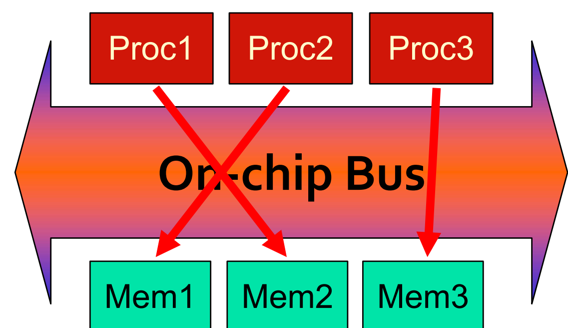
3D-MPSoCs: architectural and design technology outlook

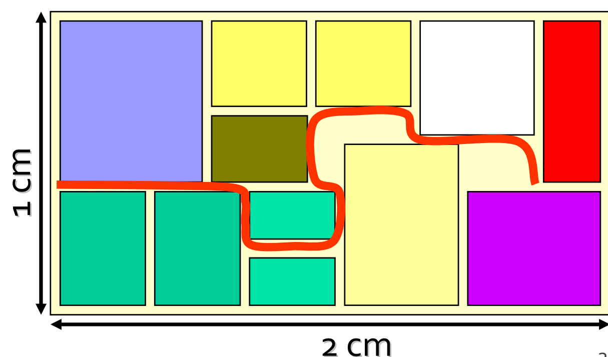
Luca Benini
DEIS Università di Bologna
lbenini@deis.unibo.it

The communication bottleneck

- Architectural issues
  - Traditional shared buses do not scale well – bandwidth saturation
  - Chip IO is pad limited
- Physical issues
  - On-chip Interconnects become increasingly slower w.r.t. logic
  - IOs are increasingly expensive
- Consequences
  - Performance losses
  - Power/Energy cost
  - Design closure issues, respins or infeasibility

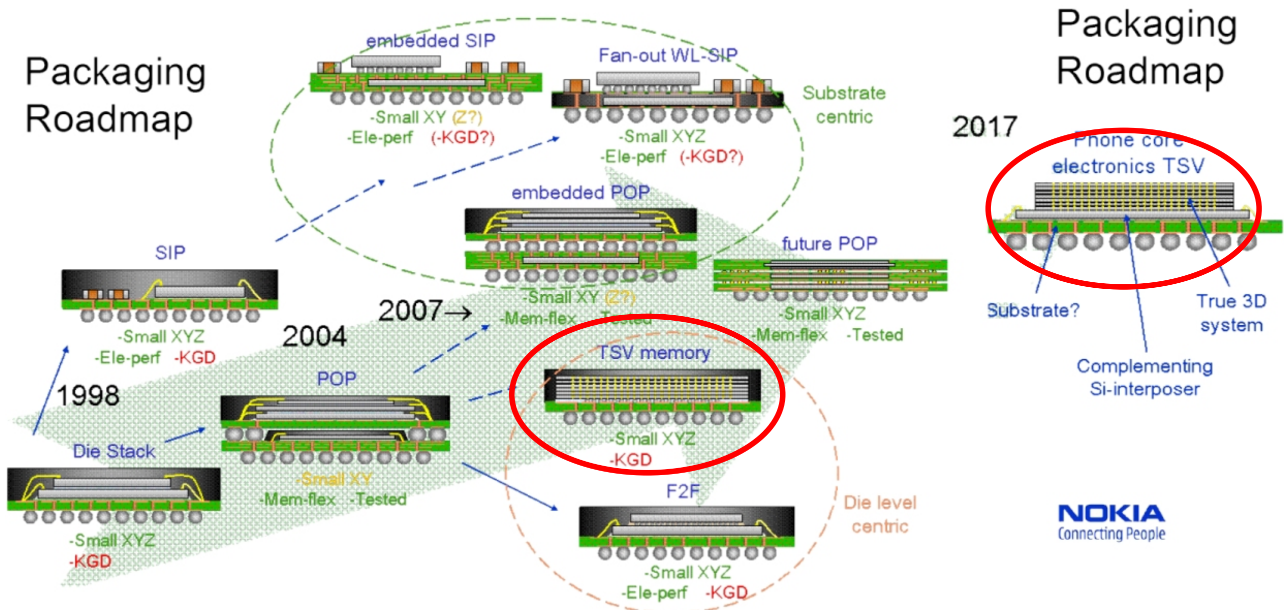
New architectures and design methods are required!





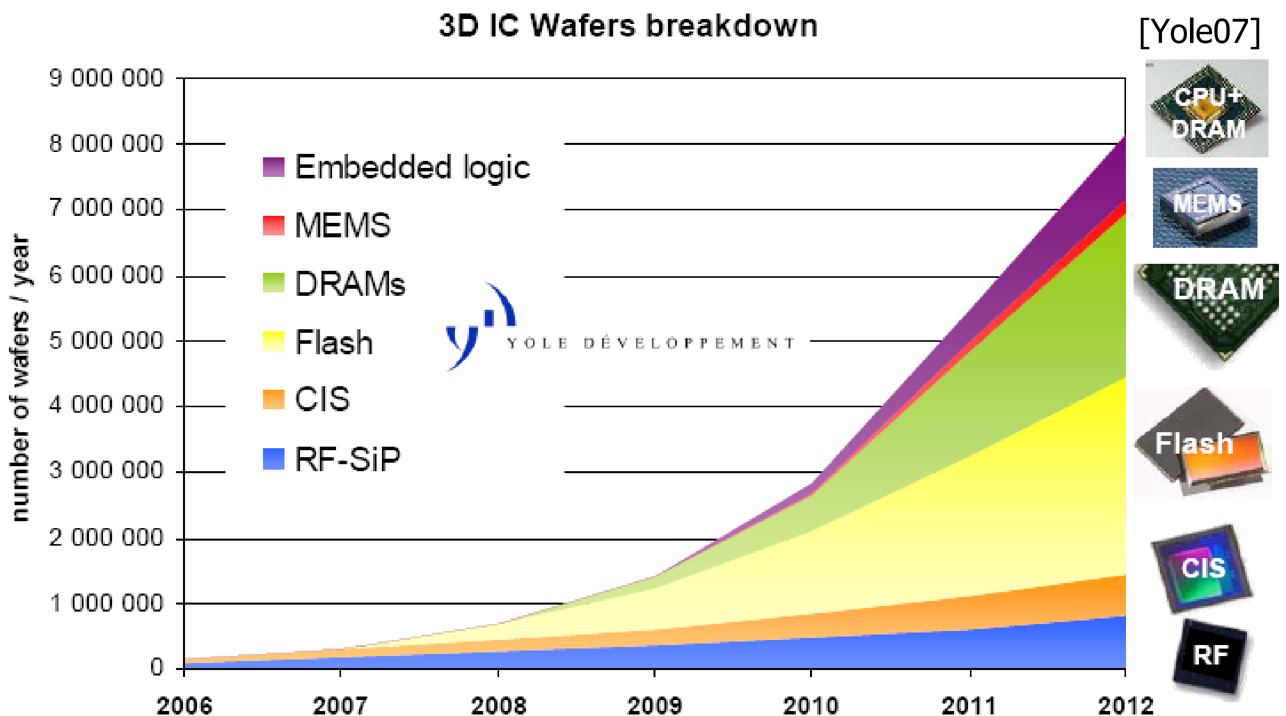
3D Integration roadmap

Coming to the rescue of communication starved 2D ICs



Through-silicon vias are at the technology bleeding edge today
Industry interest is growing: http://www.emc3d.org/

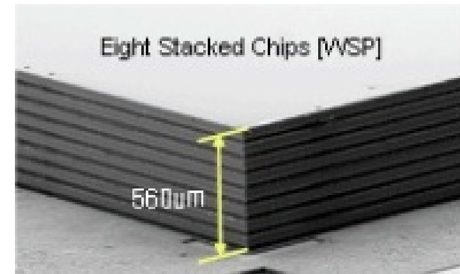
TSV market outlook



Less than 50,000 (est.) wafers to be fabbed with TSV in 2007

3D TSV-based Integrated Circuits

- Promises
  - Reduce average length of on-chip global wires
  - Increase the number of devices reachable in a given time budget
  - Greatly facilitate heterogeneous integration (e.g. logic-DRAM stacks)
- Challenges
  - Modeling and characterization
  - Reliability
  - Architecture, design & design technology implications

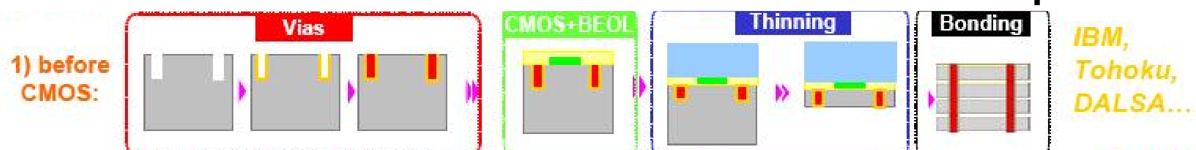


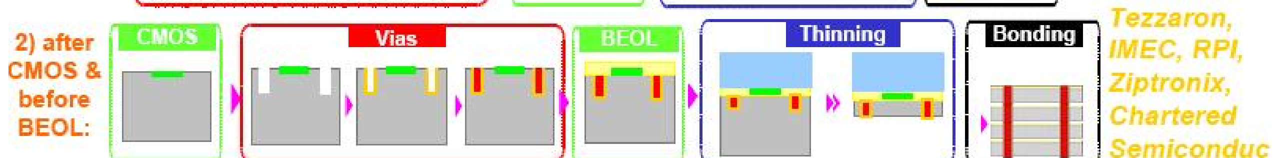
Samsung Wafer Stack Package (WSP) memory

Immature in products – significant effort in understanding architecture, design technology, system implications

