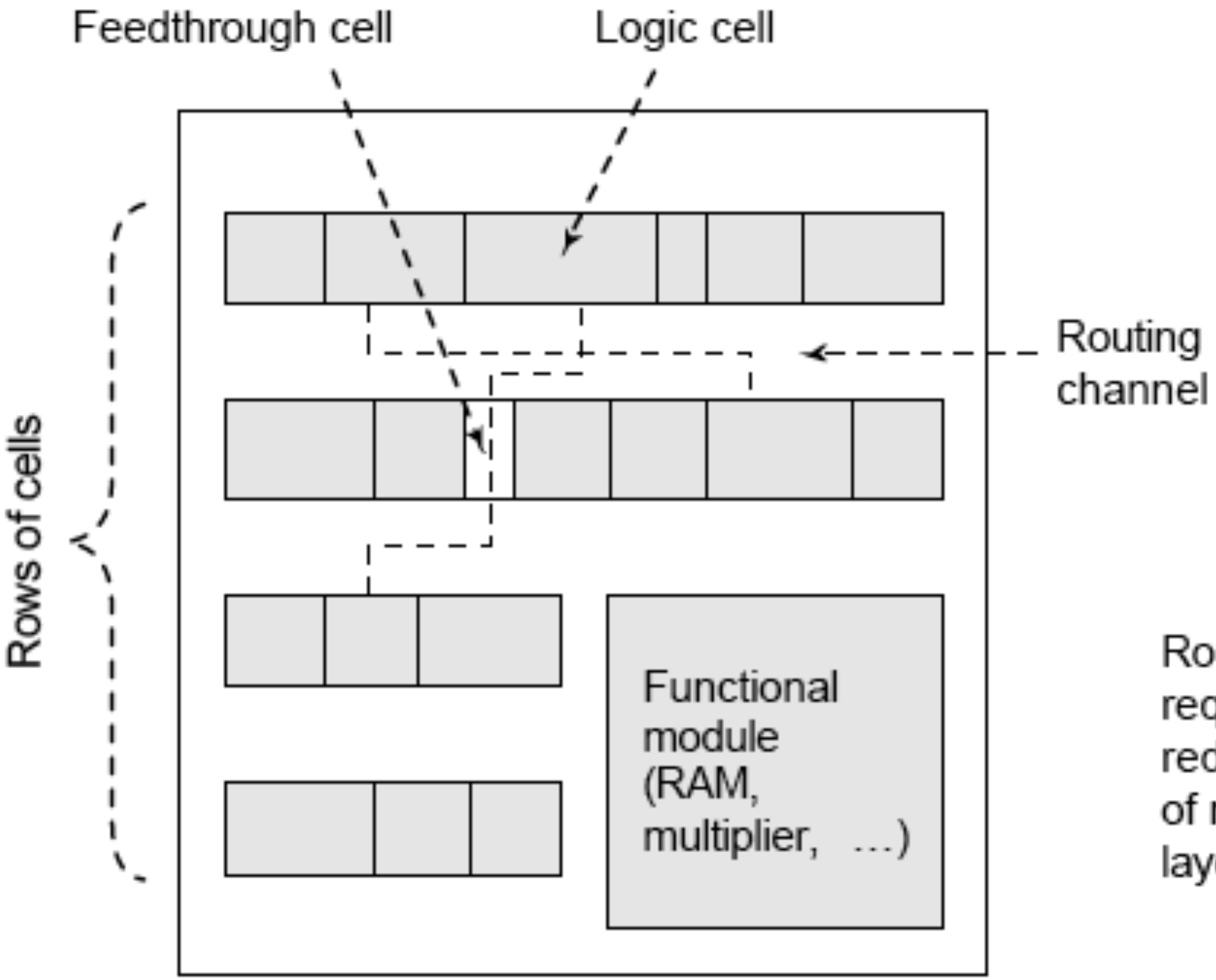
Standard Cell - Example



Path	1.2V - 125°C	1.6V - 40°C
$In1—t_{pLH}$	$0.073+7.98C+0.317T$	$0.020+2.73C+0.253T$
$In1—t_{pHL}$	$0.069+8.43C+0.364T$	$0.018+2.14C+0.292T$
$In2—t_{pLH}$	$0.101+7.97C+0.318T$	$0.026+2.38C+0.255T$
$In2—t_{pHL}$	$0.097+8.42C+0.325T$	$0.023+2.14C+0.269T$
$In3—t_{pLH}$	$0.120+8.00C+0.318T$	$0.031+2.37C+0.258T$
$In3—t_{pHL}$	$0.110+8.41C+0.280T$	$0.027+2.15C+0.223T$

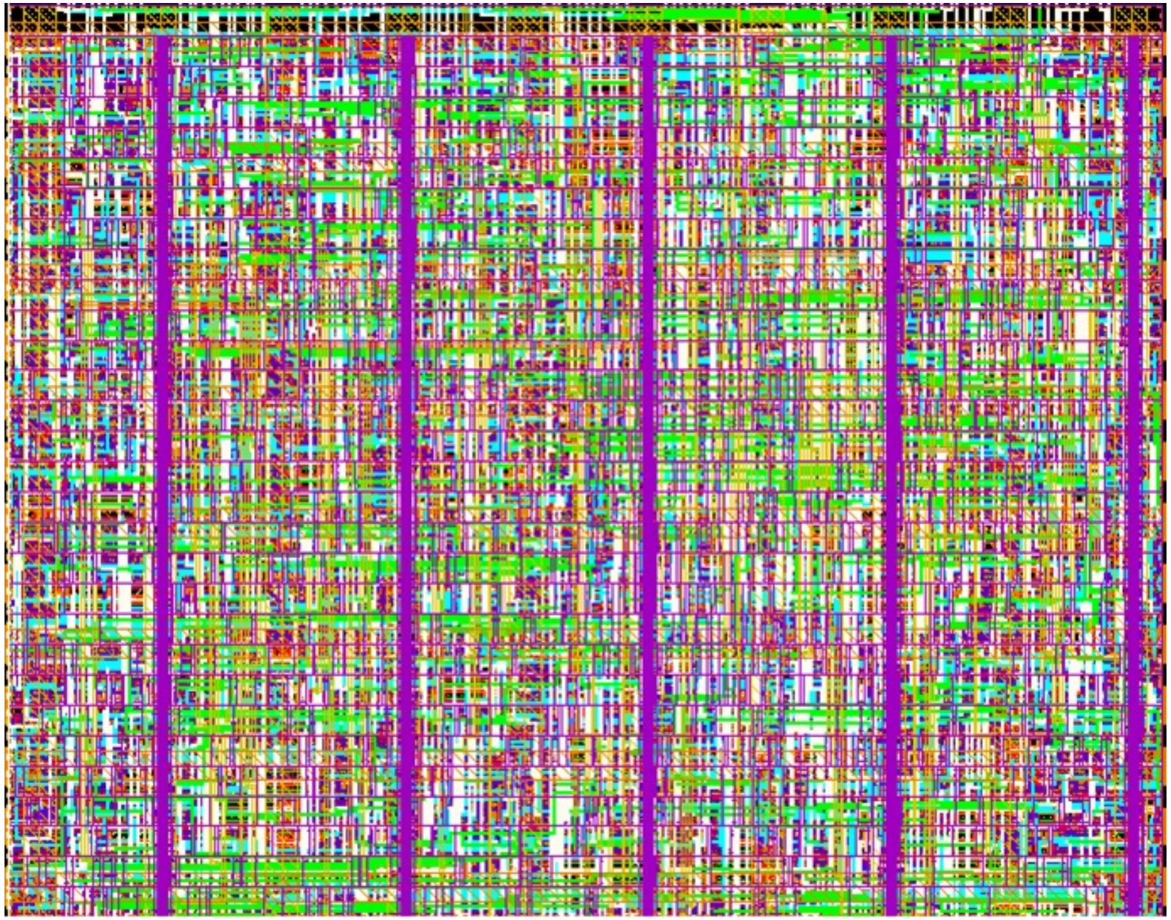
3-input NAND cell
(from ST Microelectronics):
C = Load capacitance
T = input rise/fall time

