Comment rendre la technologie de l'intégration 3D avec ces TSVs coûteux possible en utilisant un NoC 3D

Abbas Sheibanyrad
TIMA, SLS
Grenoble, France

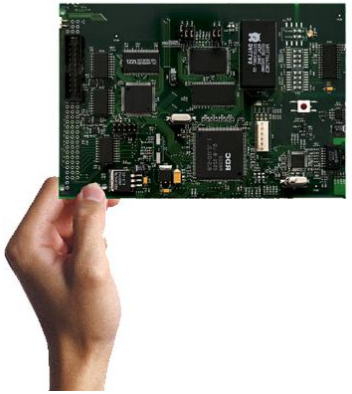
Outline

- 3D-Integration and Asynchronous NoCs
- Vertical Link Serialization
- Vertically-Partially-Connected 3D-NoC
- Conclusion

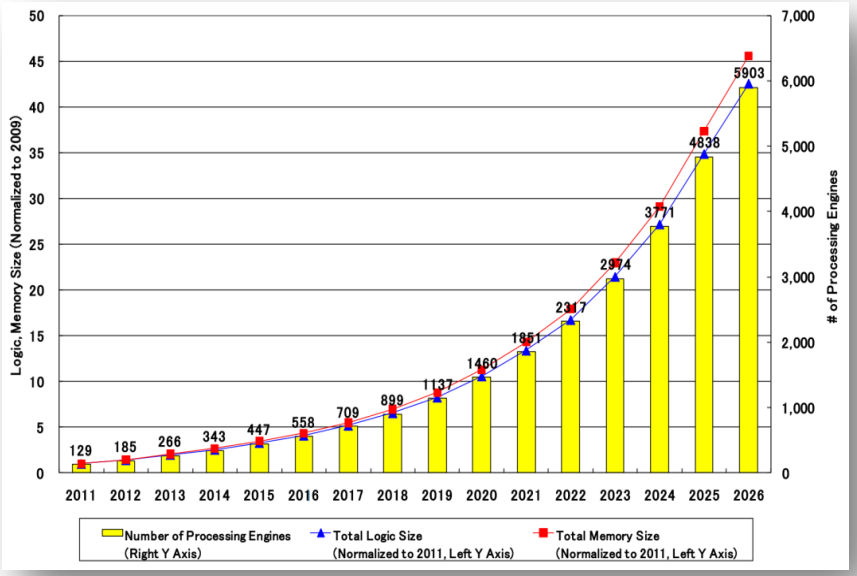
Technology Evolution !

- Evolution of the fabrication technology
  - Integration of systems with billions of transistors in only one chip
  - Hundreds and even thousands of components

System-on-Board



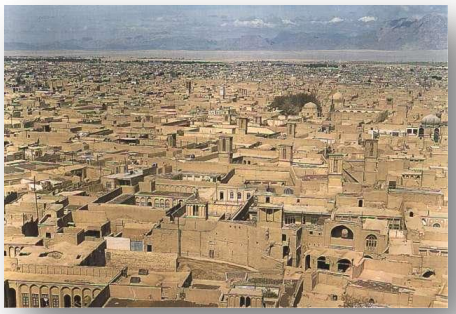
System-on-Chip



- Multi-Core Systems
  - Performance
    - Thermal issues and clock skews limit the increase of clock frequency
    - Parallelism
  - Power Consumption
    - Simple Cores
    - Power down the idle cores
  - Rapid Design
    - Reuse of repeated tiles integrated into a common infrastructure

… when the population grows!

- As the population grows, there is a tendency to build vertically rather than horizontally
  - Increase the density
    - The land becomes more and more expensive
  - Decrease the length and the number of long paths
    - The average time and energy of moving from one point to another becomes unaffordable