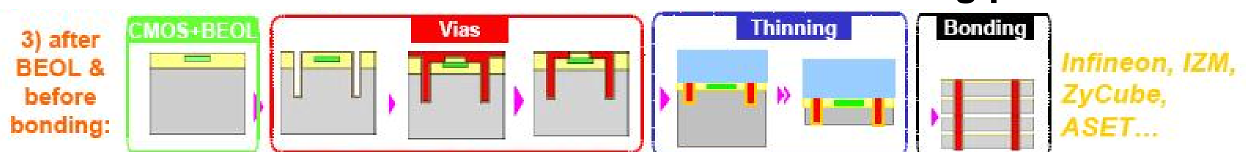
TSV Technology Options

- **Via first: TSVs realized before CMOS or before BEOL process**



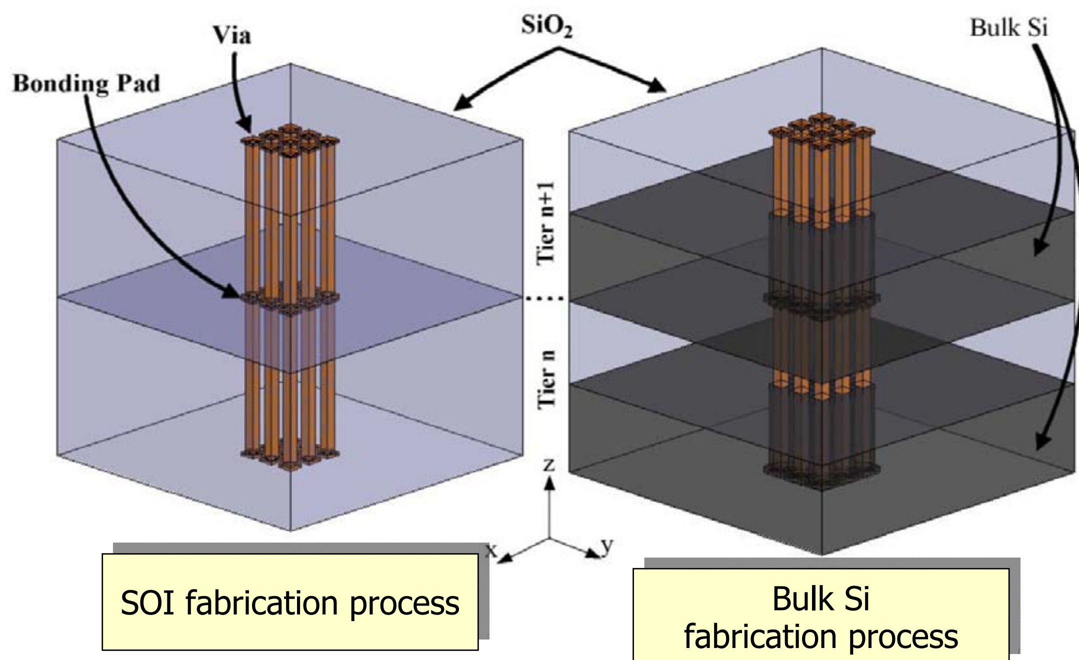


- **Via last: TSVs realized after BEOL or after bonding process**



[YOLE07]

Understanding TSV technoology



Modeling Resistance

$$R = \frac{\rho \cdot l}{\sigma}$$

- With typical vertical via parameters,

$$\frac{\rho}{\sigma} = 1.18 m\Omega / \mu m$$

- ~50 times smaller than for a typical 1.5mm Metal 8 horizontal wire in 0.13µm technology

Modeling Capacitance

$$\overline{\overline{C}} = \begin{pmatrix} C_{1,1} & -C_{1,2} & \ldots & -C_{1,n} \\ -C_{2,1} & C_{2,2} & \ldots & -C_{2,n} \\ \ldots & \ldots & \ldots & \ldots \\ -C_{n,1} & -C_{n,2} & \ldots & C_{n,n} \end{pmatrix}$$

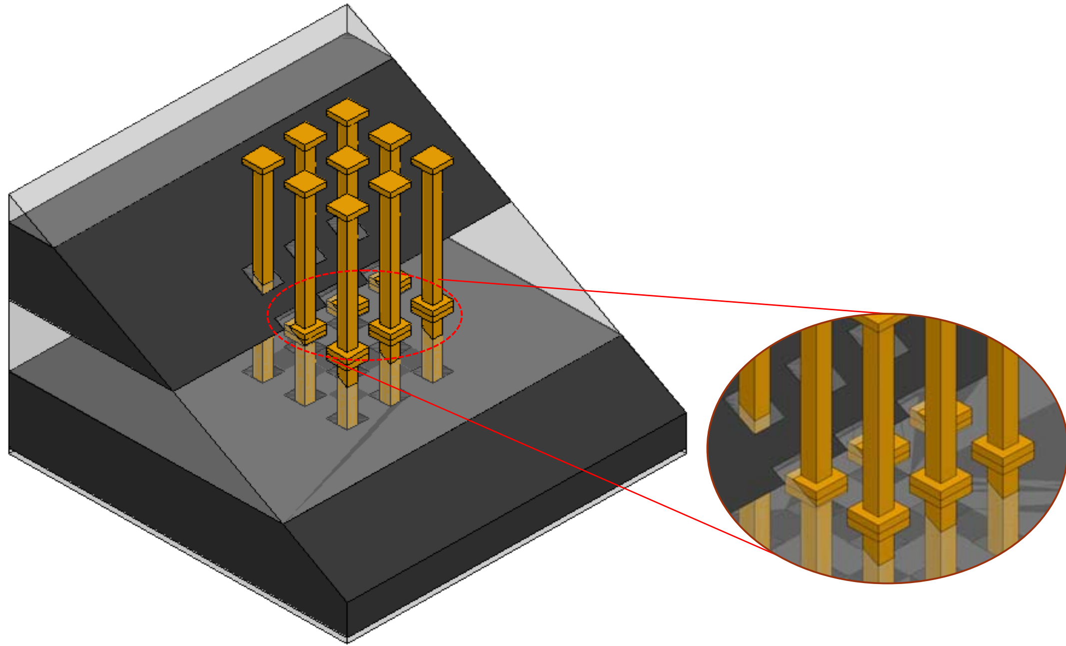
$$C_{i,i} = C_{i,0} + C_{i,1} + \ldots + C_{i,i-1} + C_{i,i+1} + \ldots + C_{i,n}$$

- Capacitance matrix summarizes ground capacitance and coupling effects
  - Bulk Si: around 23fF = 11 inverter loads (mostly towards ground)
  - SOI: around 10fF = 5 inverter loads (mostly towards other vias)
- For a whole via, capacitance is ~10 times smaller than for a typical Metal 2/3 horizontal wire of 1.5mm in 0.13μm

TSV performance

- Delay is given by combination of parasitics
  - Horizontal wire to via base
  - Via delay (includes R of bases)
  - Horizontal wire from via top
  - Load
- For a whole via of 50μm, delay is 16/18.5ps (SOI/bulk)
- For a 1.5mm horizontal link, delay is around 200ps

So far so good, but…

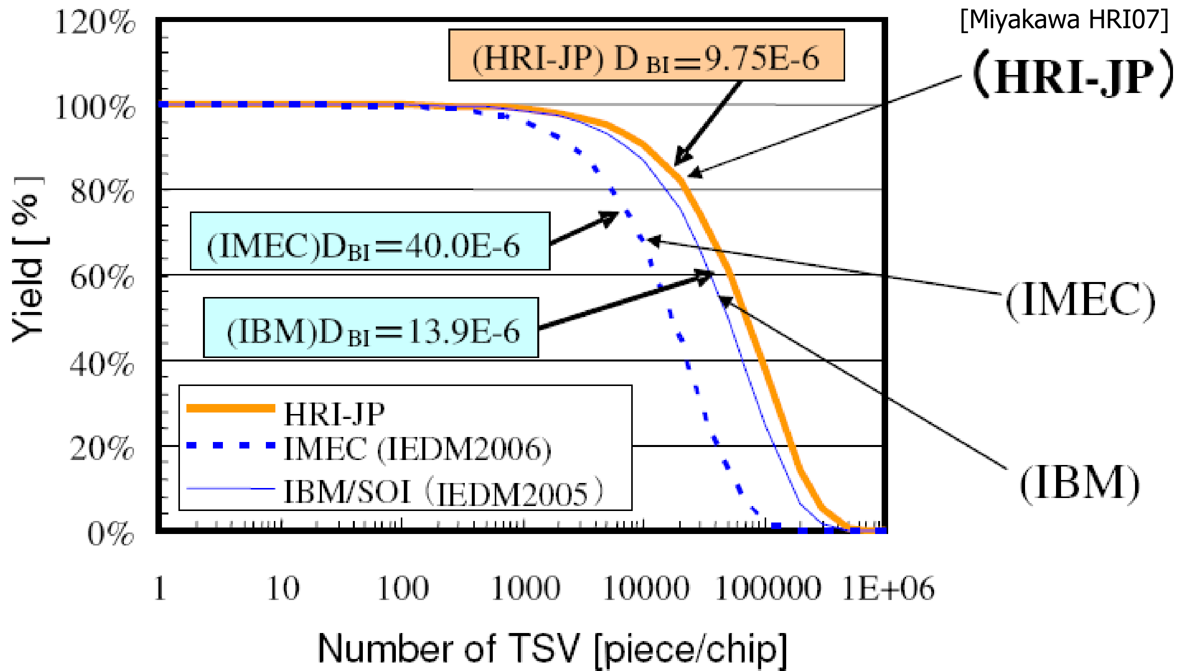


- >10um pitches seem to be realistic
- Not all TSVs can be used for signals
  - Power supply, clock, thermal vias

TSV reliability losses

- Main failure mechanisms (fabrication)
  - Misalignment
  - Voids formation during Bonding phase
  - Dislocation and defects of Copper grains
  - Oxide film formation over Cu interface
  - Partial or full Pad detaching due to thermal Stress
- Thermal dissipation is much harder in 3D stacks, thereby further increasing the risk of temperature-related failures

