

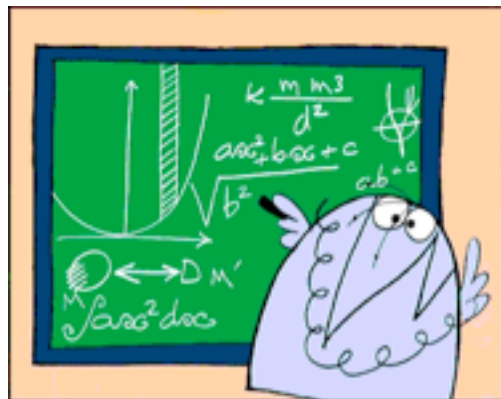
ELEC-H-541

Advanced manufacturing technologies and packaging

Part 02

Dragomir Milojevic

dragomir.milojevic@ulb.ac.be



Today:

1. Overview of solutions to scaling limitations
2. 3D Placement & Route
3. Congestion issues
4. Thermal modeling
5. Mechanical modeling
6. Practical application
7. Practical work

1. Overview of solutions to scaling limitations

(not mutually exclusive)

At macroscopic scale

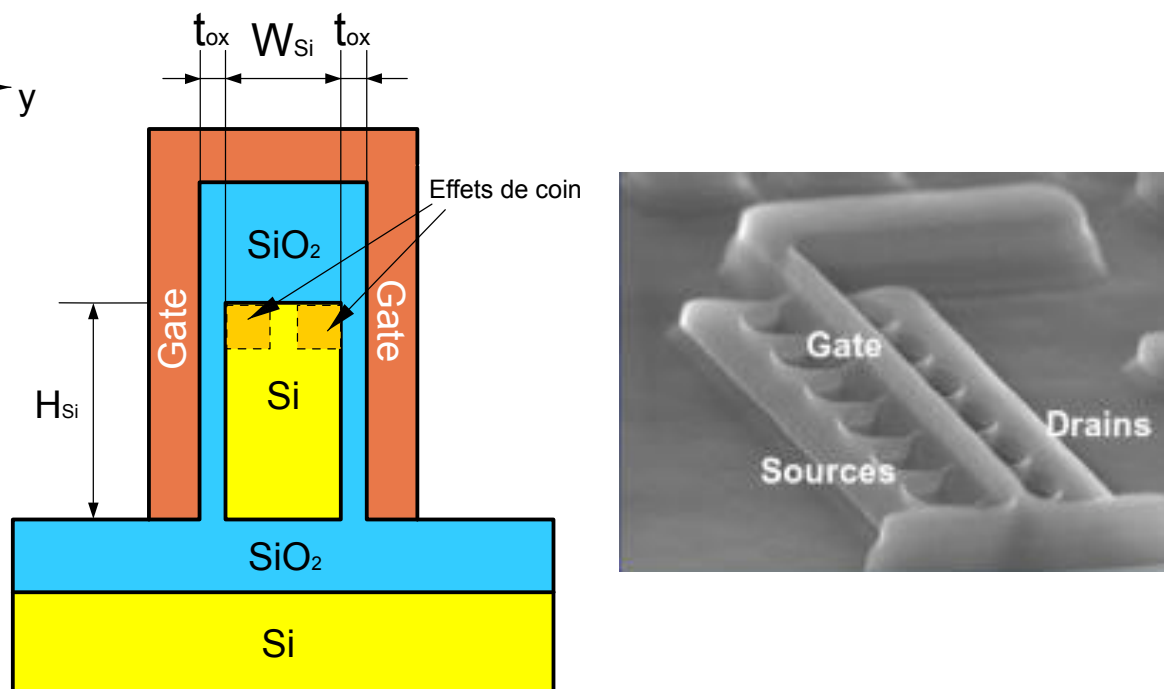
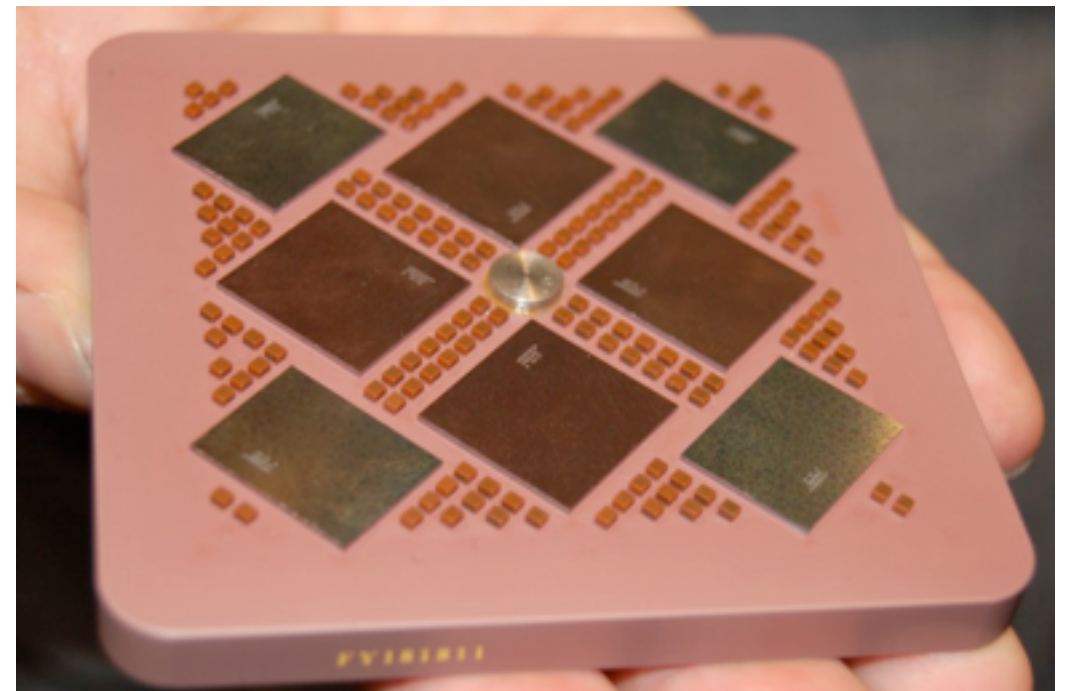
Multi-die integration within the same package

System-in-Package (SiP),

System-in-a-Package or

Multi-Chip Modules

= multiple ICs in the same package
horizontally or vertically



Intel, 2011

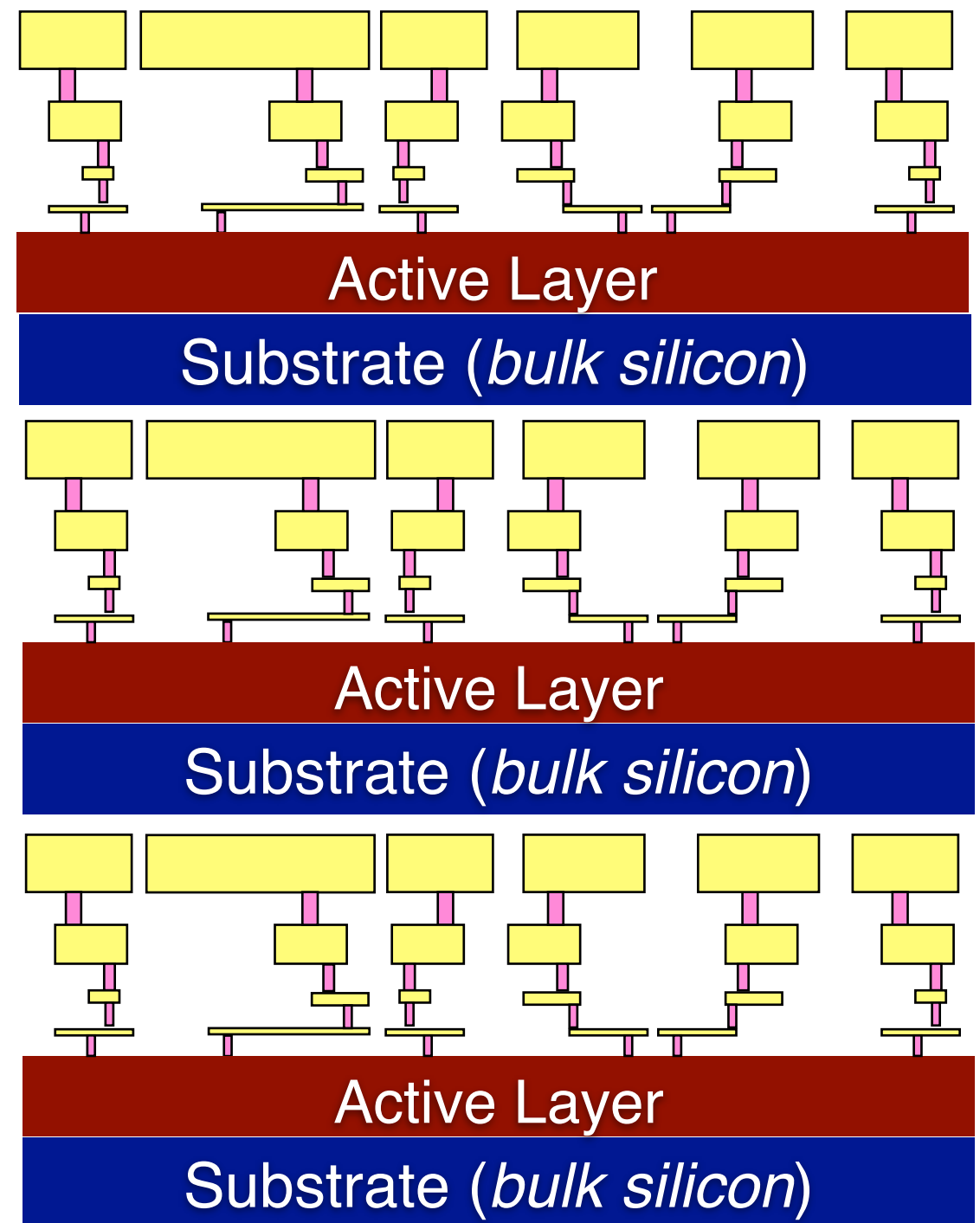
1/2 power dissipation

~35% more speed

... and using the unused dimension !

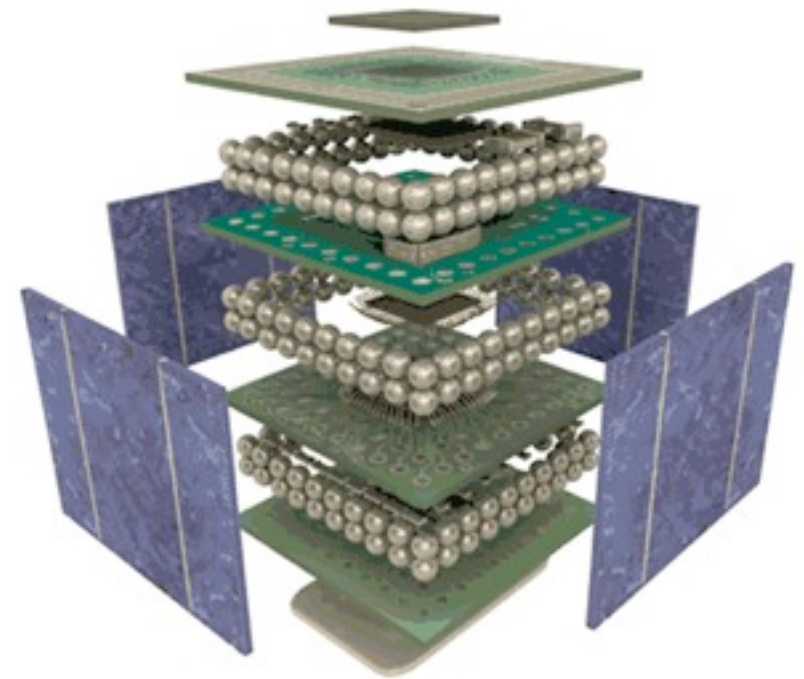
3D Integration

- CMOS is **planar** technology (sounds like 2D)
- In 2D 3rd dimension (z) is used for Metal and Vias only, not for active devices
- How to exploit the 3rd dimension?
- **Three-dimensional integrated circuit (3D IC)** is a chip in which two or more layers of active electronic components are integrated **both vertically and horizontally into a single circuit**
- Semiconductor industry is pursuing this technology in many different forms, but not yet widely used



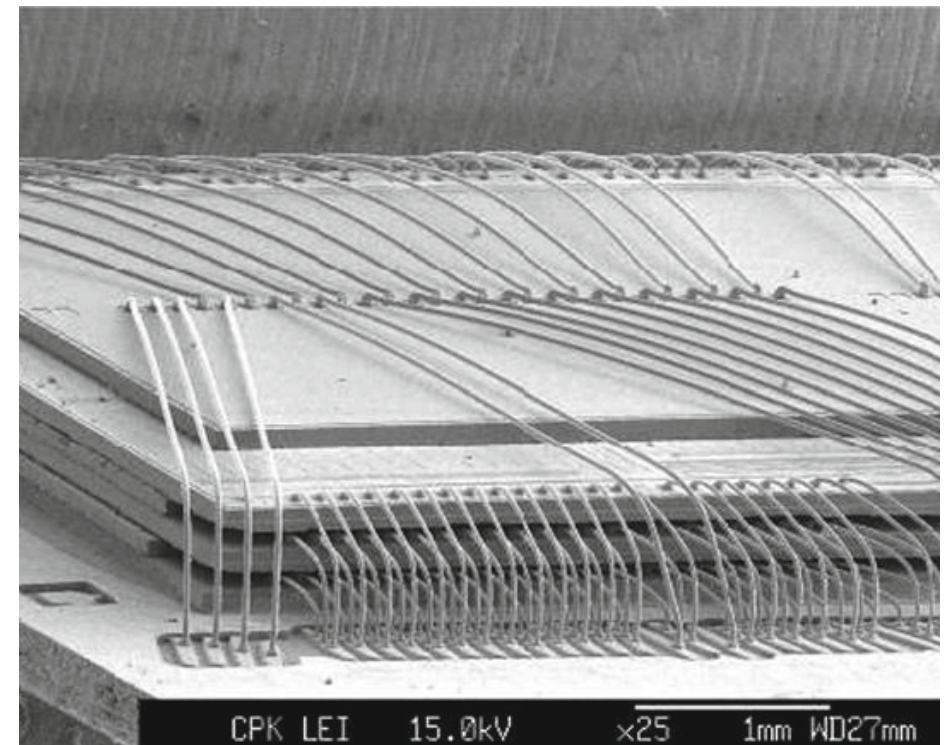
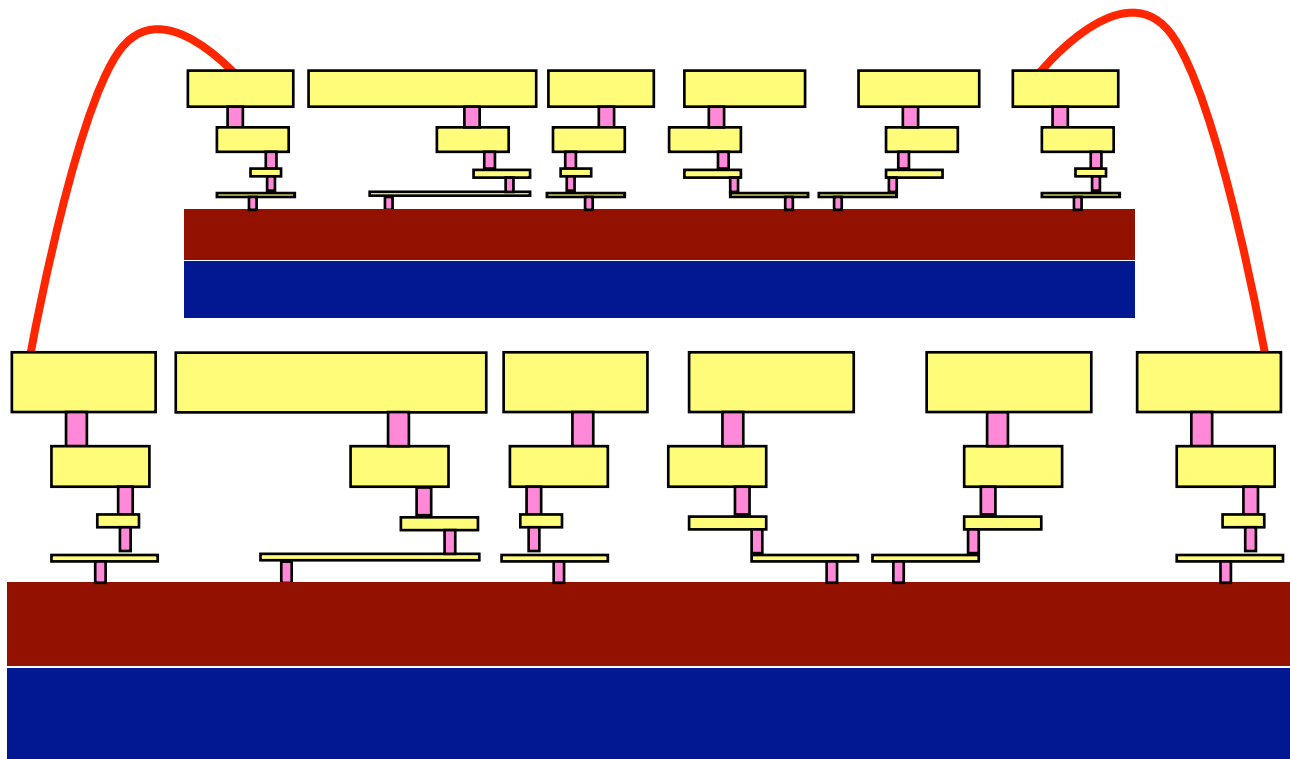
MCM why ?

- Obviously multiples dies previously mounted at PCB level can be now integrated in the same package
 - ✦ Better performance of inter-die connections
 - ✦ Lower power of the interconnect ...
 - ✦ Overall lower power
 - ✦ Higher integration density
 - ✦ Heterogeneous
- But this is too coarse: e.g. integrate memory with logic in the same package (look at the previous example)
- This is not enough ... we want to look into best possible ways to interconnect different dies in the same package
→ **this will depend on the tech. used for die interconnect !**



How to enable 3D MCM ?

Wire Bonding



Peripheral routing, huge pitch,

→ **Limited** N° of connexions with **bad** performance

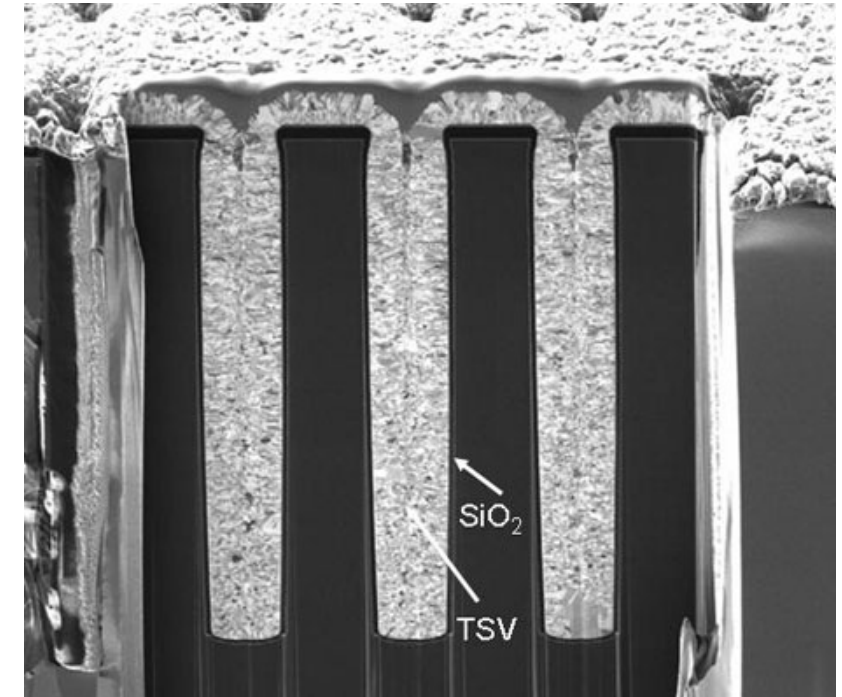
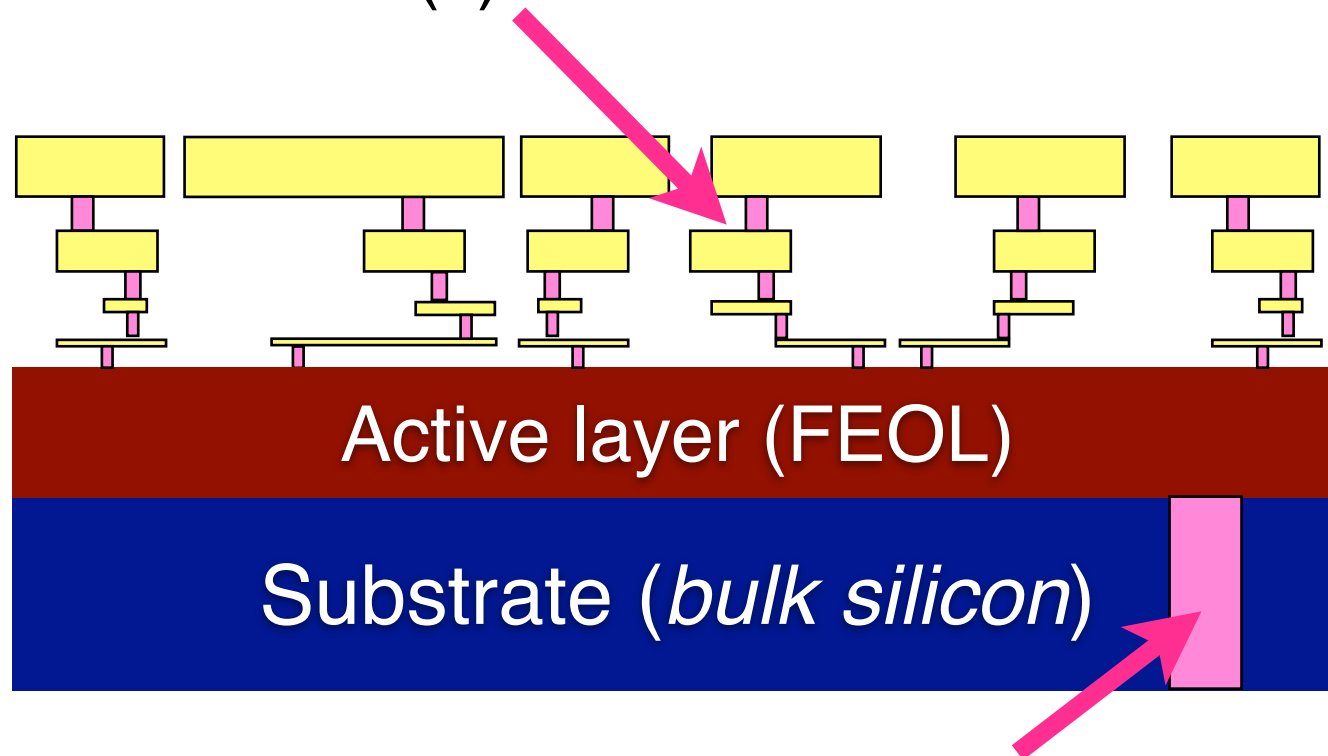
THIS IS NOT VIABLE !

(although one of the iPhones used this ...)

Realistic 3D → key enablers

1. Through Silicon Vias

Classic Via(s)



Through Silicon Via(s)

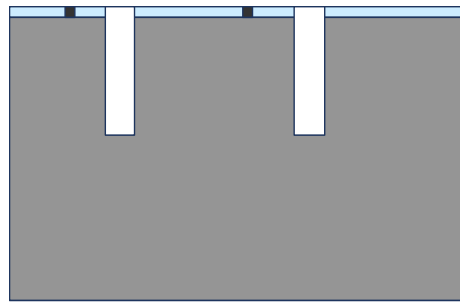
Direct routing, small *pitch* ($<10\mu\text{m}$),
→ **Huge** number of **fast** connexions

VIABLE !!!