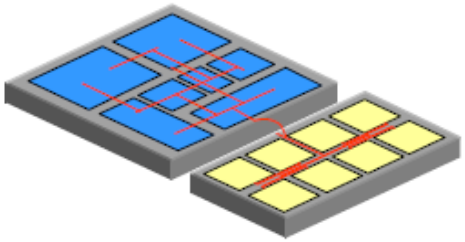
Three-Dimensional Integration

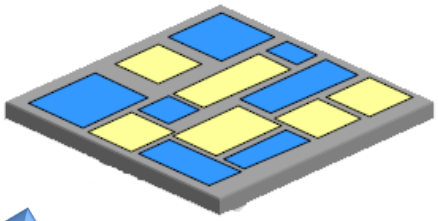
System-on-Board

no External Connections

System-on-Chip

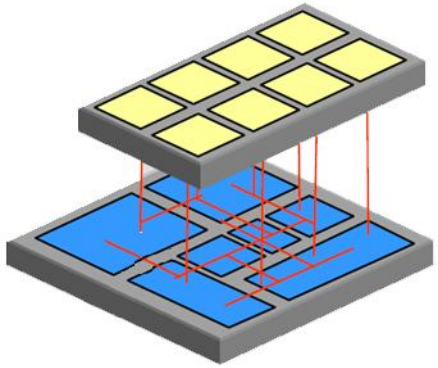
Large Die



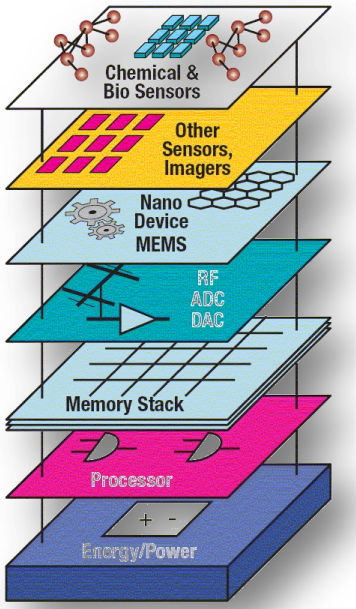


Short and Direct Connections

Small Dies

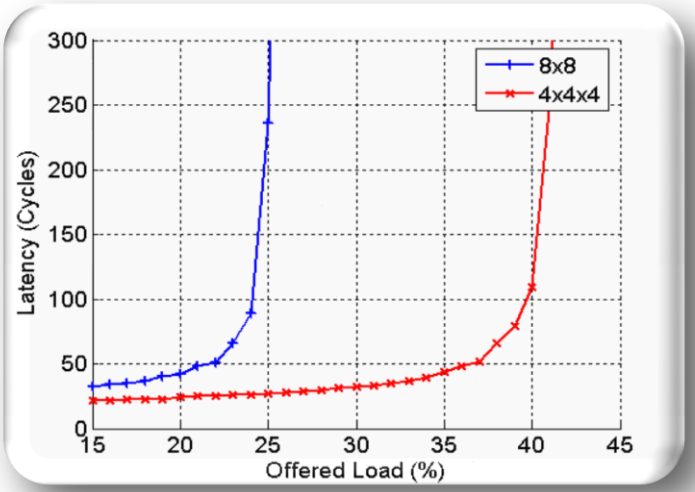


* Possibility of Integrating Different Dies with Different Technologies in a Same Package



*James Lu

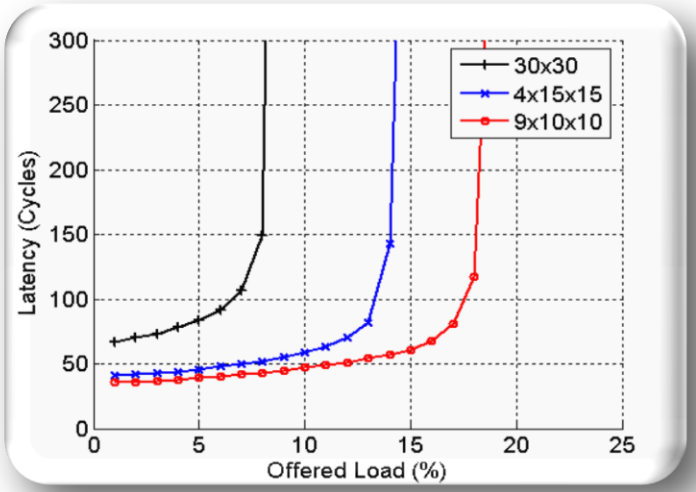
Contribution of the Third Dimension

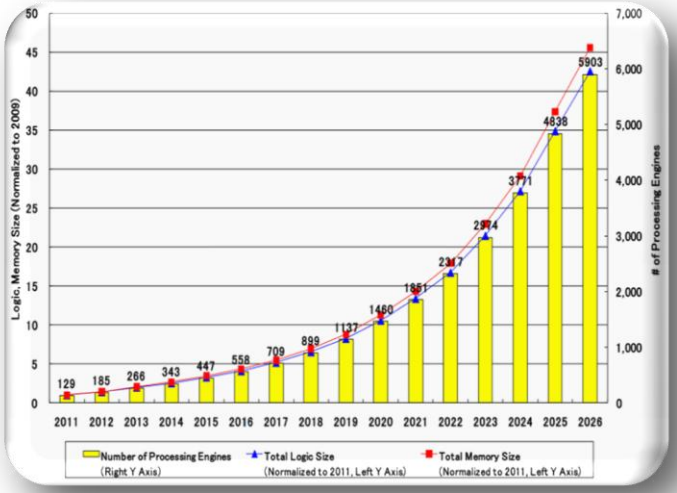


	Number of Nodes	Switch Degree	Network Diameter	Number of Channels	Number of Vertical Channels	Number of Bisection Channels	Load of the Busiest Channels [1]
2D-Mesh	$N = n^2$	5	$2\sqrt{N}$	$6N - 4\sqrt{N}$	0	$2\sqrt{N}$	$C \times \frac{1}{4}\sqrt{N}$
3D-Cube	$N = m^3$	7	$3\sqrt[3]{N}$	$8N - 6\sqrt[3]{N^2}$	$2N - 2\sqrt[3]{N^2}$	$2\sqrt[3]{N^2}$	$C \times \frac{1}{4}\sqrt[3]{N}$

[1] Assuming uniform destination distribution and dimension-ordered routing, C is the average load injected to the network by each node

How Many Layers?



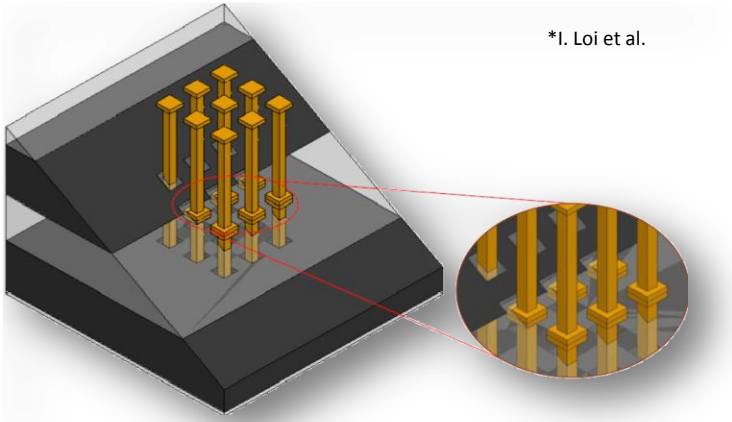


	Number of Nodes	Switch Degree	Network Diameter	Number of Channels	Number of Vertical Channels	Number of Bisection Channels	Load of the Busiest Channels [1]
30x30	900	5	60	5280	0	60	C × ¼ × 30
4x15x15	900	7	34	6510	1350	120	C × ¼ × 15
9x10x10	900	7	29	6640	1600	180	C × ¼ × 10