TSV yield



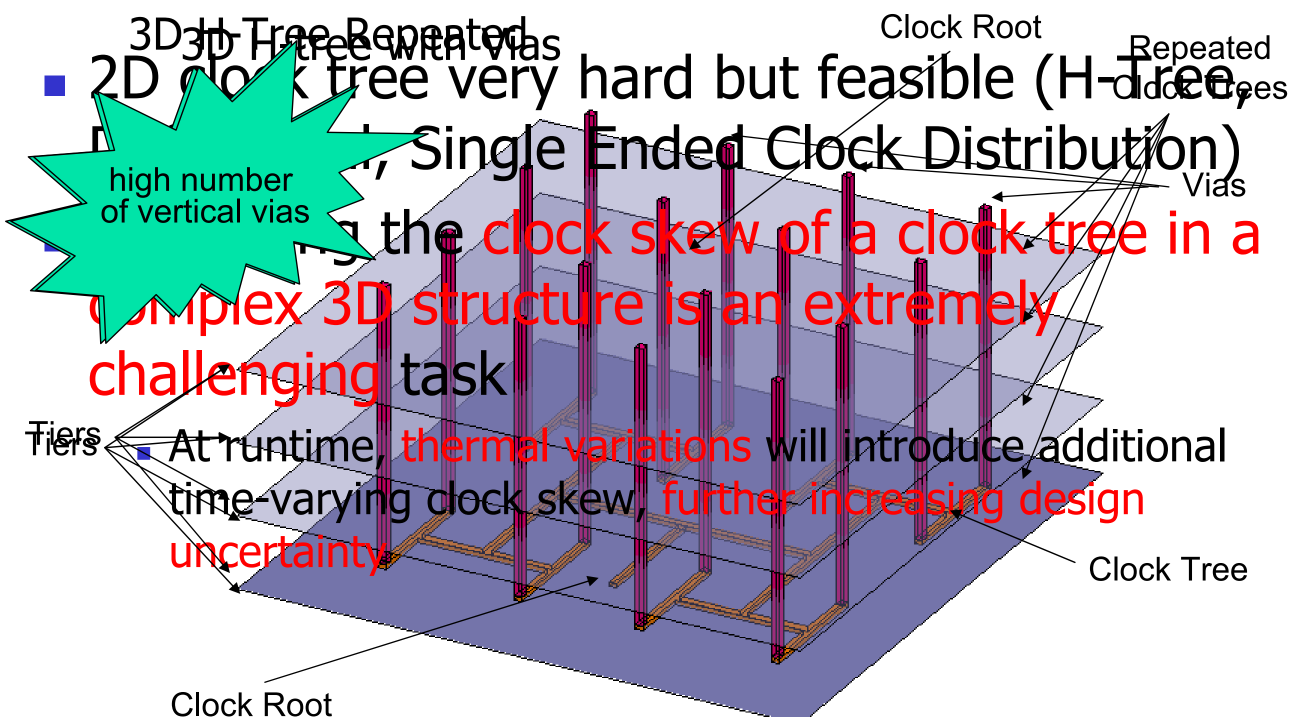
$Y=\exp(-D_{BI} * N_{BI})$

D_{BI}: defect frequency

N_{BI}: Number of TSVs

Clock Distribution in 3D

- 2D clock tree very hard but feasible (H-Tree, D[illegible]l, Single Ended Clock Distribution)
- [illegible]ng the clock skew of a clock tree in a complex 3D structure is an extremely challenging task
  - At runtime, thermal variations will introduce additional time-varying clock skew, further increasing design uncertainty

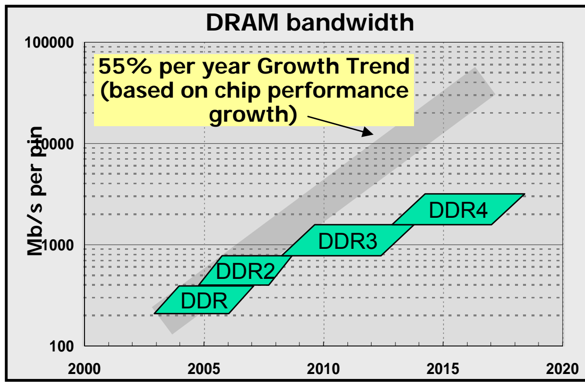


Summing up…

- Good power and speed
- Area overhead is significant
- Reliability not ideal (fabrication and aging)
- Synchronization is hard (skew minimization across layers)
- Therefore:
  - Cost and design effort are not trivial
  - Not just another dimension for wiring (as of today)
  - Need a sistematic way to deal with non-ideality

A medium-term vision

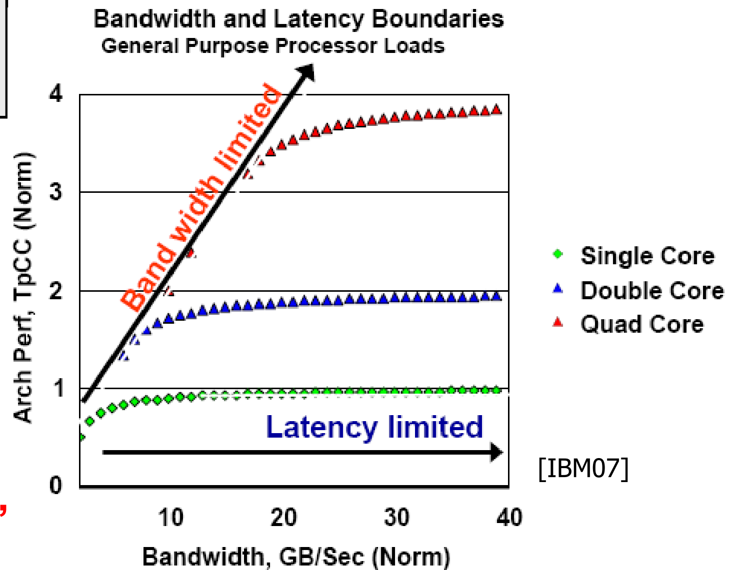
Do We Really Need It?



- Multi-core logic performance is back on track, but …

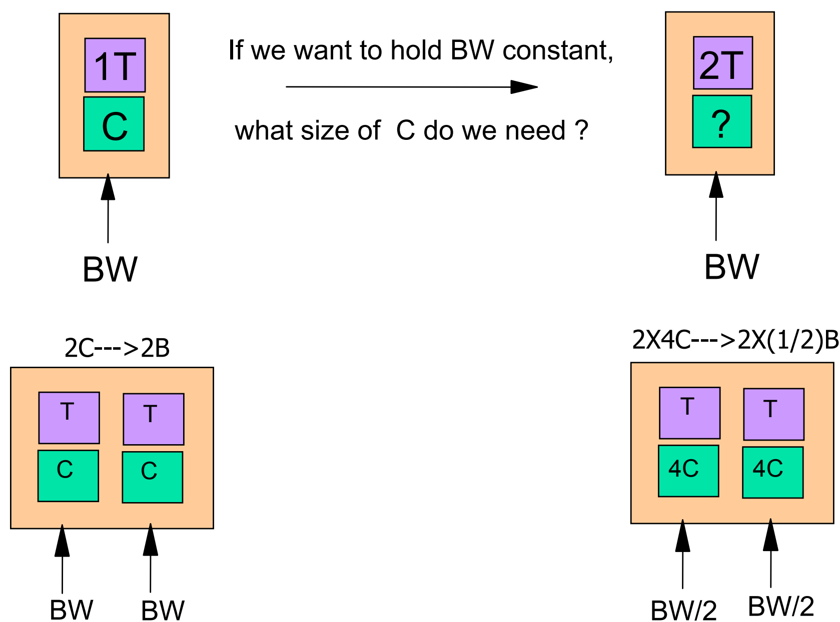
- Multi-core are bandwith-hungry:
  - Limited caches
  - Multi-threading
  - Virtualization

"The Bandwidth Challenge"



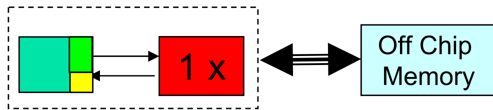
[IBM07]

Caches to reduce bandwidth requirements?

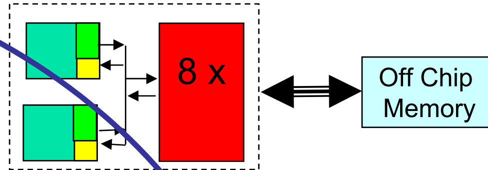


Doubling threads at constant bandwidth requires 8x cache!

Scaling #cores with constant BW

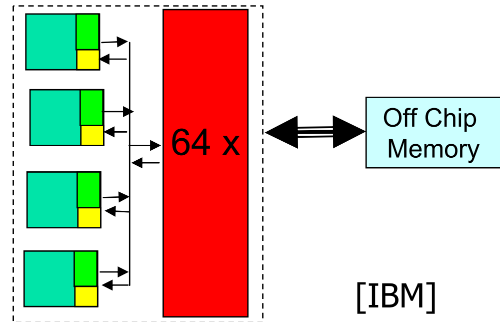


Using Cache size to accommodate increasing thread traffic is VERY expensive – using BW can be cheaper!!



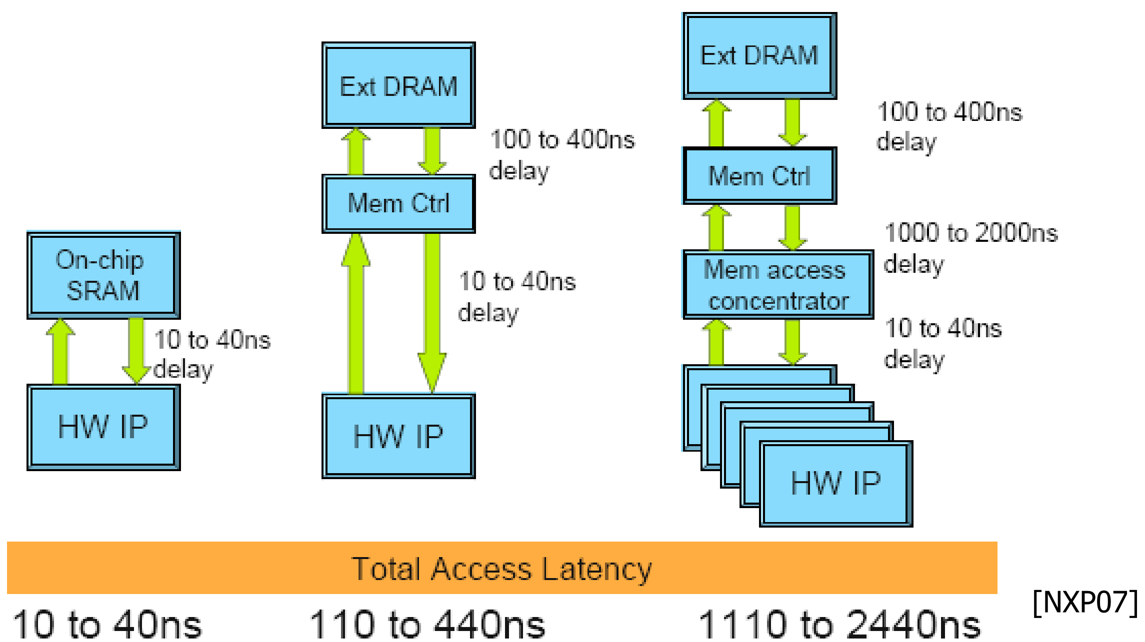
2x increased traffic drives 8x cache size (constant memory bandwidth)

4x increased traffic drives 64 x cache size (constant memory bandwidth)

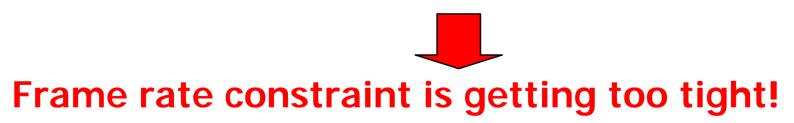


[IBM]

What about Embedded MPSoCs?