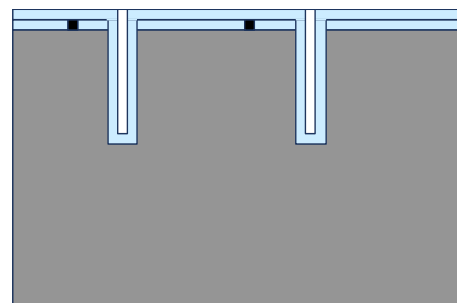
TSVs Processing

Via first processing with wafer thinning

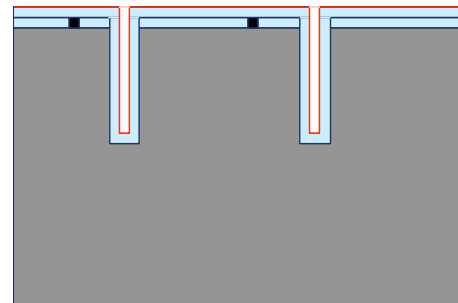
1. TSV Manufacturing



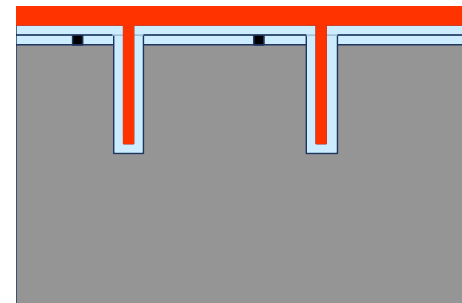
Deep silicon etching



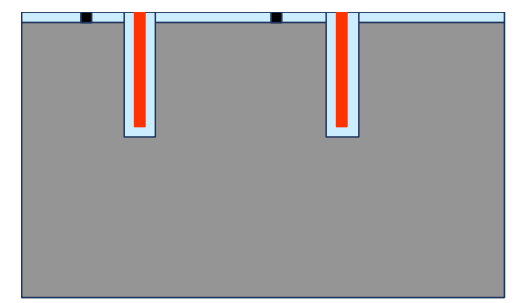
Via oxide deposition



Cu seed deposition

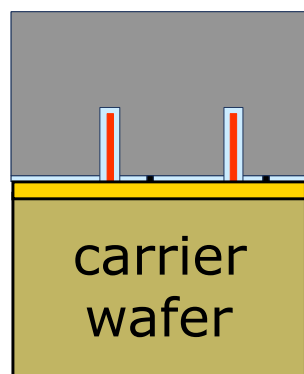


Cu Plating

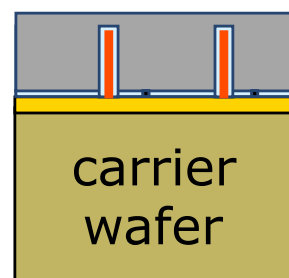


Chemical Mechanical Polishing

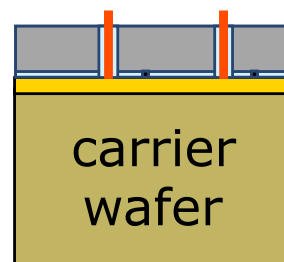
2. Wafer Thinning and Bonding



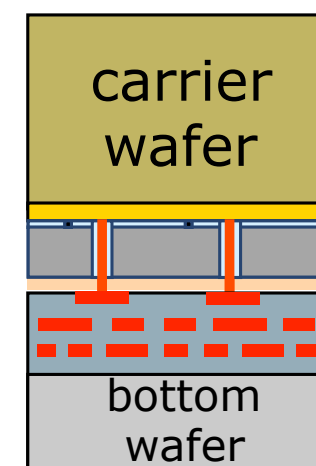
Temporary carrier bonding



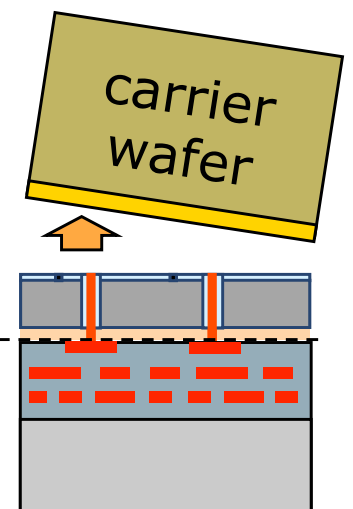
Back side thinning



Exposed Cu nails



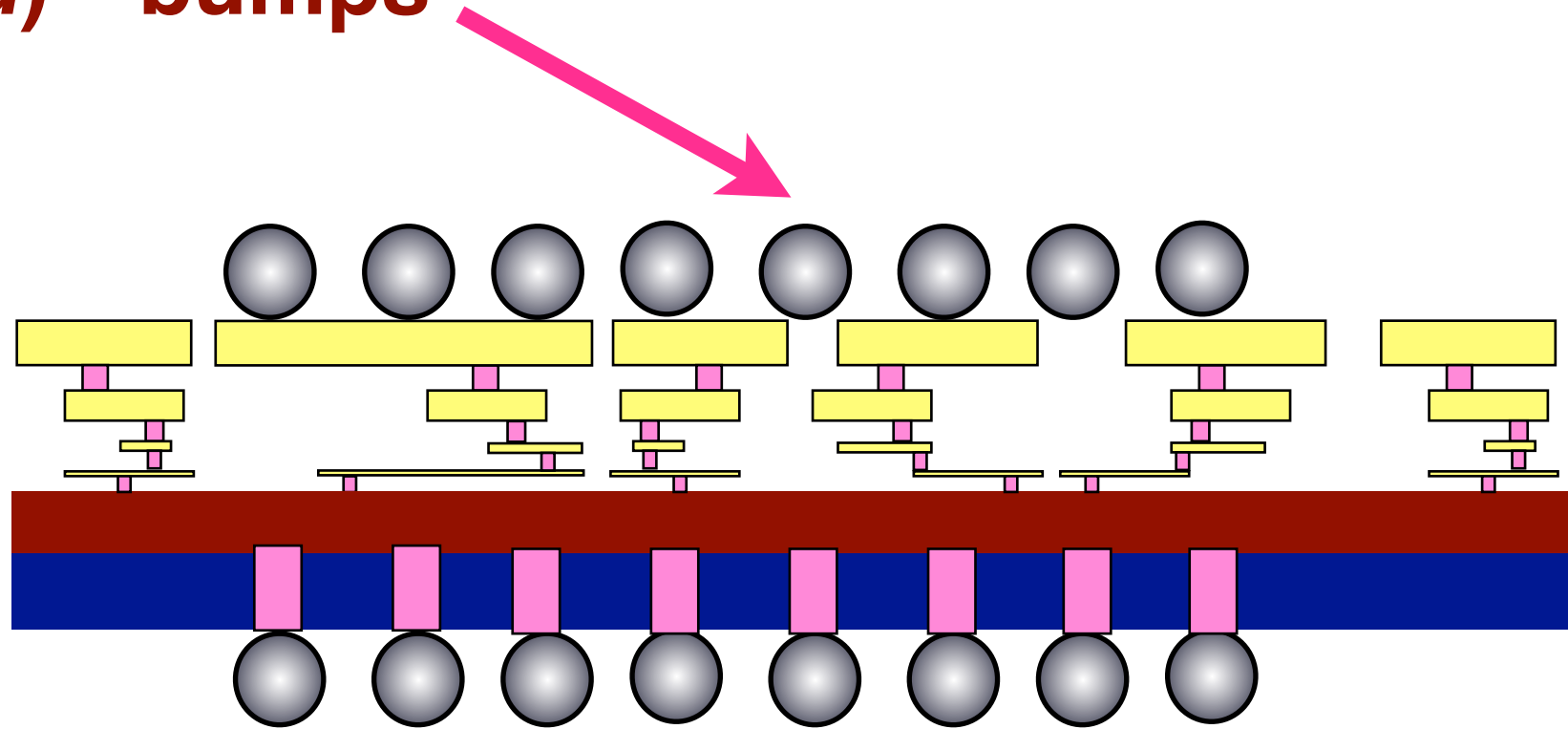
Permanent bonding



Temporary carrier de-bonding

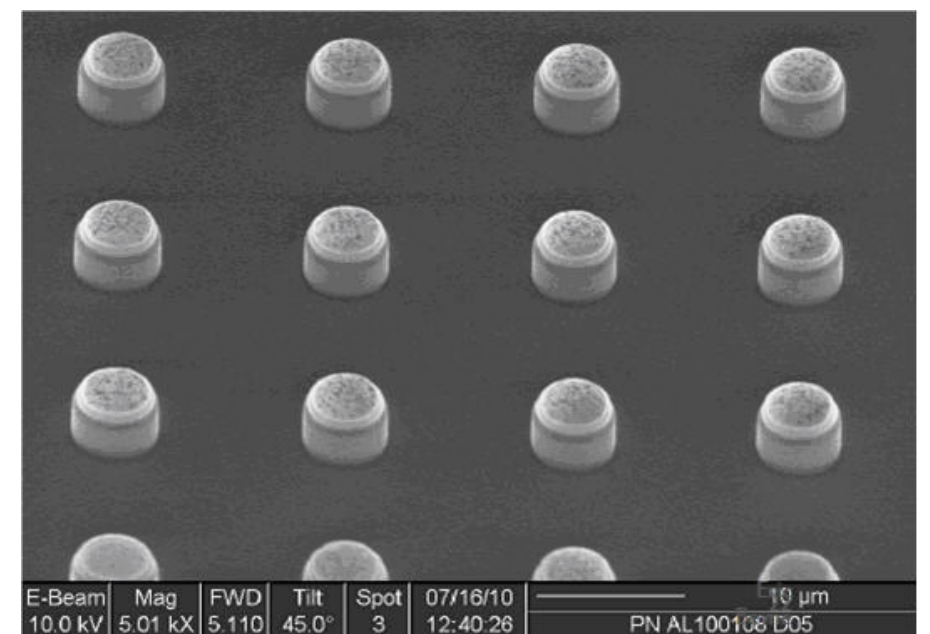
Realistic 3D → key enablers

2. Micro(μ)—bumps



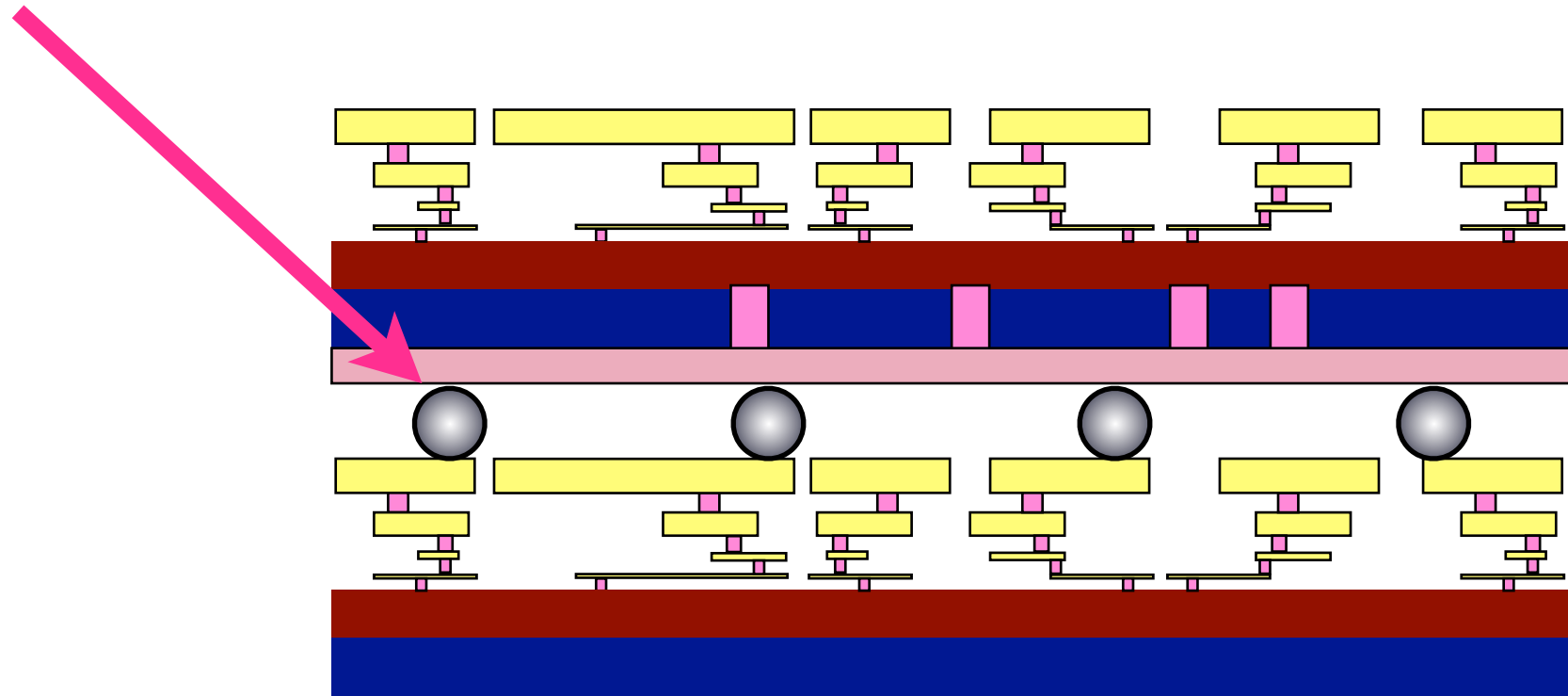
They will establish the actual
inter-die connection,
pitch $\sim 30\mu\text{m}$ (aggressive $10\mu\text{m}$)

VIABLE !!!



Realistic 3D → key enablers

3. ReDistribution Layer — RDL

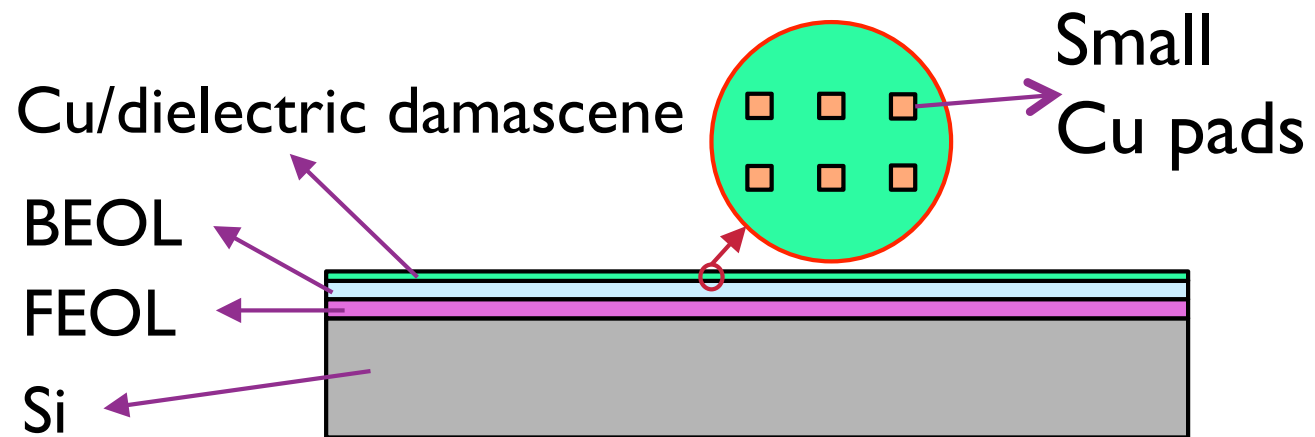


It is a metal layer that can be placed on the backside of the existing die → it will allow to route TSVs and μ bumps

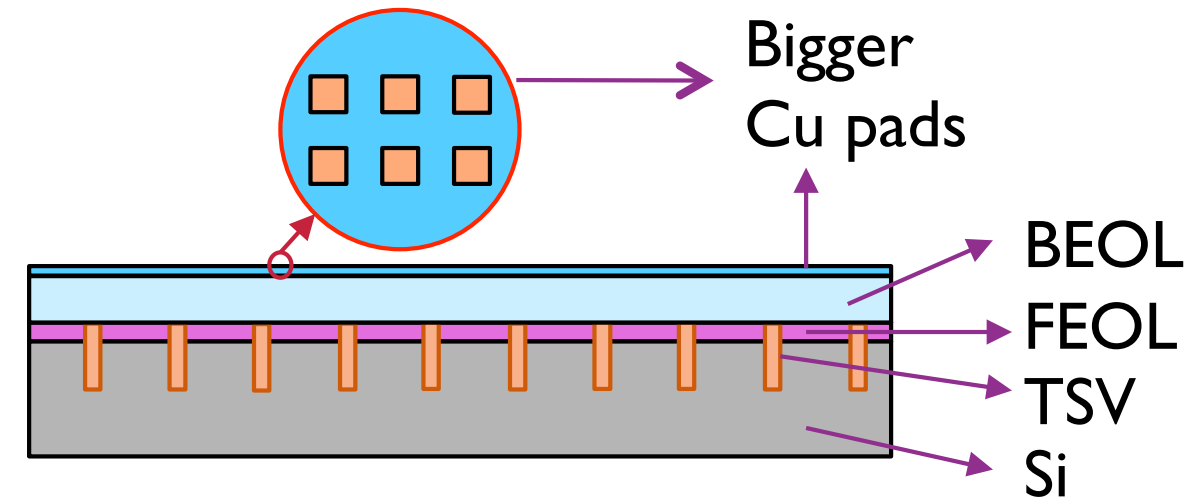
TSVs and μ bumps do not have necessarily to be aligned → more freedom for the placement & route of the TSVs on the top die.

Realistic 3D → key enablers

4. Cu Cu bonding 1/2



- No TSV & regular FEOL
- Contact pads below $1 \times 1 \mu\text{m}^2$
- Full or limited back-end interconnect stack, depending on application

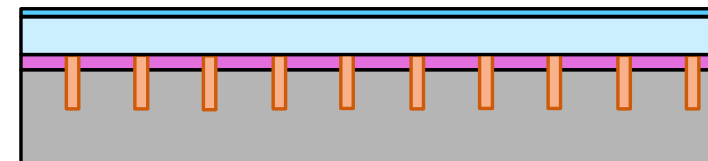
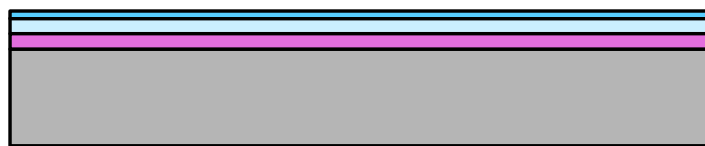


- Via-middle TSV (after FEOL)
- Contact pads below $4 \times 4 \mu\text{m}^2$
- Full or limited back-end interconnect stack, depending on application

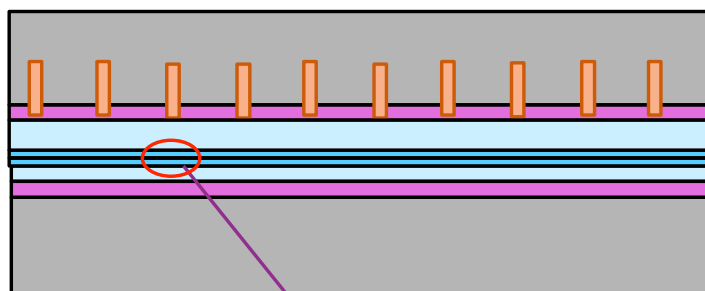
Realistic 3D → key enablers

4. Cu Cu bonding 2/2

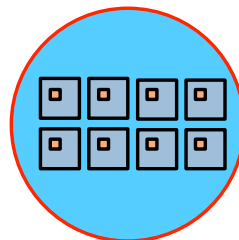
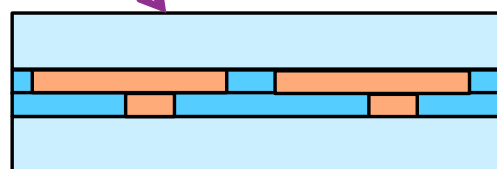
N+1 — Advanced process



W2W bonding:
BEOL-to-BEOL interconnect



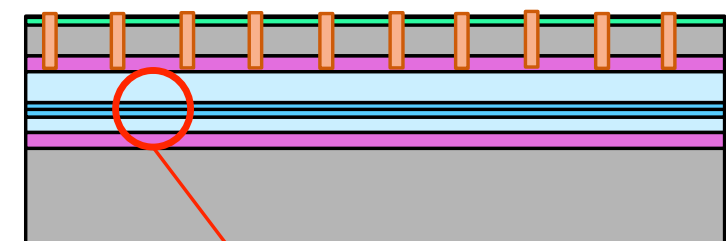
Aligned and
bonded Cu
pads (eg. 5μm pitch)



thinning

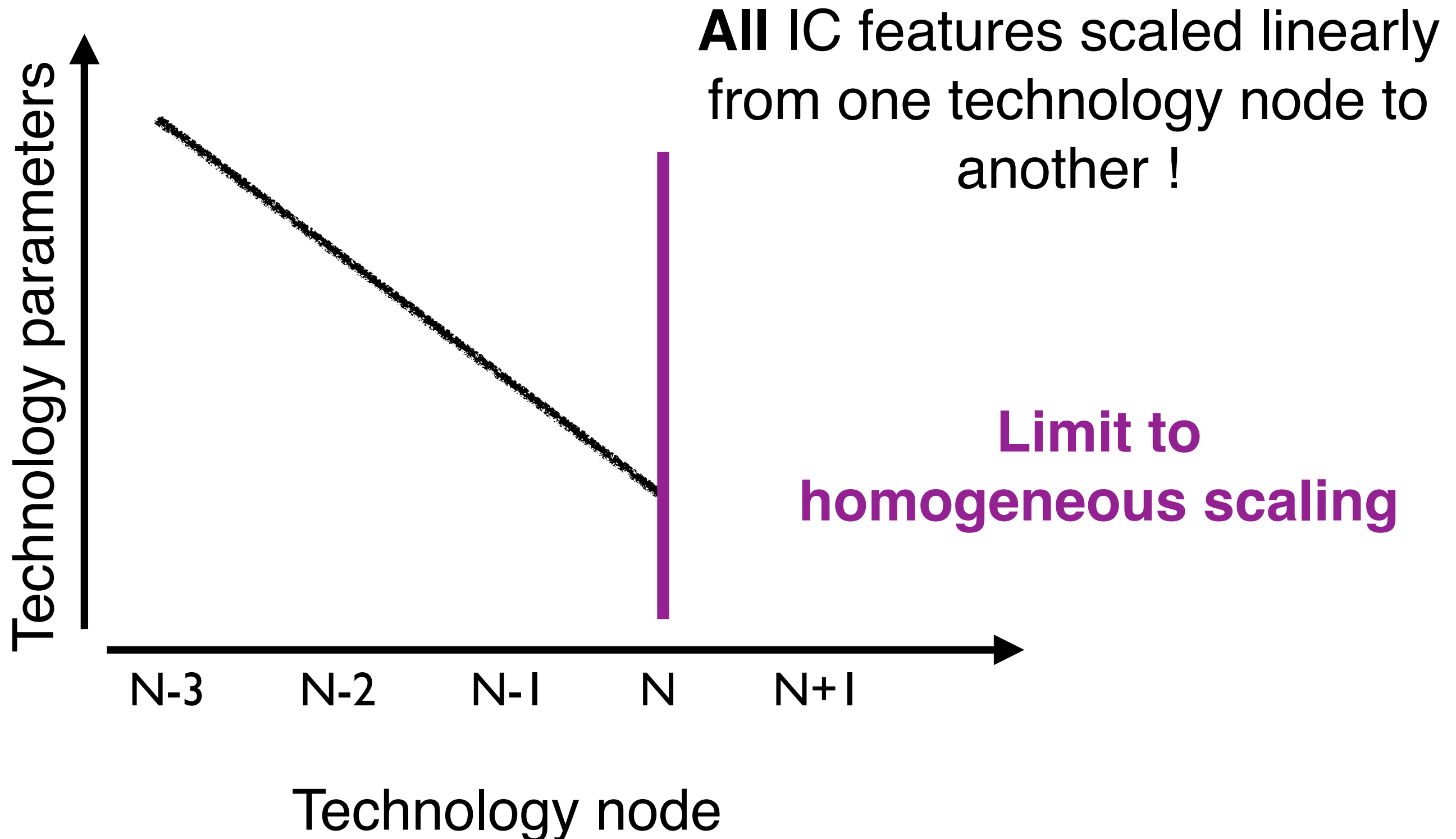


TSV exposure and backside
passivation + CMP

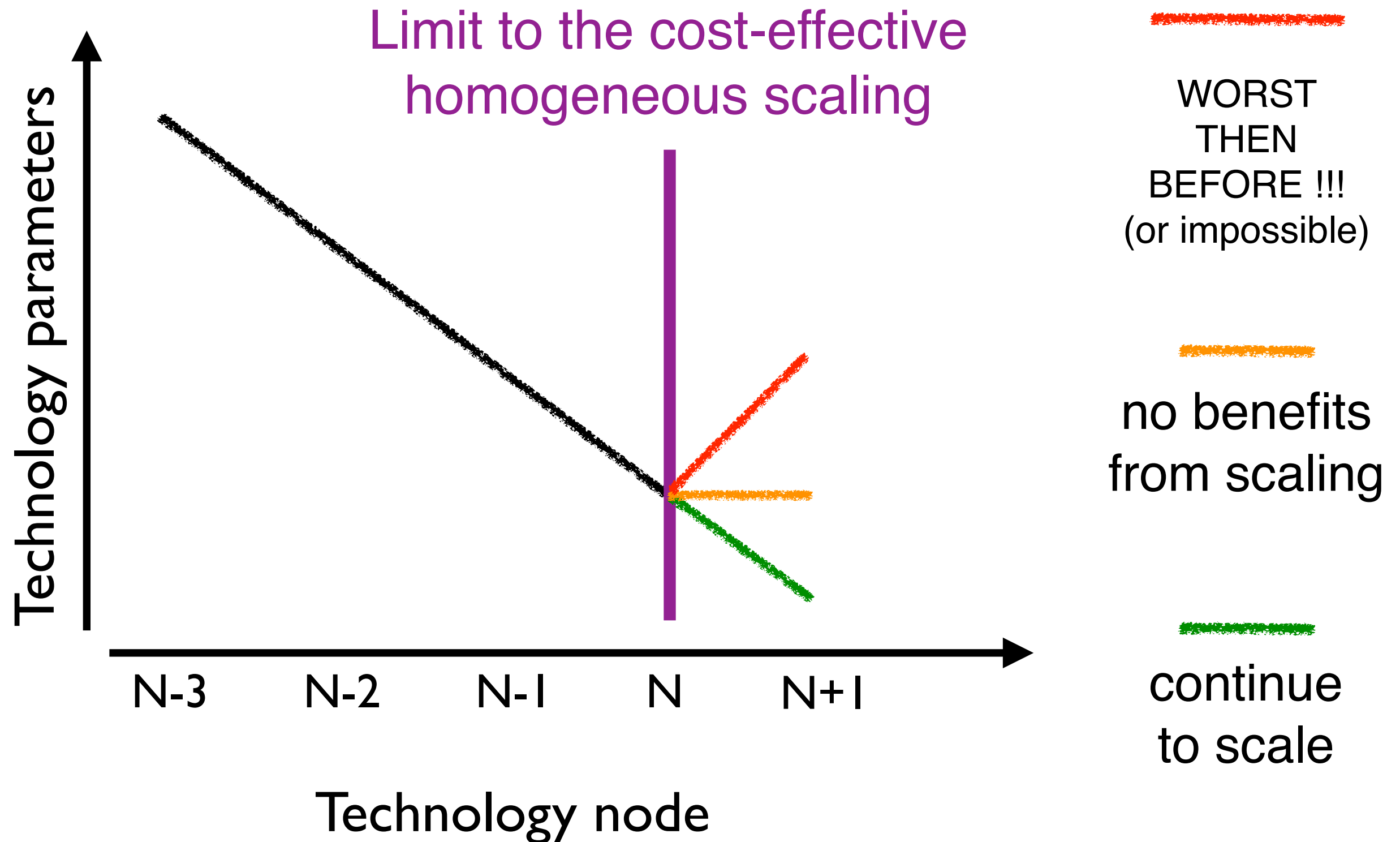


Common Back-end

PAST & PRESENT: Homogeneous scaling



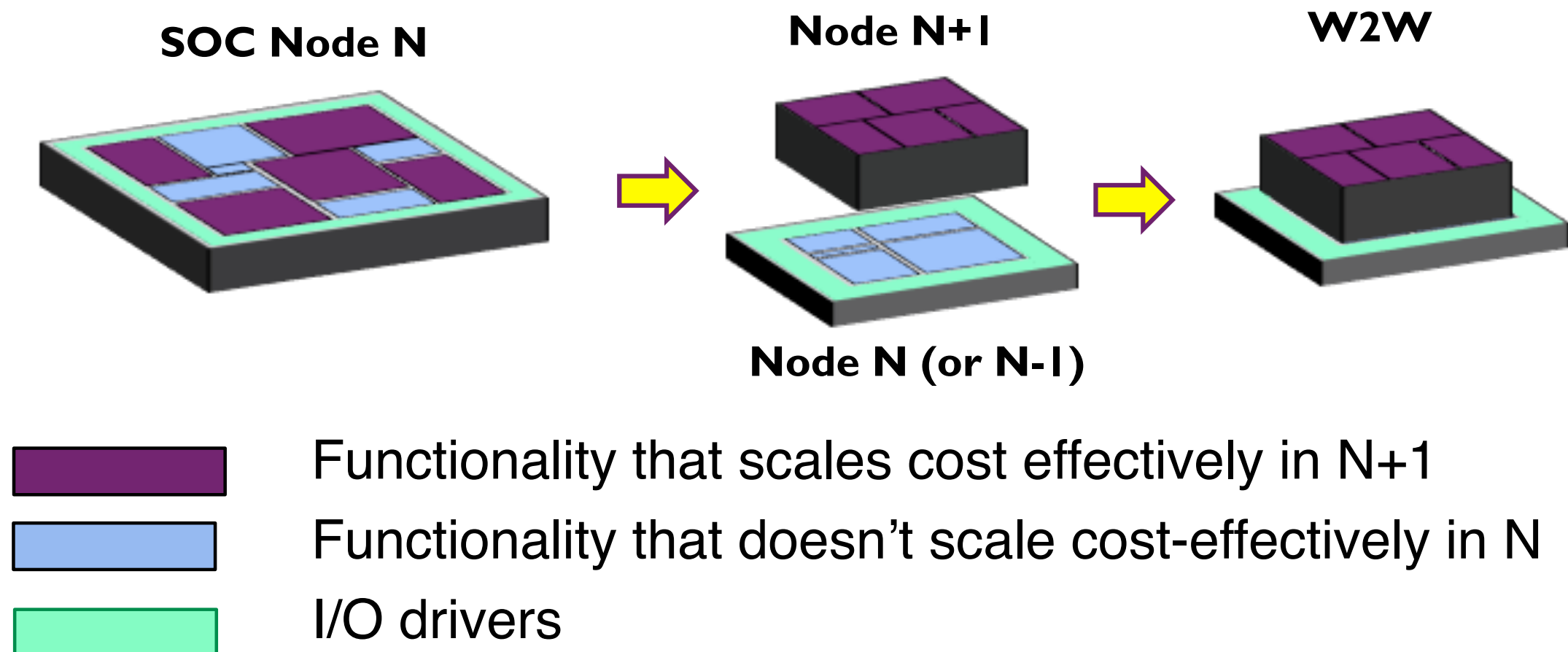
FUTURE: HETEROGENEOUS SCALING ?



Next generation ICs

Combination of :

- ♦ **advanced process** (N+1): only cost-effective functionalities (possibly no IO, ESD, long interconnect buff. etc.)
- ♦ **standard process** (N): implementing all functionalities

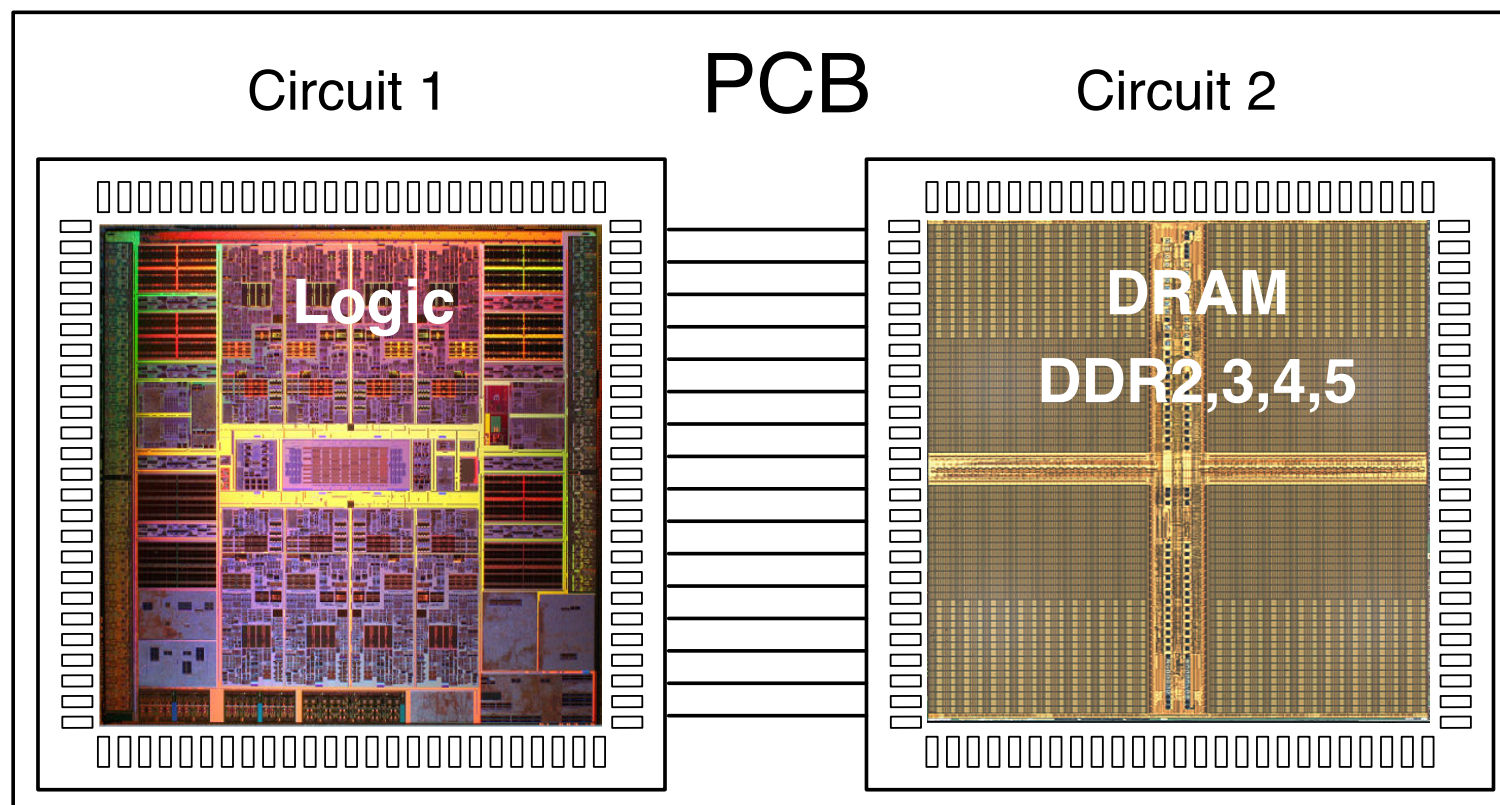


HETEROGENEOUS IC : Properties

- Very dense inter-die connectivity : **40k connexions/mm2**
 - ➡ The stack doesn't look like two separate dies
 - More like one circuit with common/shared BEOL
 - Enables fine grained partitioning
 - Migration at component (memory, ALU, etc.) or even device (IO, ESD, long interconnect buffer, etc.) level becomes possible
- Enables heterogeneous integration with very advanced technology nodes (that will maybe not have certain features, like IO, ESD, long interconnect buffer, etc.)
- Might be interesting cost wise ...