Cell-based Design (or standard cells)



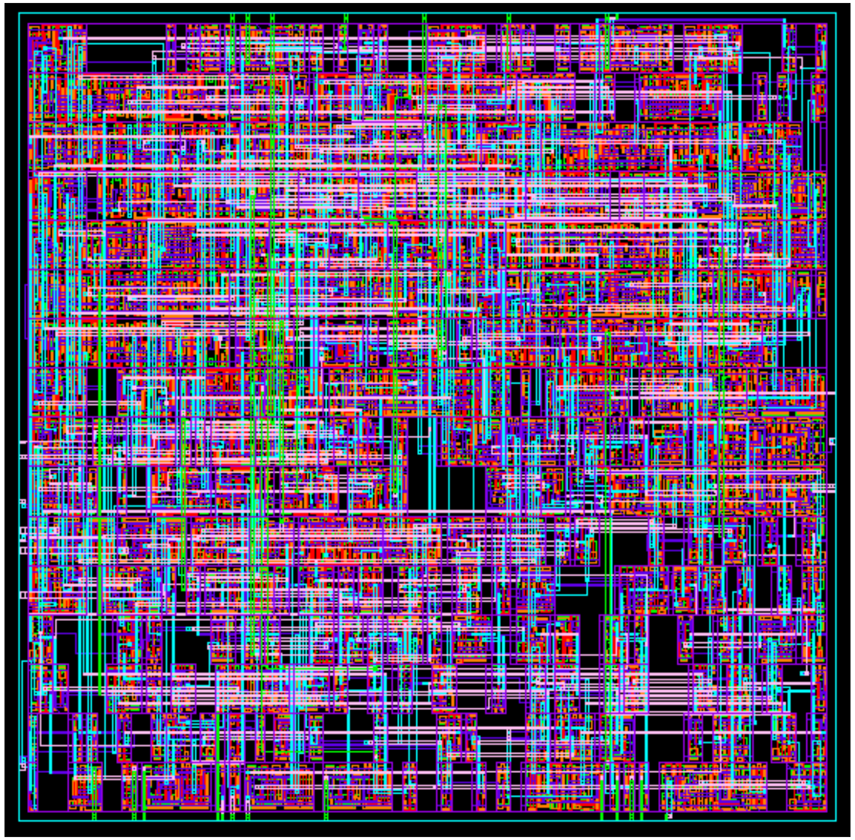
Routing channel requirements are reduced by presence of more interconnect layers

Standard Cell – The New Generation

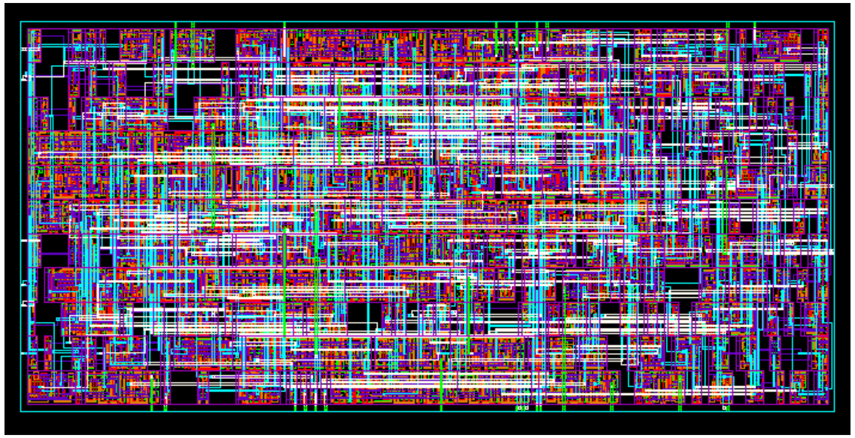


Cell-structure hidden under interconnect layers

"Soft" MacroModules

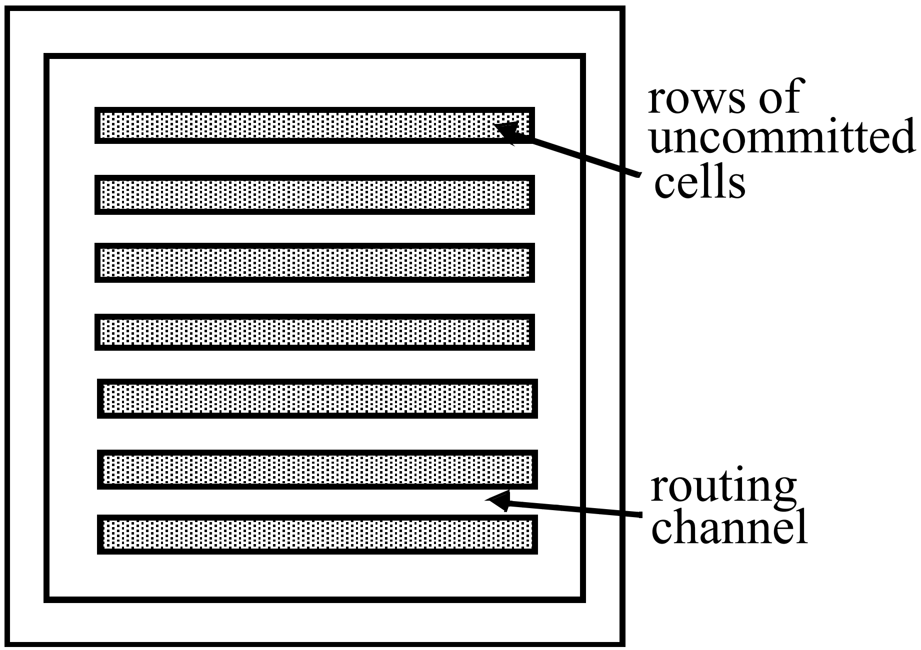


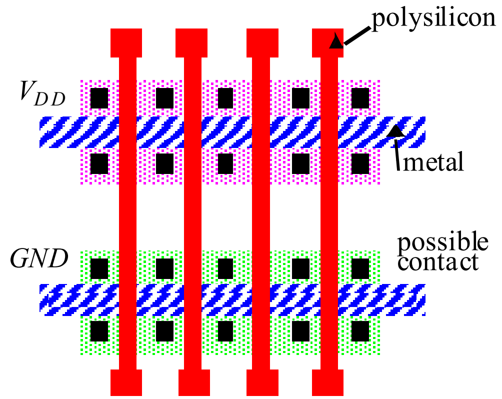
```
string mat = "booth";
directive (multtype = mat);
output signed [16] Z = A * B;
```