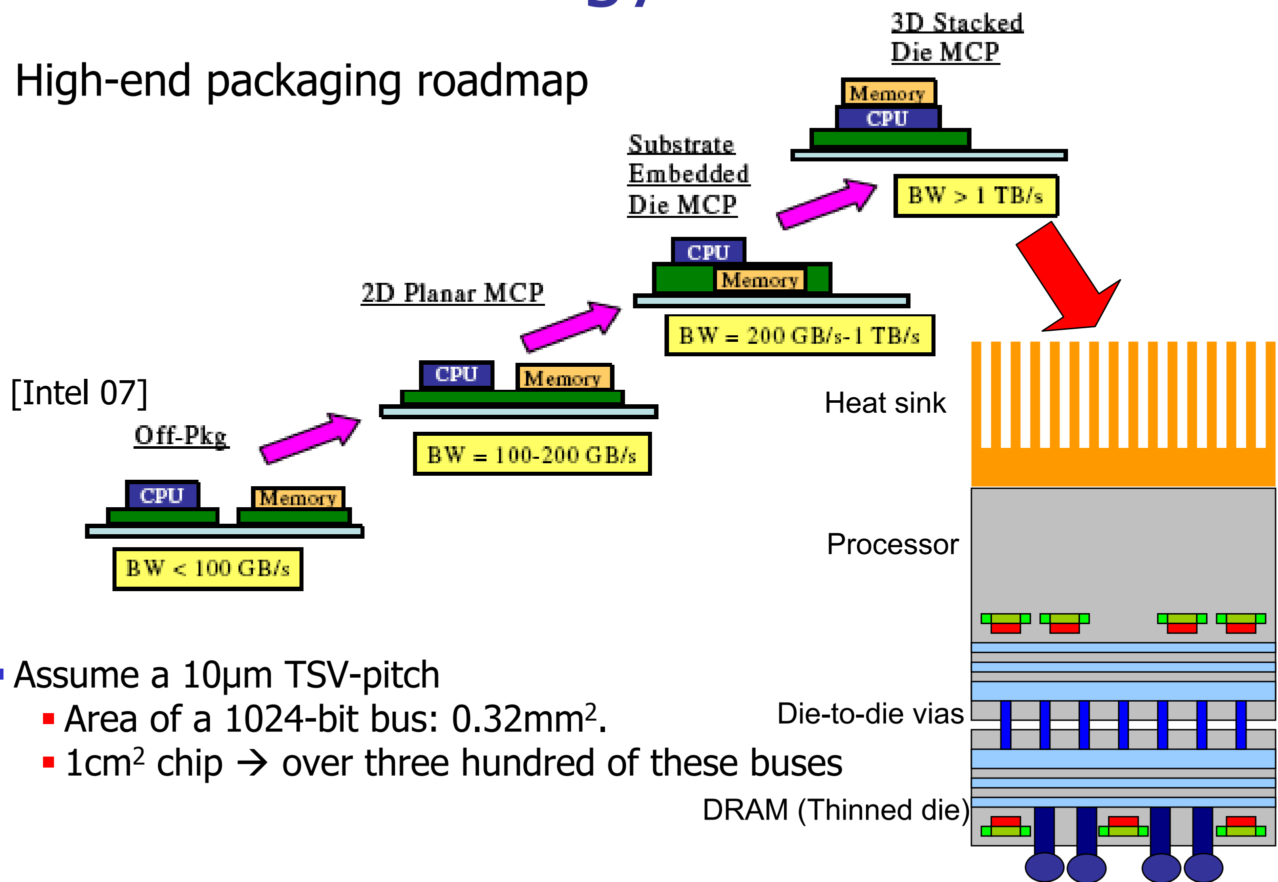
[NXP07]



Frame rate constraint is getting too tight!

3D-IC Technology to the Rescue!

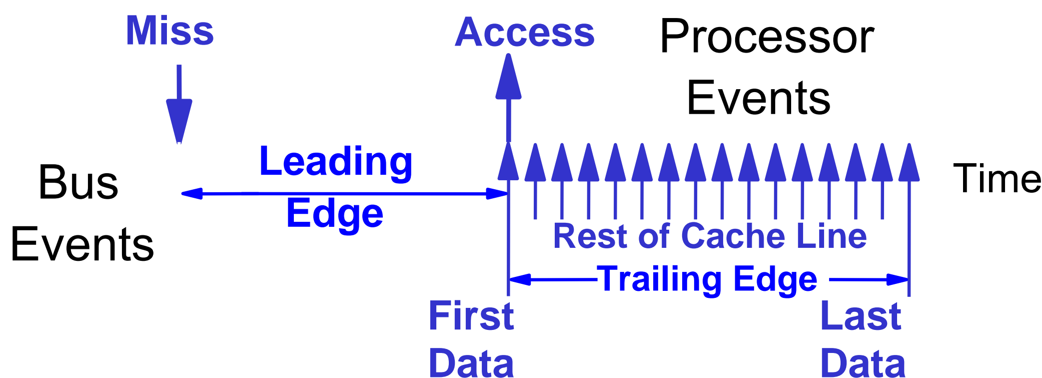
High-end packaging roadmap



- Assume a 10µm TSV-pitch
  - Area of a 1024-bit bus: 0.32mm^2.
  - 1cm^2 chip → over three hundred of these buses

Is Bandwidth All We Need?

Penalty for handling a cache miss:



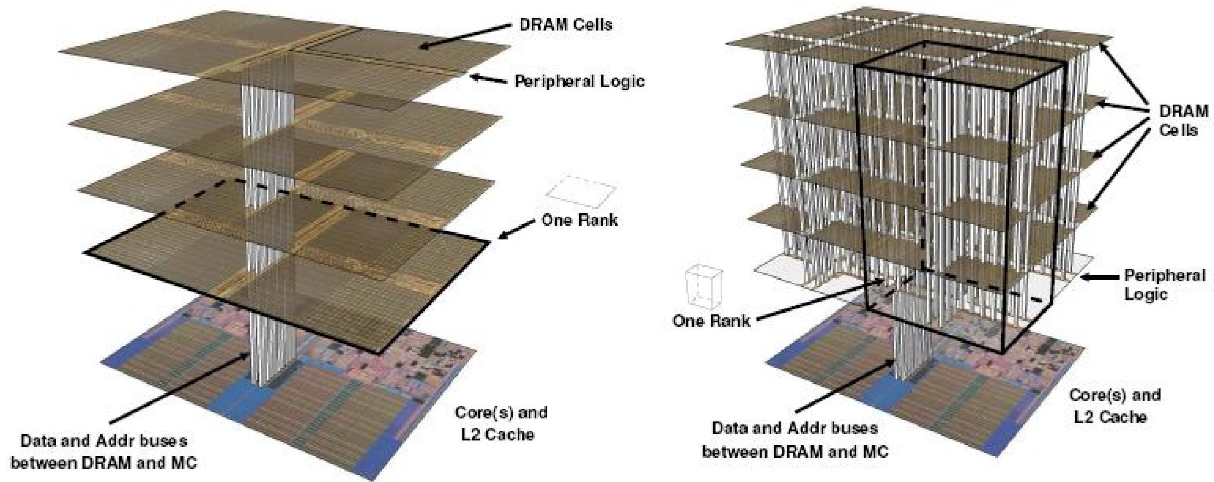
Miss Penalty = Leading Edge + Effects(Trailing Edge)

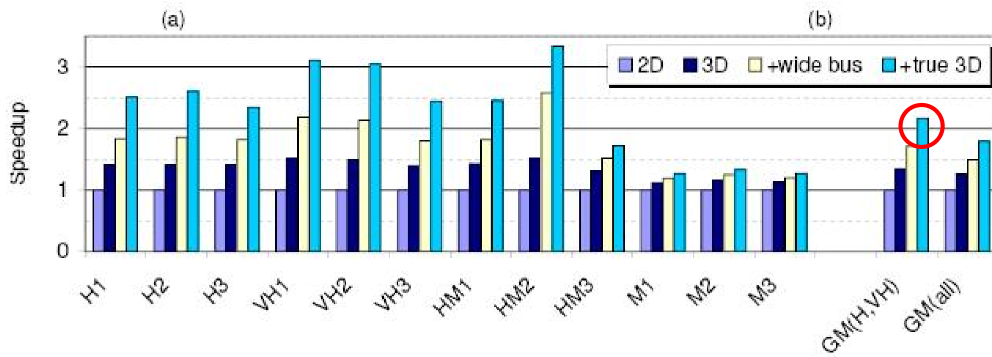
Where

Trailing Edge Effect	= (Line Size / Bus Width) x ($F_{(\mu P)}$ / $F_{(Bus)}$)
Bus Utilization	= (Trailing Edge / Intermiss Distance)

But what about leading edge effects?

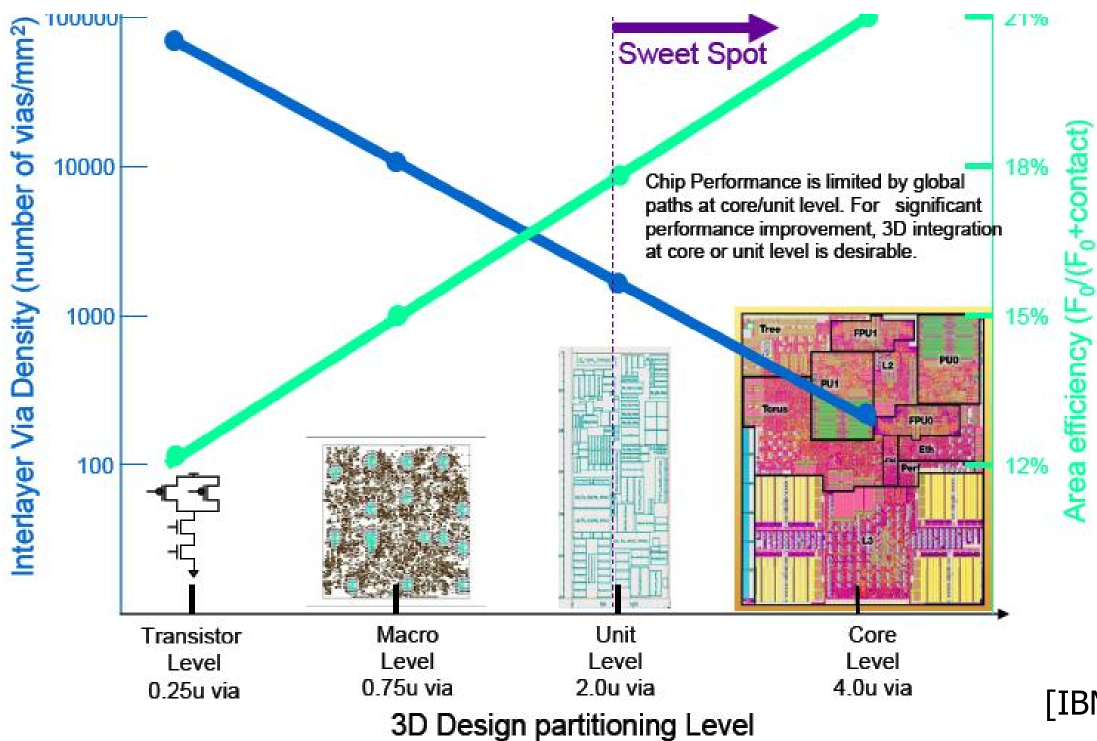
Revisiting DRAM organization?