Key enablers: bottom line

- **Inter die connection density increases !!!**
- Allow functional block with high IO count (big number of pins) to be moved in another die
- Blocks can be either from the existing design or from the outside of the package (PCB) → Example off-chip DRAM



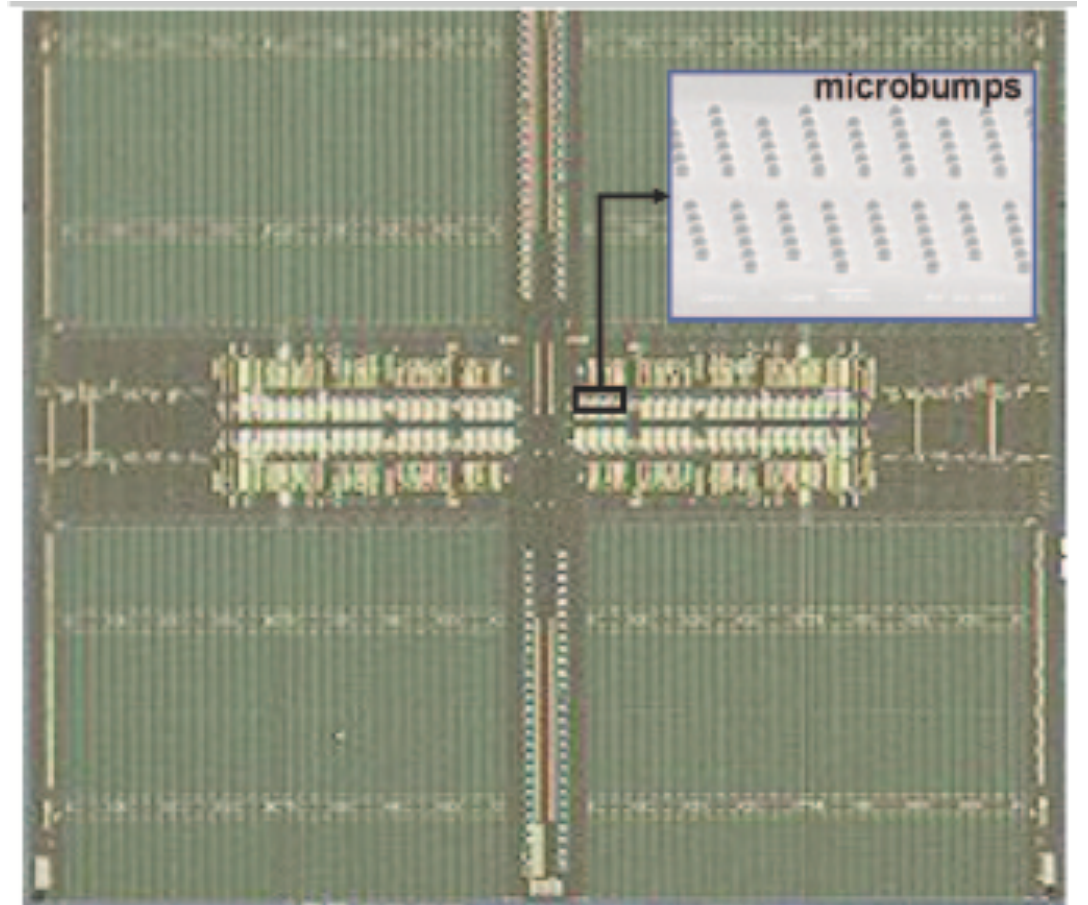
PCB wires

huge capacitive load to logic circuitry
→ big drivers that are area and power HUNGRY

New perspectives

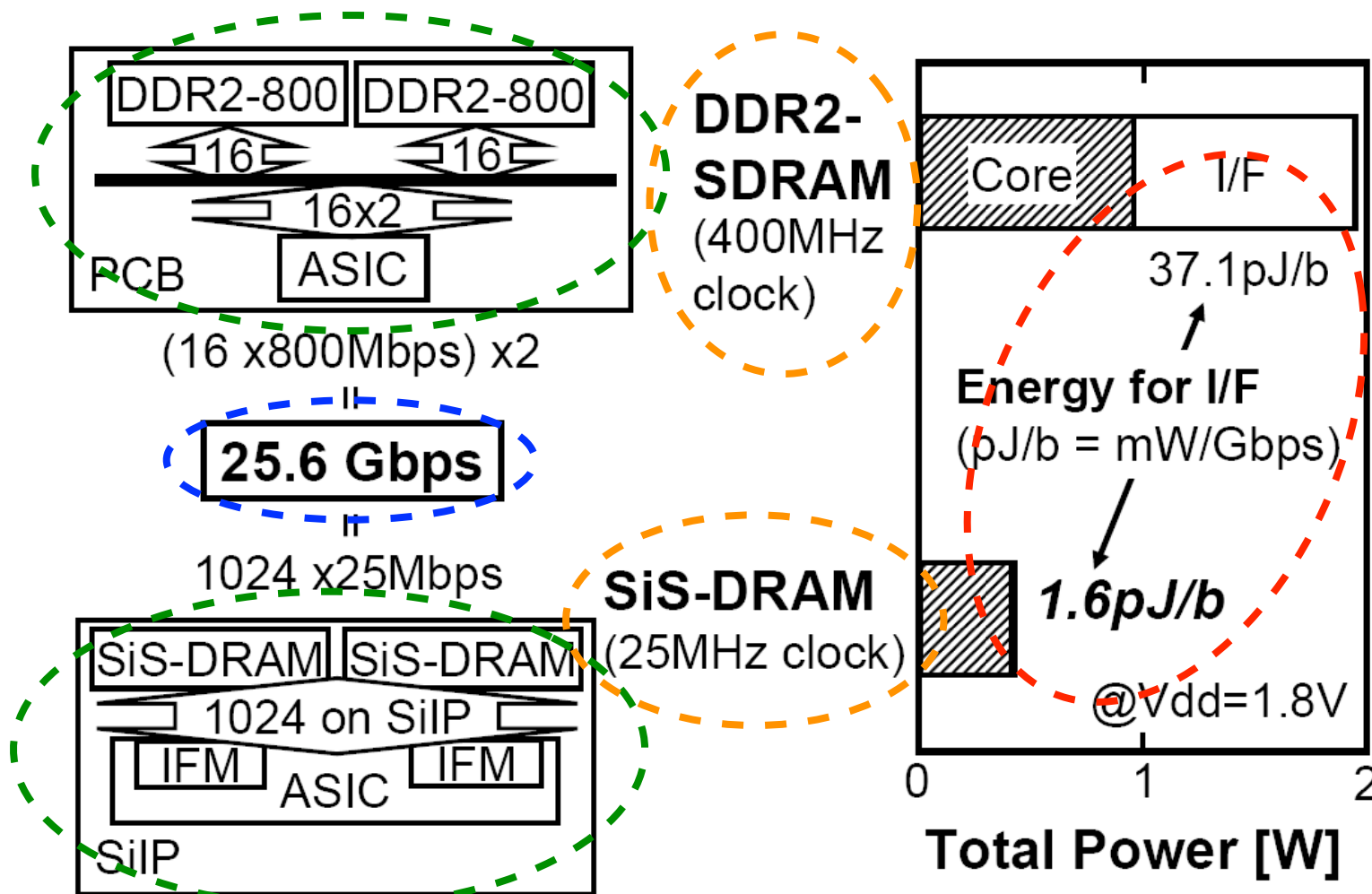
- If IO is cheap, why do not increase the datapath width?
- Until today cost is main blocker for this approach
- With 3D integration, this is not true any more
- Birth of new approaches
→ **Wide IO DRAMs**
instead of 64
→ **1200 bit wide data bus !!!!!**
- Less load capacitance mean smaller drivers, less area, power so better access to DRAM (bottleneck anyhow from system perspective)

SAMSUNG
WideIODRAM



Example of power savings

- The N° of TSVs do not influence the cost of manufacturing (10 or 10000 TSVs) - impact only on the array but @5μm diam. and 10μm pitch this is not an issue any more, but impact on the DESIGN:



Increasing the data path resulting in:

$$16x < F$$

$$= BW$$

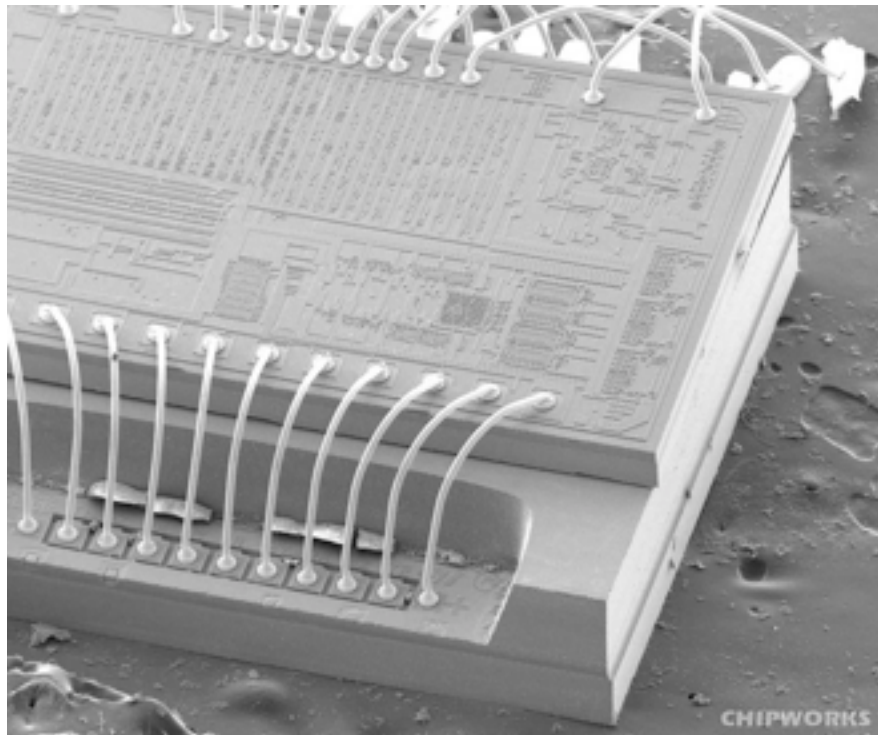
$$23x < P$$

K. Kumagai, C. Yang, et al., "System-in-Silicon Architecture and its Application to H.264/AVC Motion Estimation for 1080HDTV", ISSCC 2006.

3D integration: some advantages

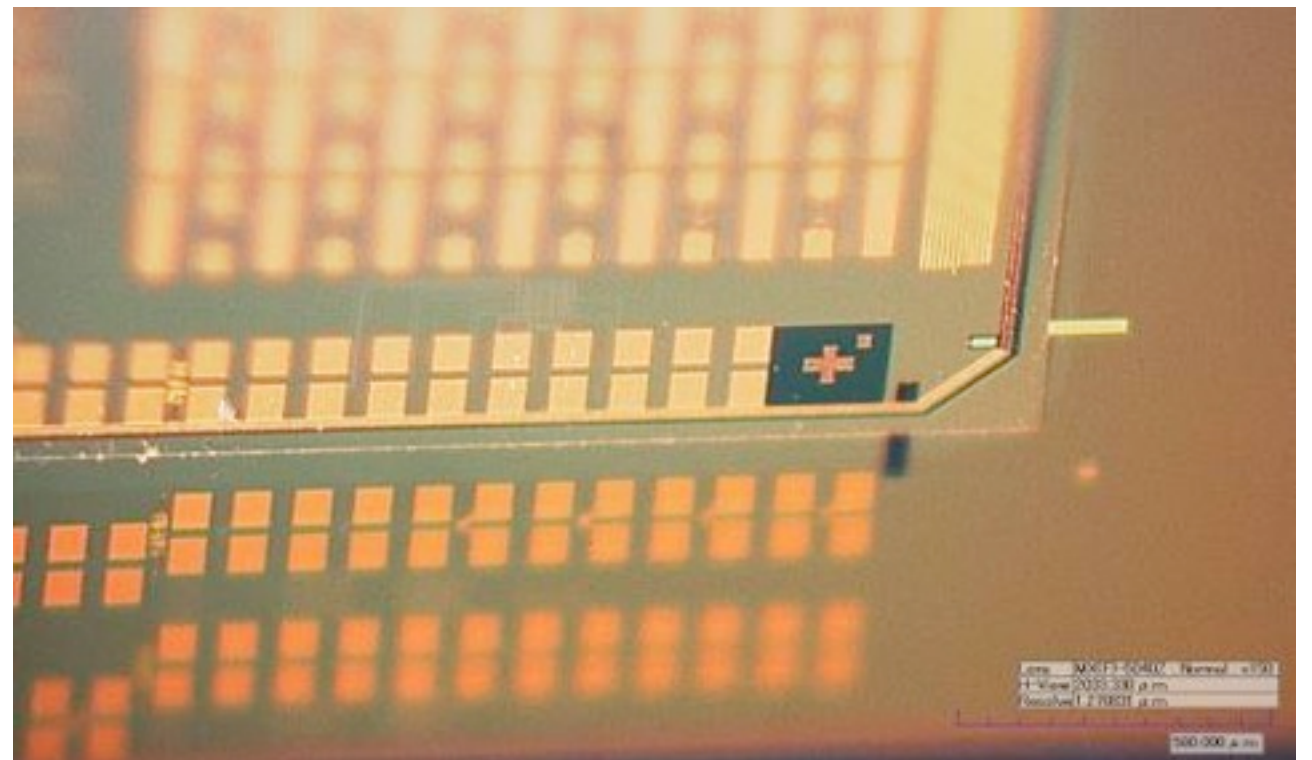
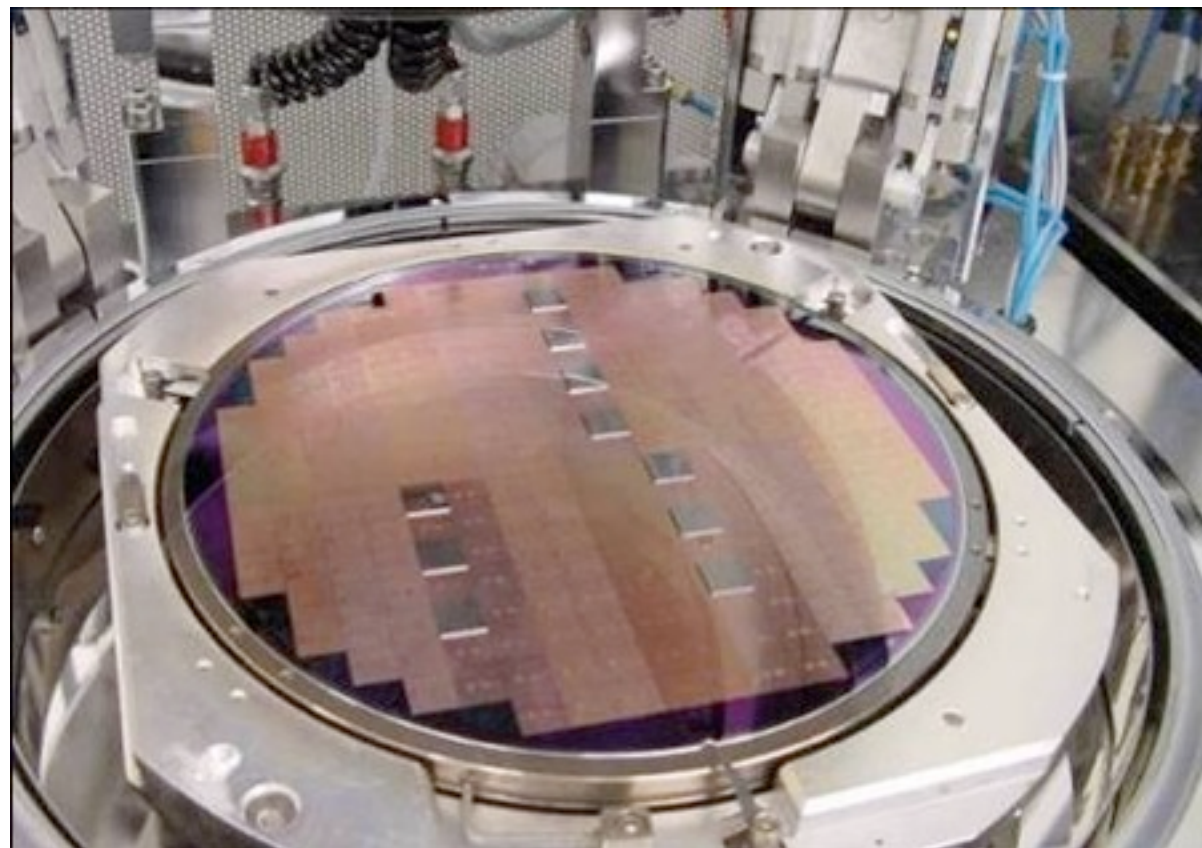
- Increased density for the same footprint and little bit bigger volume → **More functionality**
- Closer, tightly coupled blocks → **Small delays**
- We can combine circuits manufactured in different technologies: memory-on-logic, logic-on-logic, **devices that don't scale with those that can scale**
→ **Integration of heterogeneous circuits**
- Huge inter-die interconnect density → big number of connections, thousands rather than dozens of inter-die connections → **Much more bandwidth**
- Enables design of new systems (e.g WideIO DRAM)
→ **New product opportunities**
- Better yield – since smaller dies → **Lower cost**

3D is real...

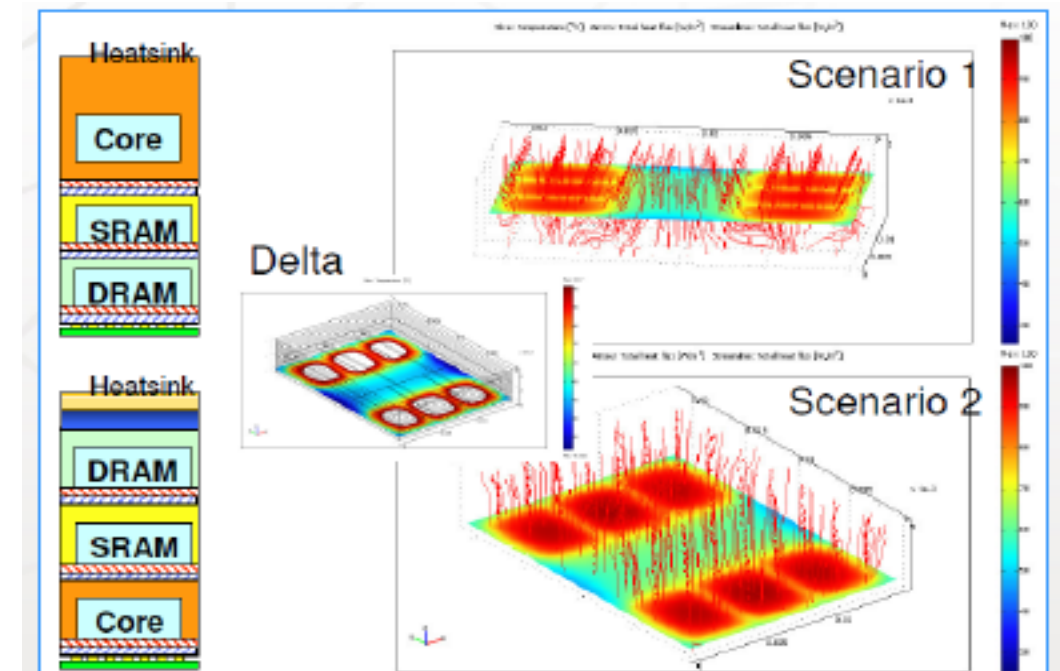
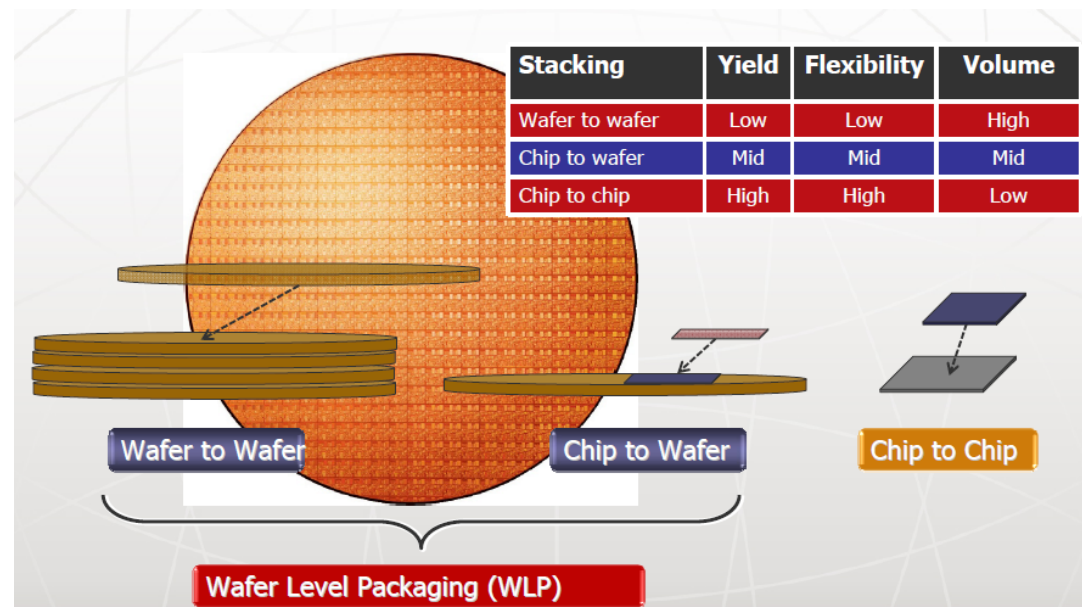


The Accelerometer

To get such a compact device, the ASIC is stacked above the MEMS structure. The MEMS structure is carefully protected in a bonded silicon lid. Cracking off the silicon lid (requiring considerable skill), we can expose the MEMS device. The top structure is the Z-axis sensor, and the bottom structure contains the X and Y sensors. This is the ASIC die used to process the tiny capacitive signals, and create a standard SPI/I2C digital interface, and several smart features such as click and double-click recognition, wake-up, and motion detection.



... but still not mainstream tech yet



Manufacturing, Test & Yield



Standardization, supply chain

Power density, peak temp.



Design flow

Advanced tech design challenges

1. How to model advanced (N+1) technology features so that they can become usable early in the design cycle?
2. How to enable EDA to make use of the new technology features (at device & package level)?
3. How to enable holistic design early in the cycle?

... and the proposed solution

1. How to model advanced (N+1) technology features so that they can become usable early in the design cycle?
 - Virtual Process Development Kit (Virtual PDK)
2. How to enable EDA to make use of the new technology features (at device & package level)?
 - Partner with EDA & make extensions to the existing tools to support these new tech features
3. How to enable holistic design early in the cycle?
 - Develop, validate and calibrate (using actual Silicon)
Compact Models used with the EDA tools above



2. 3D Placement & Route

View of the complete flow

Design Planning

- ♦ Purpose: Perform deep design exploration to fix specs early in the design flow cycle
- ♦ Tool: Atrenta's SpyGlass Physical3D® extended for 2.5/3D heterogeneous design
- ♦ Fast to enable many iterations

Compact models

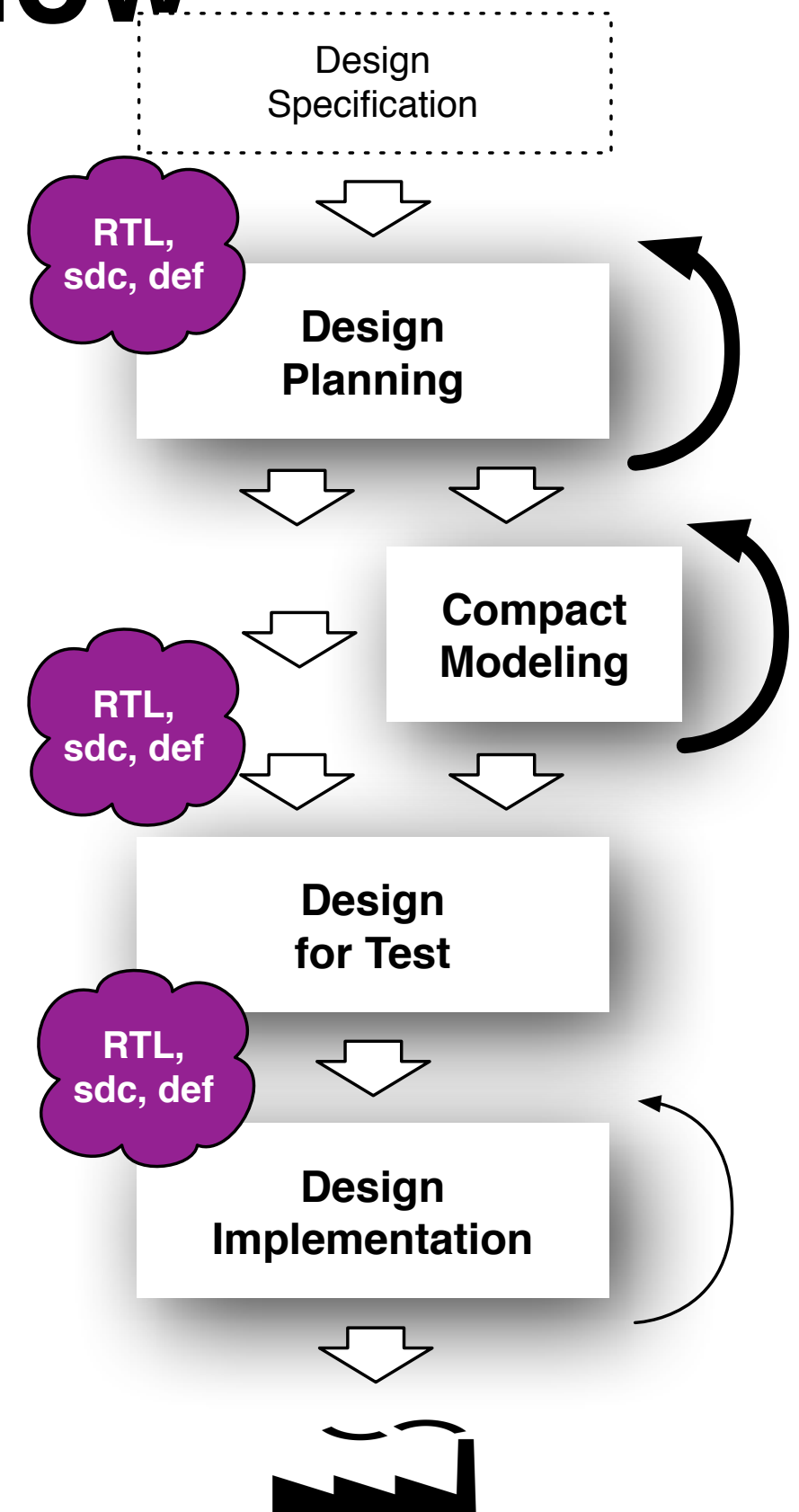
- ♦ Purpose: Check for other design properties — thermo-mechanical, delay, cost
- ♦ Tools: IMEC's CTM CMM CDM CCM
- ♦ Fast & accurate since validated using silicon

Design for test

- ♦ Automated addition of DfT structures

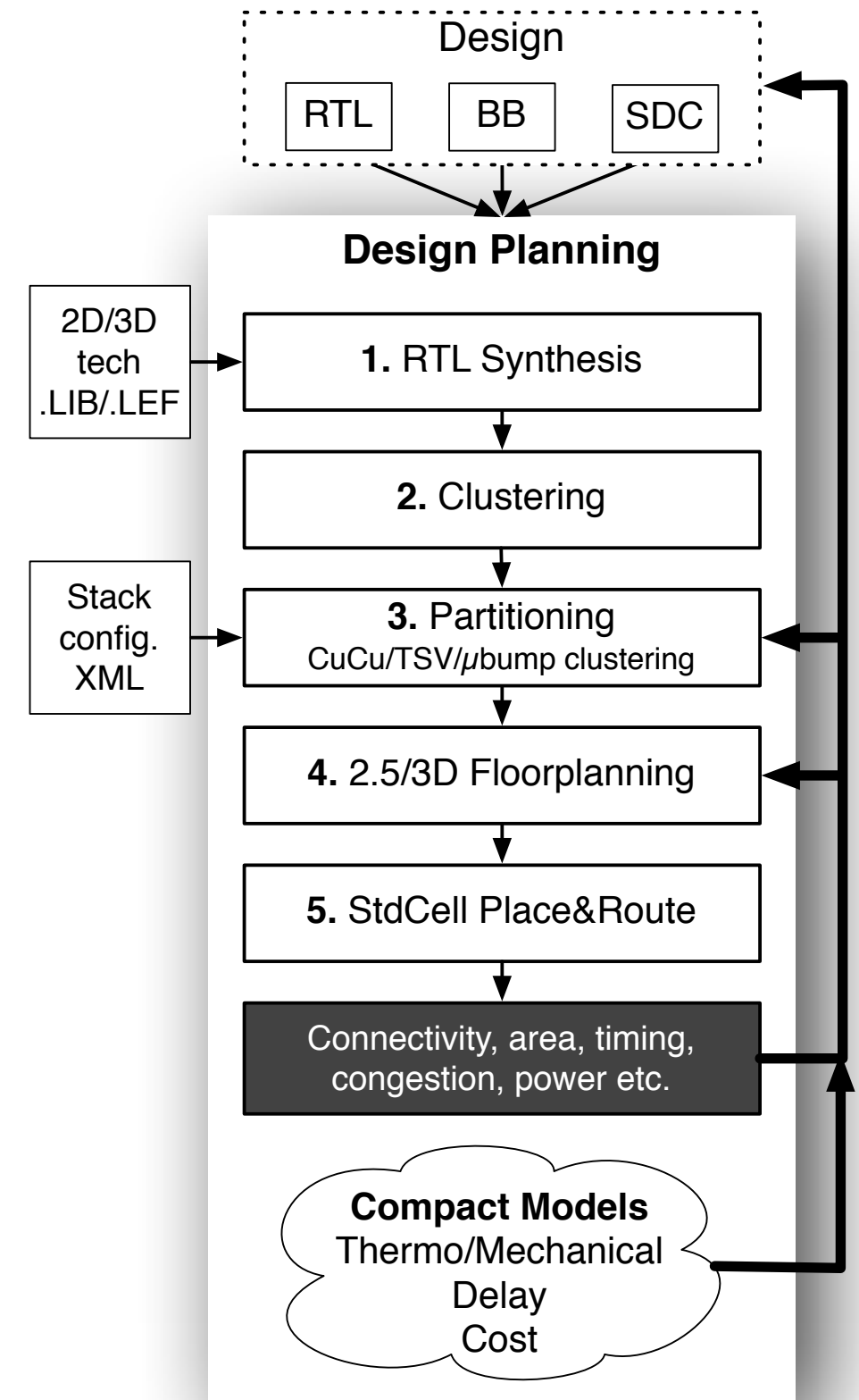
Design implementation

- ♦ Purpose: Generate the actual GDSII with minimum number of iterations
- ♦ Tools: Any industry standard back-end flow (since standard file interface all along the flow)



Design Planning: few properties

- Incomplete design specification support (RTL + BlackBox)
- Industry std. constraints (.sdc)
- Fully technology aware flow (.lib/.lef)
- Fast synthesis and physical clustering
- Flexible stack configuration (.XML)
- Automated gate-level netlist partitioning
- Automatic inter-die net extraction
- Support for TSV/ μ bump clustering, P&R, technology features exploration
- Std cell placement & front/RDL routing
- Links to Thermo-Mechanical/Delay/Cost Compact Models



Design Planning Steps

1. Synthesis (& Flattening)

a) Design input

- Technology input (.lib/.lef → OpenAccess)
- RTL (or gate-level) supporting black box view
- Design constraints supplied using .sdc