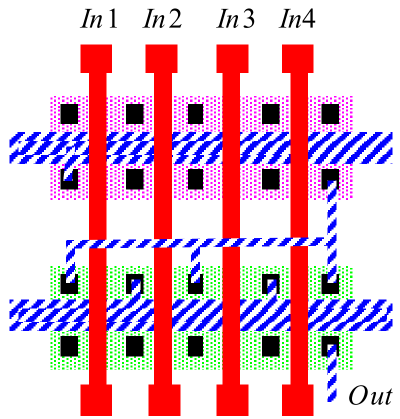
Synopsys DesignCompiler

Gate Array — Sea-of-gates



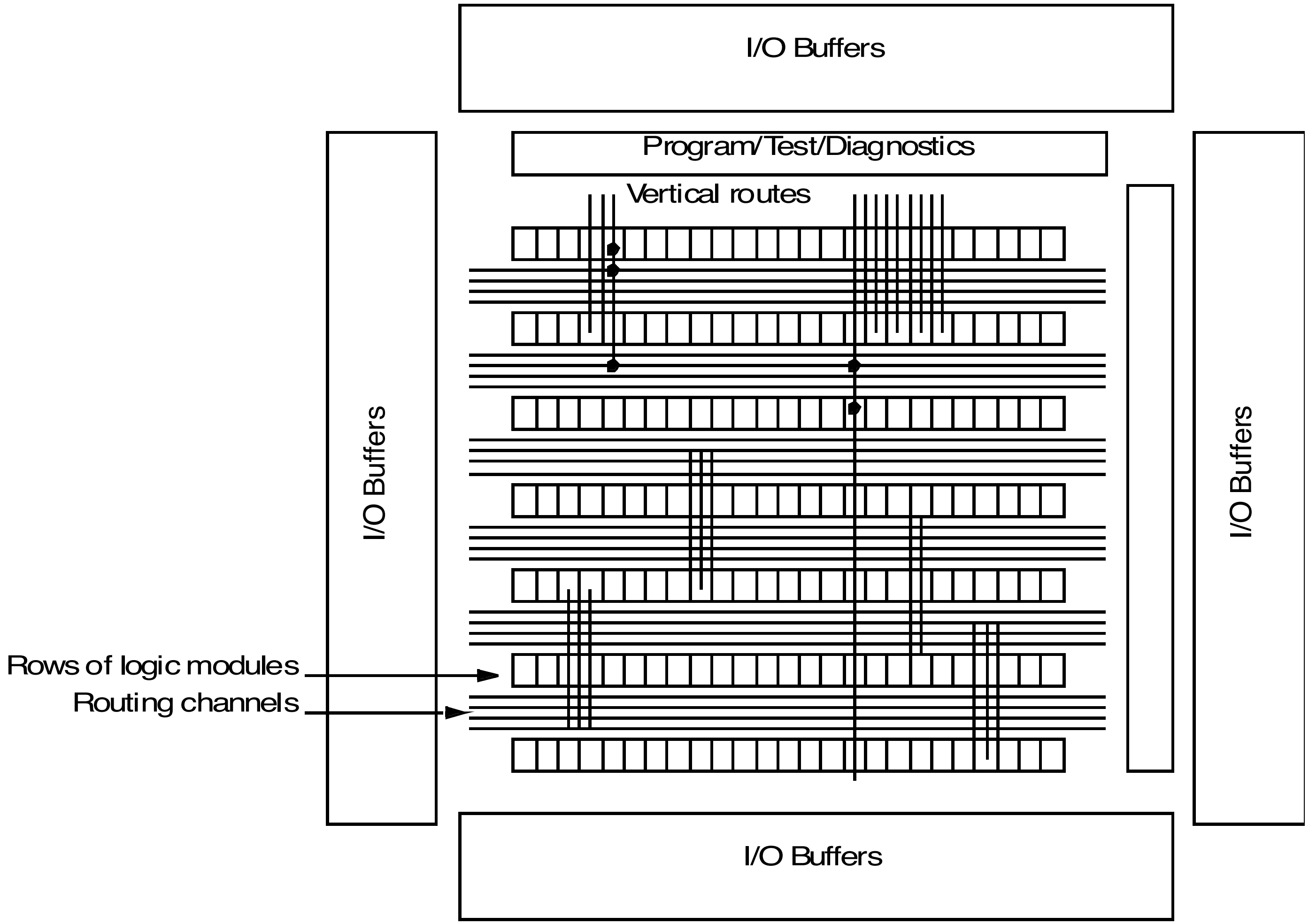


Uncommited Cell



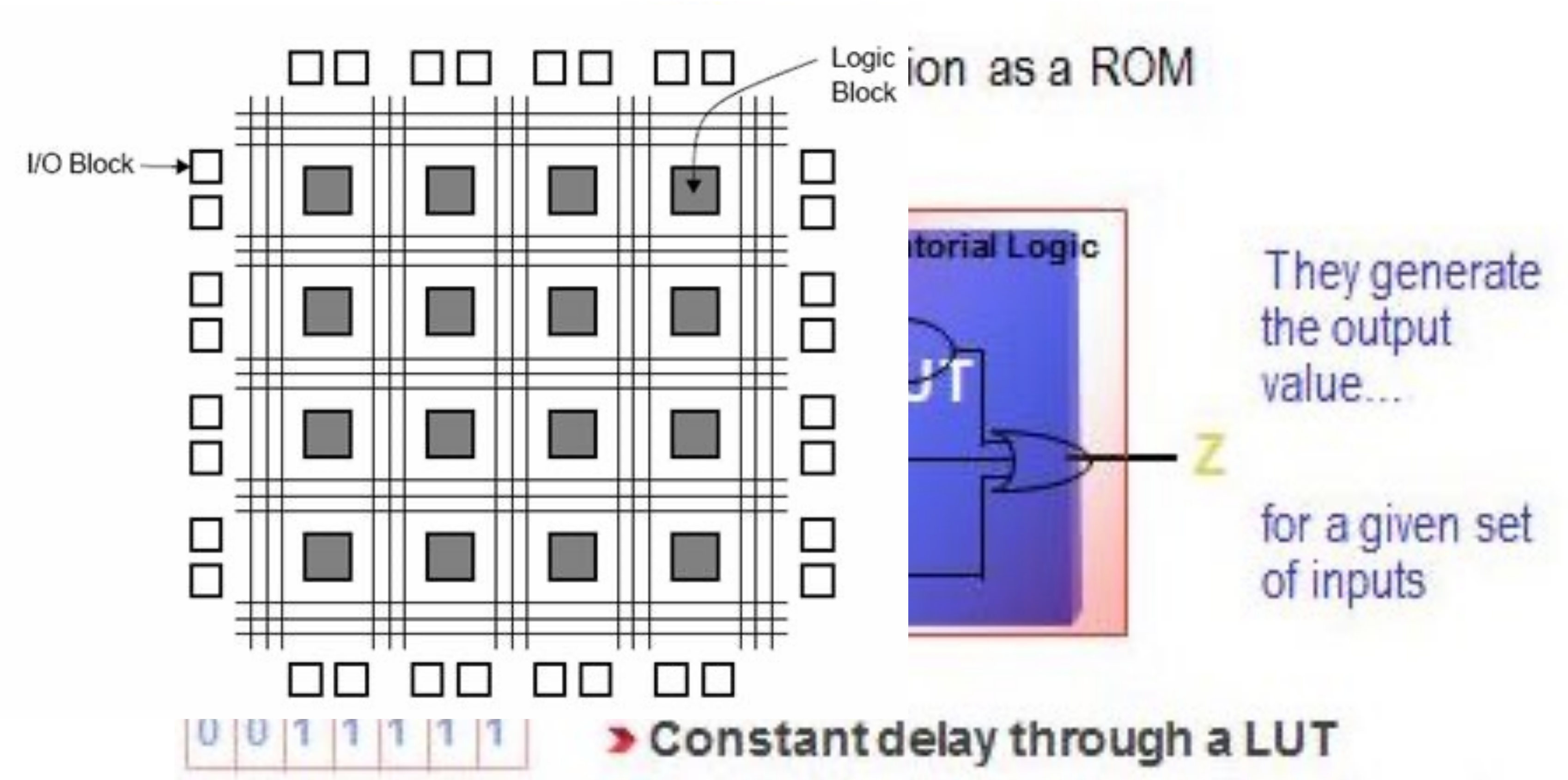
Committed Cell (4-input NOR)