[Loh08]

A Technological Reality Check...

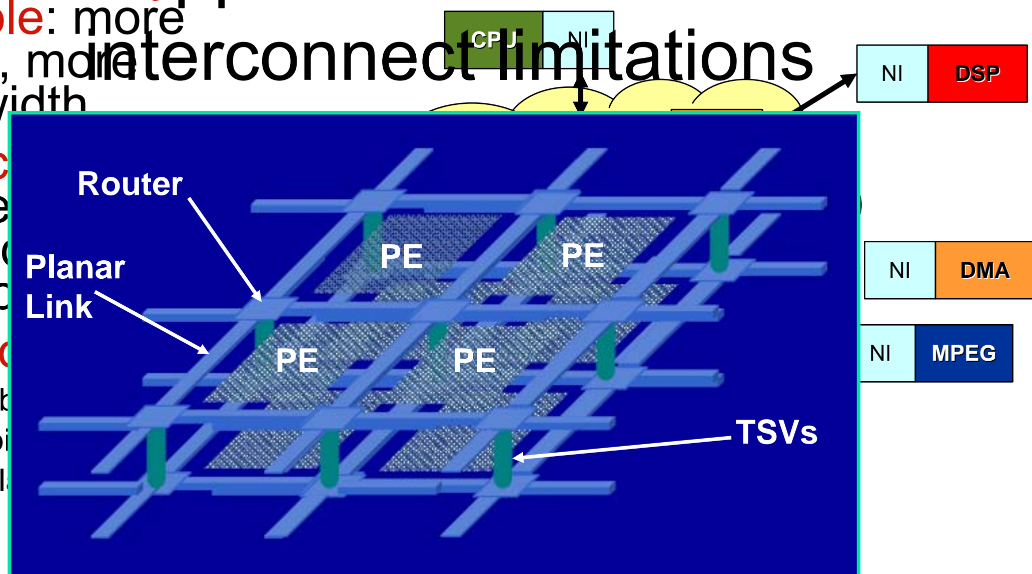


[IBM07]

A long-term vision

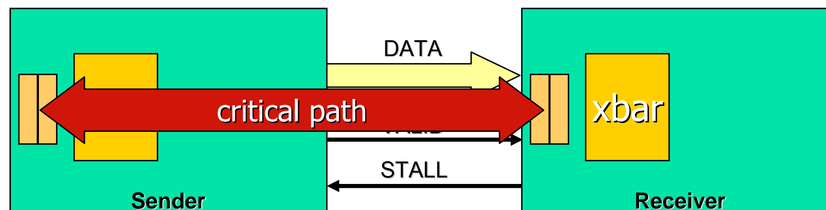
3D Network On Chip

- Module-level interconnect
- Architecturally scalable: more nodes, more bandwidth
- Physic segme neede (predic
- Ideal fo
  - Scalab
  - Low pi
  - Modul

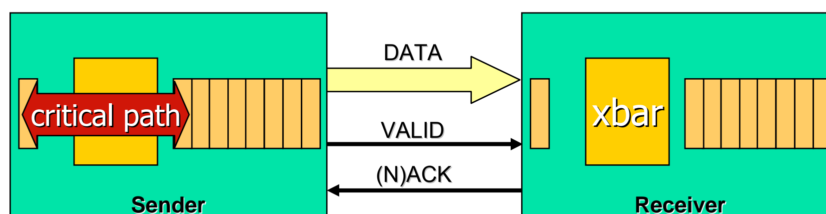


Reference Switch Architectures

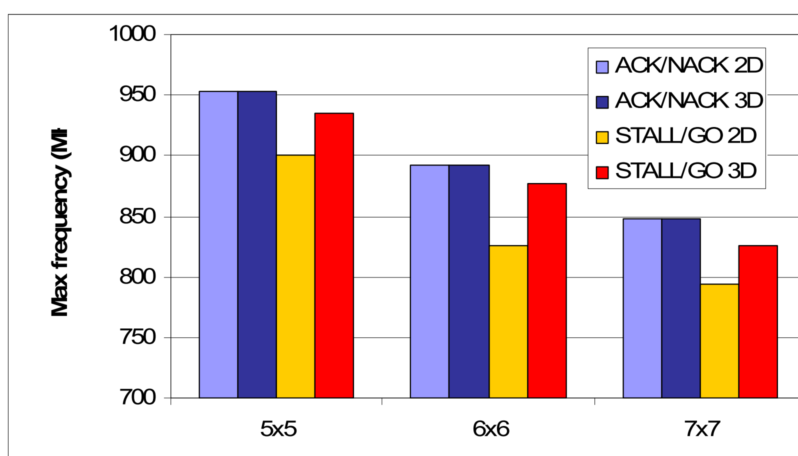
- Based on the xpipes NoC library
- STALL/GO flow control



- ACK/NACK flow control



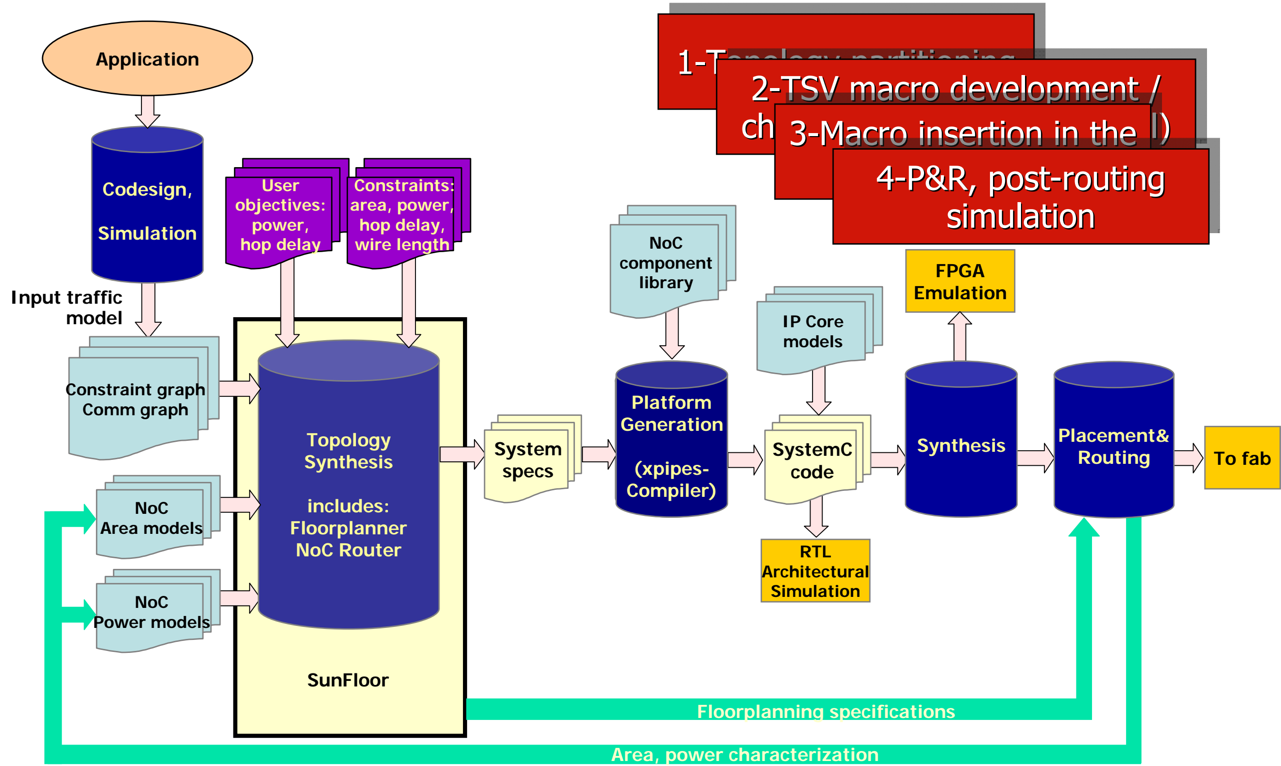
Impact on Max Frequency



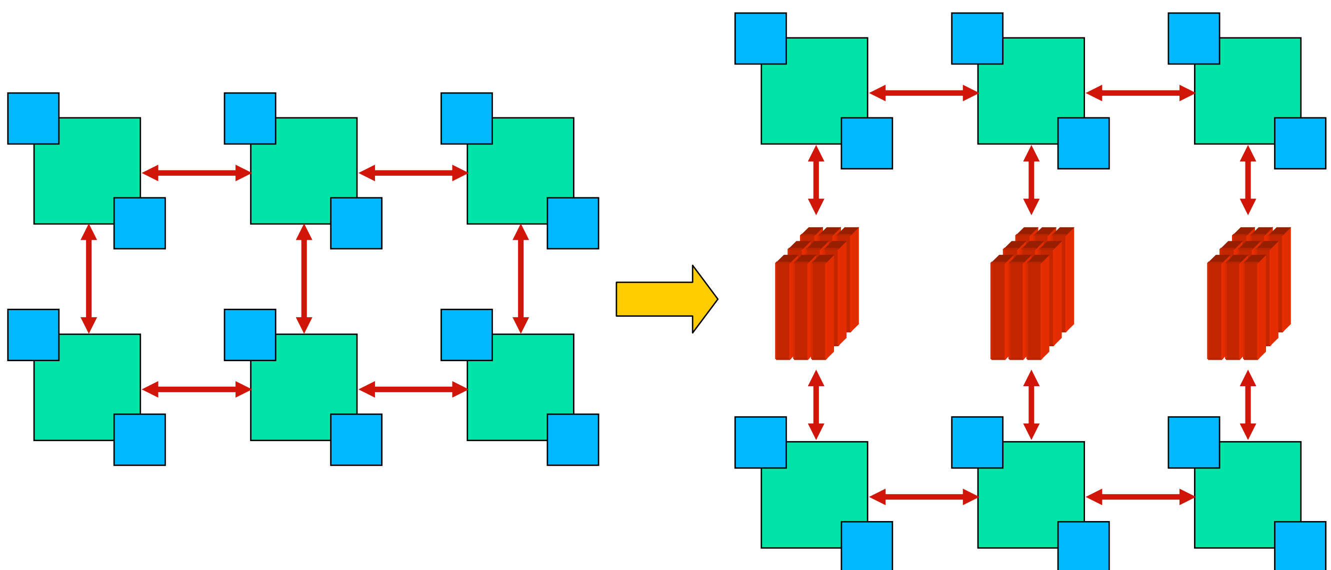
UMC 0.13µm
H links: 1.5mm long
V links: 50µm long

- ACK/NACK: critical path is within the switch module, no difference to be noticed
- STALL/GO: critical path includes the link, up to 50 MHz frequency gain