[1] Assuming uniform destination distribution and dimension-ordered routing, C is the average load injected to the network by each node

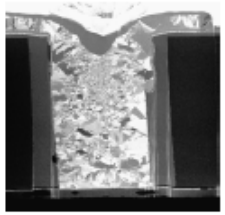
Through-Silicon-Via

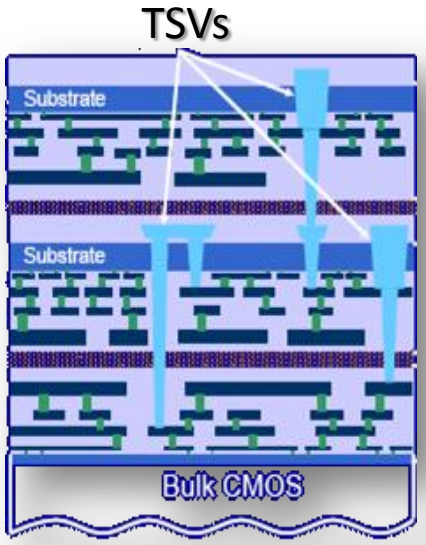
- The most promising Technology of Vertical Interconnection
  - Low Resistance and Capacitance
    - High Bandwidth
    - Low Power Consumption



- Via-First (higher density of TSVs)
  - Diameter ≈ 5 µm
  - Pitch ≈ 10 µm
  - Depth ≈ 20-50 µm

- Via-Last (lower cost of the process)
  - Diameter ≈ 35 µm
  - Pitch ≈ 50 µm
  - Depth ≈ 40-150 µm



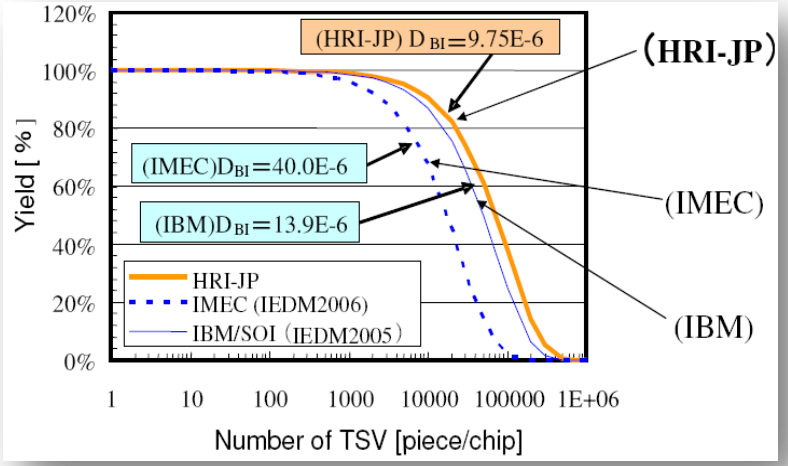


…but, is there any problem ?

- Large area overhead because of large TSV pitch, mainly due to the large pads to compensate misalignment of dies

- Important risk of failure due to several additional fabrication steps (a potential reduction on the Yield)
  - Misalignment
  - Dislocation
  - Void formation
  - Oxide film formation over Copper interfaces
  - Pad detaching
  - Defects due to temperature
  - …

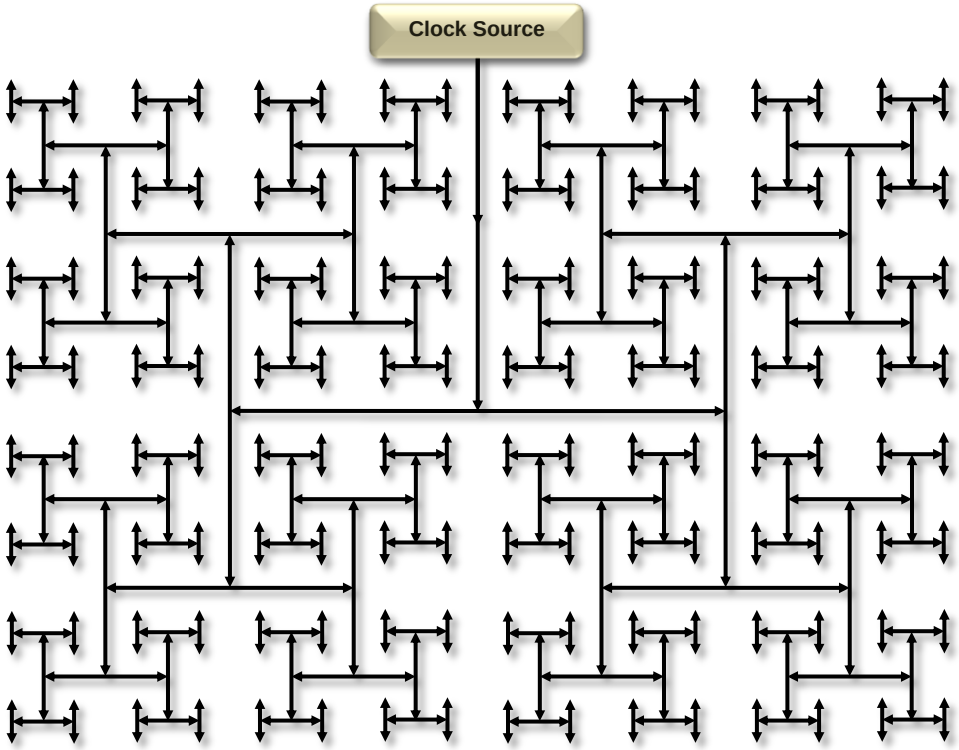
The Three-Dimensional Integrated Circuits are limited by the number of TSVs to be exploited