Field-Programmable Gate Arrays
Fuse-based



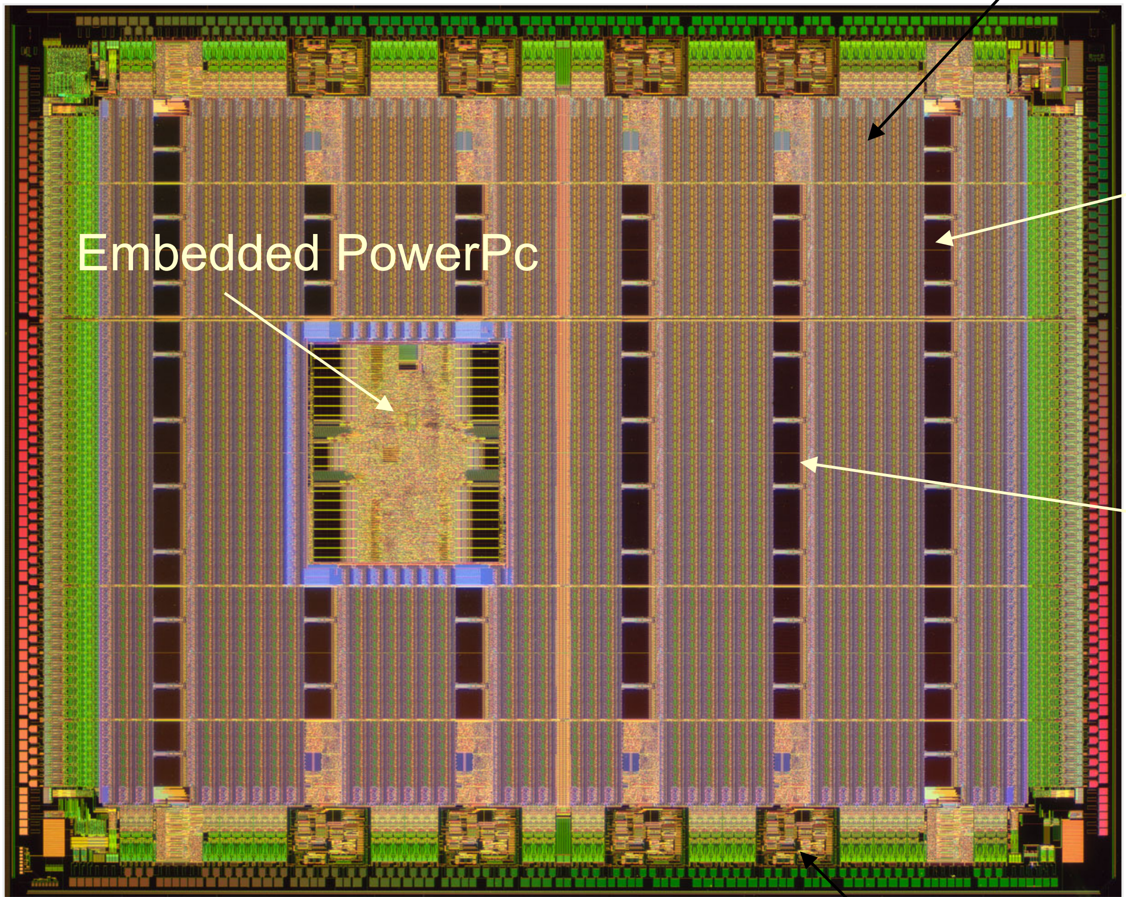
Standard-cell like floorplan

Field Programmable Gate Arrays
Xilinx LUT Architecture



Courtesy Xilinx

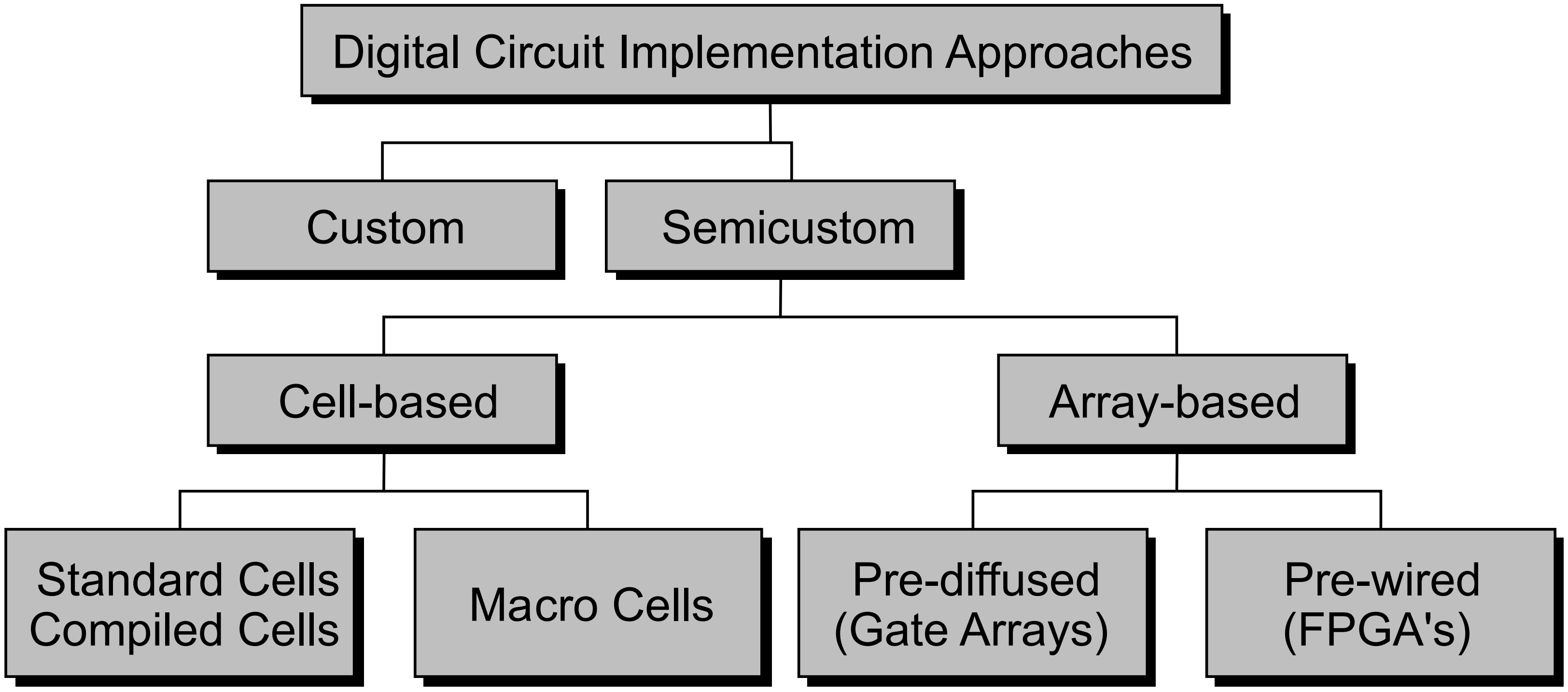
Heterogeneous Programmable Platforms



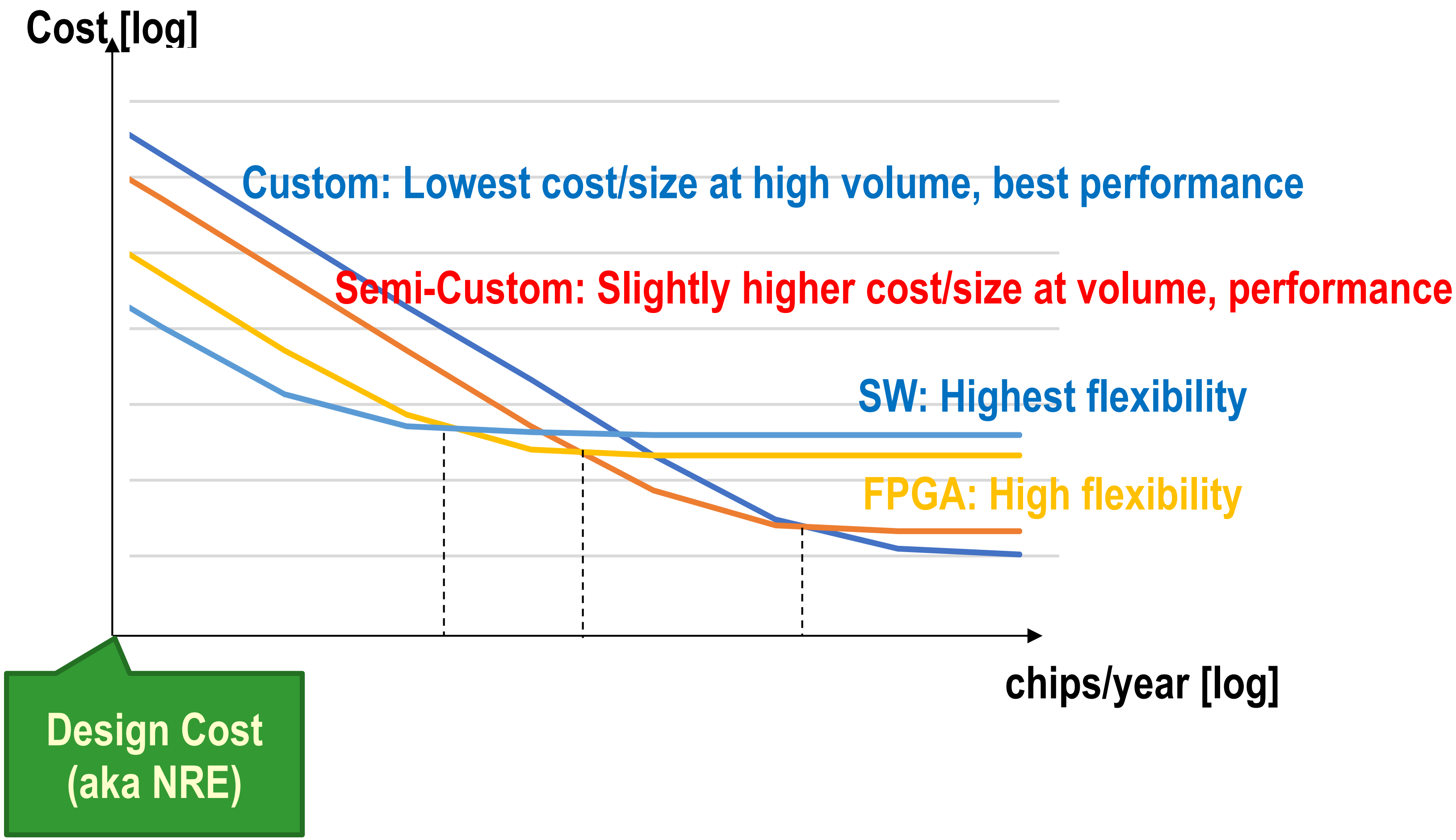
Xilinx Virtex-II Pro