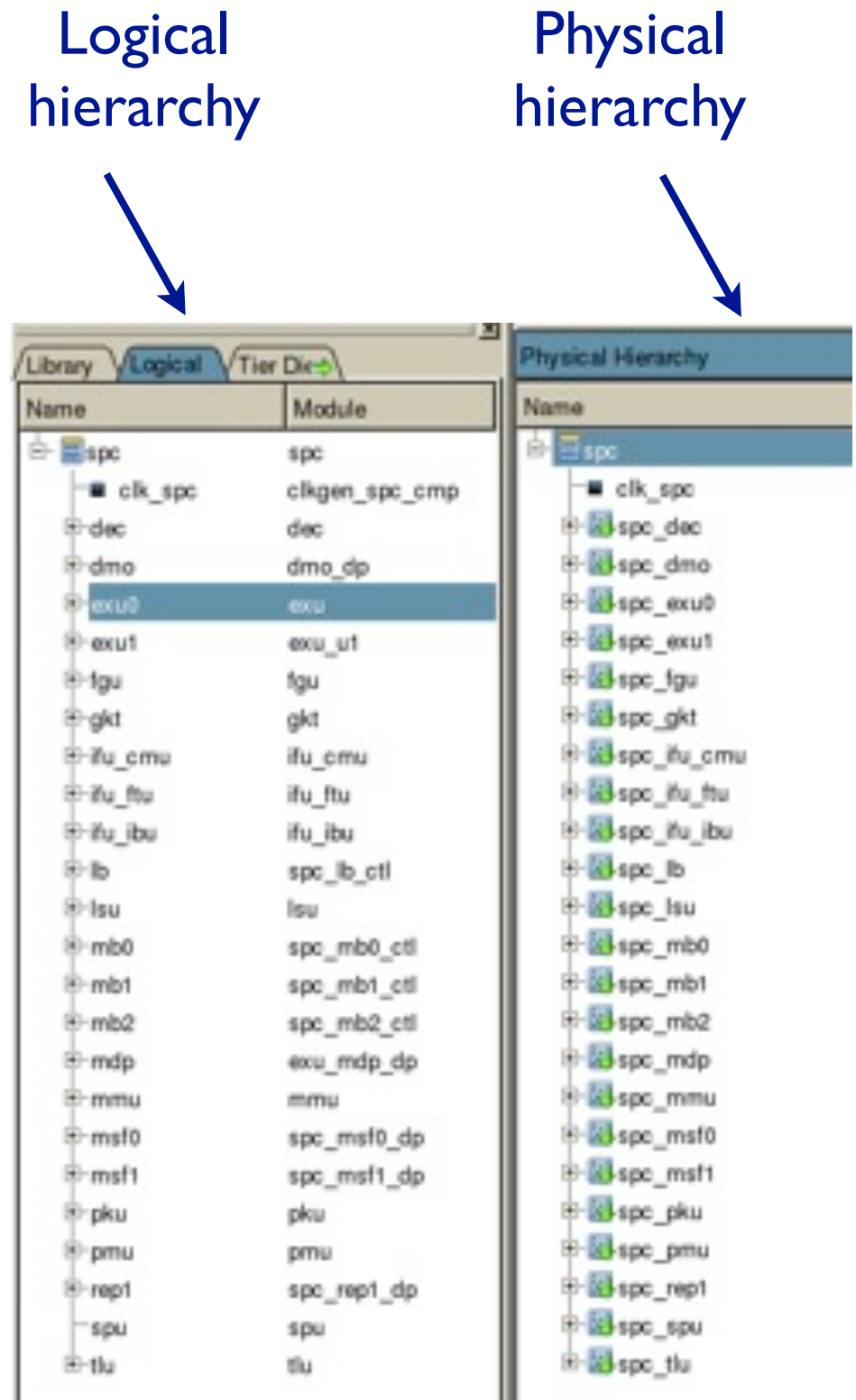
b) Synthesis

- Creates abstracted data flow operators (to accelerate synthesis process)
- Generates internal area/timing for these
- Support for Design Ware components

c) Flattens initial logical view to produce physical model of the design

**Produces physical hierarchy :
flattened gate-level netlist &**

**In parallel: original logical view
→ new hierarchies could be explored**



Design Planning Steps

1. Synthesis

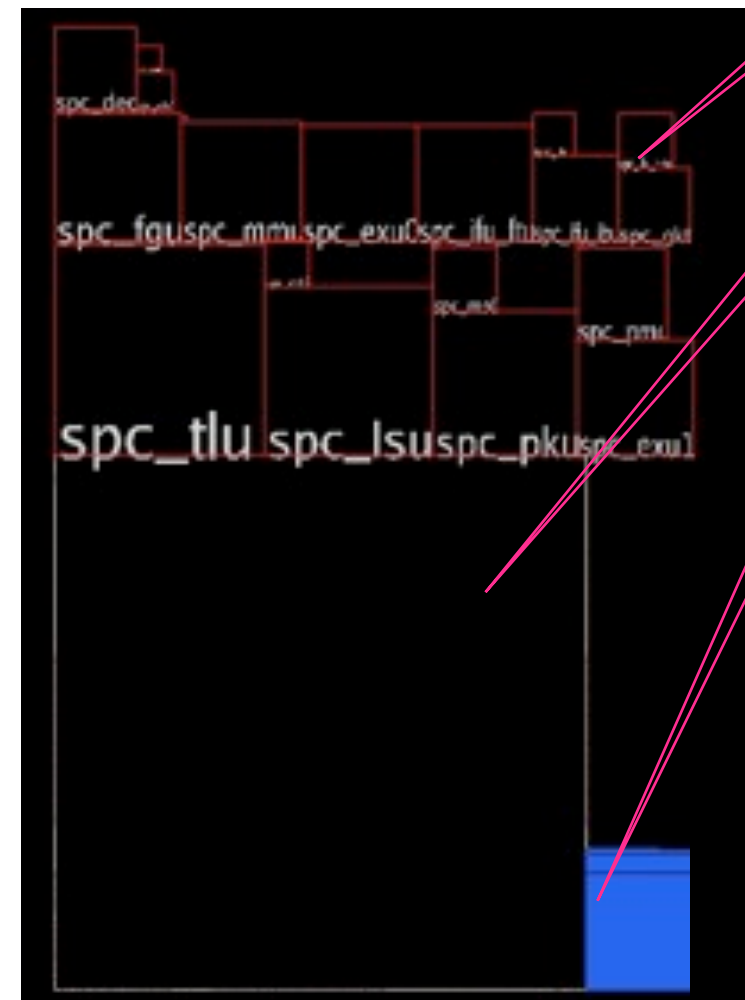
2. Clustering

- a)** Logical hierarchy is not necessarily good for physical implementation
- b)** Automated physical hierarchy generation from flattened netlist
- c)** Support for semi-automated hierarchy generation to include user defined exploration

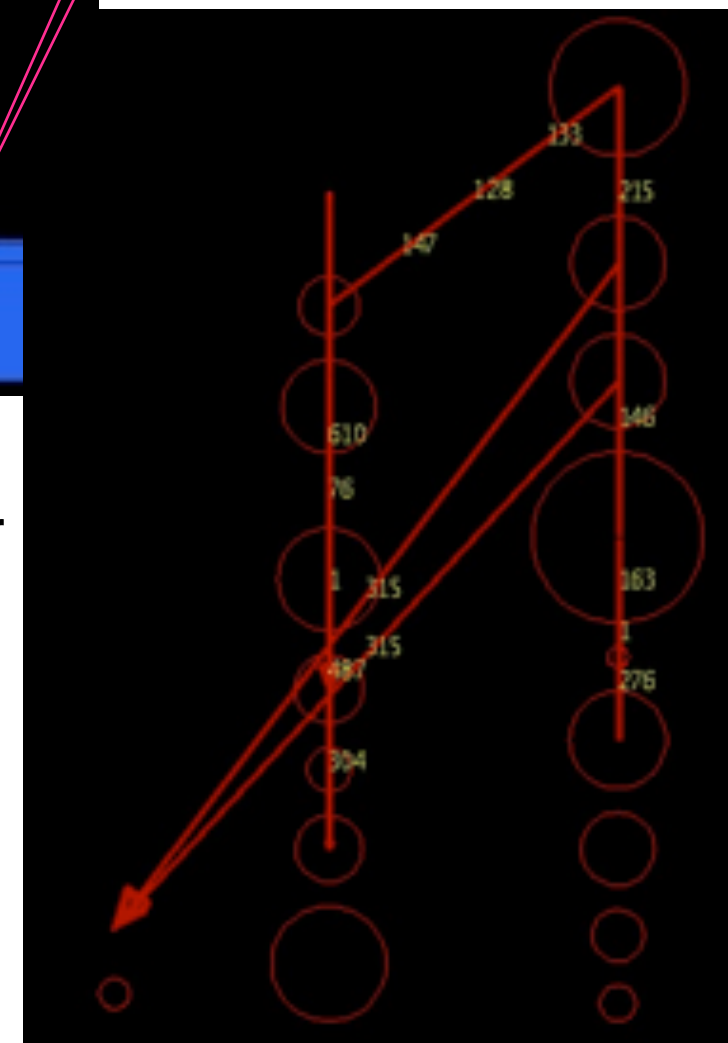
- **Produces clustered view of the design to simplify and improve floorplanning, place&route**

-Support for design analysis

- **Simplifies multi-die design partitioning**



Clusters connectivity diagram → each cluster is a circle, arrows indicate the weight of the connectivity (here arrows show are those with > 300 nets)



Clusters

Floorplan area

Standard cells

Black boxes

Design Planning Steps

1. Synthesis

2. Clustering

3. Design Partitioning

- a) Stack configuration (XML)
- b) Component assignment to Dies on Tiers (@ cluster/comp. level)
- c) Automatic generation of 3D Physical Components
- d) Optional TSV/ μ bump clustering
- e) Explicit placement of ubumps

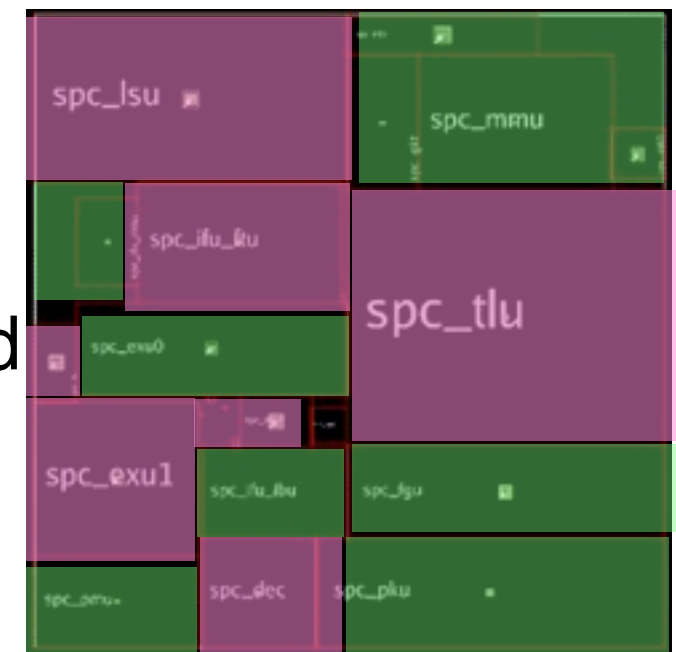
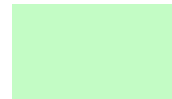
Produces partitioned gate-level netlist and creates appropriate models of inter-die nets

2D Design to be partitioned into 2 dies :

Die0



Die1



```
<tier llx="0" lly="0" z="20" urx="19000" ury="19000" name="Tier0" >
  <die name="Die0">
    <techno tech="tcbn40lpwupuc" />
    <type interposer="false" back_side_rdl="true" />
    <dimension llx="0" lly="0" urx="19000" ury="19000" />
    <origin x="0" y="0" />
    <utilization util="0.7" aspect_ratio="1.0" use_util="true" />
    <orient face_up="false" />
    <rdl layers="2" width="1.0" pitch="2.0" />
    <tsv width="5" pitch="5" />
    <bump pins width="20" pitch="20" />
  </die>
</tier>
```

```
load_xml_file -file $(TCL)/3D_face_to_face.xml
add_to_die -inst_list wDRAM_i -tier Tier1 -die Die1

add_to_die -inst_list ccu ccx dbg0 dbg1 dwu efu esd_core0 esd_core1 esr fsr_bottom fsr_left fsr_
0 12t1 12t2 12t3 12t4 12t5 12t6 12t7 mac mcu0 mcu1 mcu2 mcu3 mio n2_clk_gl_cust n2_revid_cust no
s tw0 tw1 -tier Tier0 -die Die0

report sgp3D.connectivity -rpt rpt/conn1.rpt
report sgp3D.area -rpt rpt/area1.rpt
report sgp3D.tier_die -rpt rpt/tier_die.rpt
report sgp3D.bump_pins -die Die0 -tier Tier0 -tcl_cmd -rpt bp.rpt1

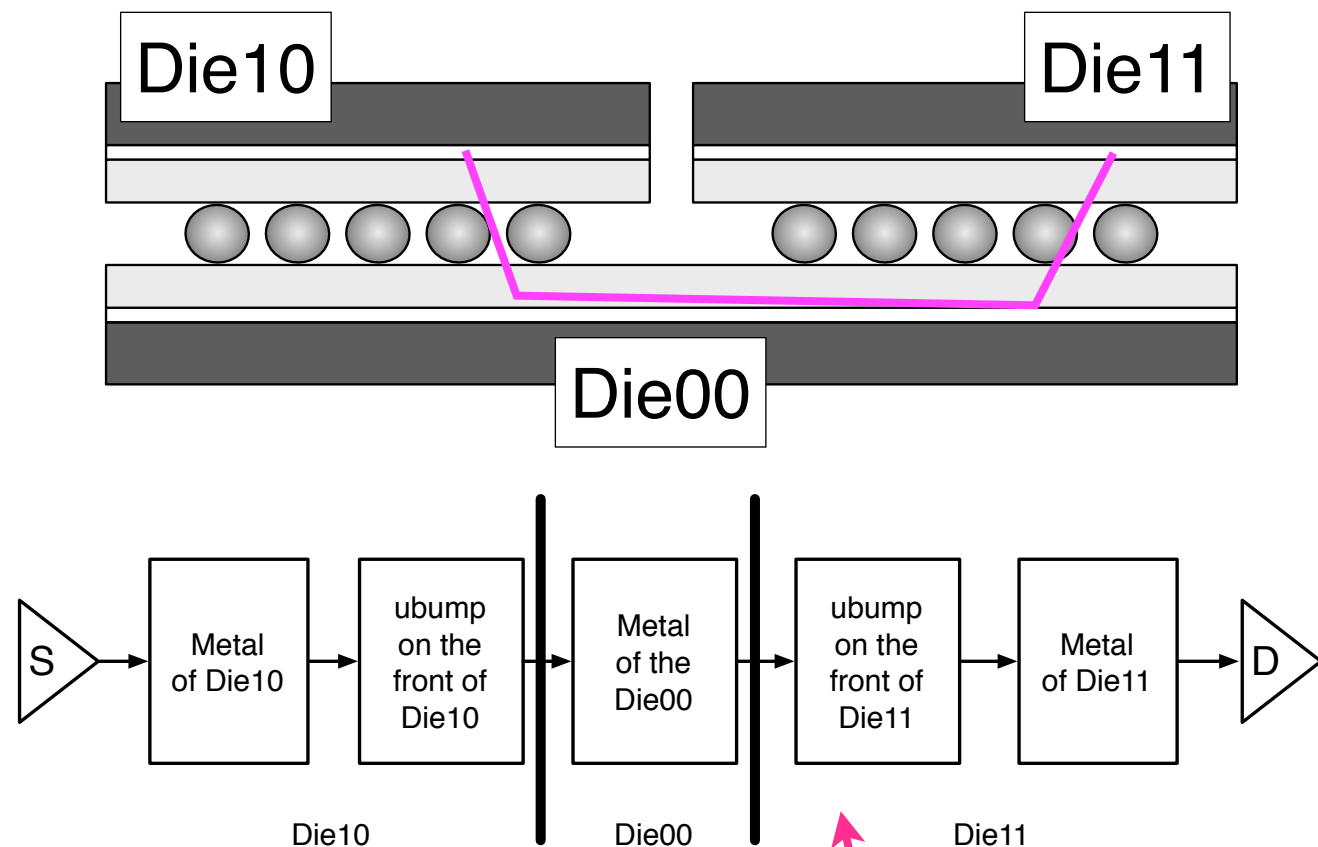
c) ## Automated creation of all 3D net components
create_3D_physical_components

## Clustering of the 3D components not to overload floorplanner with all small 3D components
## TSVs on the bottom die
d) source $(TCL)/Cluster_TSVs_logic_3D.tcl
## ubumps on the memory die
source $(TCL)/Cluster_ubumps_wDRAM_3D.tcl

e) ## Place explicitly the position of the ubumps of the top die
source $(TCL)/Placeubumps_wDRAM_3D.tcl
```

Automatic generation of Inter-die nets

Face-to-face interposer

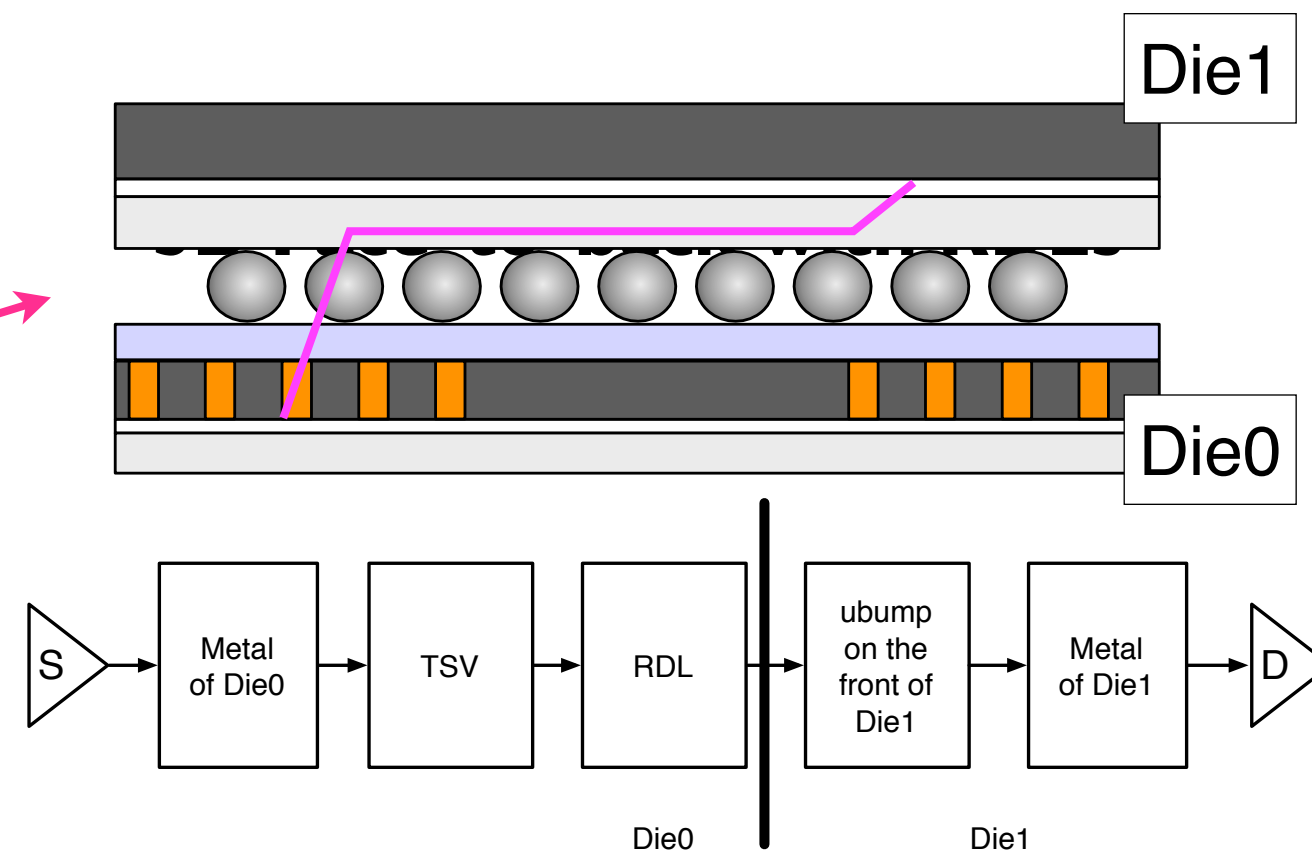


- Stack config defined through xml file
- The xml file is read before partitioning
- During partitioning this info is used to generate appropriate physical model of the net

Two identical nets, two completely different physical nets → this needs to be automated !

Generation of these 3D components → one command:

`create_physical_3D_components`



Design Planning Steps

1. Synthesis

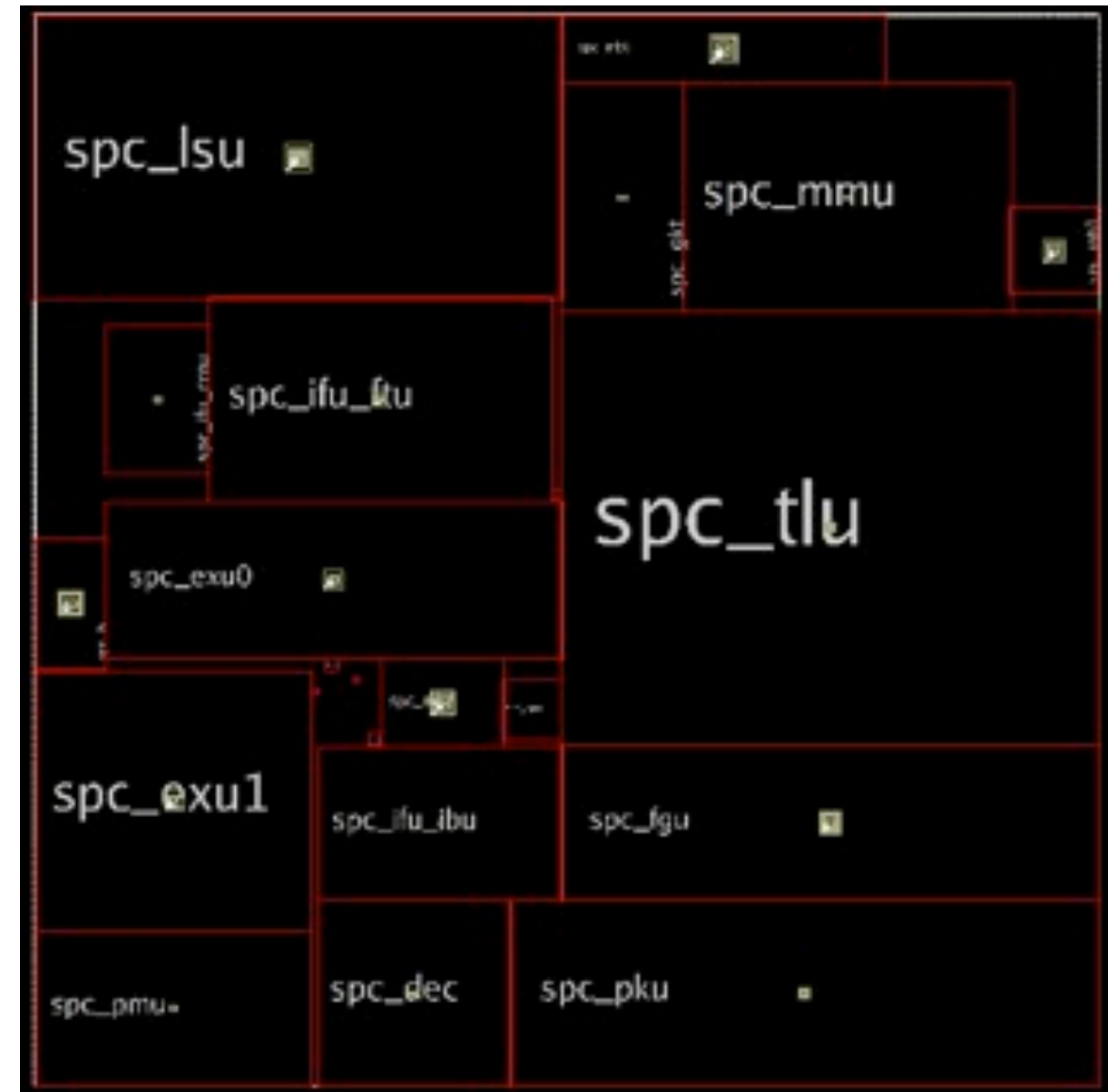
2. Clustering

3. Design Partitioning

4. 2/2.5/3D Floorplanning

- a) On per die basis, supporting an arbitrary number of tiers and dies on tiers
- b) Simultaneous (true 3D) floorplanning can be obtained with appropriate constraints generation (currently manual)
- c) Fast: minutes, few tens of minutes, depending on design complexity and constraints

OpenSPARC core



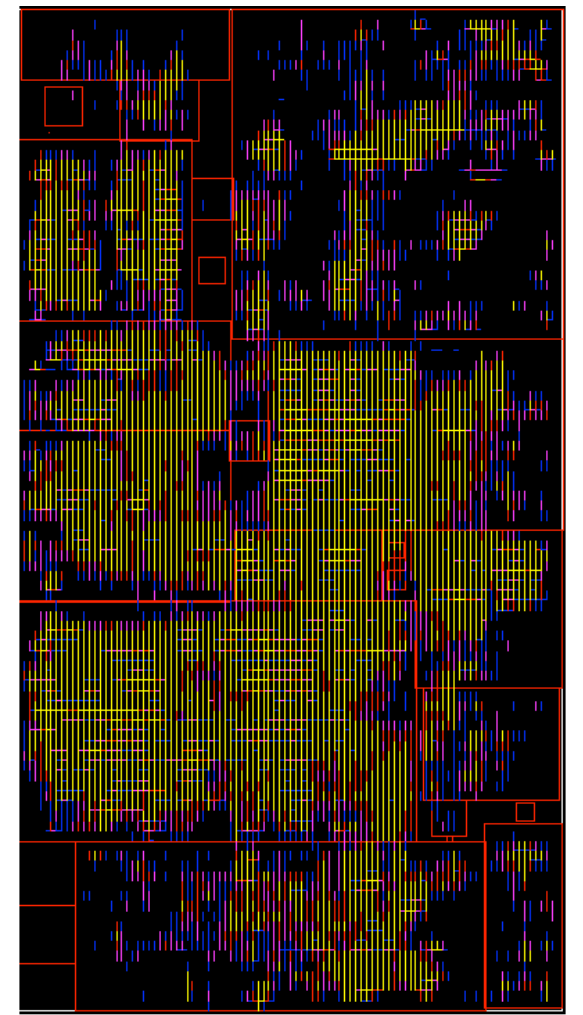
Design Planning Steps

1. Synthesis
2. Clustering
3. Design Partitioning
4. 3D Floorplanning
5. Placement & Routing
 - a) Standard front side
 - b) Back-side routing with flexible technology input : on the fly number of the RDL number & width/spacing specification
 - c) Congestion analysis

Produces layout and all associated metrics (geometry, delay, power,...)



3. Congestion

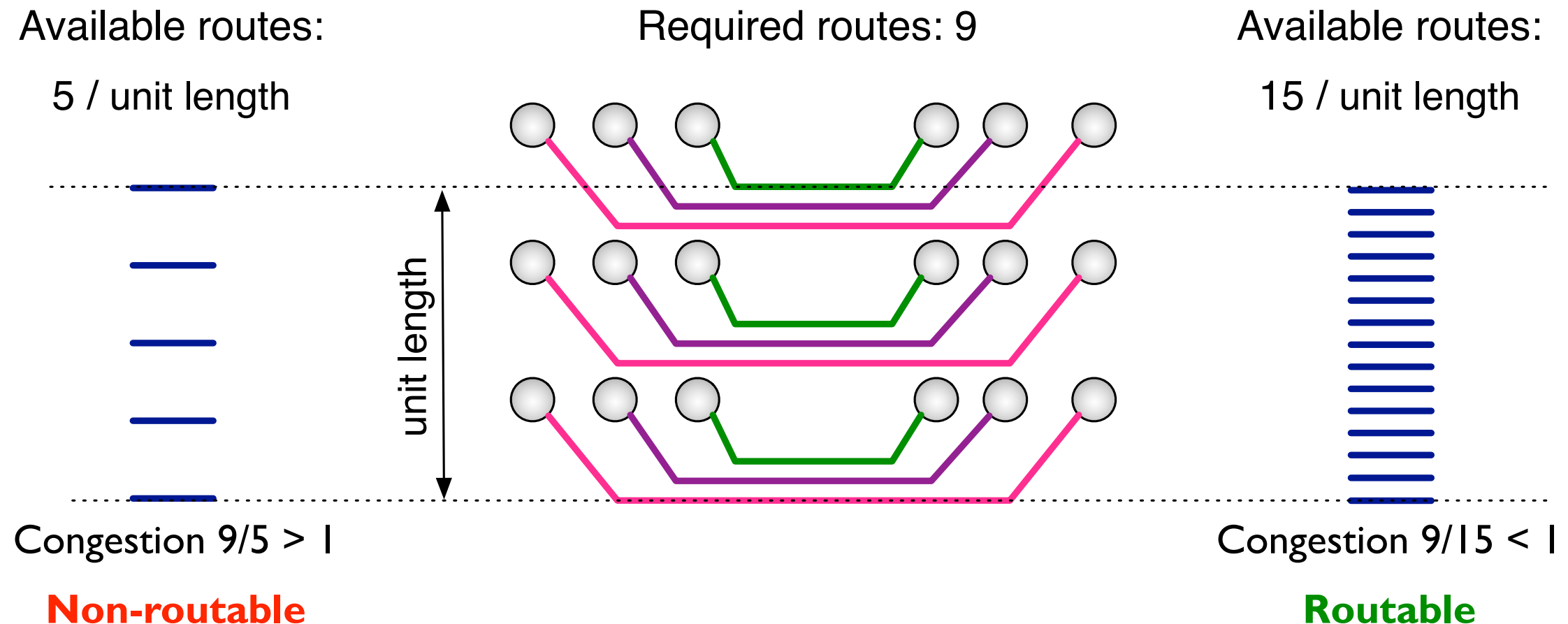


Congestion analysis: Why doing it?