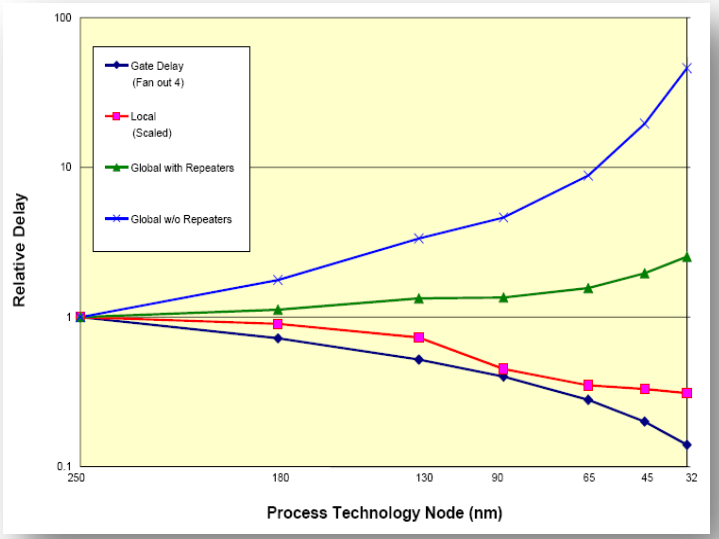


*C. Seiculescu et al.

Clock Distribution

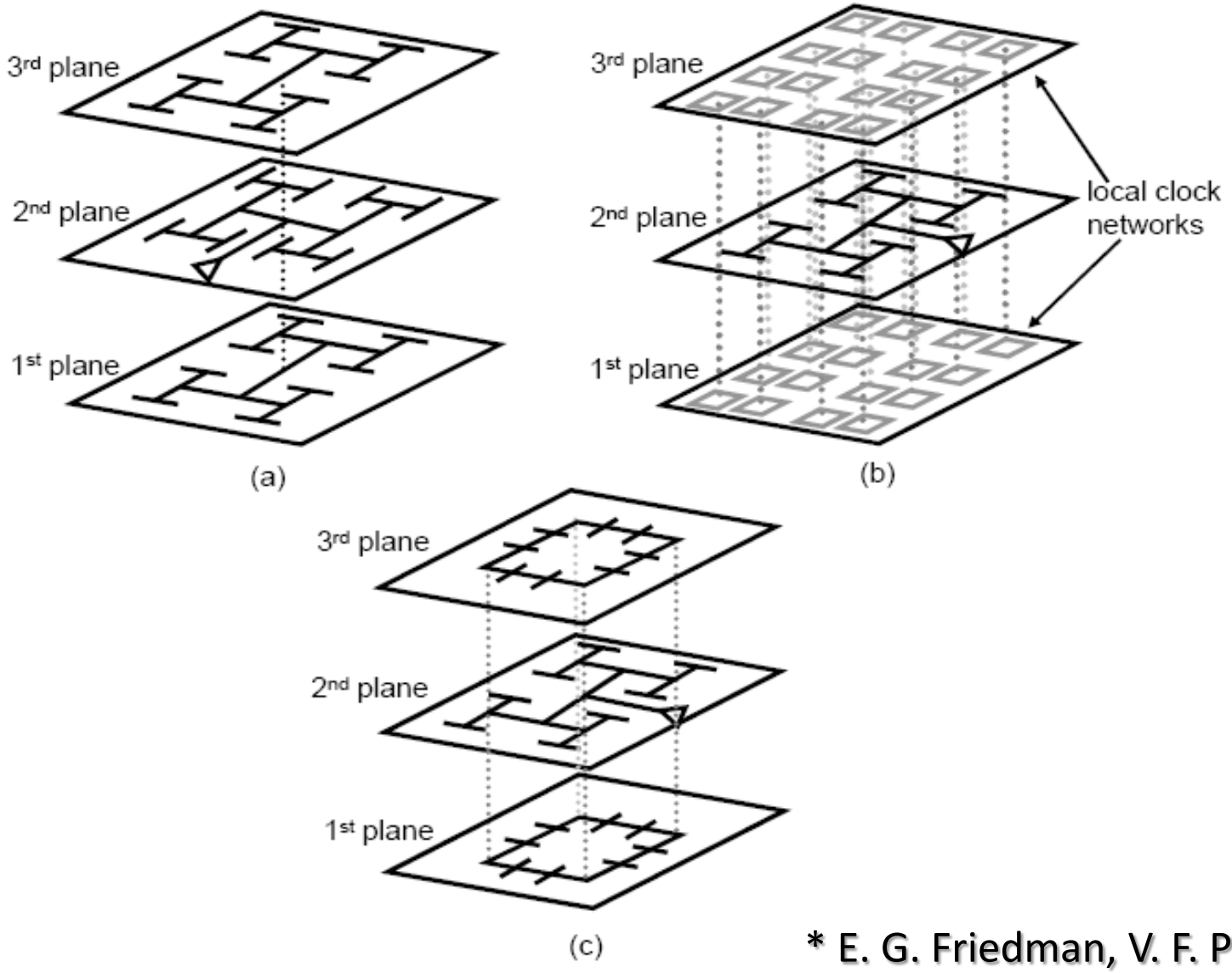
- Deep Submicron Technologies
  - Aggravation of physical problems
  - Predominant effect of long wires on delay and consumption





- Nightmare of Global Synchronization
  - Impossible Global Distribution of a single clock signal over a chip
  - Clock skew claiming a larger relative part of the total cycle time
  - The clock distribution network demanding increasing portions of the power and area budget
  - Fabrication Process Variation
  - Temperature Variation