Reference NoC Design Flow

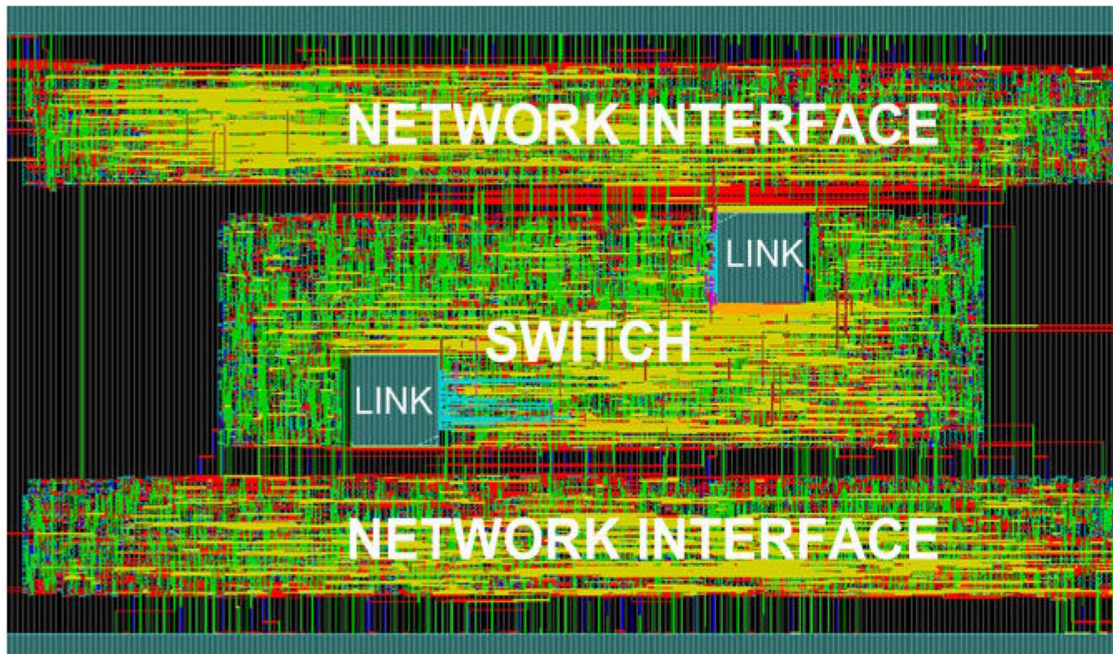


Example Topology



- 2x3 q-mesh is split into two 1x3 layers
- Layers are connected by vertical vias

Bottom Layer Layout



- Vertical links are laid out as floorplan obstructions

Area Overhead of TSV Bundle

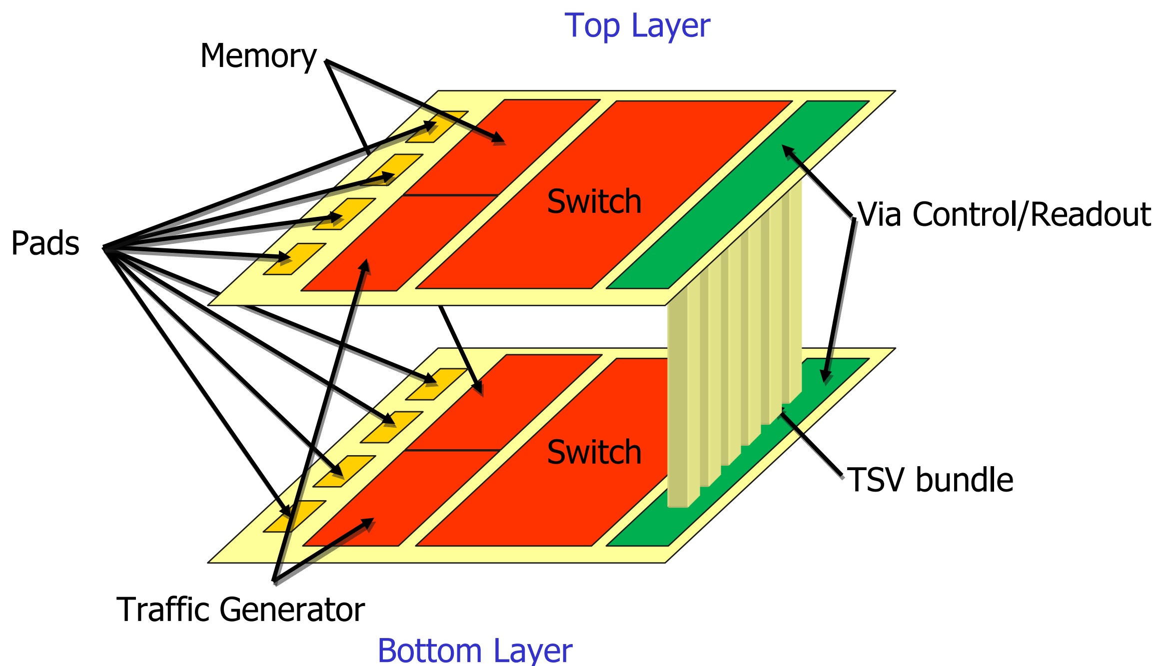
- For a vertical bidirectional link, needed wires:

$$2 \cdot (5 + DataWidth)$$

- At minimum diameter/pitch, each via's overhead is 64 μm^2 (12 NAND2 equivalents)
- For a 6x6 xpipes switch with 28-bit *DataWidth*, overhead is
  - ACK/NACK: 6% of switch area
  - STALL/GO: 9% of switch area (switch is smaller)
- But since frequency is higher, buffering could be reduced (saving area)

3D NoC Test Chip

Taped out in february – Joint project IMEC – INOCs - UNIBO

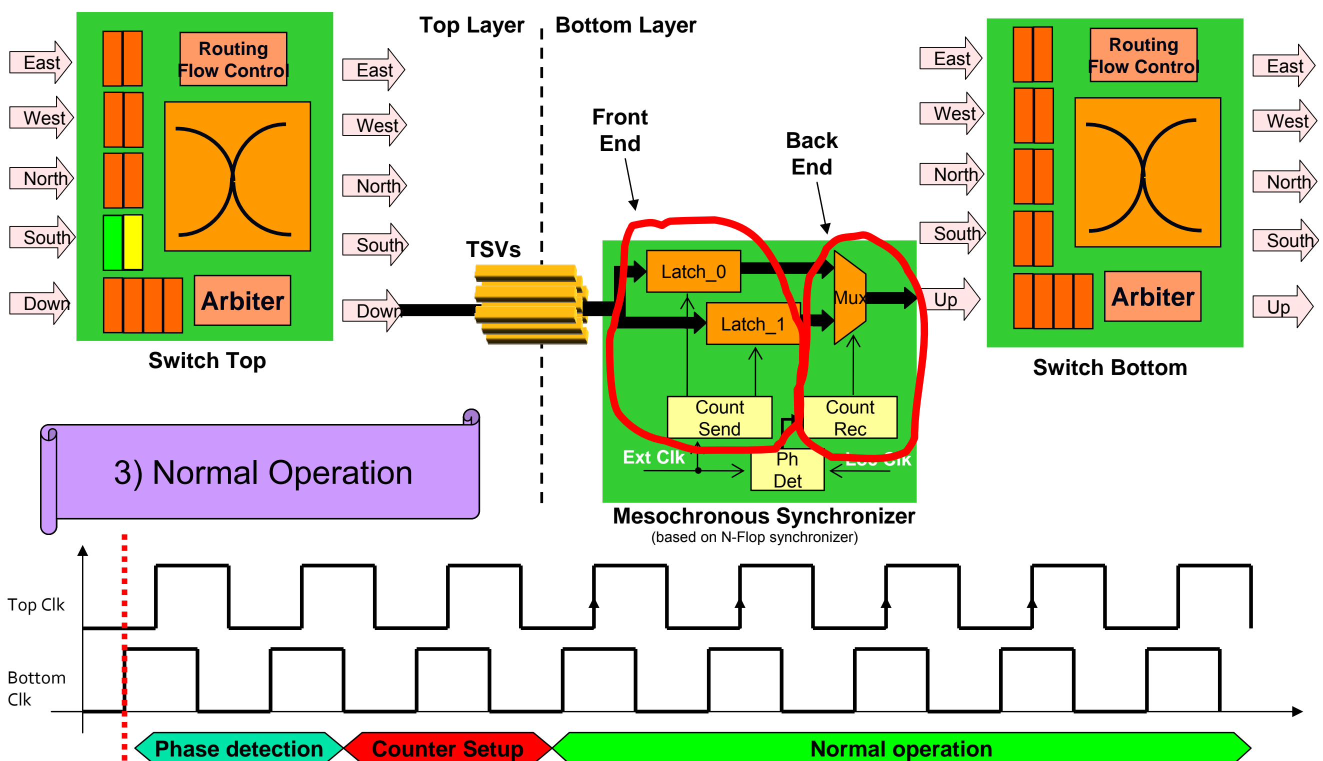


Insight to Be Gained

- Demonstrator: 3D NoC
  - If functional tests are passed
- Electrical performance of 2D vs. 3D interconnects
  - By injecting 2D-only or 3D traffic, and pushing frequency until still operational
- Yield analysis / redundancy policies for 3D vias
  - By querying via status by JTAG
- Skew analysis for 3D clock distribution
  - By skewing layer clocks on purpose and checking max achievable operating frequency
- Power analysis of 3D vias / NoCs
  - By monitoring power in continuous operation mode