- Area estimations of the design based on: gate count and area products are NOT accurate since they do not take the routing into account
- The routing will add extra space depending on:
 - ✦ Technology parameters,
 - ✦ design complexity,
 - ✦ physical implementation strategy
- **Congestion** is directly related to
 - ✦ Design cost (area)
 - ✦ Effort spent to optimize design (on both functional and physical implementation aspects)

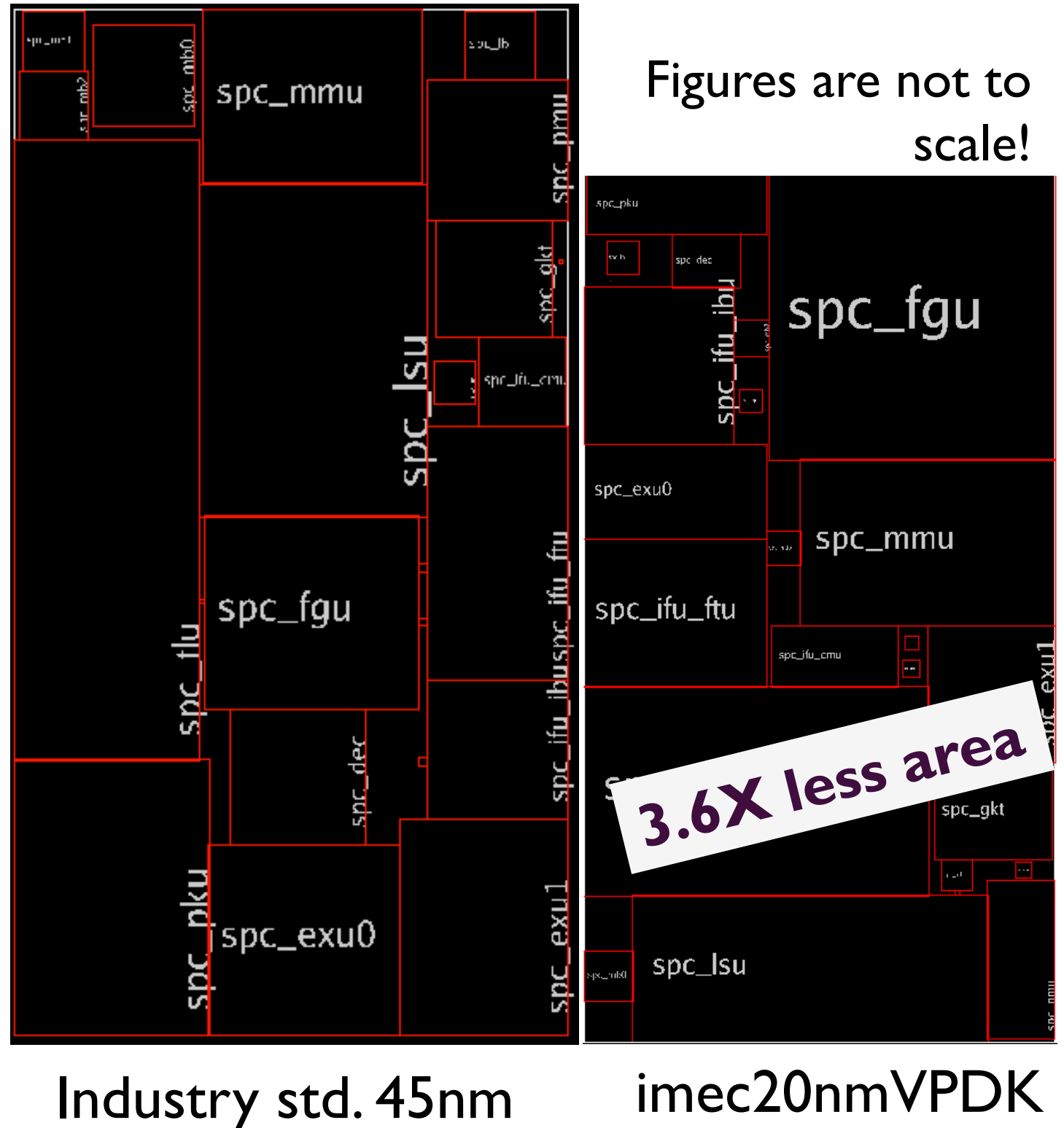
Congestion defined

- Congestion is defined as ratio between: the number of routes **demanded** by the design & number of routes **supplied** by the technology
- If the **demand** > **supply** → congestion, design not routable
- If **demand** < **supply** → OK



2D illustration: OpenSPARC core

- ▶ Design OpenSPARC
 - Core only
- ▶ Synthesis, floorplan place & route with industry proven 45nm and VPDK 20nm
- ▶ Congestion analysis
 - Ratio between required and available routing resources for a grid of a fixed size
 - Global congestion score
 - > 10 non-routable
 - < 10 routable



2D: OpenSPARC core

Impact of new tech on routability

75%

88%

63%

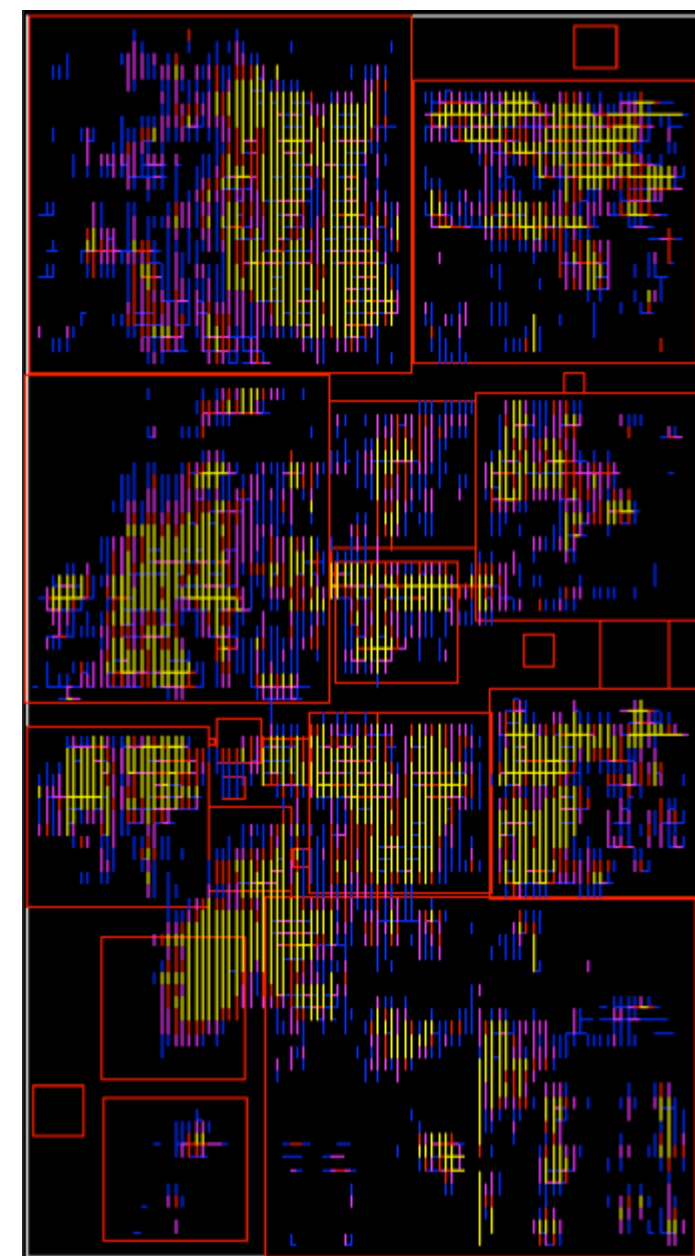
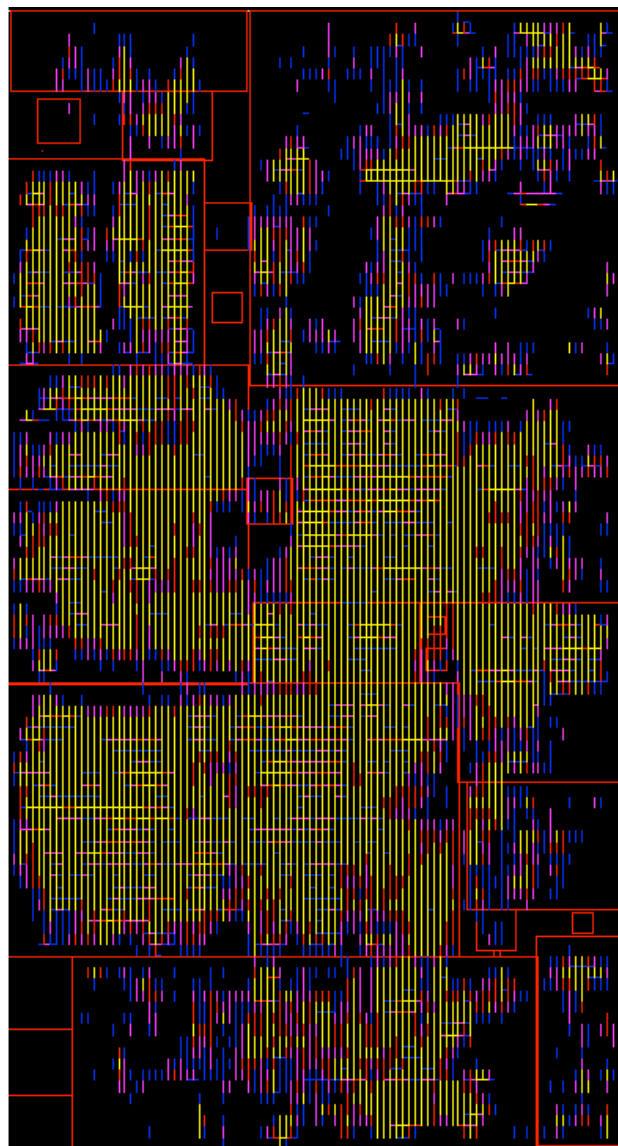
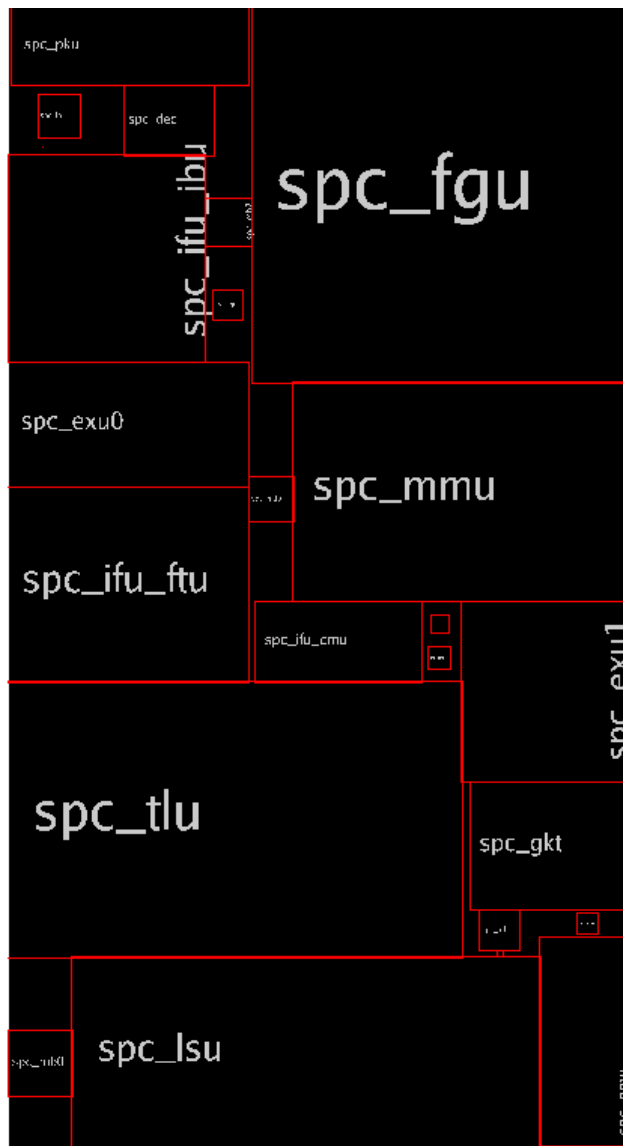
100%

Floorplan using
20nm IMEC
Virtual PDK

9 metal layers
Congestion
score= 15

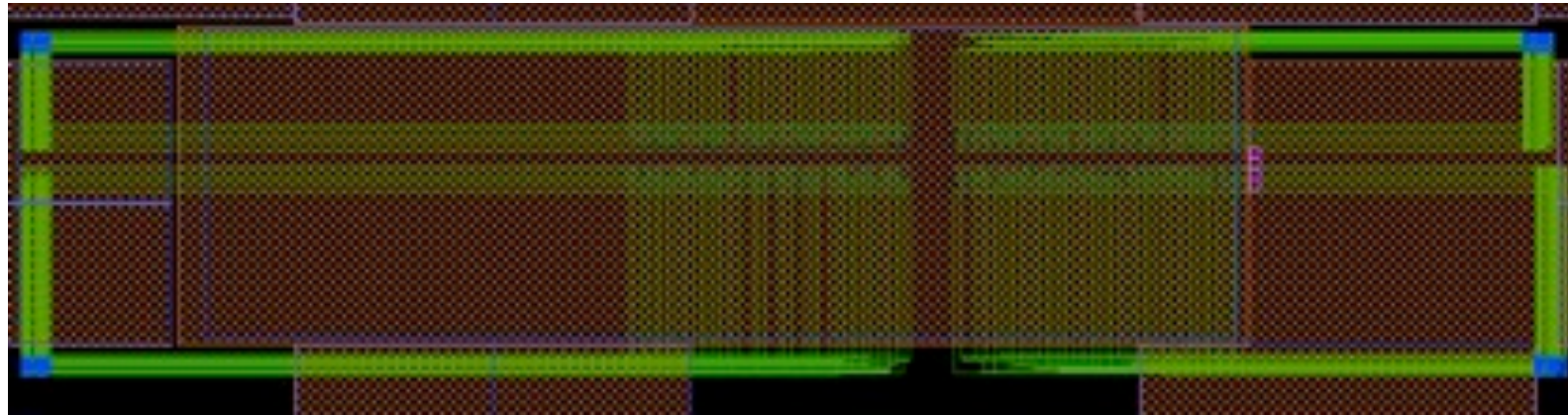
11 metal layers
Congestion
score = 10

9 metal layers
+ 20% area
Congestion
score = 8



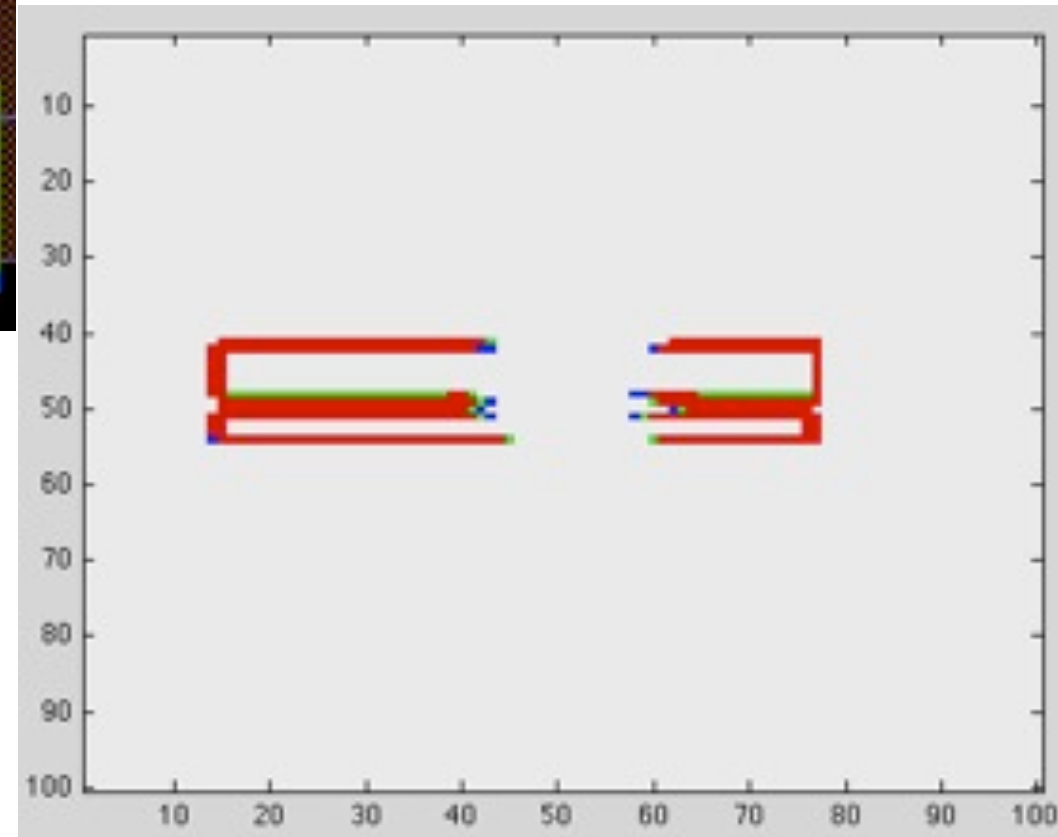
Back side congestion

B in ASCII map
(or RED in color map)
Congestion more then
120%



Congestion map (as ASCII or color map):

- For 2 RDL layers and 10 μm pitch
unroutable !
- For 2 RDL layers and 3 μm pitch no problems

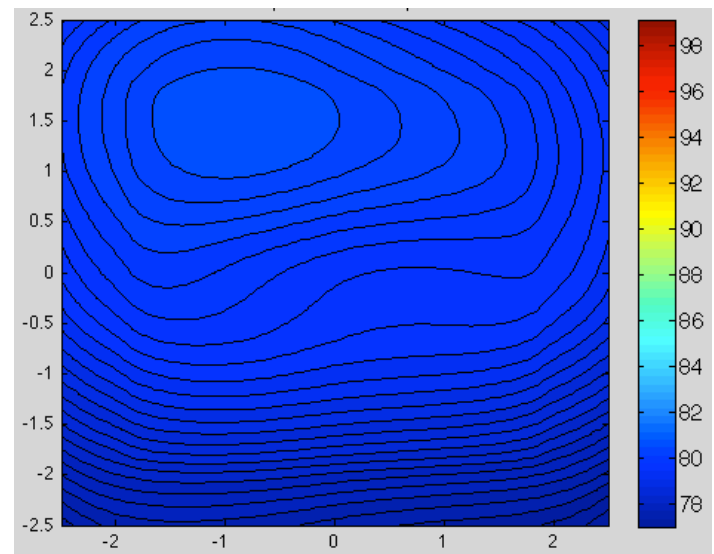
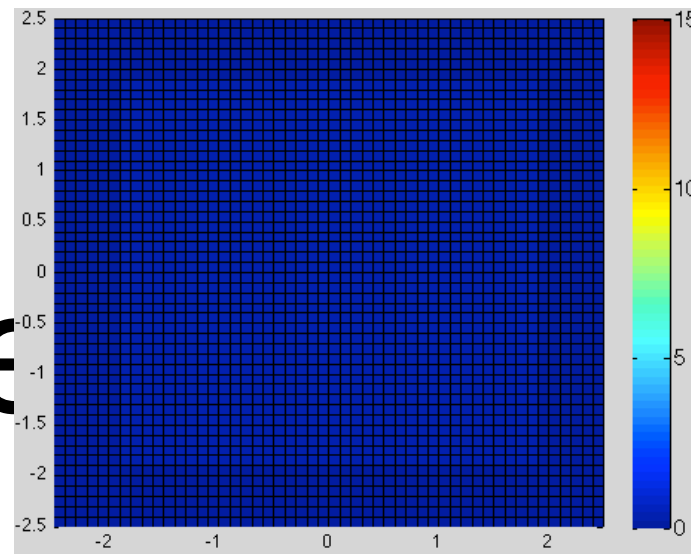
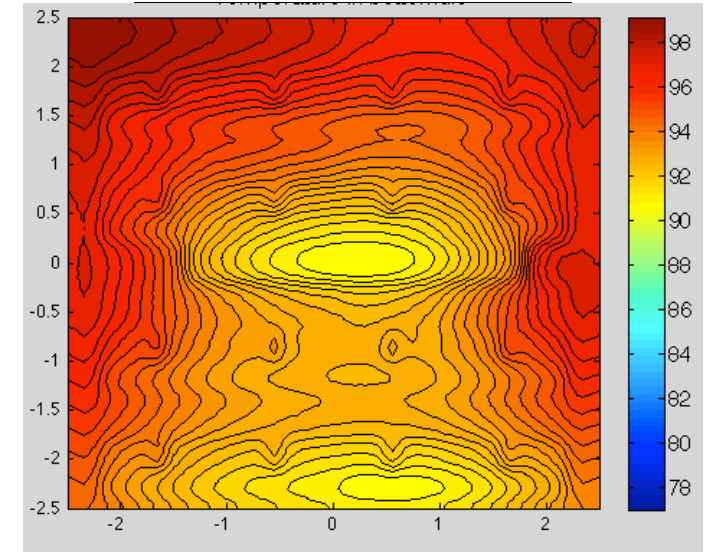
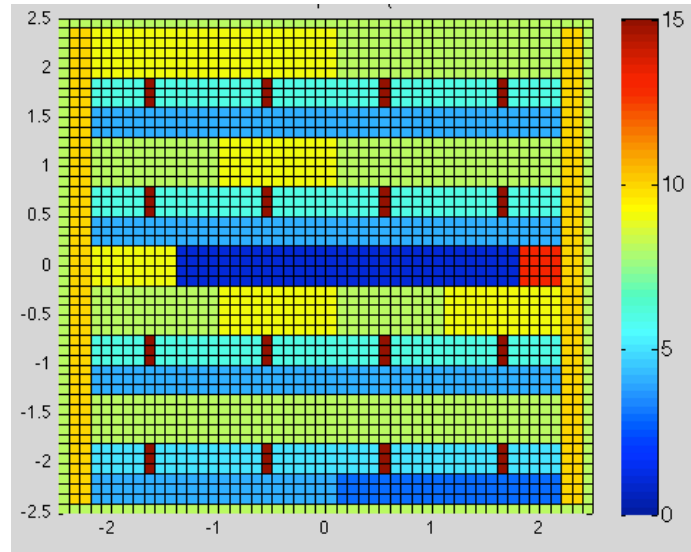


```

.....BBBBBBBBBBBBBBBBBBBBBBBBBBBBVV.....VBBBBBBBBBBBBBBBBB.....
.....BBVVVVVVVVVVVVVVVVVVVVVVVVVVVVBBVxx.....xVBBBVVVVVVVVVVVVVVB.....
.....VB.....B.....
.....VB.....B.....
.....VB.....B.....
.....VB.....B.....
.....VB.....B.....
.....BBvvvvvvvvvvvvvvvvvvvvvvvvvvVV.....xxVBBVBvvvvvvvvvvvvvvVB.....
.....BBBBBBBBBBBBBBBBBBBBBBBBBBBBBVvx.....vVVVBBBBBBBBBBBBBBBB.....
.....BBBBBBBBBBBBBBBBBBBBBBBBBBBVvx.....xvVVVVVVVVVVVVVV.....
.....BBBBBBBBBBBBBBBBBBBBBBBBBBBVvx.....xvBBBBBBBBBBBBBBBBBBBB.....
.....VB.....BB.....
.....BB.....BB.....
.....xBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBVV.....vVBBBBBBBBBBBBBBBBBB.....
.....
.....

```

4. The model



Motivation: Thermal issues are scary, especially in logic-on-logic 3D-stacks

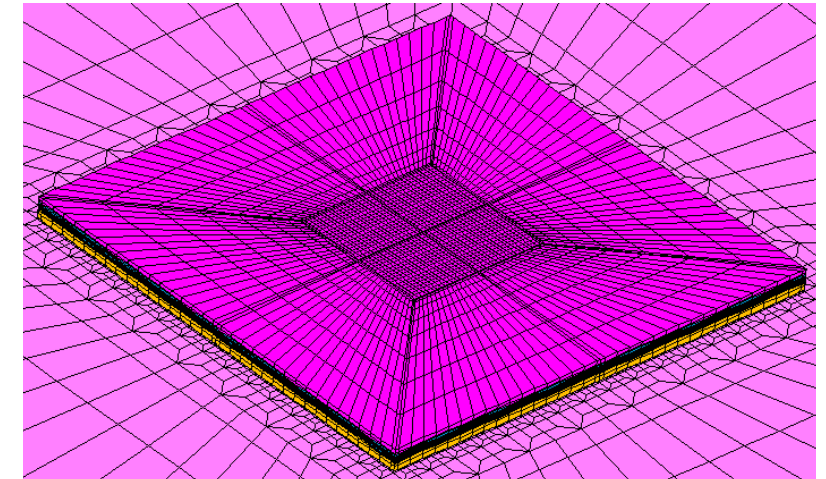
- Vertical superposition of heat-dissipation devices in a 3D stack will have a strong impact on the maximum temperature reached in the dies
- This could be solved:
 - ✦ At **system level** — correct floorplanning (impact on development cost) and architecture definition
 - ✦ Using appropriate **cooling solution** (impacts package/cooling cost)
- Obviously has to be checked early in the design cycle

**Need for fast thermal assessment tool
to co-optimize system architecture, (partition)
floorplan and cooling solution**

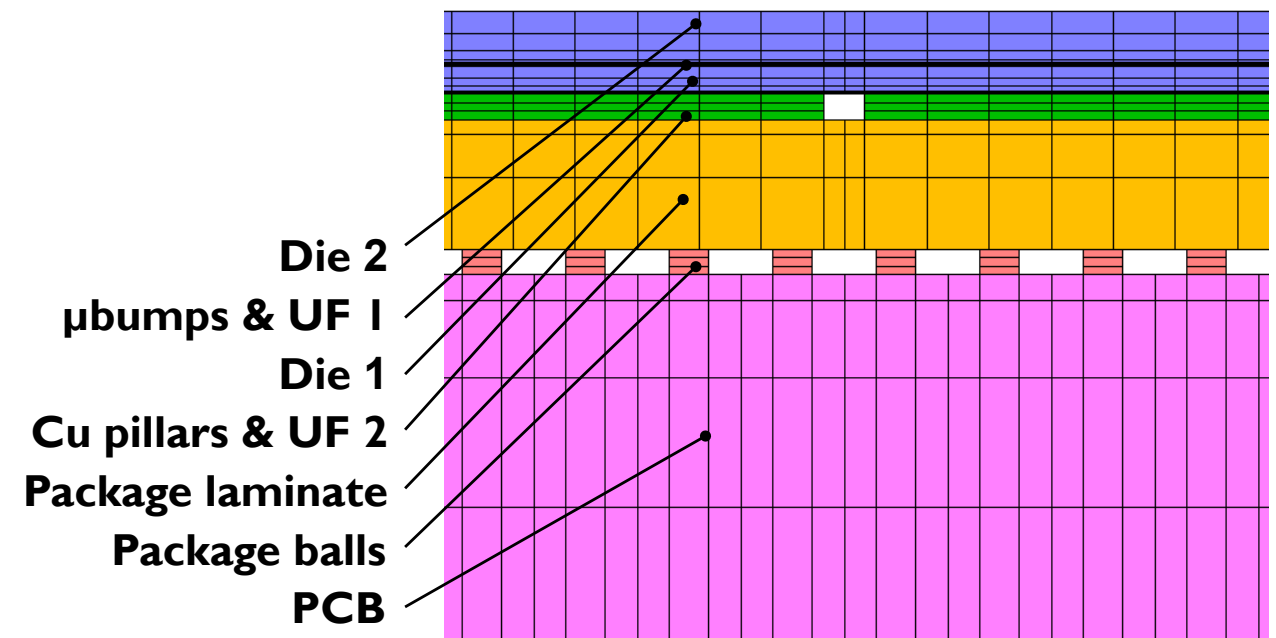
Followed approach

- Based on Finite Element modeling
- Built and simulated using industry standard software
- Steady state thermal analysis
- **Input**
 - ▶ Power dissipation
 - ▶ Material properties
 - ▶ Geometry
 - ▶ Boundary condition (cooling)
- **Output**
 - ▶ Temperature distribution
 - ▶ Thermal performance

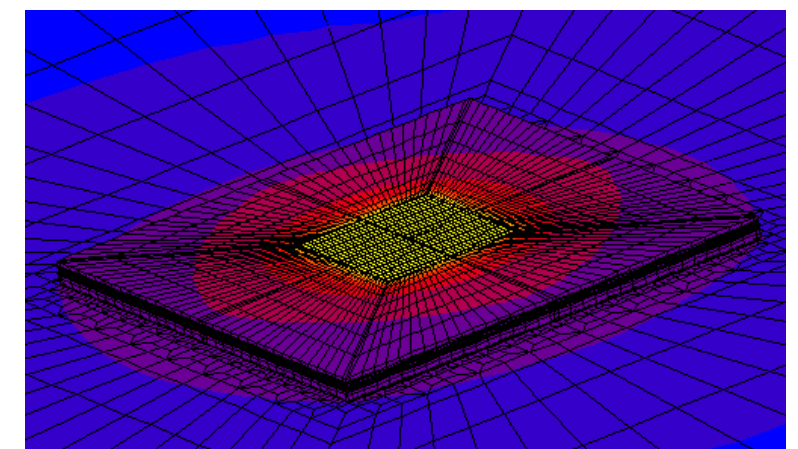
Finite element model



Stack cross section



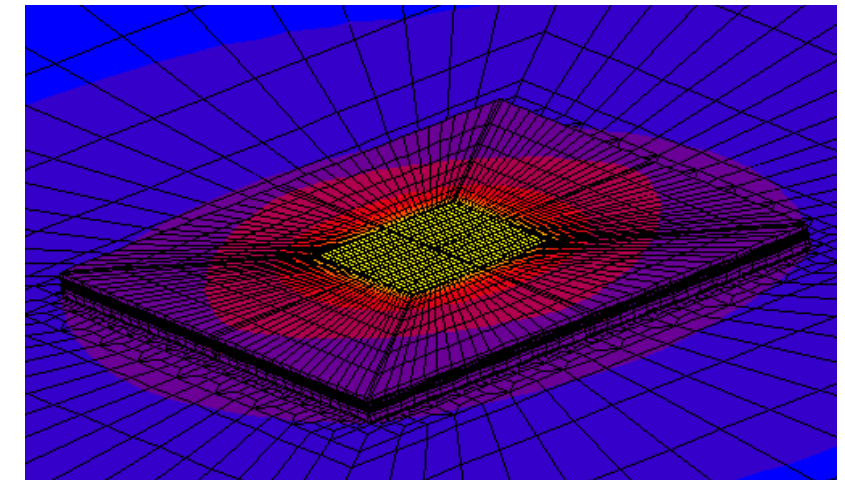
Thermal output



Finite element model vs. Compact Thermal Model

- **Finite Element (FE) modeling**

- **Advantages:** accurate, can be used for exhaustive analysis at expense of time



3D Stacked VP — Thermal Modeling

- **Disadvantages:** slow, requires some time to set-up, requires 3rd party license in order to run

- **Compact Thermal Model**

→ **Simplified**

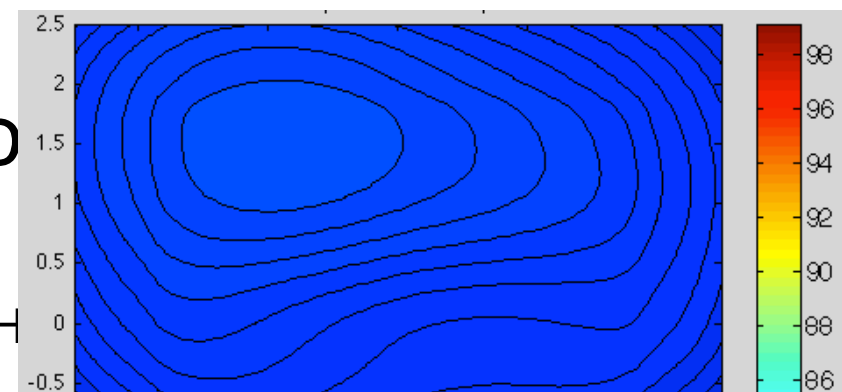
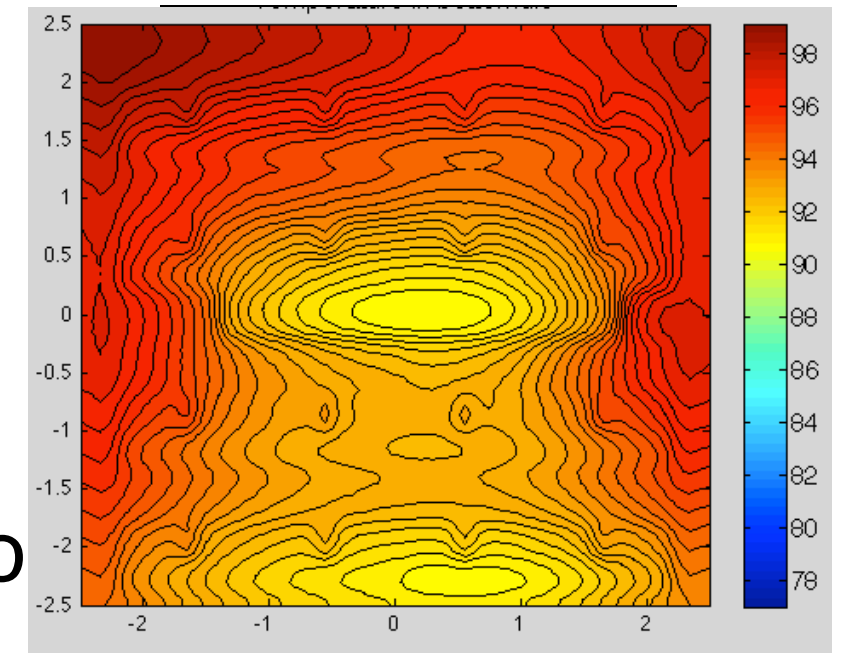
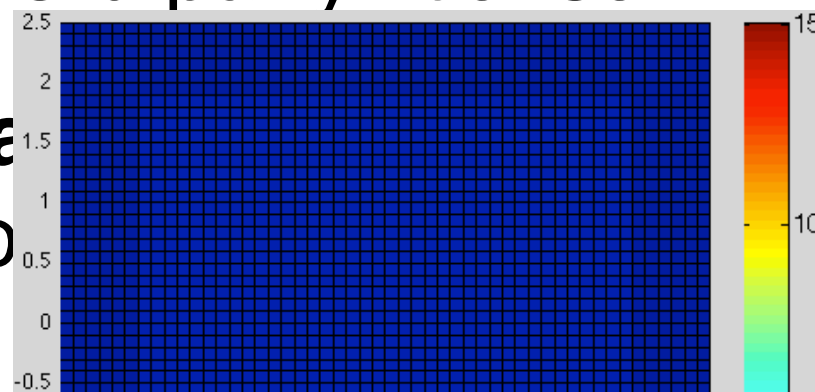
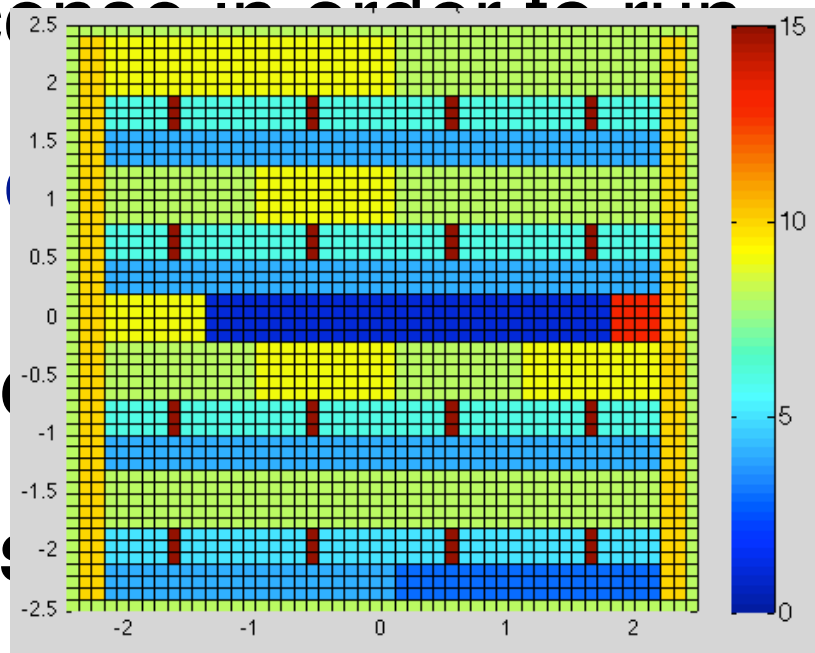
of the FE model