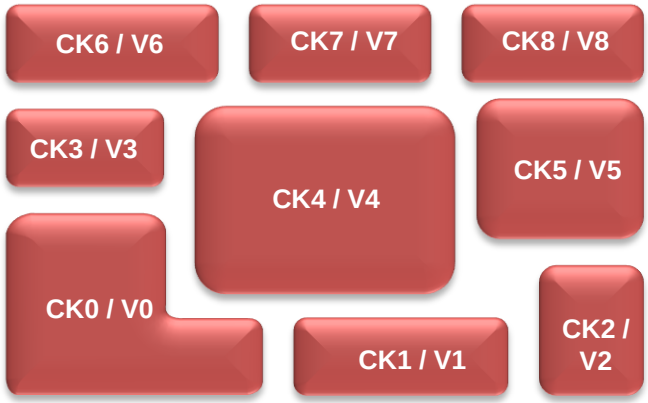
Clock Distribution in 3 Dimensions

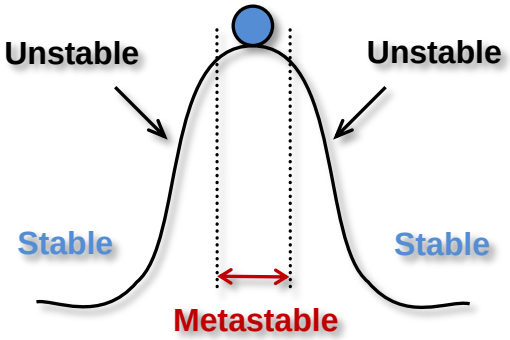


* E. G. Friedman, V. F. Pavlidis

GALS/DVFS always demanded !

- Reducing the Problem to a number of smaller Sub-Problems
  - Several Independent Clock/Voltage Clusters
- Networks-on-Chip are the most Structured Approaches
  - The Network is the asynchronous global part of the system
  - The subsystems are the synchronous local parts of the system



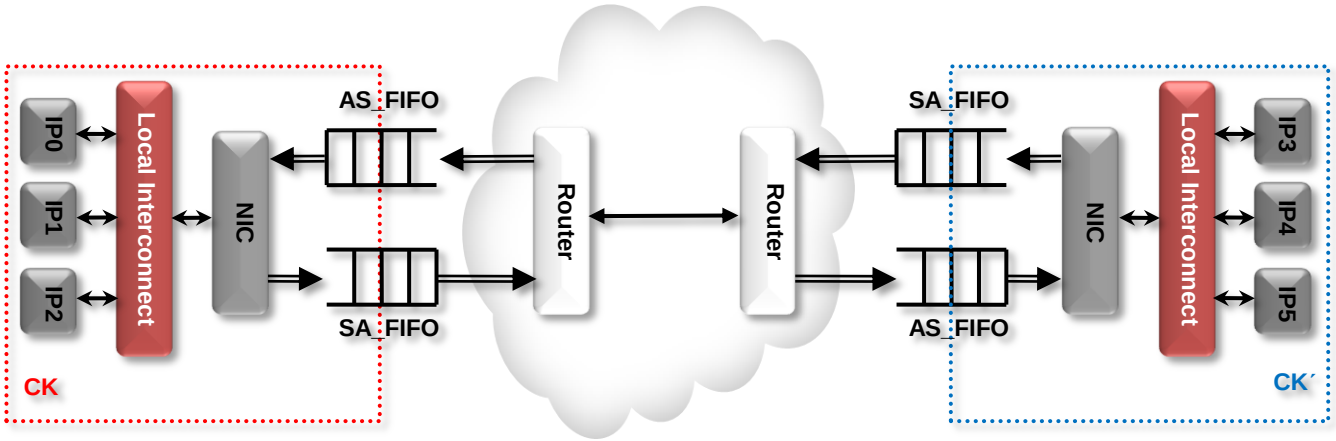


… but, how can two separately clocked domains communicate in a reliable manner ?

- Metastability, an unavoidable state of a bistable system, is the major problem of the GALS architectures

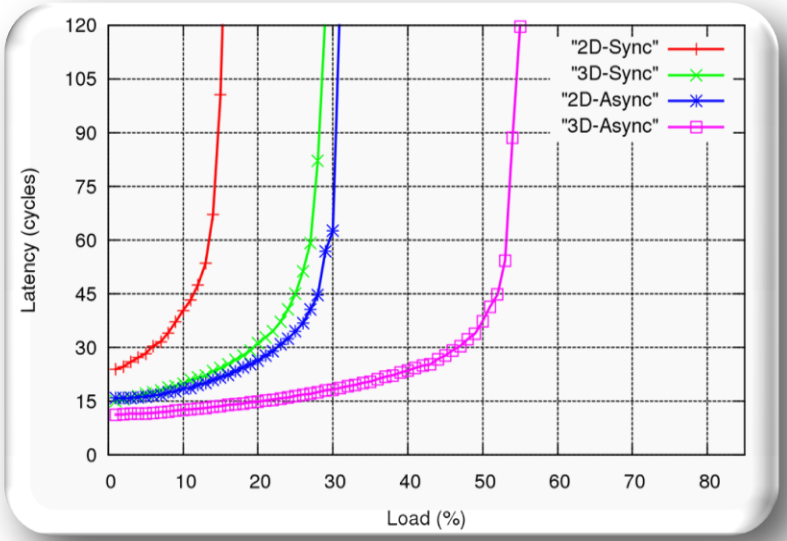
Asynchronous NoC (e.g. ASPIN)