Courtesy Xilinx

Integrated Circuit Design Styles



The Basic Trade-Offs



Credit: Stanford EE292A Lecture 1

Moore's law

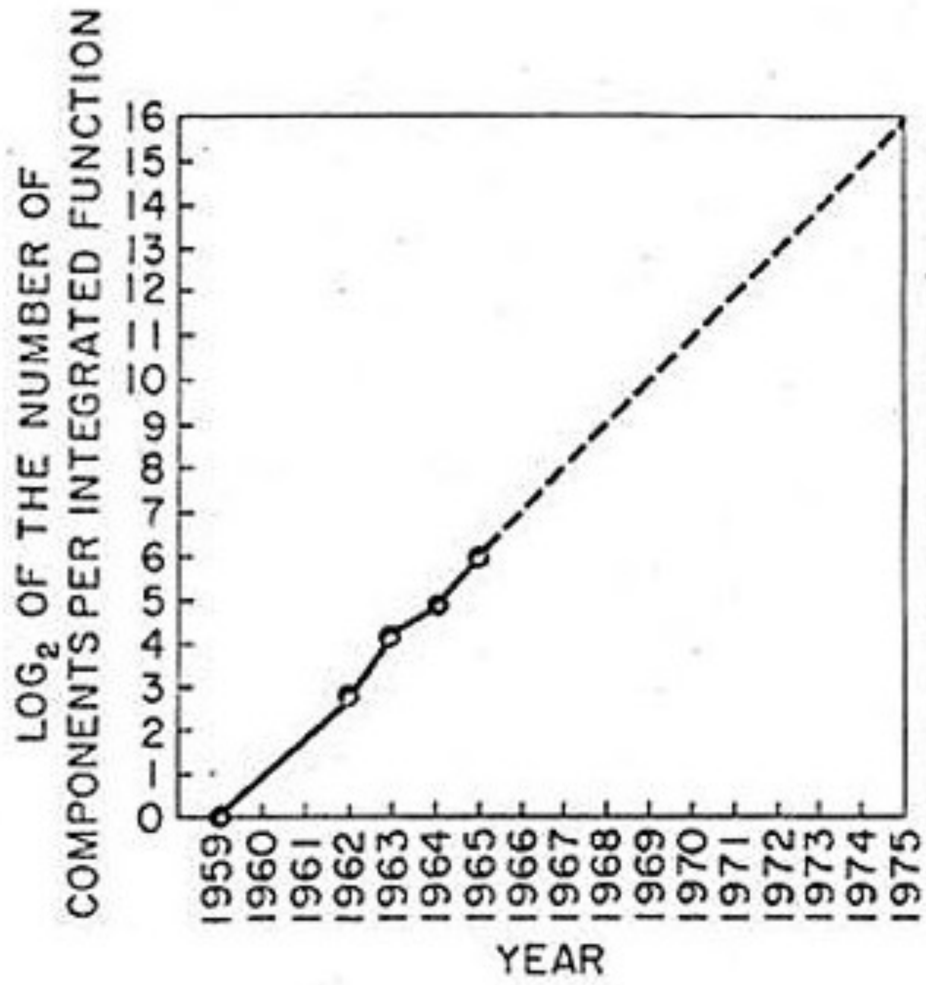
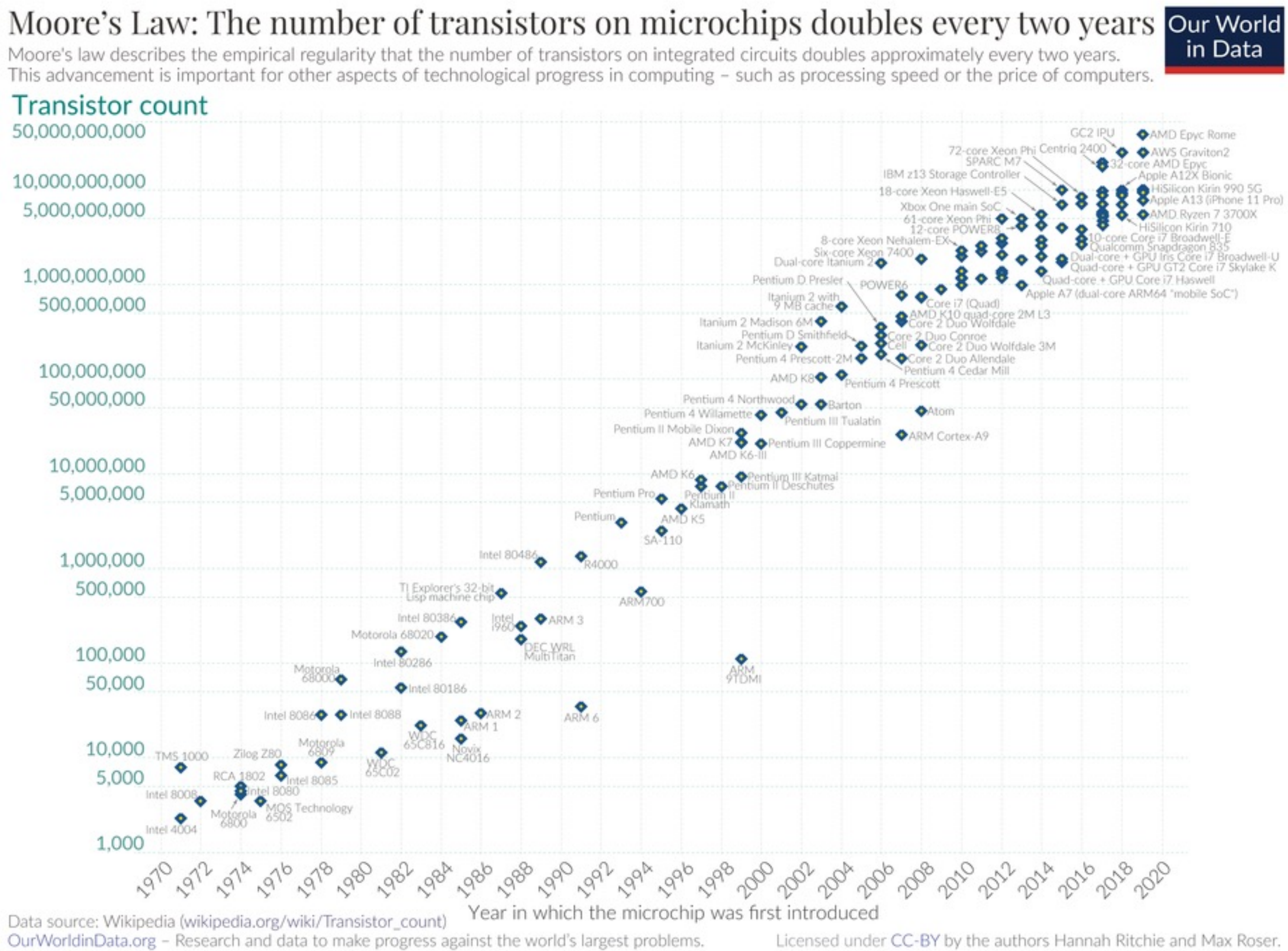


Fig. 2 Number of components per integrated function for minimum cost per component extrapolated vs time.