- **Advantages:**

no need for 3rd party license

- **Disadvantages:**

configuration



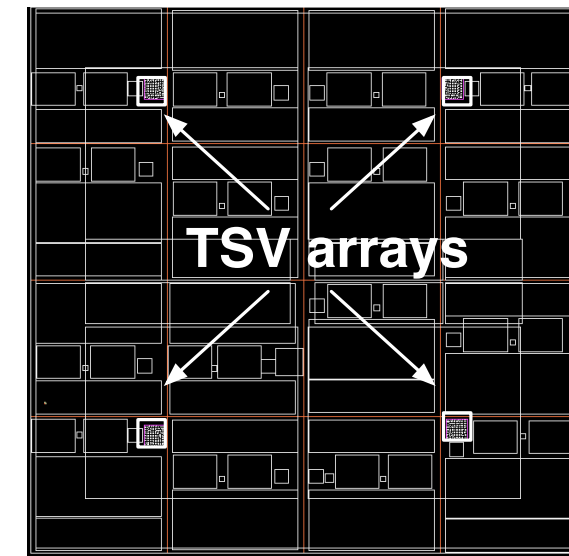
and only o

gomiir/ELEC-H

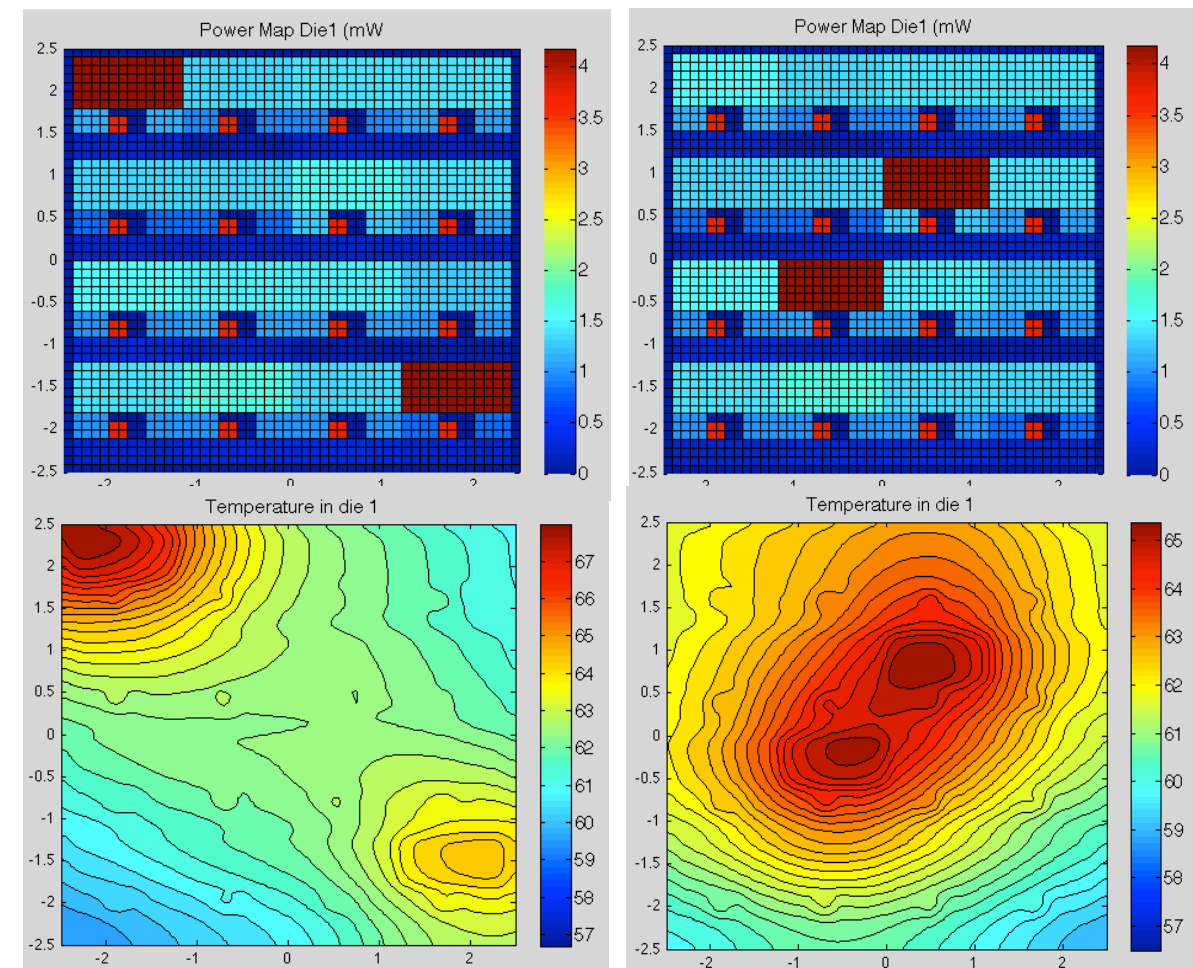
Compact Thermal Models: usage illustration

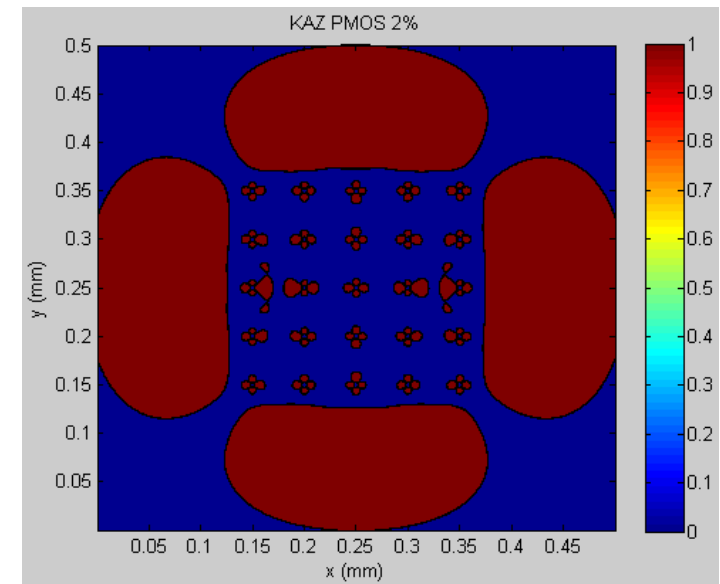
- WideIO DRAM memory on the top of the SoC logic die: low-power multi-processor system
- Design flow applied to extract:
 - ▶ Layout
 - ▶ Power dissipation of different components
- For a given stack configuration, thermal maps are computed to identify hot-spots and max stack temperature

SoC Layout
(logic die only)



Power density maps with
corresponding thermal profiles
for 2 different power scenarios





5. Mechanical modeling

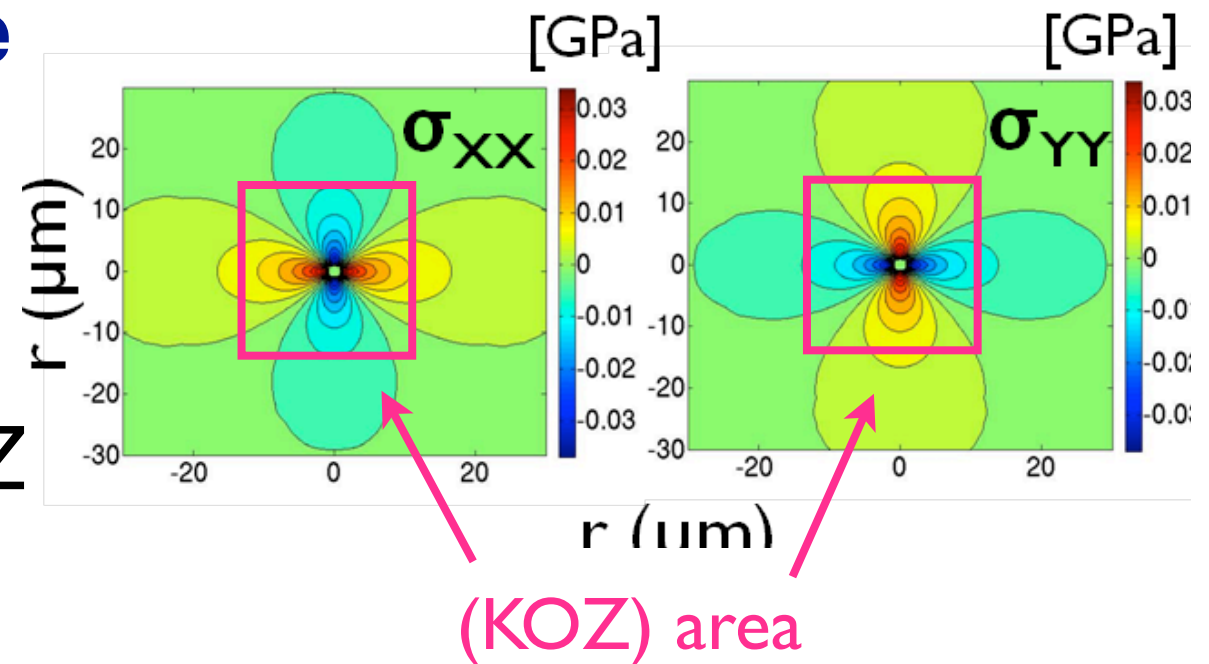
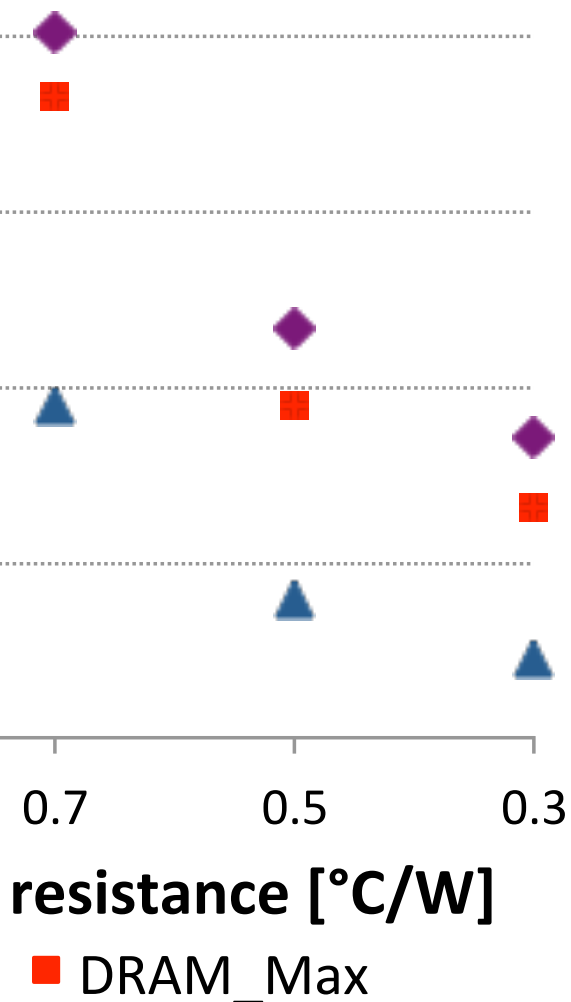
Compact Mechanical Modeling: Motivation

materials for TSVs, μ bumps, package
it Coefficient of Thermal Expansion (CTE)

CTE $\rightarrow$ shrinking & bending of the silicon
tress at μm , mm scale depending on distance

device performance

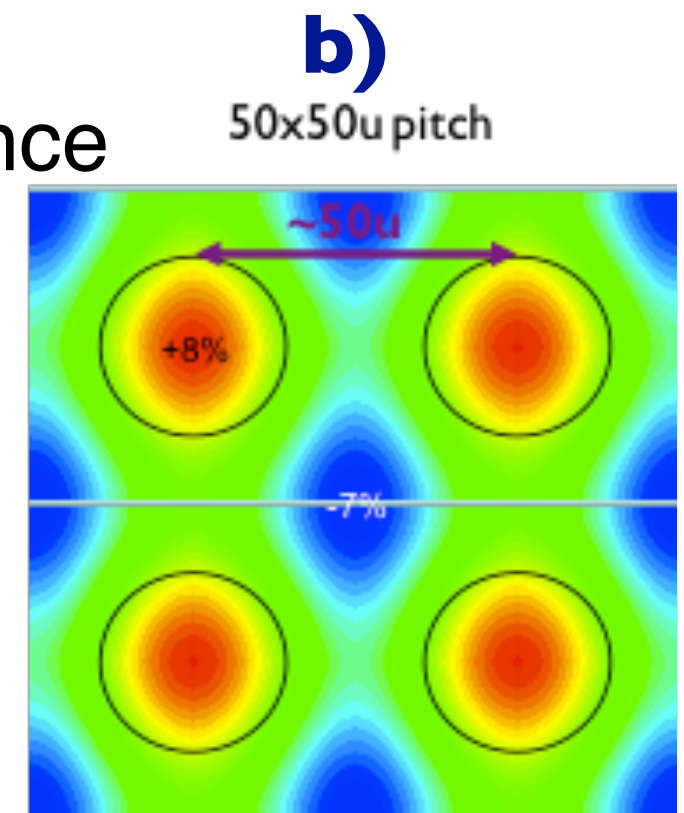
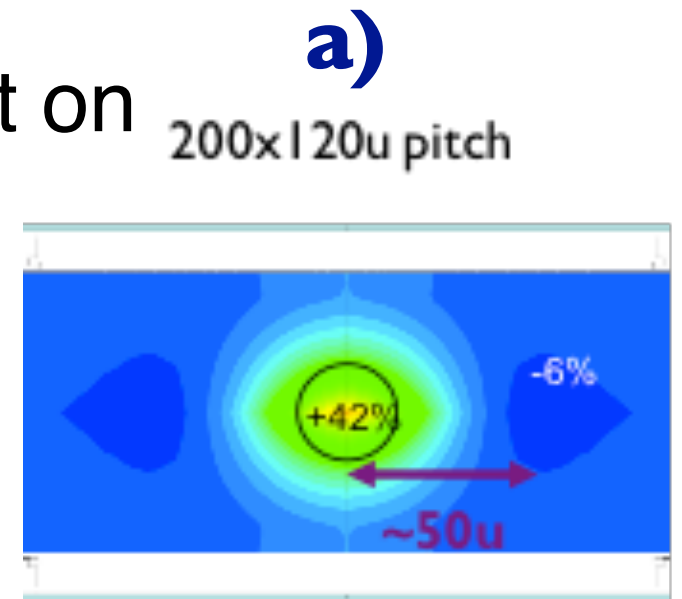
ven performance
of the device
ut Zone (KOZ) area
re devices inside KOZ



**estimation of the KoZ area is necessary
to produce meaningful layout !**

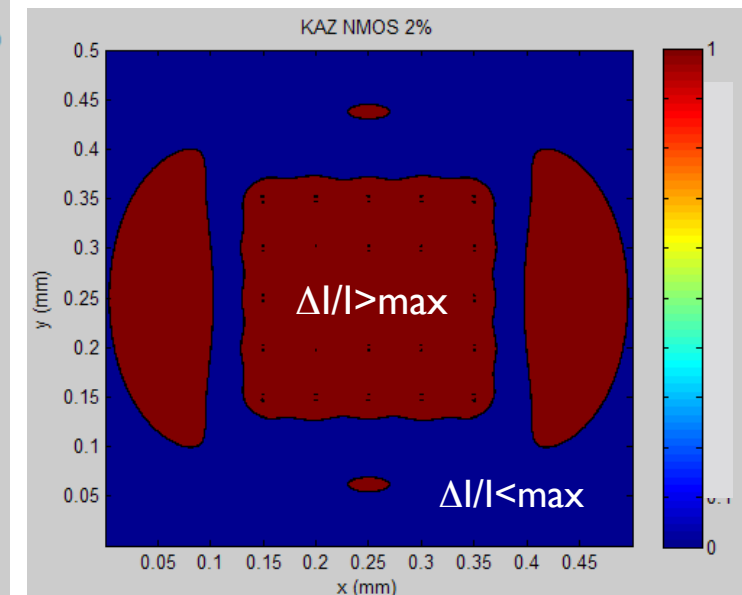
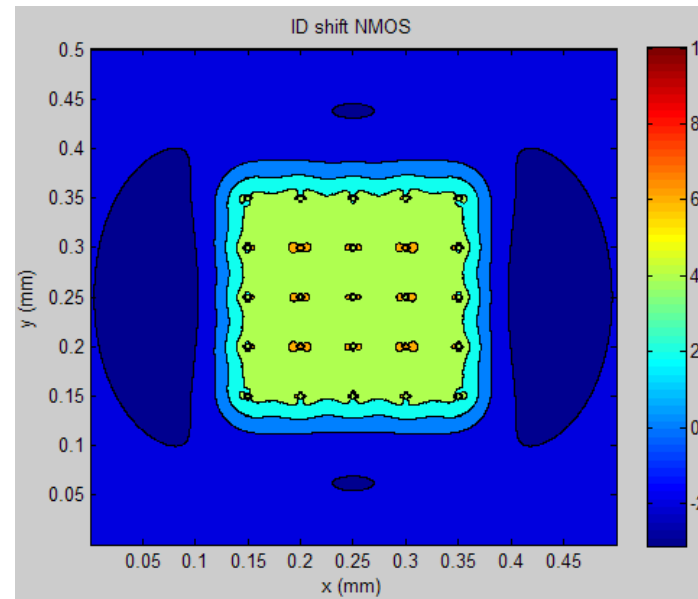
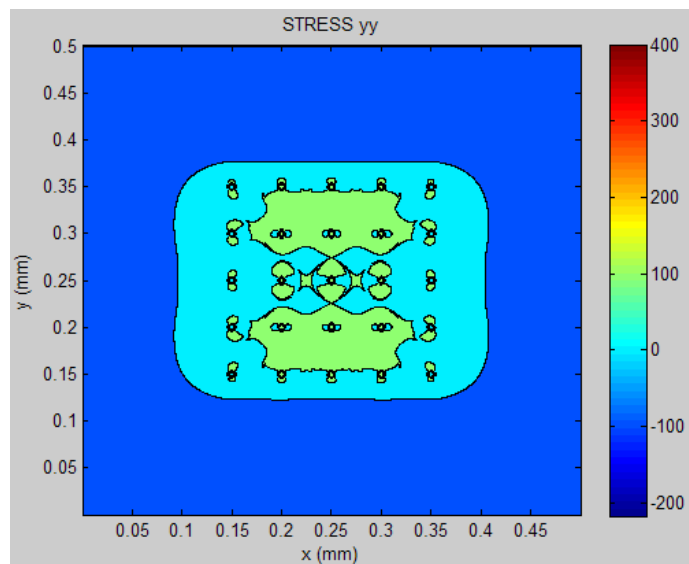
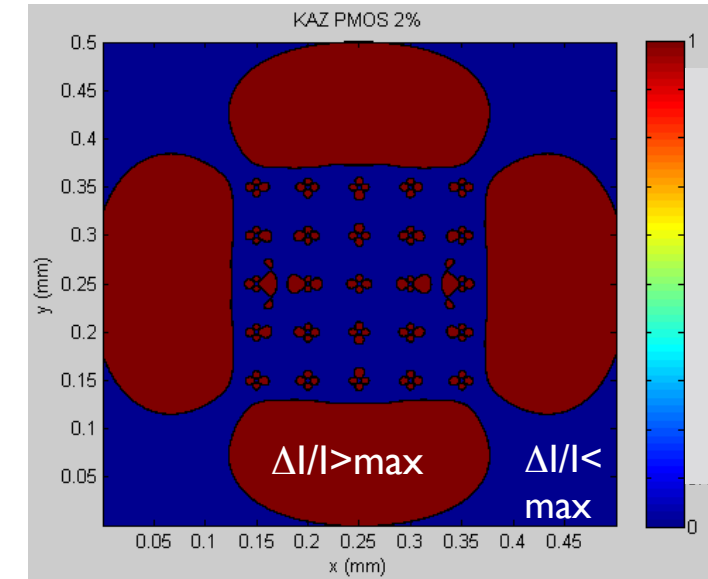
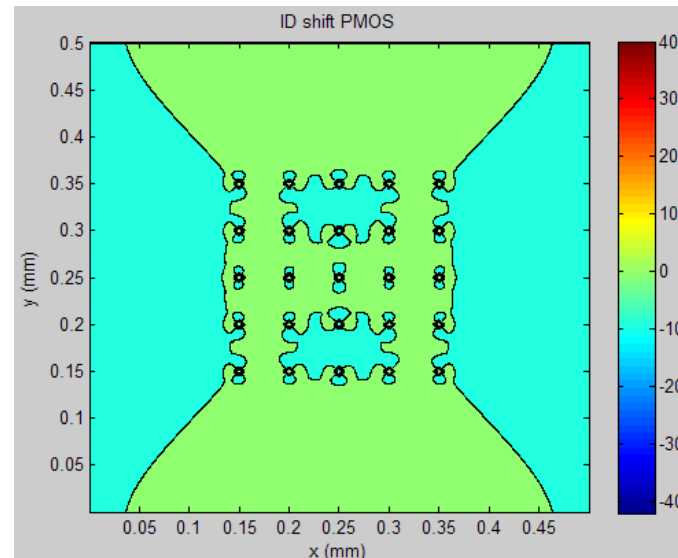
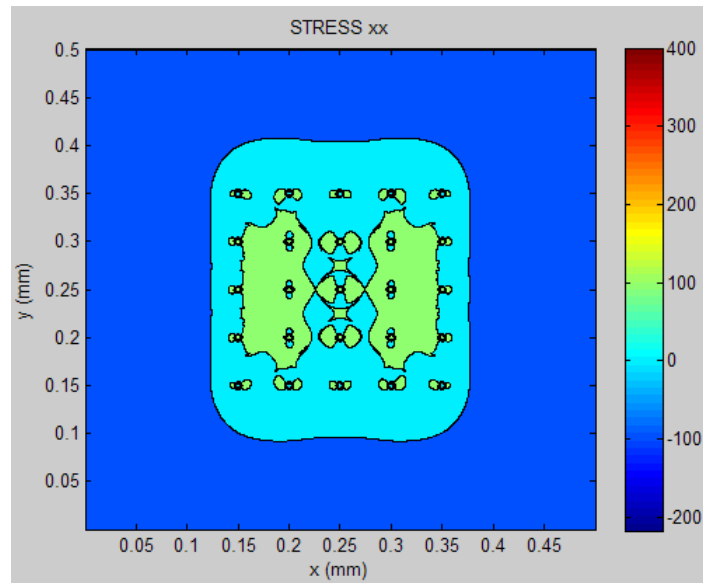
Compact Mechanical Modeling : Illustrations

- Examples of stress profiles and their impact on μ bump array planning : partitioning and placement
- Two different organizations:
 - ✦ **a) μ bumps at High-pitch**
 - ◉ μ bumps at $200\mu\text{m}$ pitch (isolated)
 - ◉ Extreme stress values are at $50\mu\text{m}$ distance
→ **Total Ion variation 48%**
 - ✦ **b) μ bumps at Low-pitch**
 - ◉ μ bump array at $50\times 50\mu\text{m}$ pitch
→ **Total Ion variation 15%**
- Low-pitch is $\sim 4\text{X}$ better ! because stress cancelation will take place



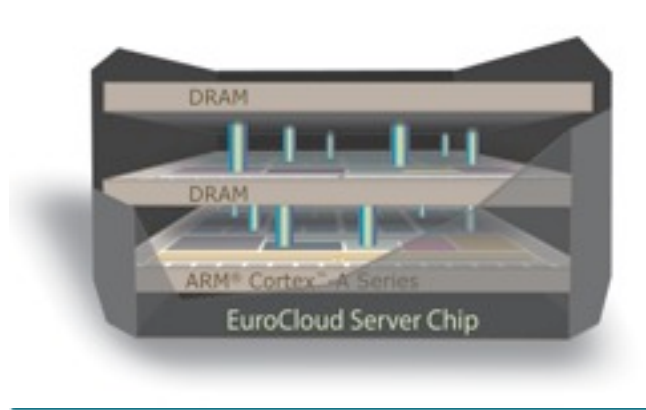
Compact Mechanical Model : Usage illustration

INPUT PARAMETERS: μ bump dimensions,
UF properties, temperature, TSV diameter, ...



OUTPUT PARAMETERS: KOZ size

Stress sensitivity (piezo-coefficients) of your transistor technology



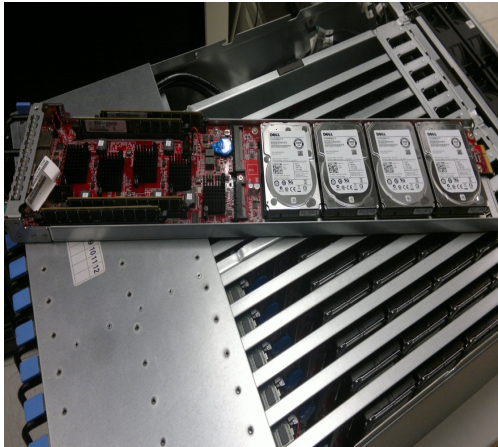
6. Practical application

3D-Integration & applications:

Scale-Out Processors

Data centers are power hungry !

Board



Rack



Building



... in all, thousands of CPUs using considerable power.

Did the BIG ones (MS, Yahoo, etc.)
became “GREEN” ?

\$\$\$ Electricity bill \$\$\$
In 2007 : 7.2 Billions US\$

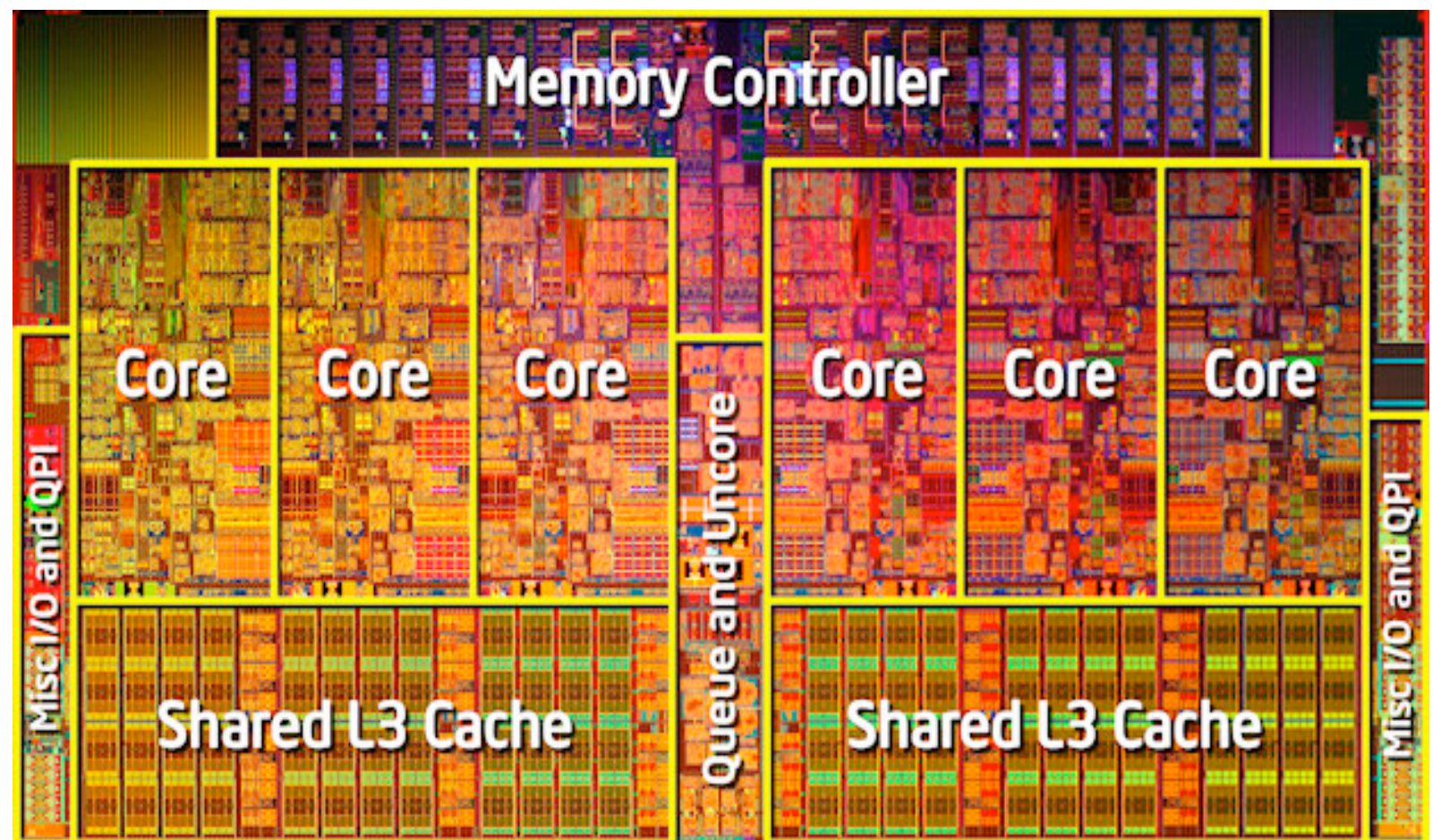


Data centers use traditional cores

- Heavily pipelined
- Bunch of FPUs
- SIMD support
- Big, shared caches
- Complex circuits, built to suit any application ...
(as long as it is not embedded)

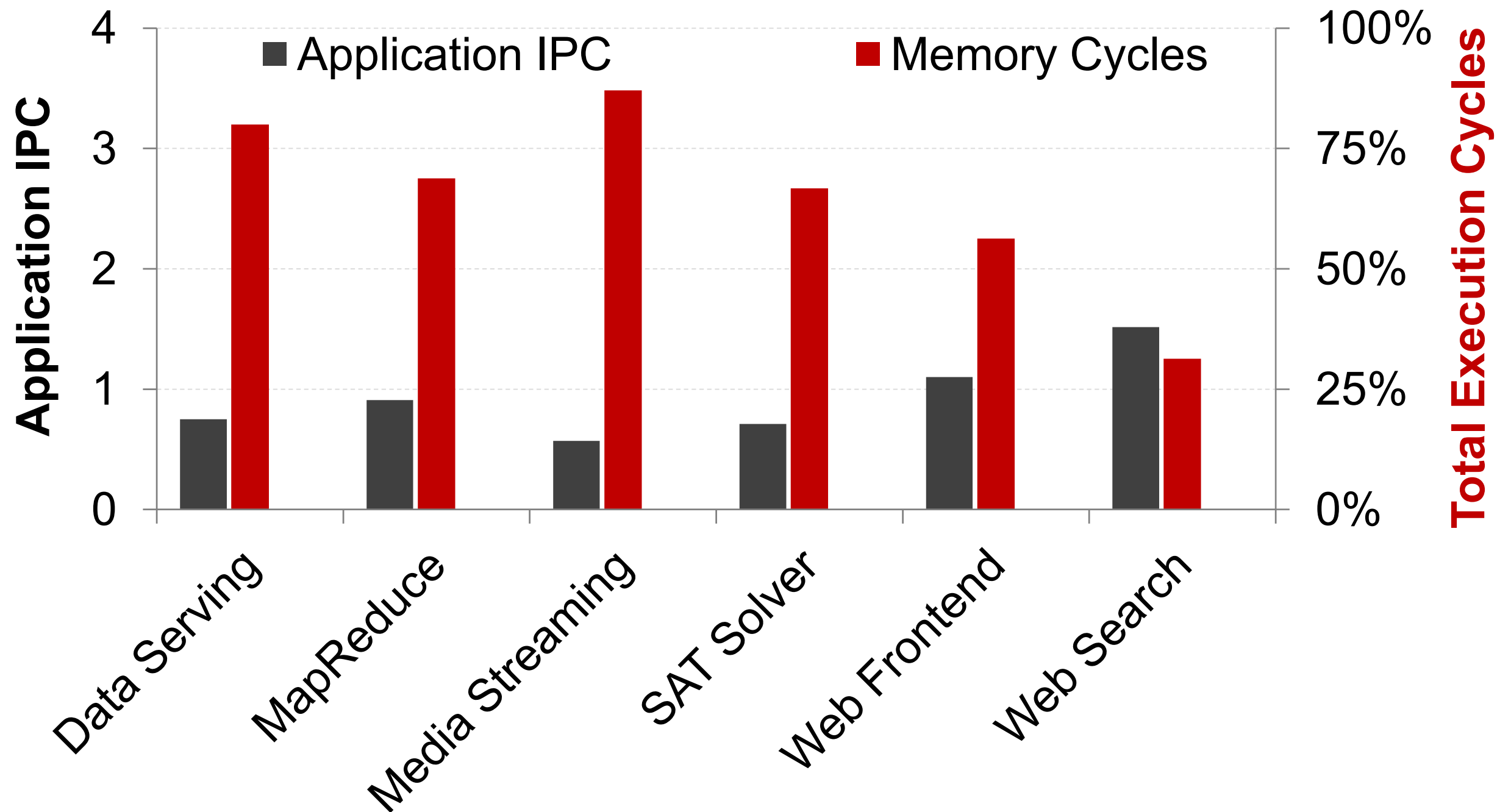
Next generation cores :
IC technology scaling
(same old story...)