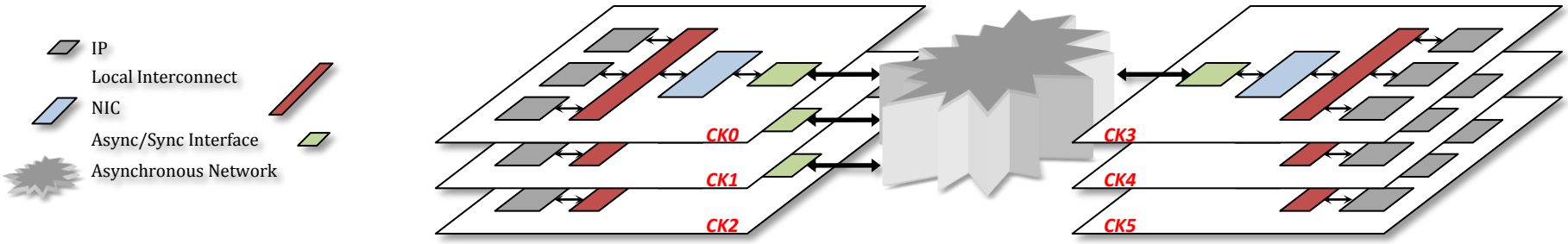
- The need of synchronization reduced to the network interfaces
  - Special FIFOs: Async-to-Sync and Sync-to-Async
  - An End-to-End latency much lower than the multi-synchronous version
- As fast as possible and independent from the rest of the circuit
  - Saturation threshold improved compared with the multi-synchronous version
- The almost zero dynamic Power Consumption in the idle state
- Scalability and Reusability in a Plug & Play fashion and independent from the size of subsystems



Asynchronous 3D-NoC

- Insensitive to Delay Variation due to Temperature Variation or Process Variation
- Exploitation of the whole high Bandwidth of TSVs
- Speed ratio of 2 as a worst-case assumption
  - Using STMicroelectronics 90nm GPLVT transistors, 400MHz as the maximum frequency of usual SoCs
  - Using the same technology, 1100 Mflits/s as throughput of an asynchronous NoC





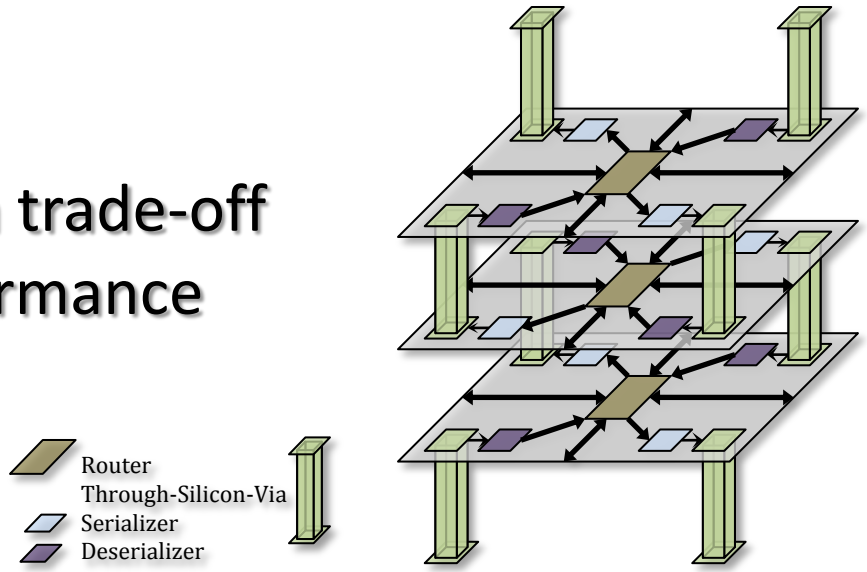
Outline

- 3D-Integration and Asynchronous NoCs
- Vertical Link Serialization
- Vertically-Partially-Connected 3D-NoC
- Conclusion