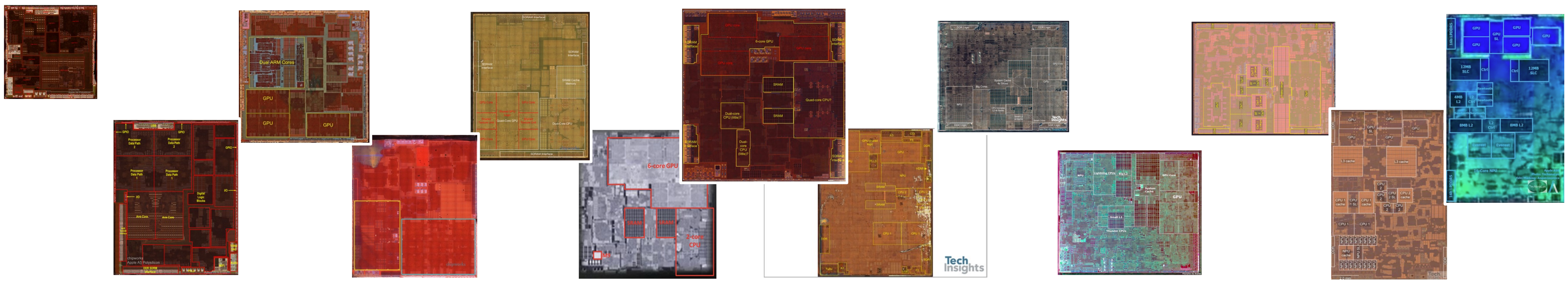
Source: Gordon E. Moore,
Cramming More Components onto Integrated Circuits,
Electronics, pp. 114–117, April 19, 1965.

End of Moore’s Law?



13 Generations of Apple Mobile System-on-Chips

Chip	A4	A5	A6	A7	A8	A9	A10	A11	A12	A13	A14	A15	A16
Year	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
Device	iPhone 4	iPhone 4s	iPhone 5	iPhone 5s	iPhone 6	iPhone 6s	iPhone 7	iPhone 8 & X	iPhone Xs	iPhone 11	iPhone 12	iPhone 13	iPhone 14
Node	45nm Samsung	45nm Samsung	32nm Samsung	28nm Samsung	20nm TSMC	16nm TSMC	16nm TSMC	10nm TSMC	‘7’nm TSMC	7nm TSMC	5nm TSMC	5nm TSMC	4nm TSMC
Area [cm2]	0.52	1.25	0.96	1.03	0.89	1.05	1.25	0.88	0.83	0.98	0.88	1.08	1.14



Die photos:
chipworks/TechInsights/Angstonomics
Data source: wikipedia

Module 2

- **Objective**
  - **Electronic design automation**
  - **Synthesis and optimization**
  - **Multi-criteria optimization**

Computer-aided design

- **Enabling design methodology**
  - **Support large scale system design**
  - **Design optimization, centering and trade-off**
  - **Reduce design time and time to market**
- **…** ***the only purpose of science is to ease the hardship of human existence*** **… [Galileo/Brecht]**

Microelectronic circuit design

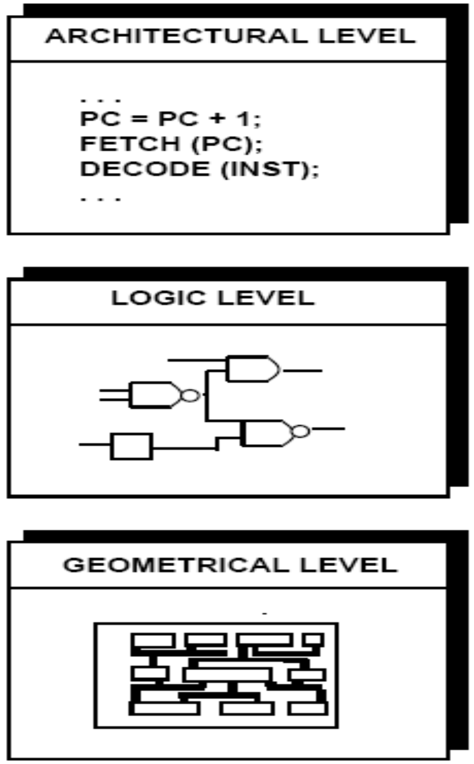
- Conceptualization and modeling
  - Hardware description languages
- Synthesis and optimization
  - Model refinement
- Validation
  - Check for correctness

Synthesis history

- Few logic synthesis algorithms and tools existed in the 70's
- Link to place and route for automatic design
  - Innovative methods at IBM, Bell Labs, Berkeley, Stanford
- First prototype synthesis tools in the early 80s
  - YLE [Brayton], MIS [Berkeley], Espresso
- First logic synthesis companies in the late 80's
  - Synopsys and others

Modeling abstractions

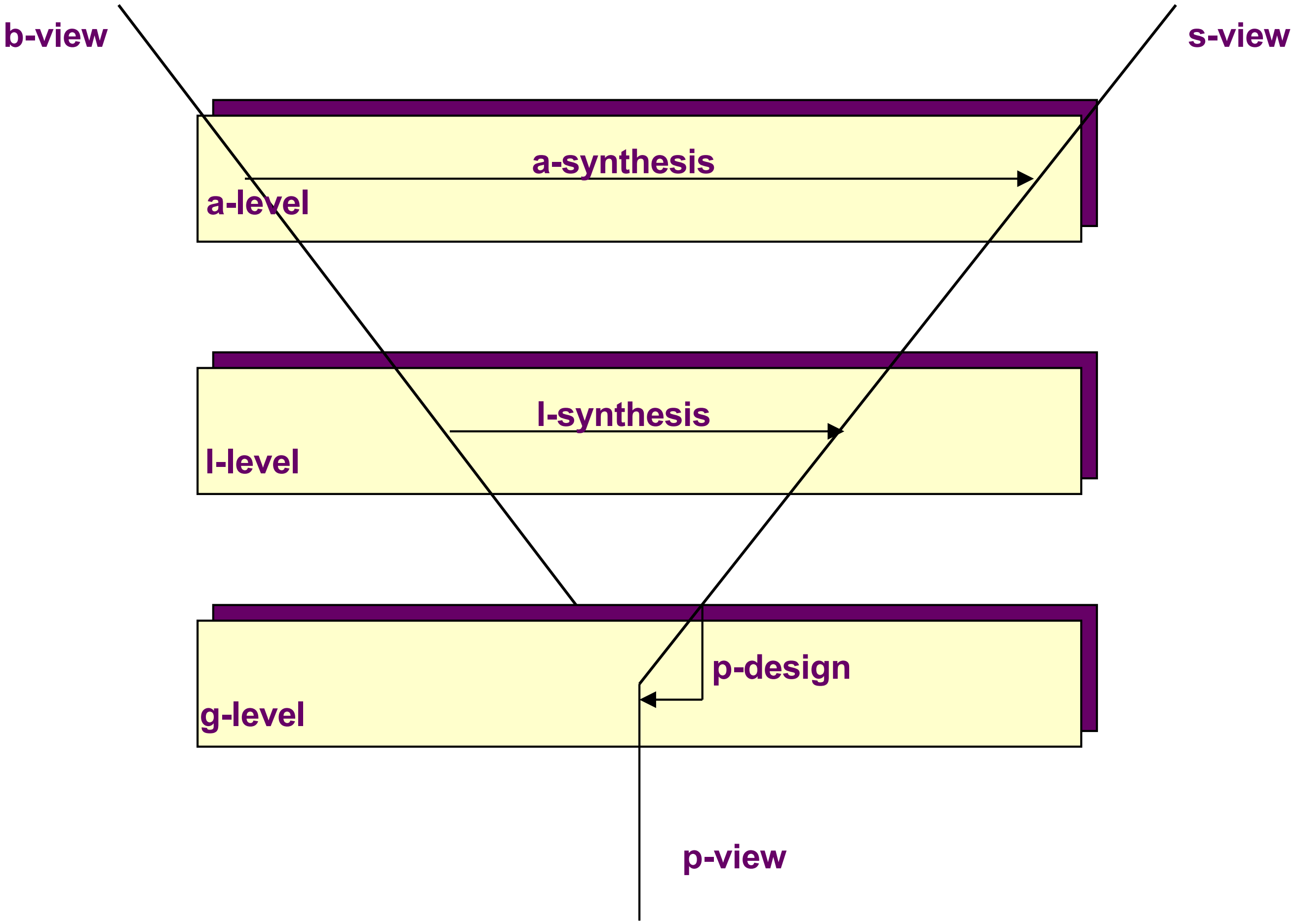
- Architectural level
  - Operations implemented by resources
- Logic level
  - Logic functions implemented by gates
- Geometrical level
  - Transistors and wires