→ **Scale-up Processors**

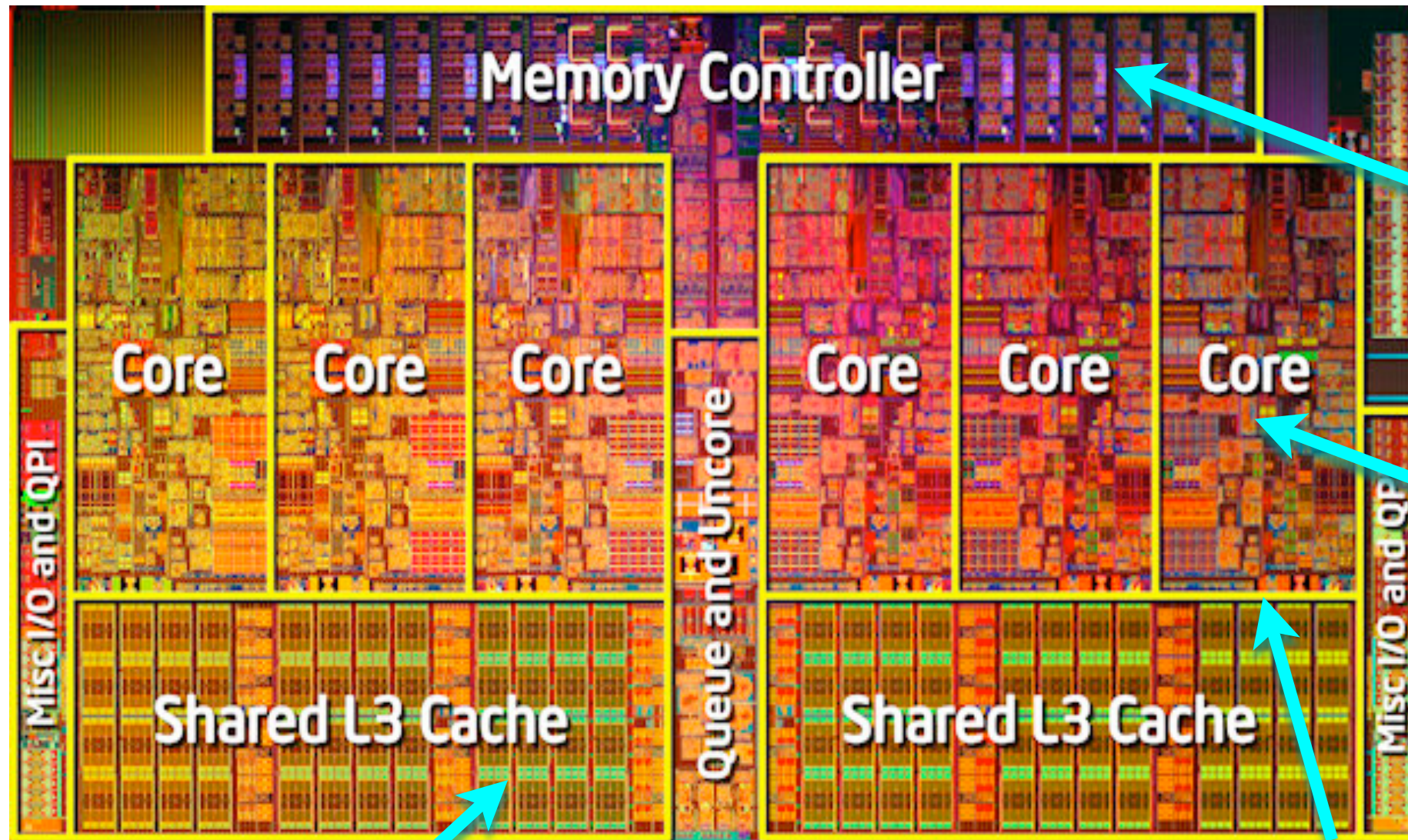


Bottleneck is memory, not CPU speed !

IPC vs. Memory cycles for typical datacenter applications



Scale-up CPUs don't fit scale-out apps* !



B/W unused!

Cores too fat!

Very little reuse

Too few cores!

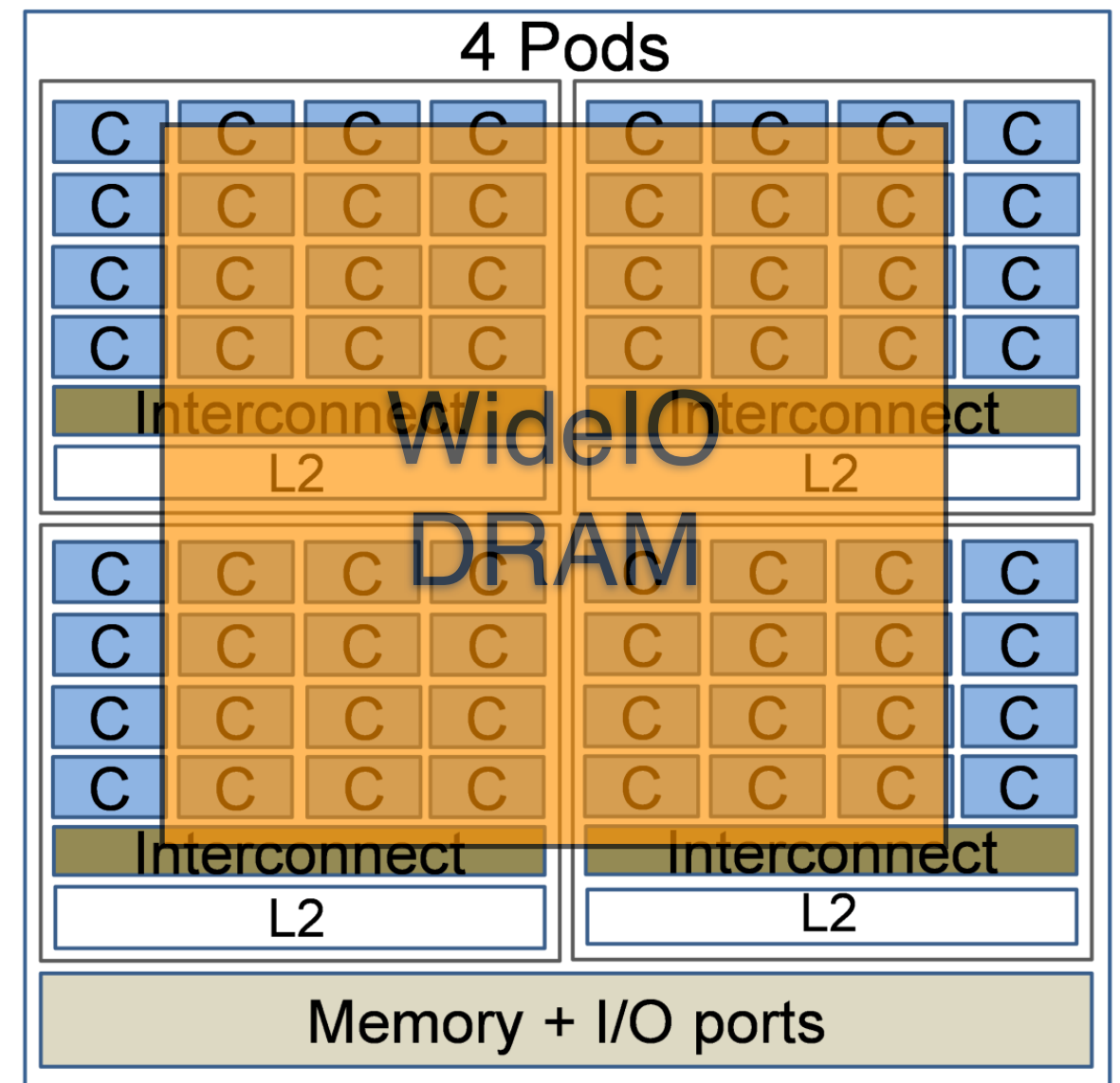
* Clearing the Clouds [Ferdman et al, ASPLOS '12]

Next gen. data-centers → scale-out CPUs* !

Many small,
low-power cores, based
on ARM-A9 architecture

Locally connected,
no global interconnect

Many, small, individual
caches, shared L2



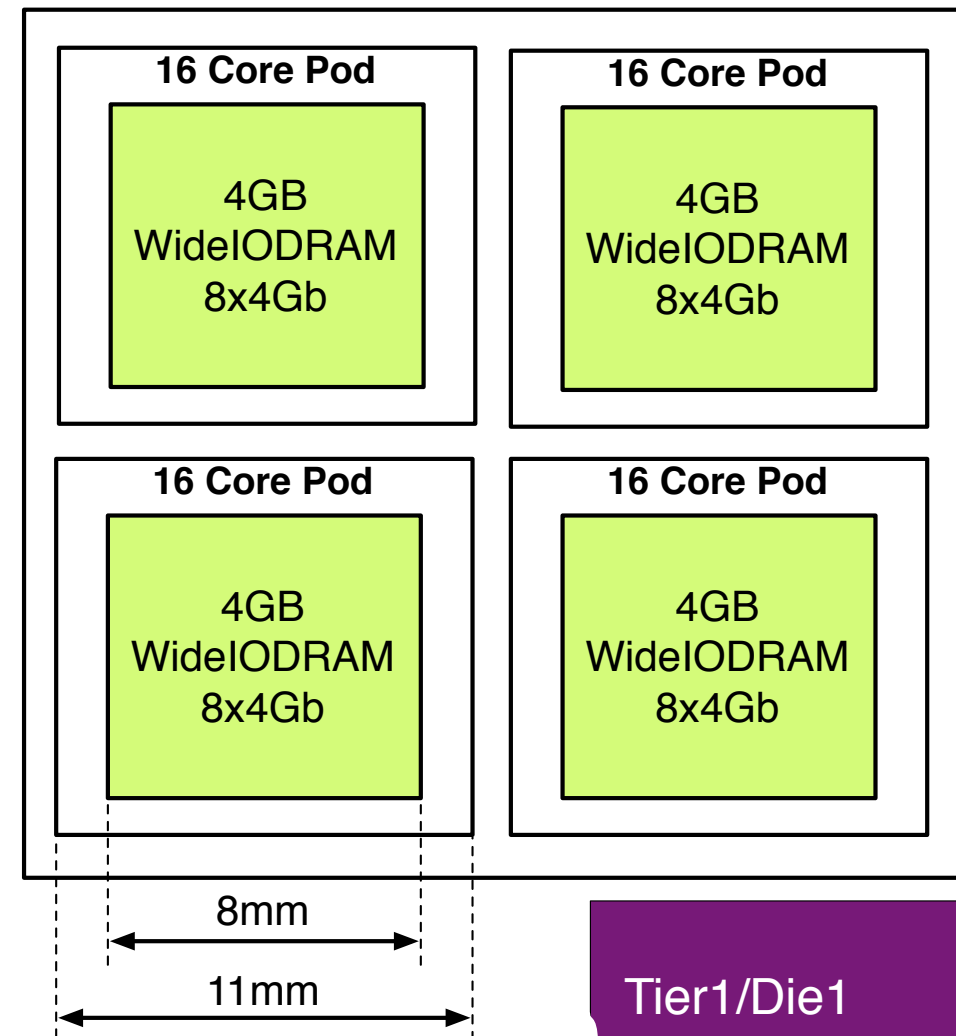
Big, WideIO DRAM as L3 cache !

Is this feasible today?

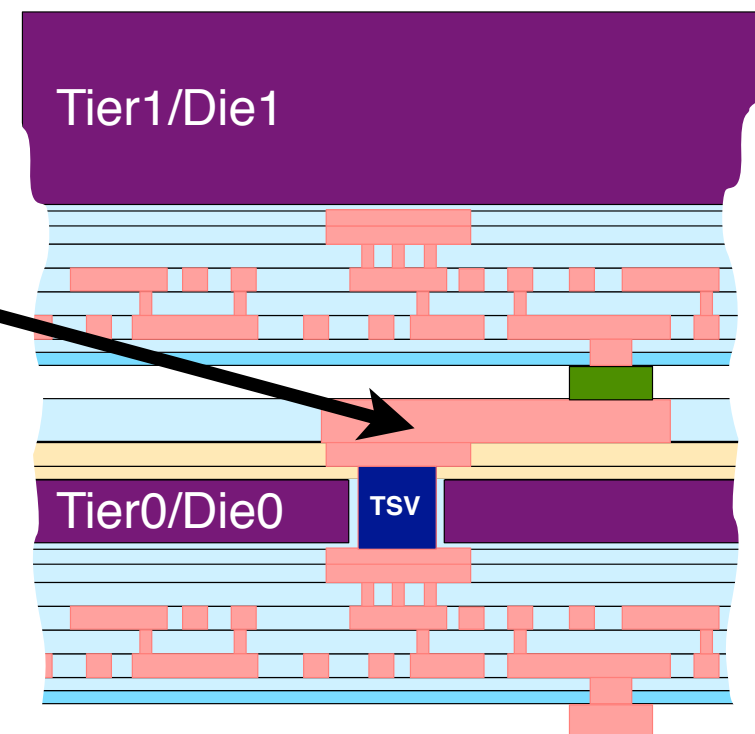
*[Lotfi-Kamran et al. ISCA '12]

Scale-out Processor: 3D-Stacked

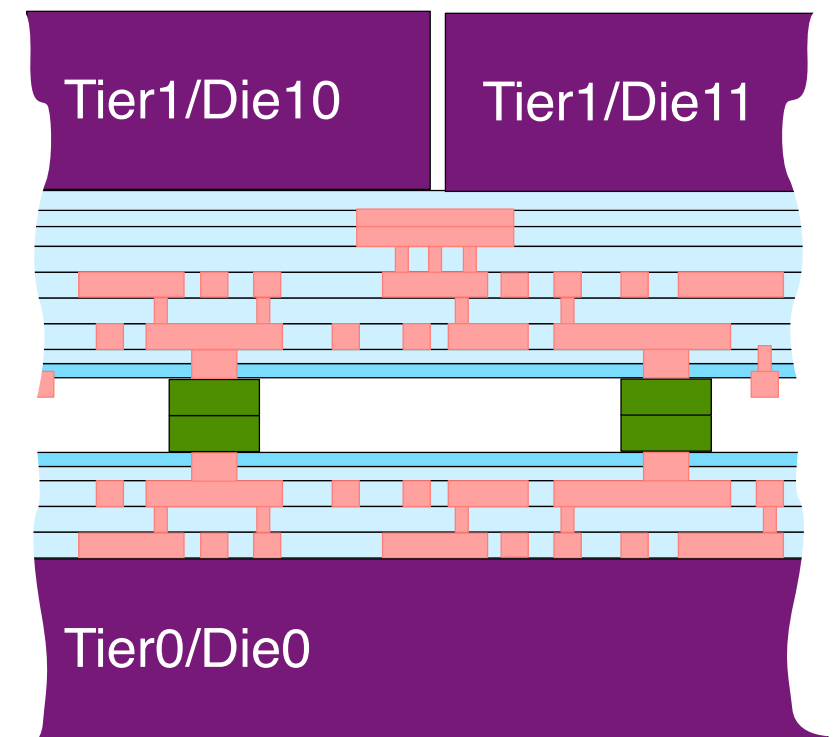
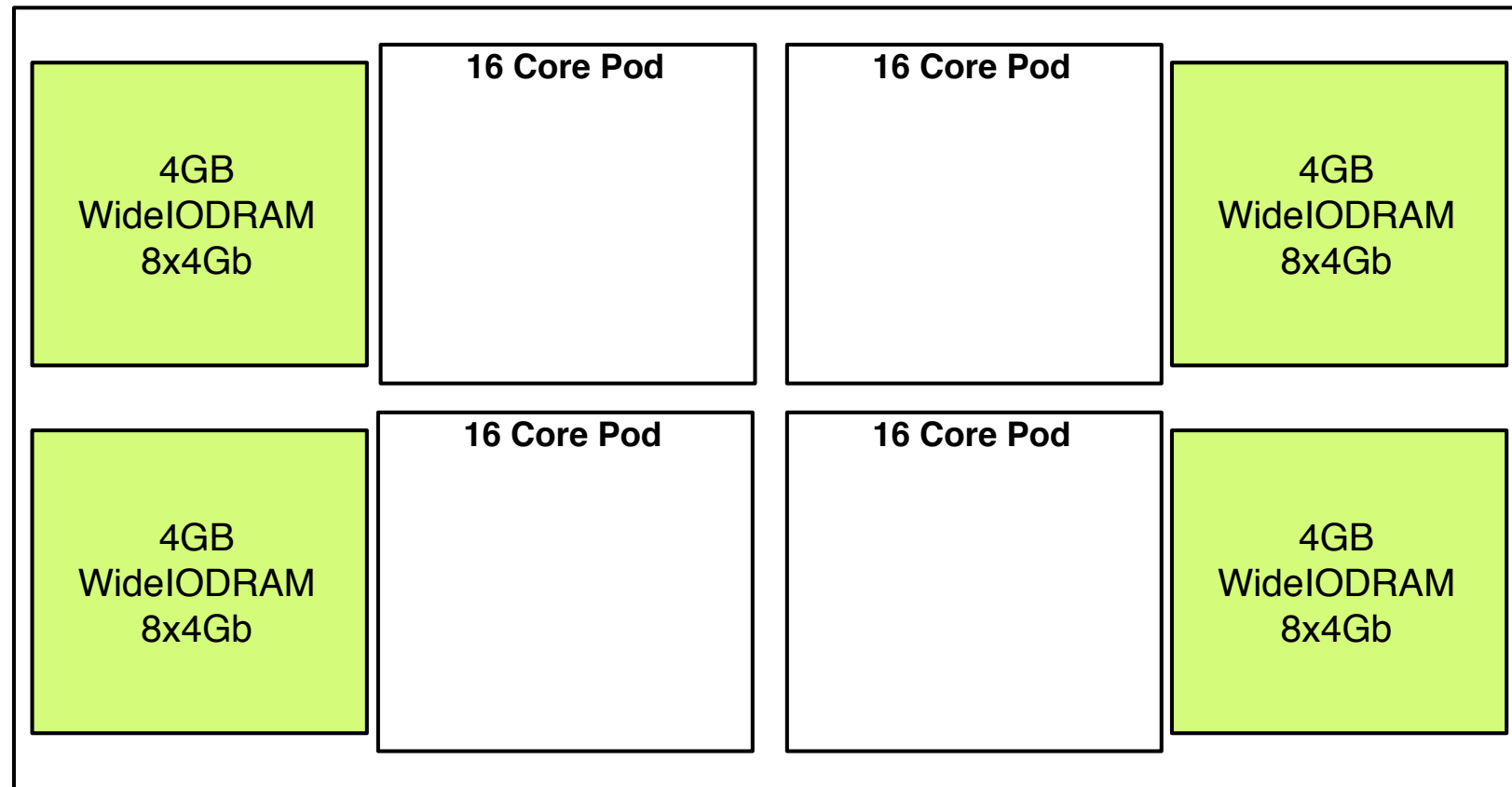
- 4 independent pods
- Each Pod
 - 1 Wide IO DRAM 4GB with 8 stacked dies of 1Gb
- 1 Wide IO DRAM = 4 channels with 128b/channel
 - 1056 IOs
- Face-to-back assuming logic die on the bottom, one WideIO on the top
- μ bumps of the WideIODRAM routed to the TSVs of the logic die using RDLs



RDL

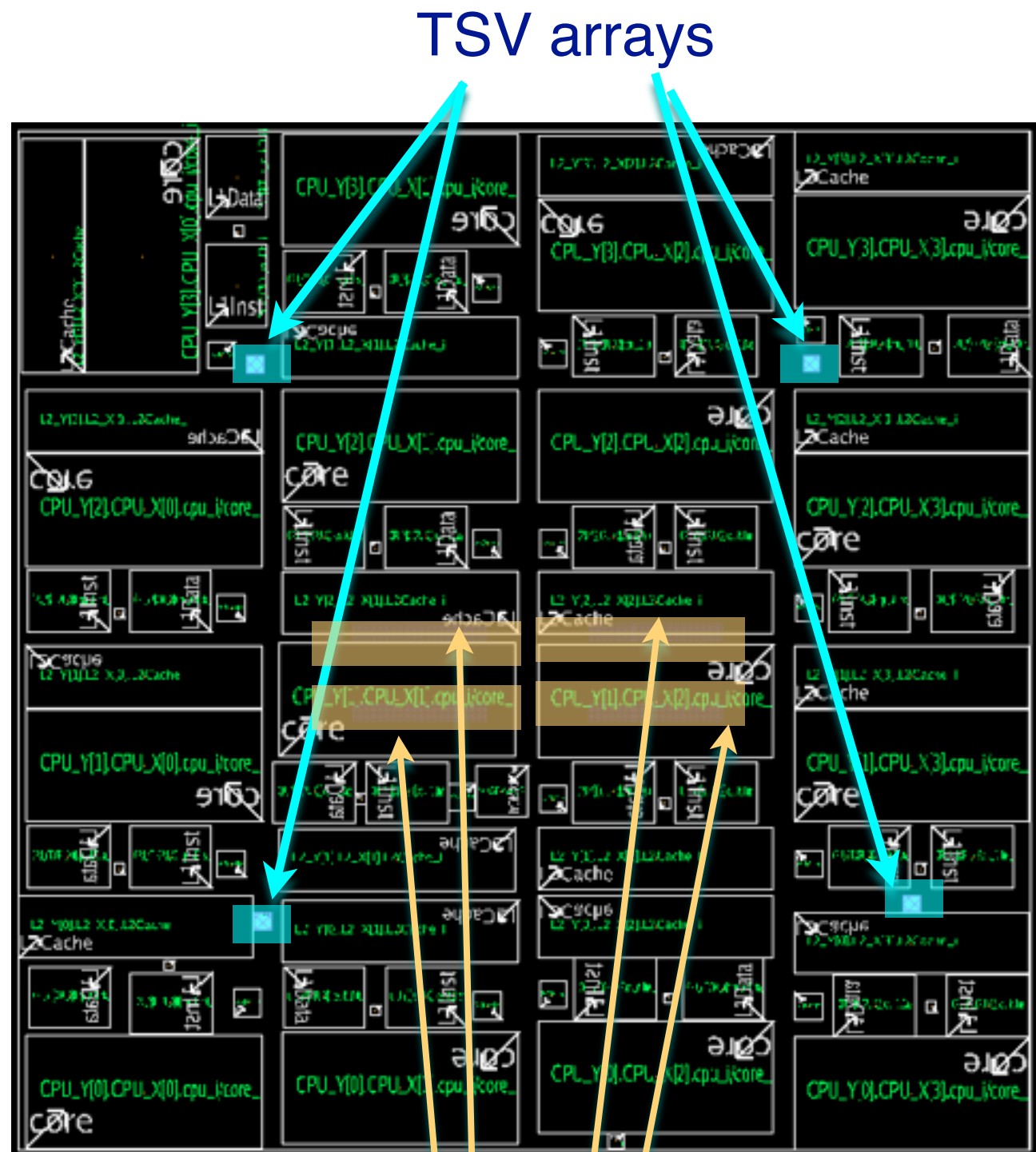


Scale-out Processor: 3D-Interposer



- Face-to-face integration using μ bumps only
- There are no TSVs in logic to memory interconnect
- Interposer designed to accommodate max reticle size

3D Stacked Scale-out Processor



TSV arrays

μ bump arrays
Logic die

1 Pod shown



WideIO DRAM

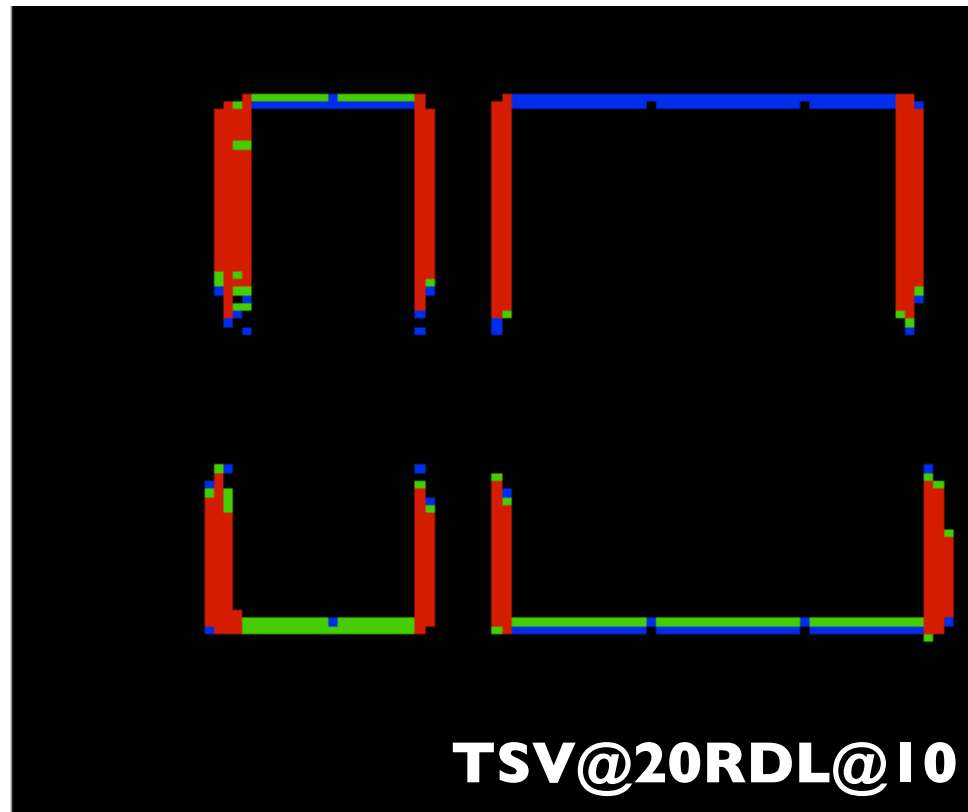
Backside routing & reporting



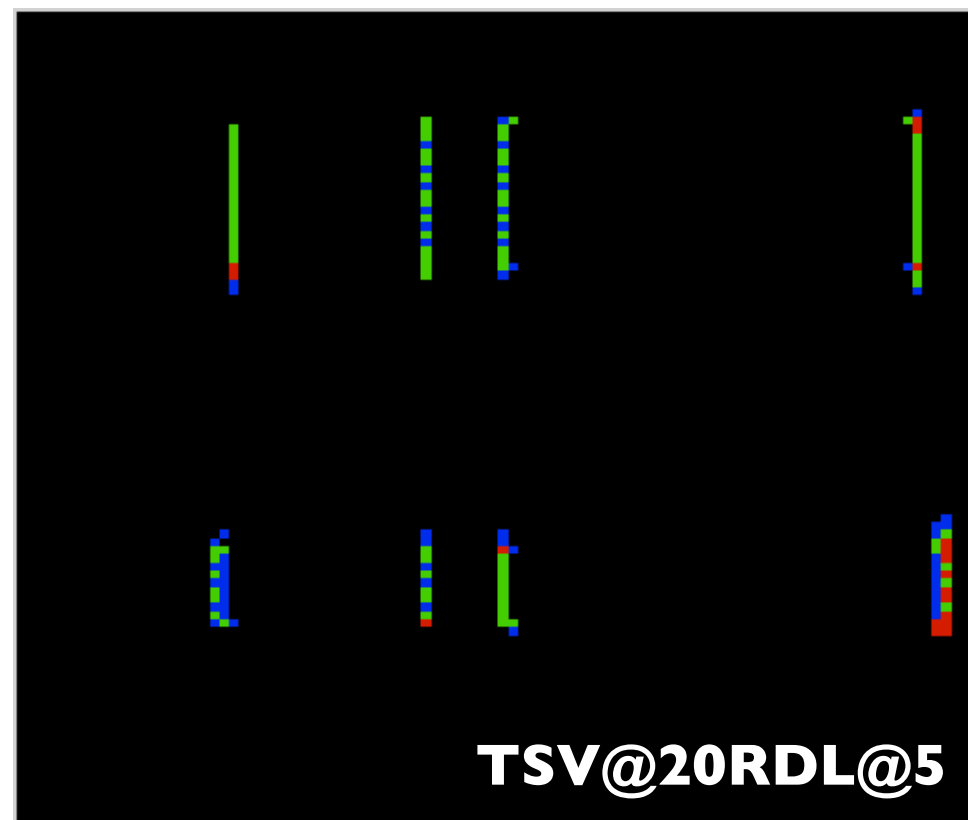
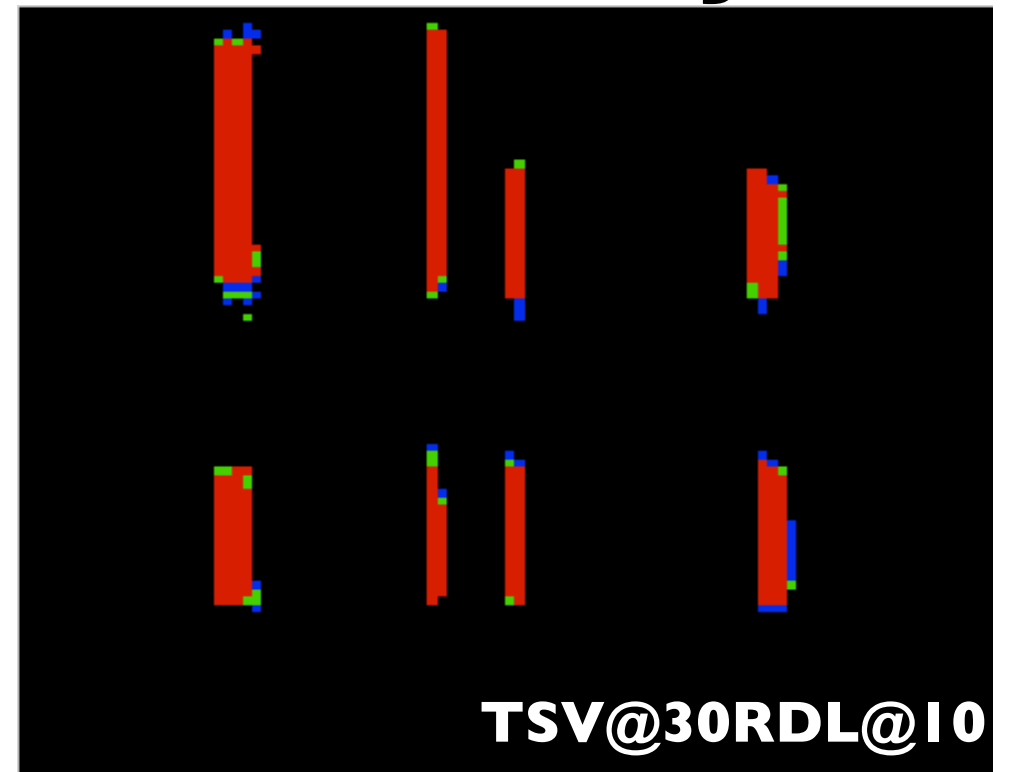
BRDL Pitch: 1.00
RDLs : 2