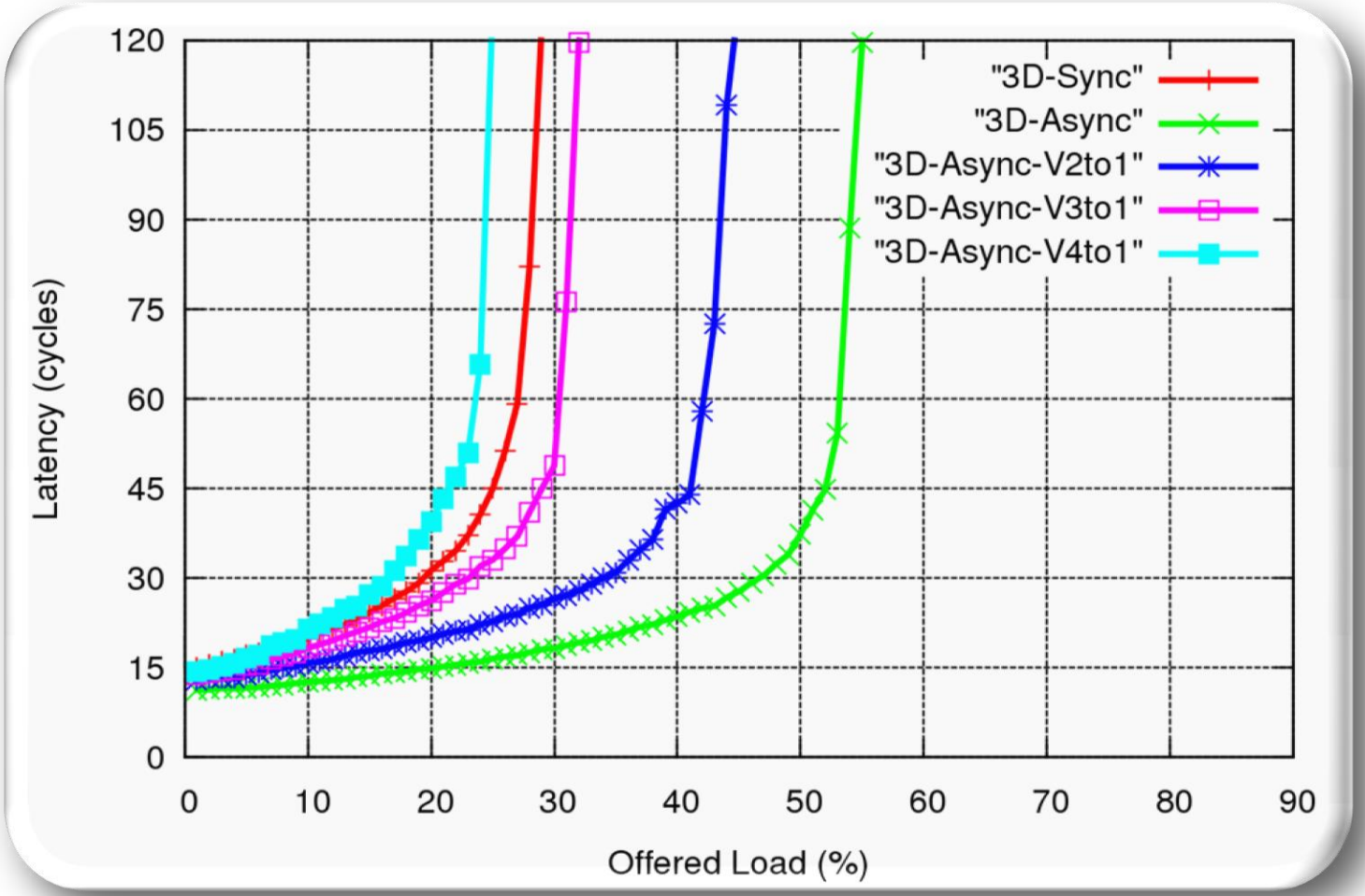
… why not Serialized Vertical Links!

- Remembering
  - Using TSVs guarantees a faster vertical data transfer with lower power consumption than horizontal links in moderate size
  - but, the Pitch of TSVs is large, and, several additional steps of TSV fabrication add a potential reduction of the Yield
  - Only a small fraction of the capacity of vertical link is exploited in a NoC
  - Large number of physical connections for each link of the router

- Serialization of data on TSVs is a trade-off between the cost and the performance



Vertically Serialized Asynchronous 3DNoC



Circuit Implementation

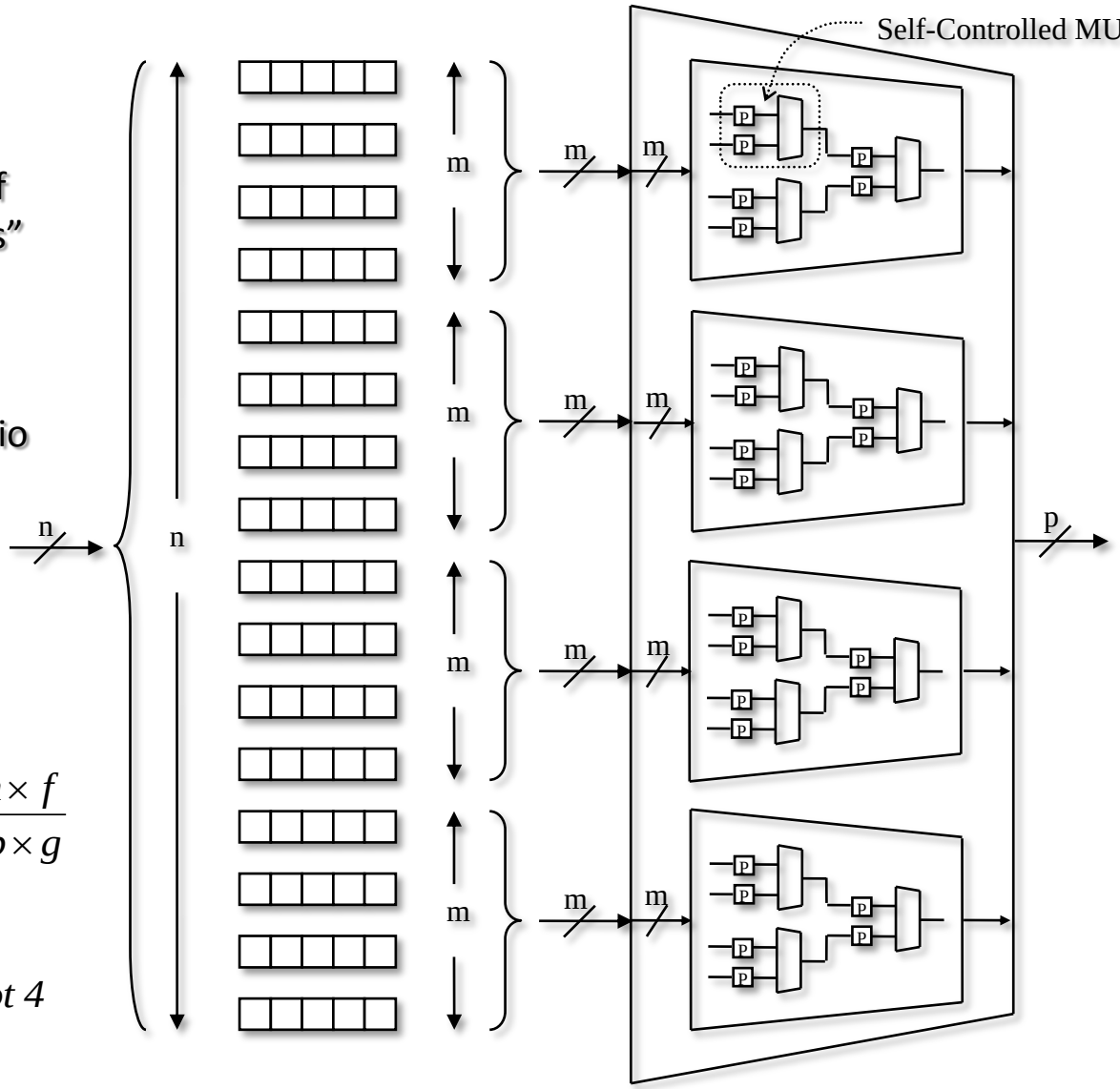
- a Serializer of n:p composed of p Serializer of m:1
  - a Serializer of m:1 is a tree of "Self-Controlled Multiplexors"