Circuit synthesis

- Architectural-level synthesis
  - Determine macroscopic structure
    - Interconnection of major building blocks
- Logic-level synthesis
  - Determine the microscopic structure
    - Interconnection of logic gates
- Physical design
  - Geometrical-level synthesis
  - Determine positions and connections

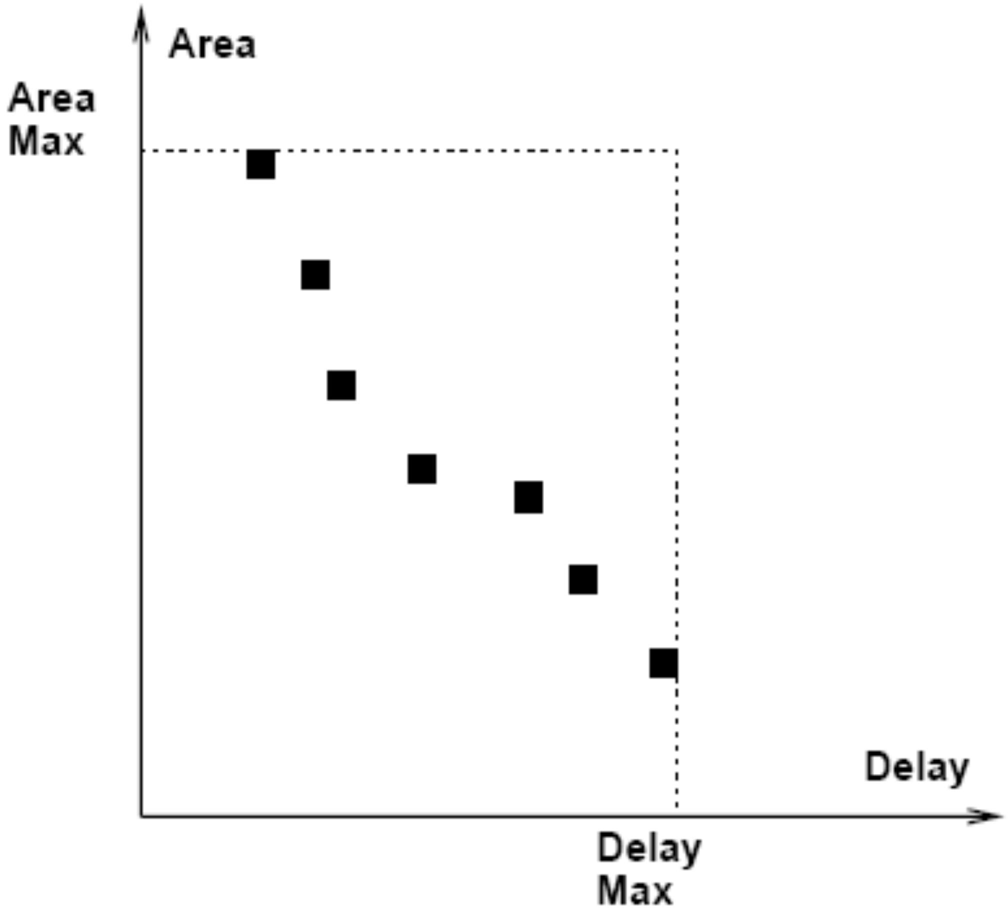
Synthesis levels



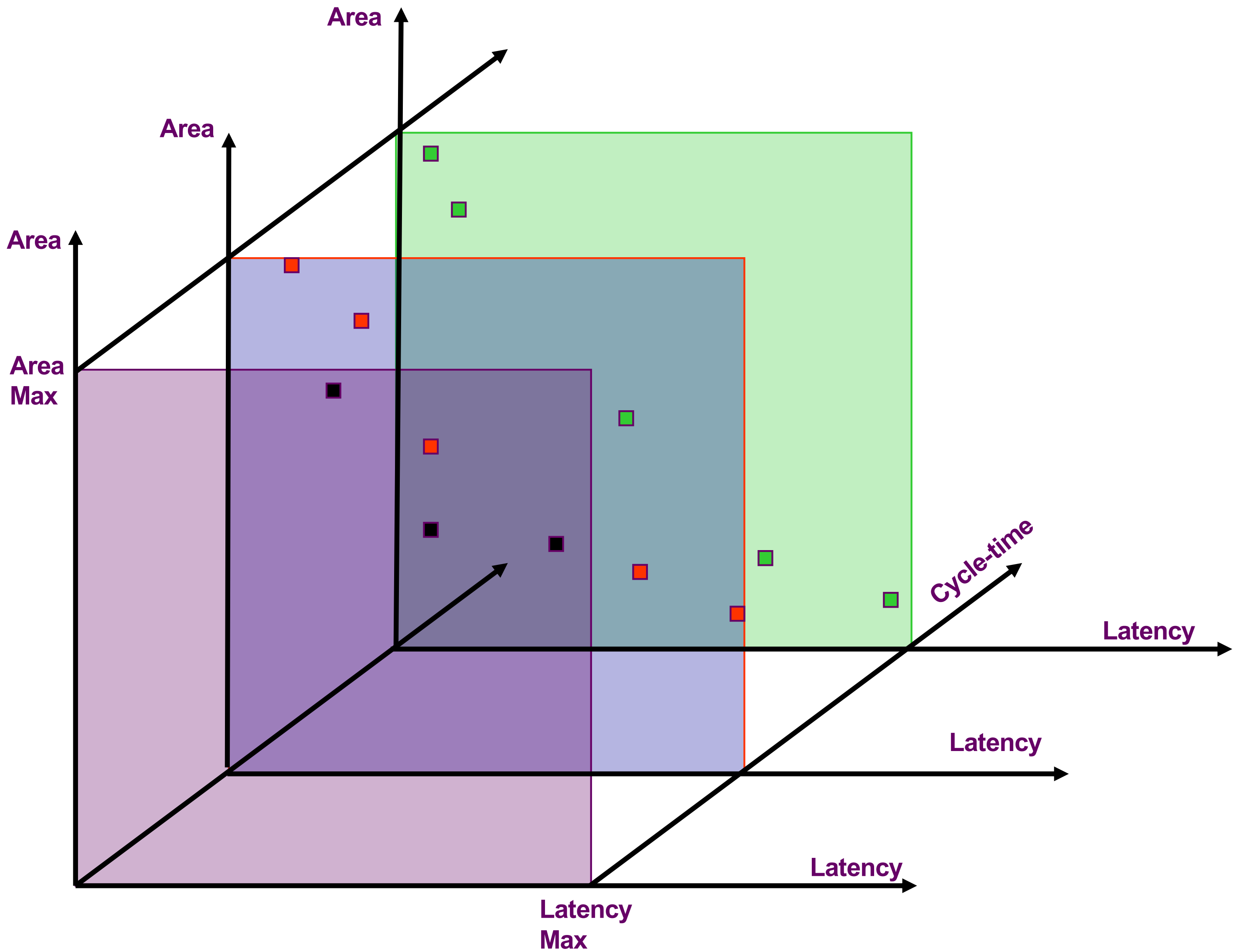
Synthesis and optimization

- Synthesis with no optimization has no value
- Optimization is the means to outperform manual design
- Objectives
  - Performance
    - Frequency, latency, throughput
  - Energy consumption
  - Area (yield and packaging cost)
  - Testability, dependability, …
- Optimization has multiple objectives
  - Trade off