Total Nets 1056
Total Routed Nets 1056
Ignored Nets 0
Failed Nets 0
Max Wire length 5206.26
Min Wire length 2775.24
Average Wire length 4079.85
Total Wire length
4308327.00

Backside routing congestion analysis



Parameters:
TSV@30,20,10μm
RDL@10,5,2μm

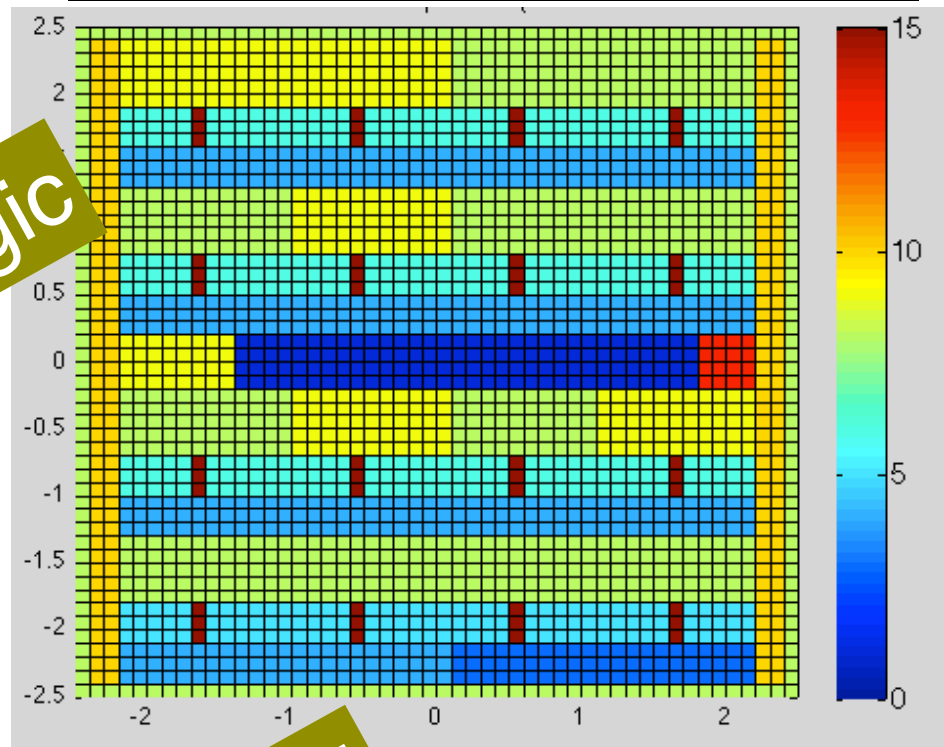


High Congestion
Medium Congestion
Low Congestion
No Congestion



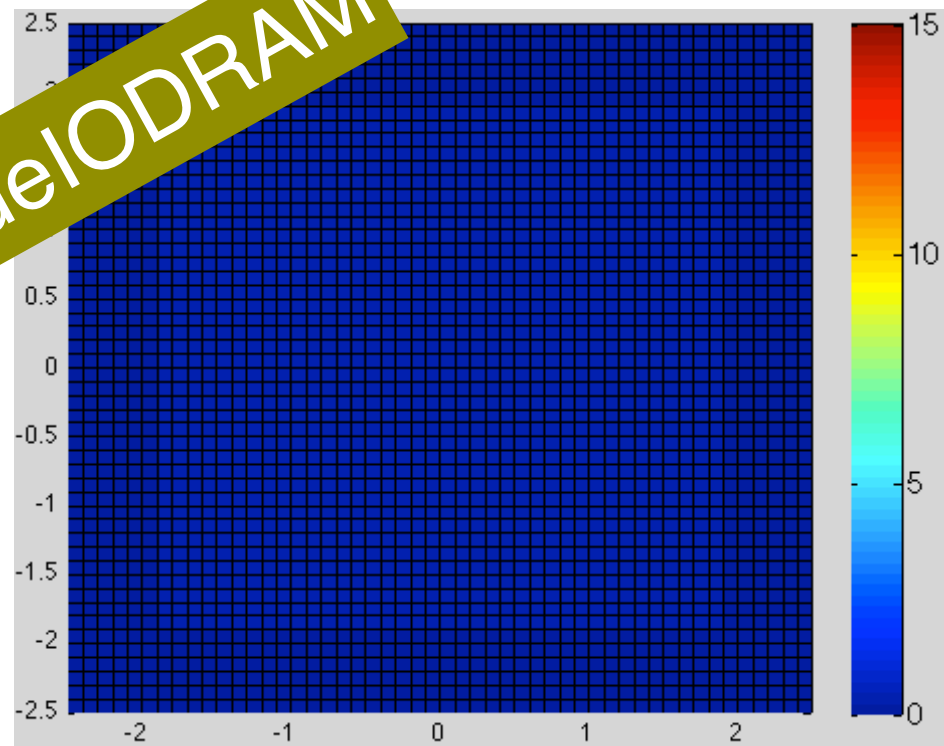
3D Stacked VP — Thermal Modeling

Power density map

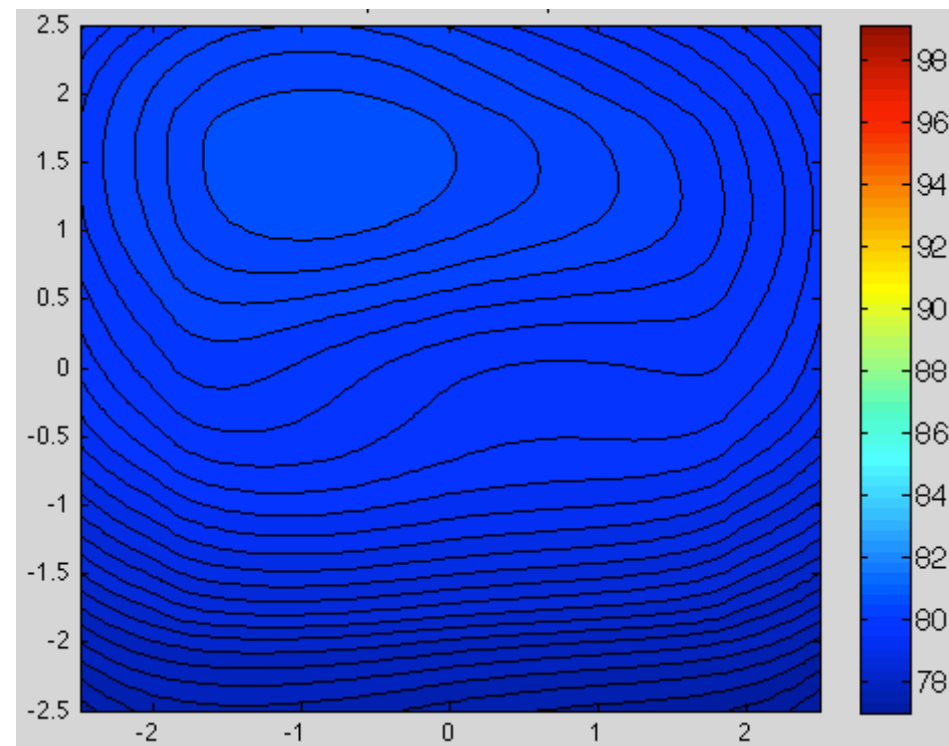
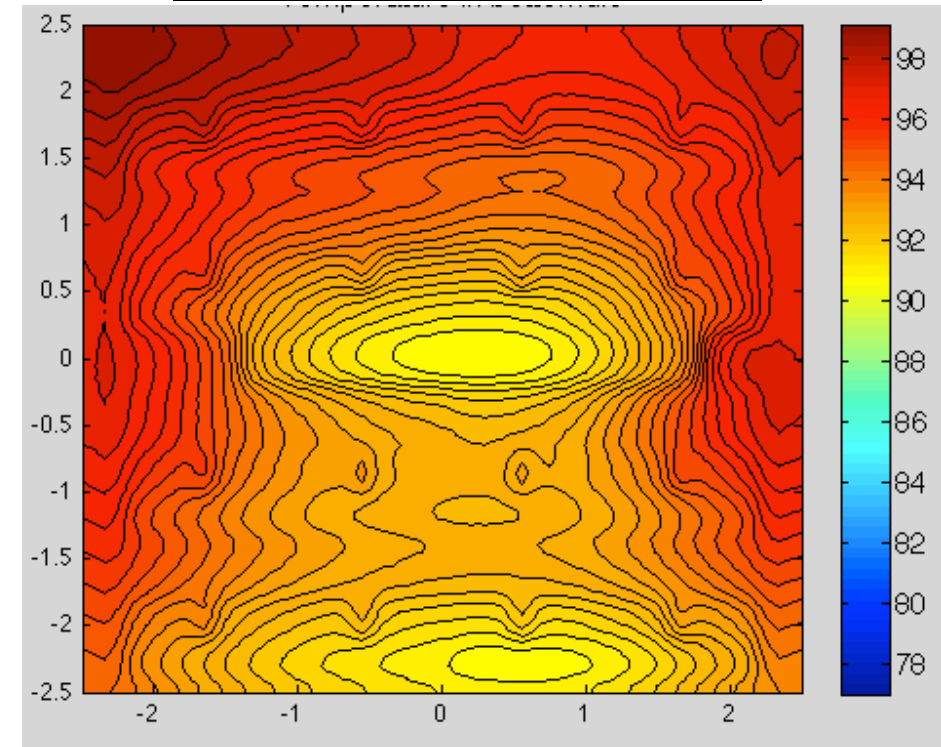


Logic

WideIODRAM



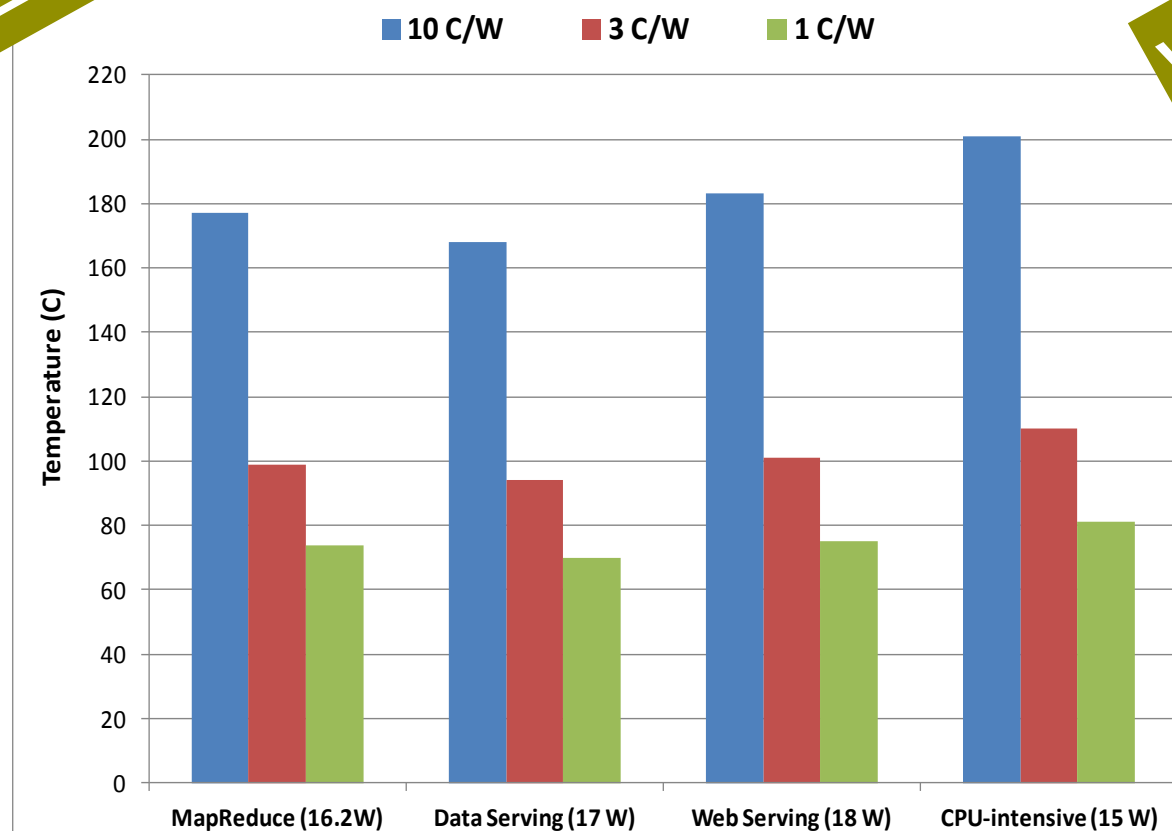
Thermal map



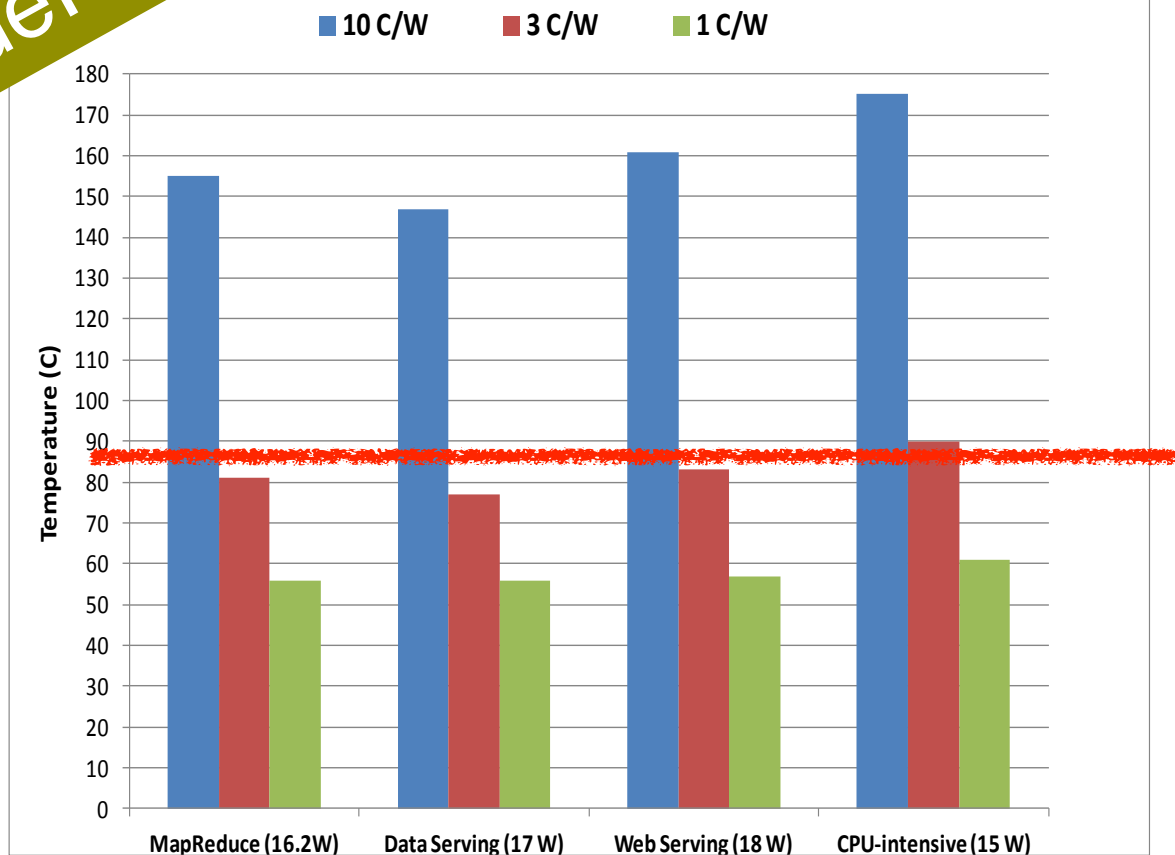
No fancy cooling solution is required !

Junction-Ambient Thermal Resistance (C/W)	Chip Package	Chip Cooling
10	30x30mm FBGA	None
3	30x30mm FBGA	Passive Heat Sink
1	30x30mm FBGA	Active Heat Sink

Logic

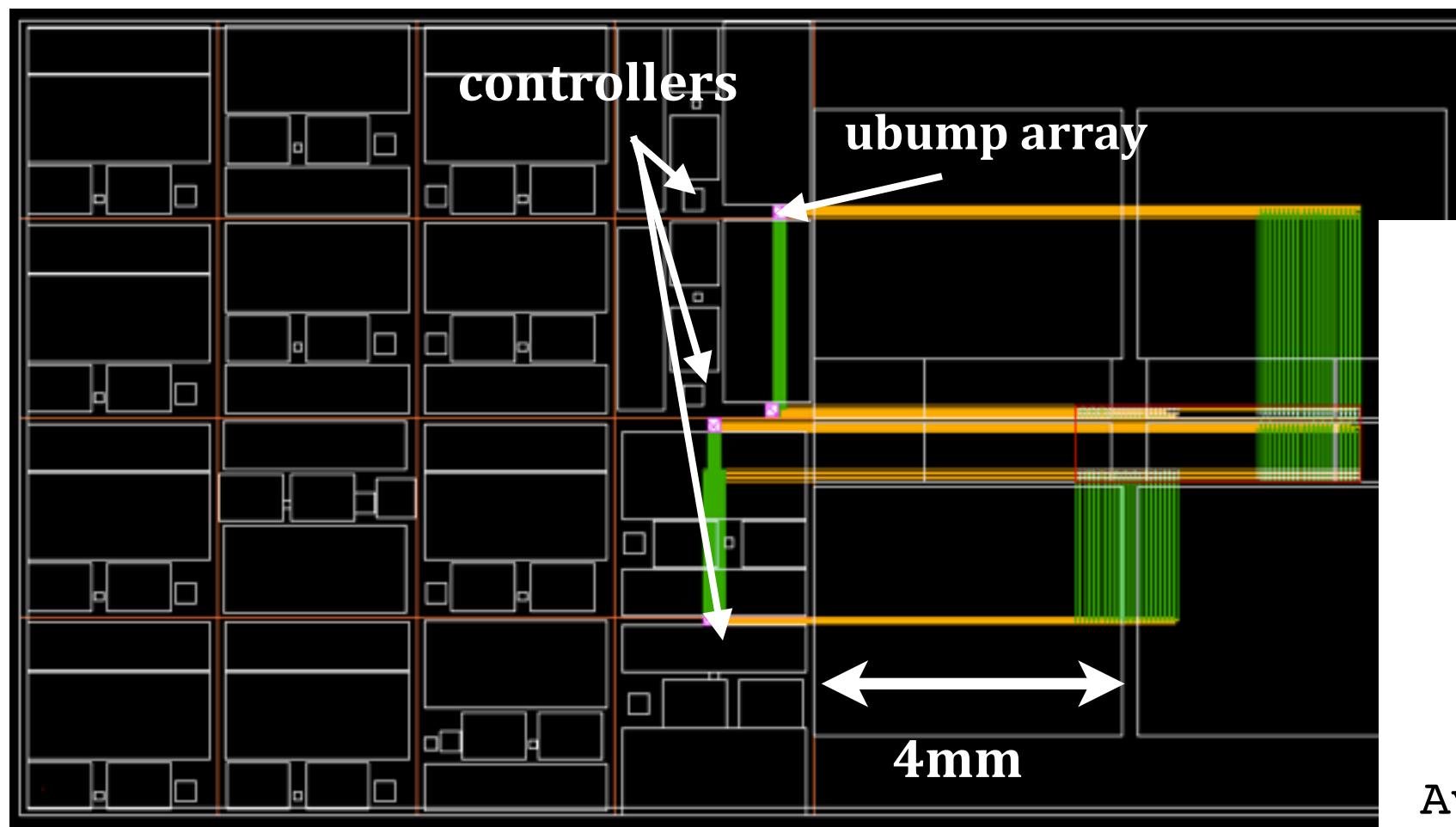


WideIO DRAM



Silicon Interposer

- We need to limit the max length of the wires on the interposer
- μ bumps have to be pushed on the edge of the logic die
- Trade-off between the Interposer and Logic die routing



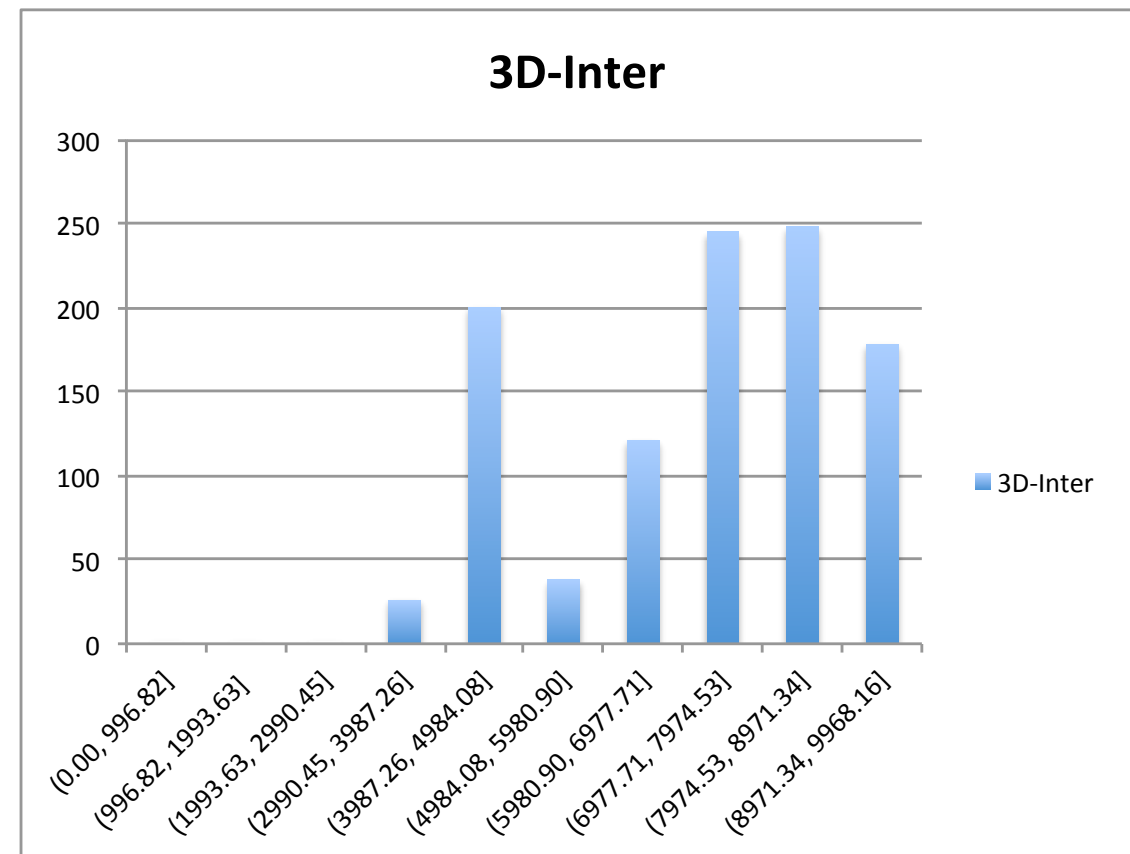
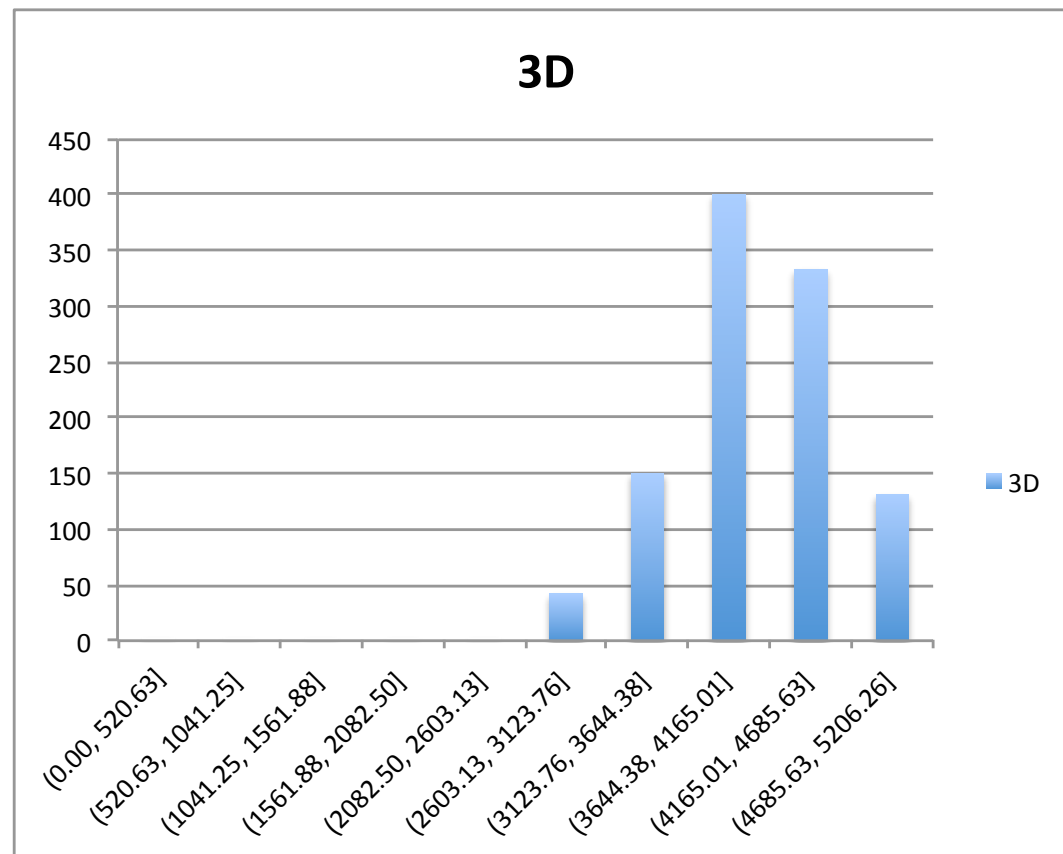
Number of RDL: 2
RDL Pitch: 1.00

Total Nets 1056
Total Routed Nets 1056
Ignored Nets 0
Failed Nets 0

Max Wire length 9968.16
Min Wire length 3801.48
Average Wire length 7216.75
Total Wire length 7620890.00

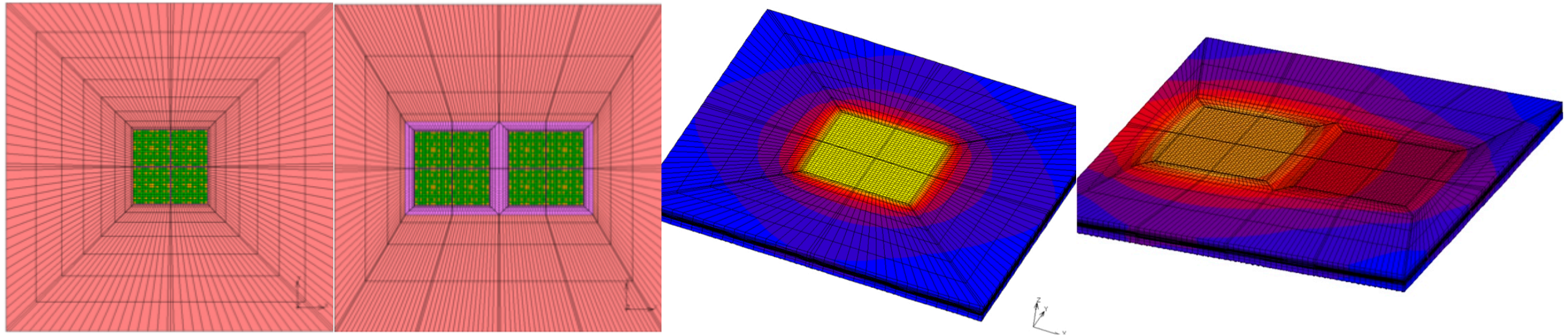
Interposer vs. 3D Stacked

Backside RDL (3D) vs. front side (3D-Interposer)
wirelength distribution



	3D	3D-Interposer
Max length	4079.85	9968.16
Total WL	4308327.00	2x 7620890.00

Thermal: 3D vs. Si-Interposer



Low-power		2.5D	3D	ΔT
Logic	2W	76	84	+8
DRAM	0.2W	64	84	+10
High-power				
Logic	20W	86	110	+24
DRAM	0.2W	37	96	+59

Natural

Forced

2X

**Better thermal behavior of the interposer.
We need to trade-off performance vs. thermal**

7. Practical work

- Now that you know more about Advanced integration, it is your turn ...

LES
SHADOKS



Suggested topics

- This is what I have in mind (you can choose):
 - ✦ Floorplanning for 3D ICs
 - ✦ Thermal floorplanning for 3D ICs
 - ✦ Post synthesis gate-level partitioning for 3D circuits for standard design flow (RD)
 - ✦ Graph theory applied to circuit partitioning (partition algorithms)
- I will provide description of each topic during lectures
- If you don't like proposed topics, make a suggestion by email by the beginning of the next week

For next week

- Do your article search (“Google is your friend”)
 - ✦ For non-free articles I can help you in getting IEEE or similar PDFs, send me the link
- Make a first round of the SoA
- Objective : come up with the overview, who, what etc.
- Make a plan
- We will then refine on person to person basis