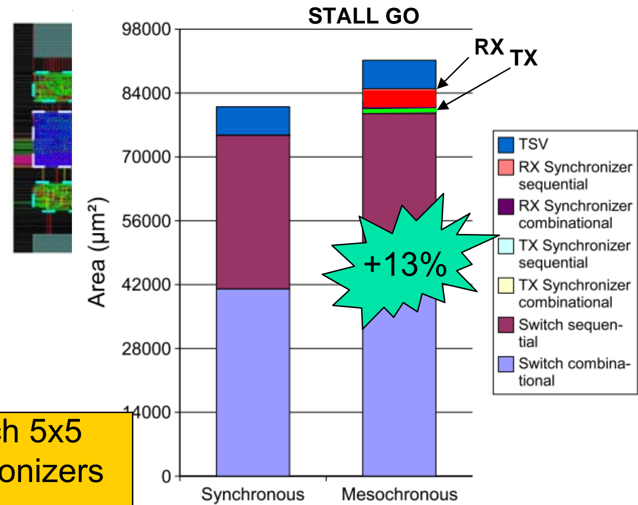
Design Challenges

Mesochronous Synchronization



Performance Analysis

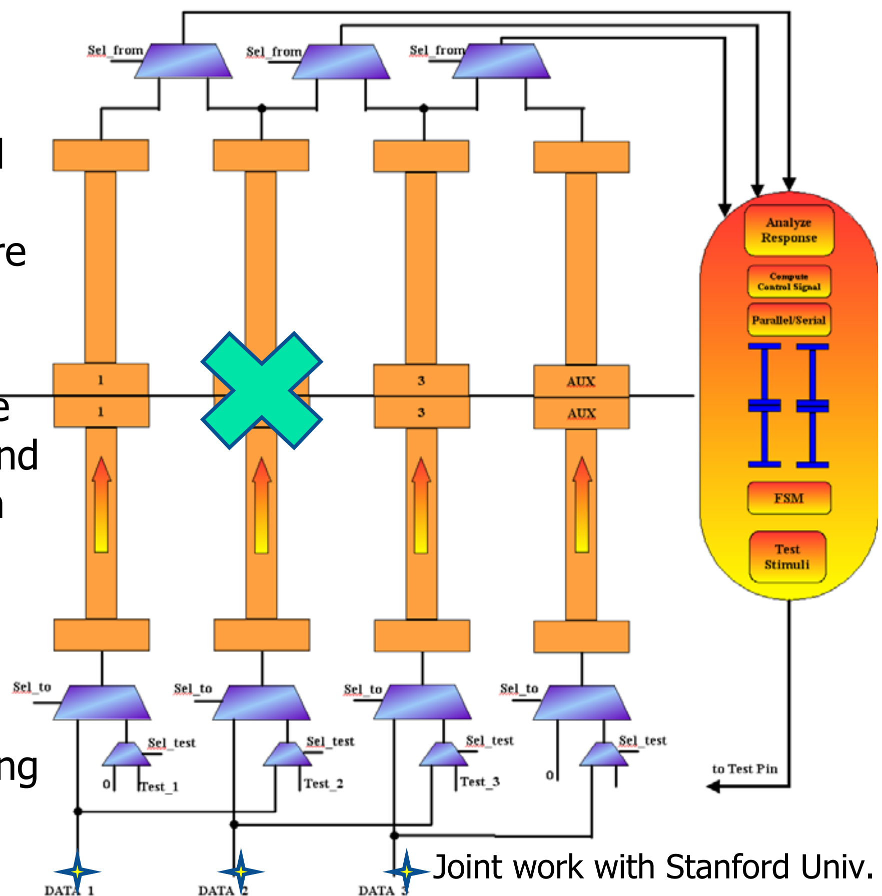
- Low-overhead mesochronous synchronizer
- Tested with 32bit flit width: small area overhead:
  - RX Synch → 3200μm²
  - TX Synch → 700 μm²
- Due to the introduced latency additional buffer resources are needed to avoid data loss or throughput penalty.
- Stall Go require at least two additional buffers



Cost for a baseline composed of 2 switch 5x5 TSVs obstruction, mesochronous synchronizers and **flow control support**

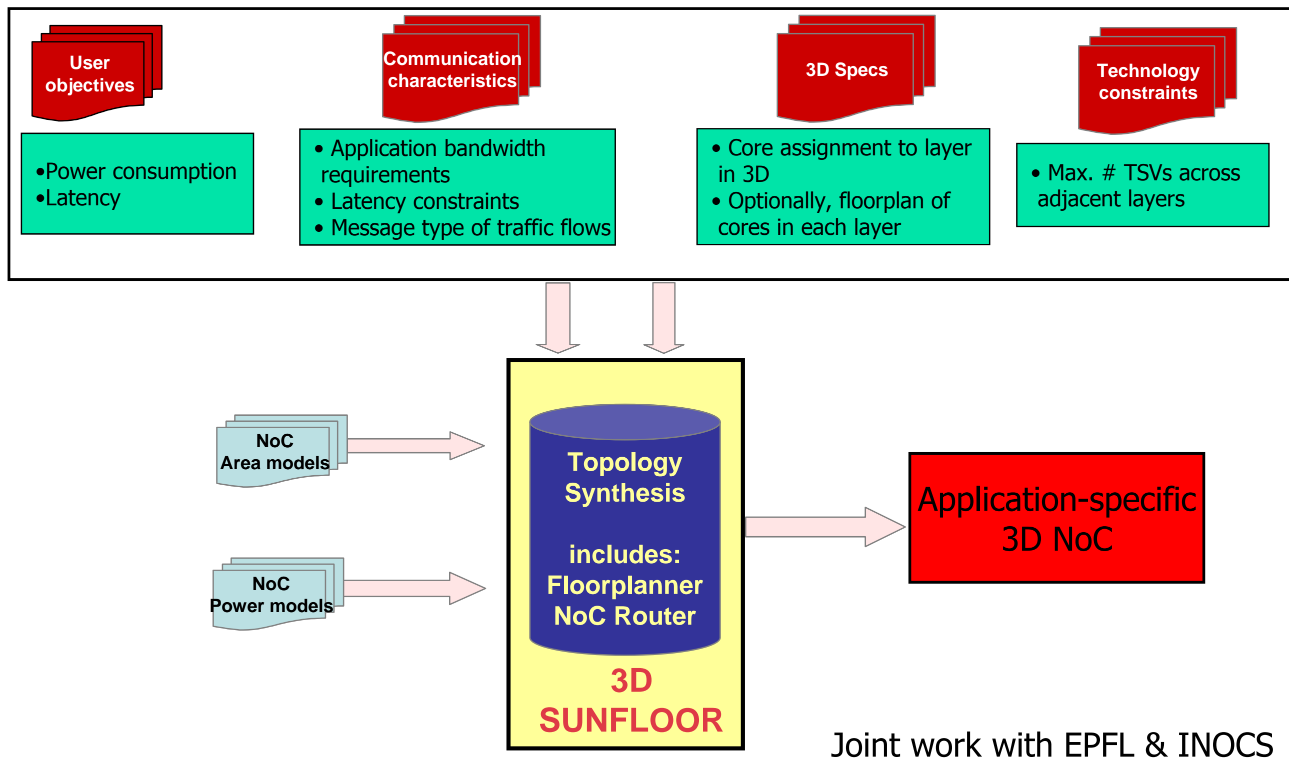
Reliability Enhancement

- TSV check on reset
- Control use dedicated Vias in order to establish which vias are corrupted.
- If 1, 2 and 3 TSVs are OK, the control set the enable signal set_to and set_from: broken path are skipped!
- Pads routing shift as show in the figure
- Need to define The handling protocol during the TSVs check



Joint work with Stanford Univ.

3D NoC Topology Synthesis



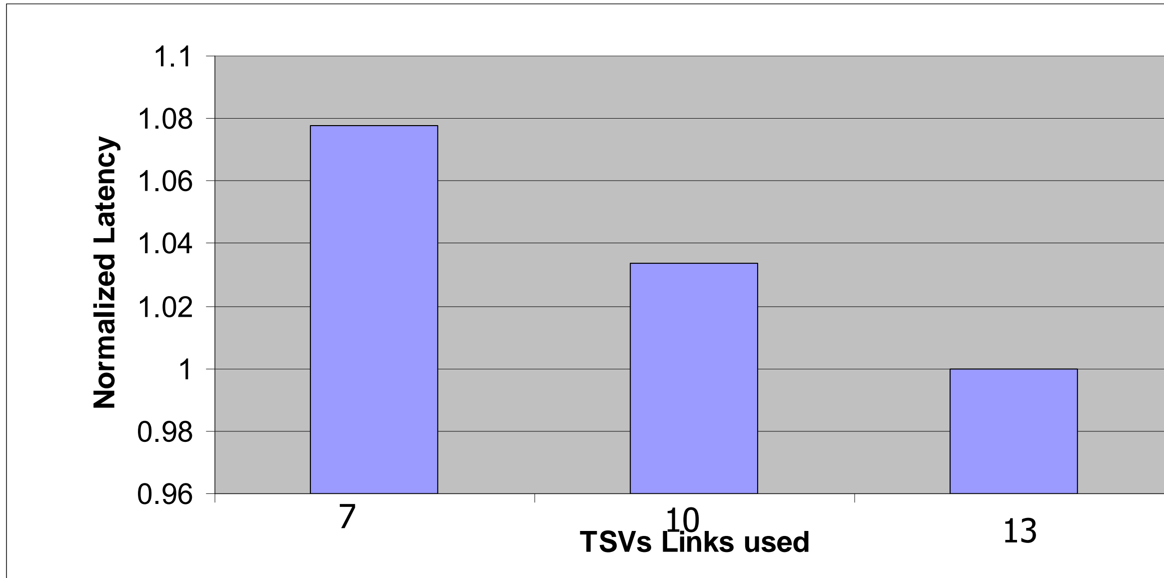
Joint work with EPFL & INOCS

Topology Synthesis Algorithm

- Features:
  - Deadlock removal (routing and message-dependent) intra and inter layer
  - Floorplan of cores, switches, NIs layer by layer
    - TSV alignment across layers is yet to be done
  - Meet frequency, TSV constraints
- Design Trade-offs
  - Vary number of TSVs → performance Vs yield analysis
  - Effect of core to layer assignment on communication
  - Effect of floorplaning on communication

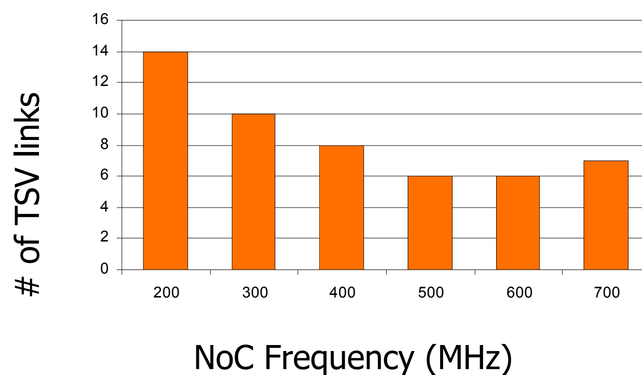
Effect of TSV Constraint

Normalized with respect to unbounded resources case (which needs 13)