Combinational circuit optimization



Optimization trade-off in sequential circuits



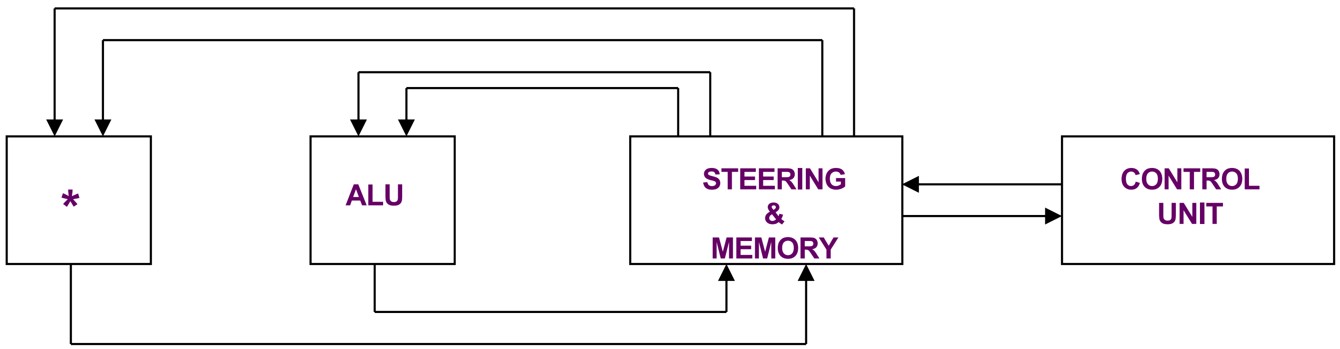
Pareto points

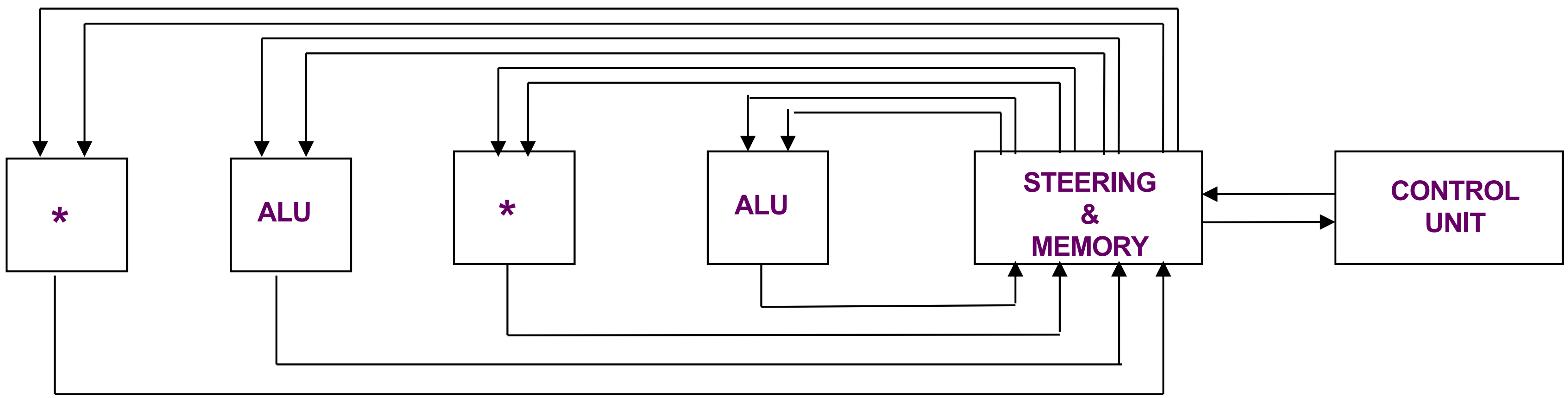
- Multi-criteria optimization
- Multiple objectives
- Pareto point:
  - A point of the design space is a Pareto point if there is no other point with:
    - At least one inferior objectives
    - All other objectives inferior or equal

Example
Differential equation solver

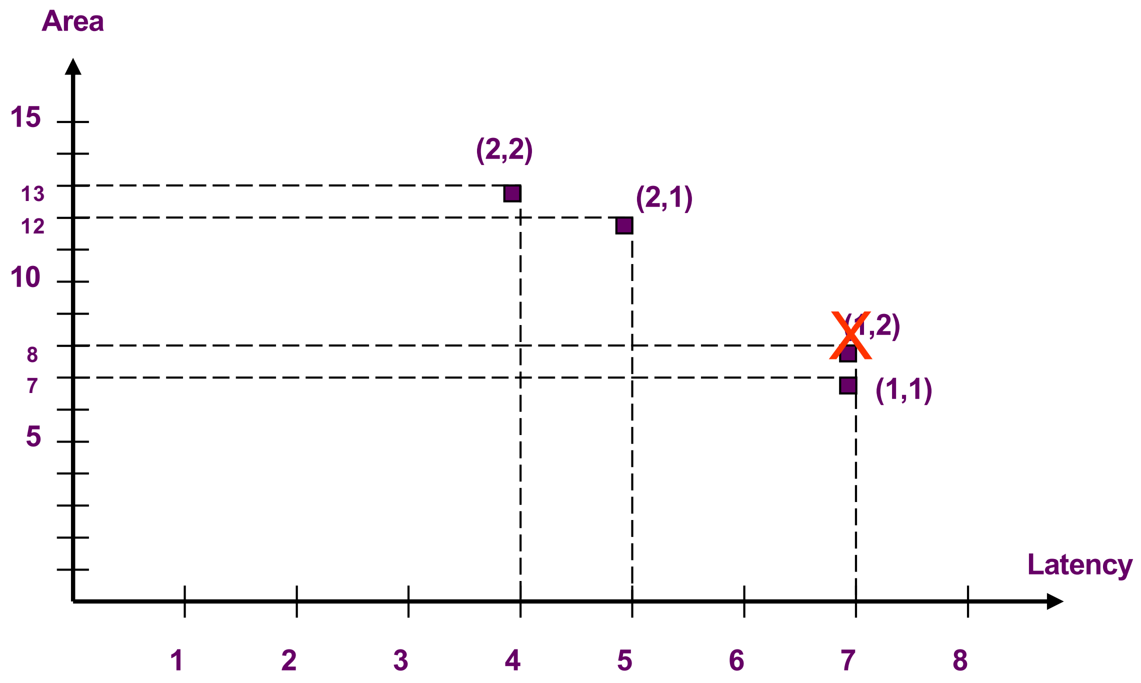
```
diffeq {
    read ( x, y, u, dx, a ) ;
    repeat {
        xl = x + dx;
        ul = u – ( 3 · x · u · dx ) – ( 3 · y · dx ) ;
        yl = y + u · dx ;
        c = x < a ;
        x = xl; u = ul; y = yl ;
    until ( c );
write ( y )
}
```

Example


*
ALU
STEERING
&
MEMORY
CONTROL
UNIT


*
ALU
*
ALU
STEERING
&
MEMORY
CONTROL
UNIT

Example



Summary

- Computer-aided IC design methodology:
  - Capture design by HDL models
  - Synthesize more detailed abstractions
  - Optimize circuit parameters
- Evolving scientific discipline