Links have 32bit payload

Effect of NoC Frequency

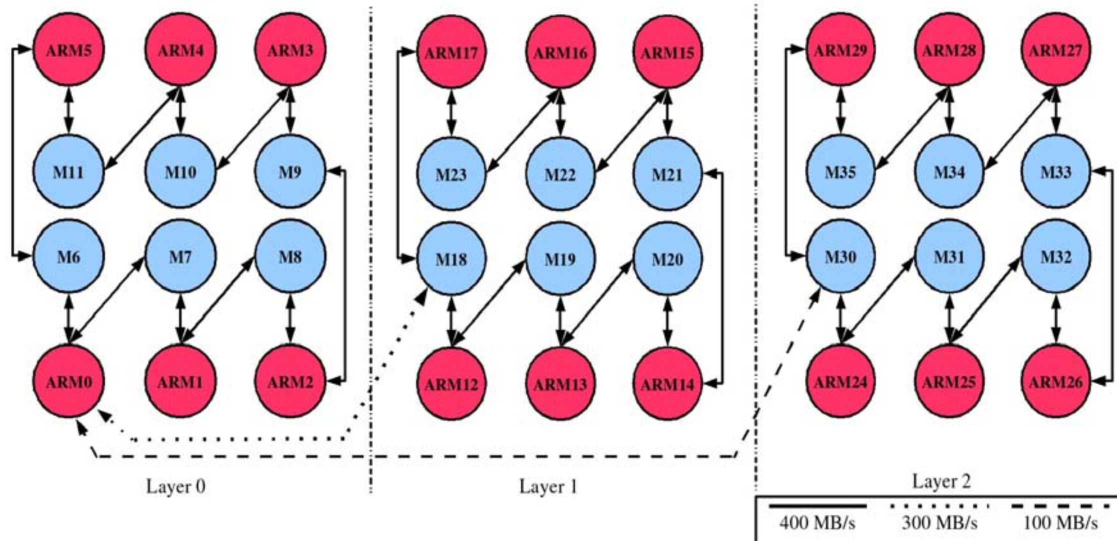


NoC Frequency (MHz)

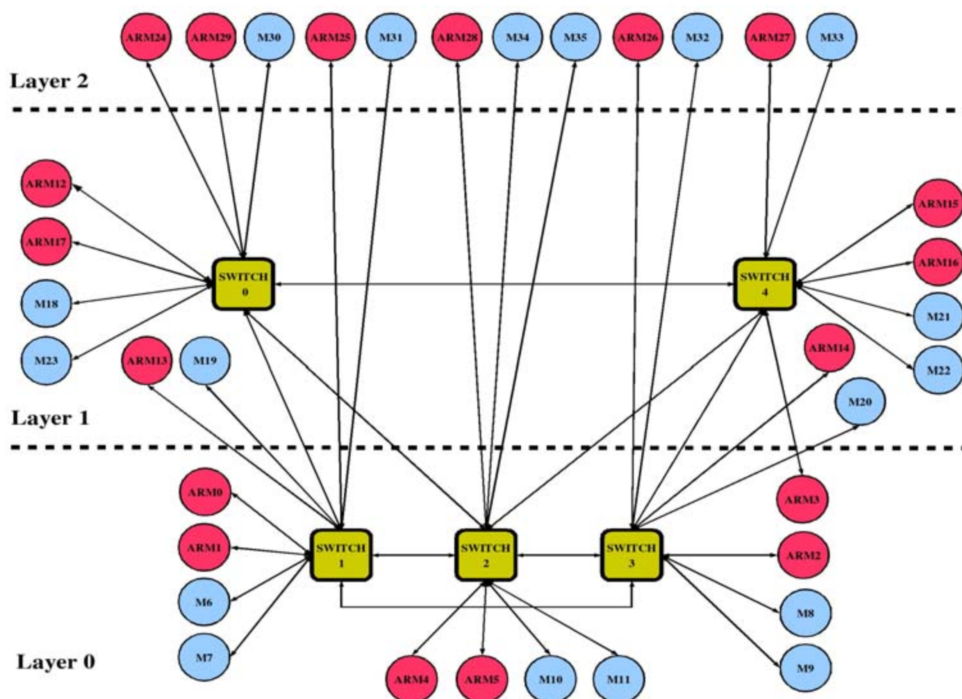
- With NoC frequency increase
  - Fewer links are needed to support same bandwidth → fewer TSVs needed
  - Smaller, more, switches are needed → more TSVs needed
  - Trade-offs to be explored

Case Study

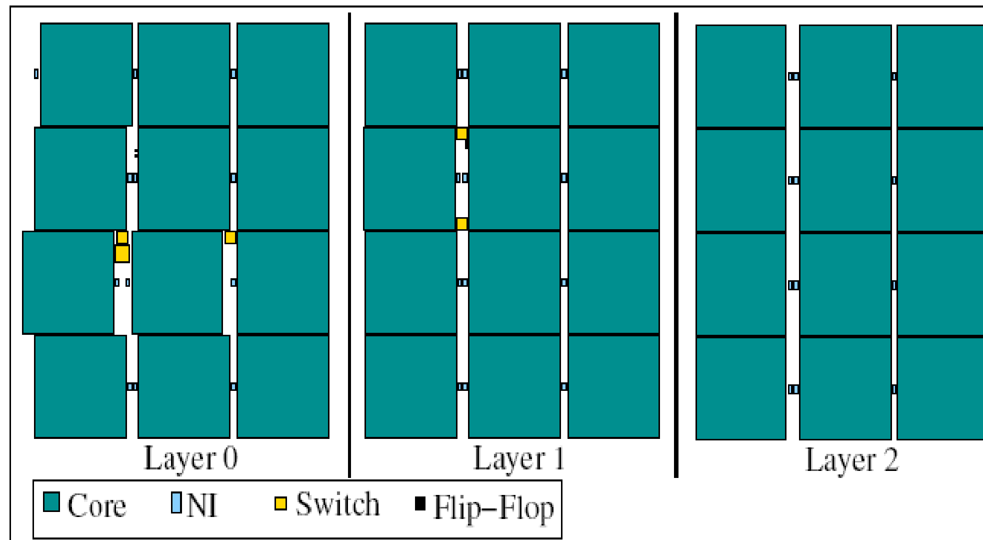
- 36 core multi-media benchmark
- Mapped onto 3 layers



Generated 3D Topology

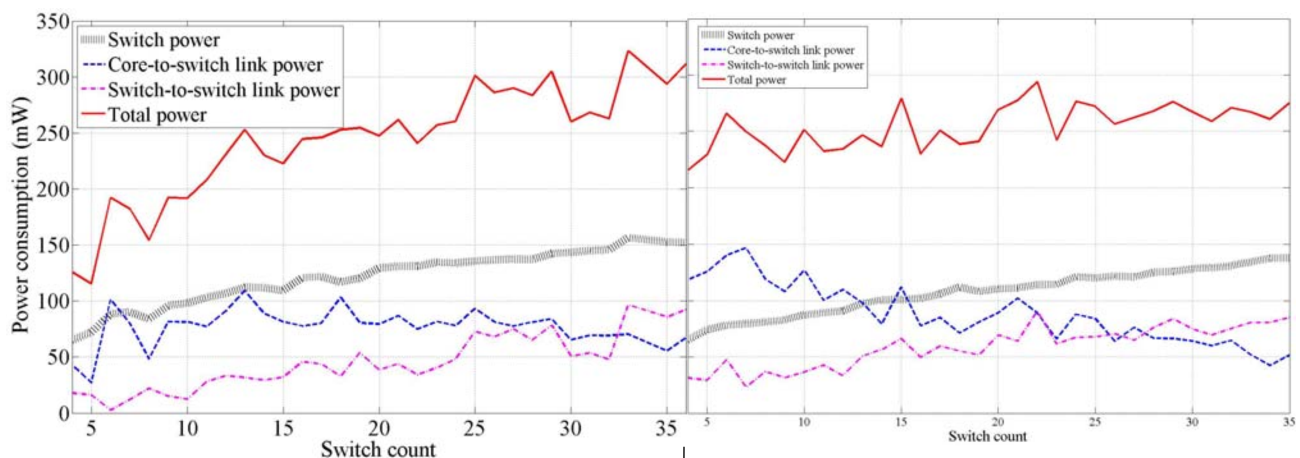


Design Floorplan



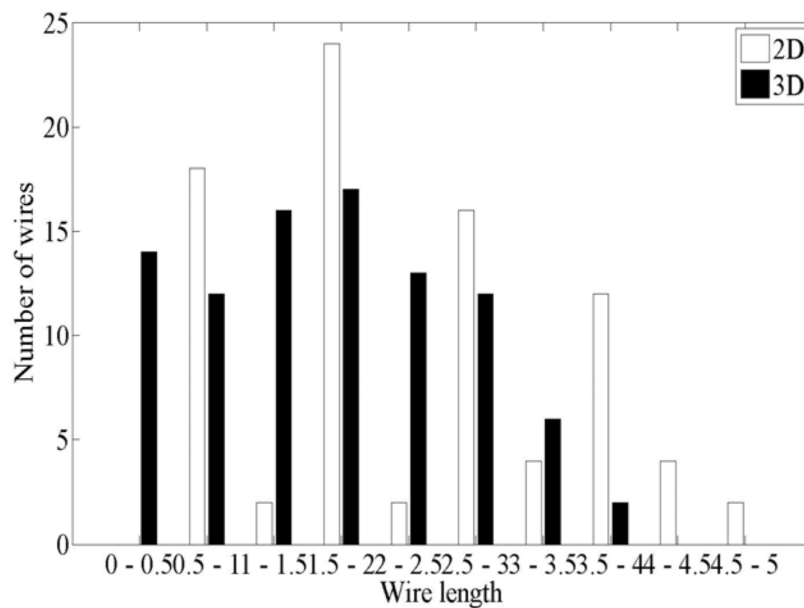
Each core is assumed to be of dimension 1mmx1mm

Comparison with 2D NoC



- Number of switches increases, total power increases
  - More Switch-to-switch wires
  - More switch power, as more hops traversed
- Longer wires in 2D, higher NoC power consumption
- Studied several benchmarks for 2D vs 3D comparisons
  - 32% lower NoC power in 3D
  - 15% lower NoC latency in 3D

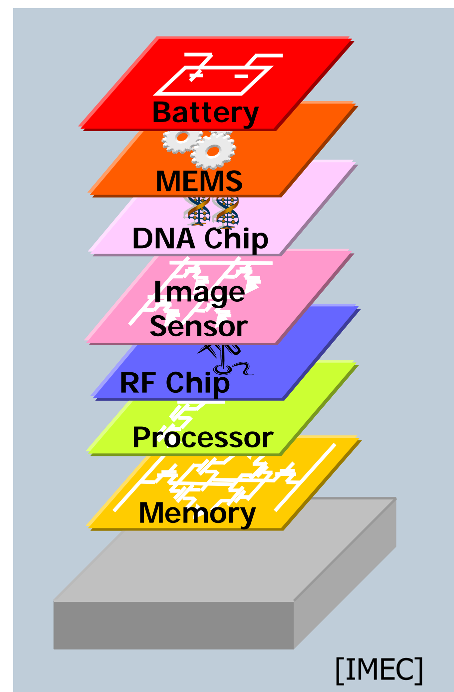
Wire Length Distribution



3D design has many shorter wires

Conclusions

- 3D IC revolution is happening
  - “Evolutionary revolution”
- Adoption in high volumes will be slow
  - Technology needs time to mature
  - Cost needs to come down
- Architectural and circuit solutions can help
  - 3D-NoC could become a strategic technology



The future…

Thank You!