$$m = Serialization\ Ratio = \frac{n}{p}$$

- *R*, The Serialization Bandwidth Ratio as the throughput cost factor
  - *f*, the transfer rate of parallel input data
  - g, the transfer rate of serialized output data

$$R = Serialization\ Bandwidth\ Ratio = \frac{n \times f}{p \times g}$$

$$R = \frac{32 \times 750 Mflits/s}{8 \times 2800 Mflits/s} = 1.07, \text{ and not } 4$$



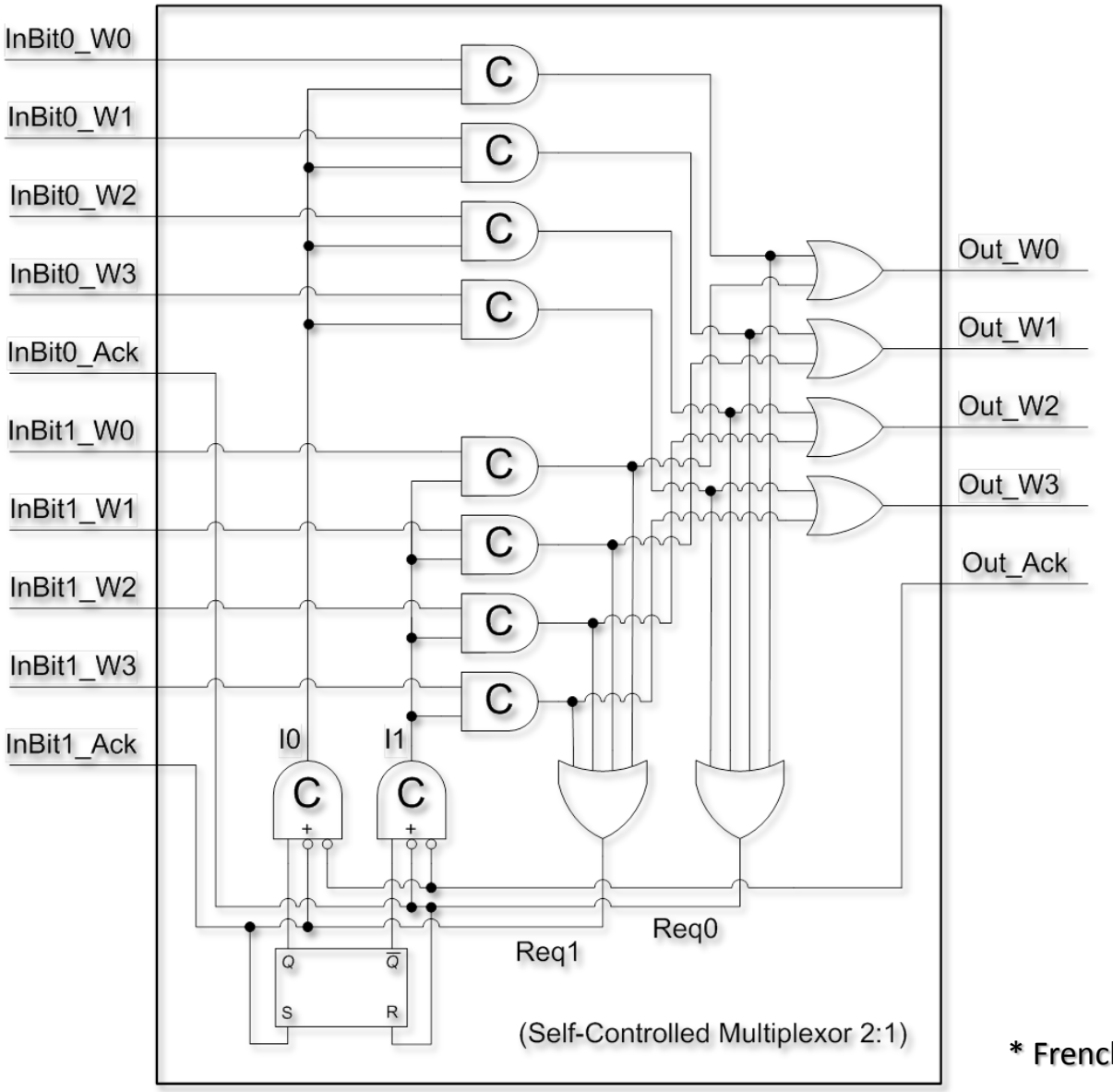
SPICE Simulation Results

- Horizontal Link Throughput: 710 Mflits/sec
  - Router Throughput : 1100 Mflits/sec
  - Inter-Core wire (2mm) delay : 125 ps
- Serialized (8:1) Vertical Link Throughput: 2080 Mflits/sec
  - Serialization Throughput: 2500 Mflits/sec
  - TSV delay: 20 ps

- Speed ratio : (710*32)/(2080*4) = 2.73 (and not 8 !)

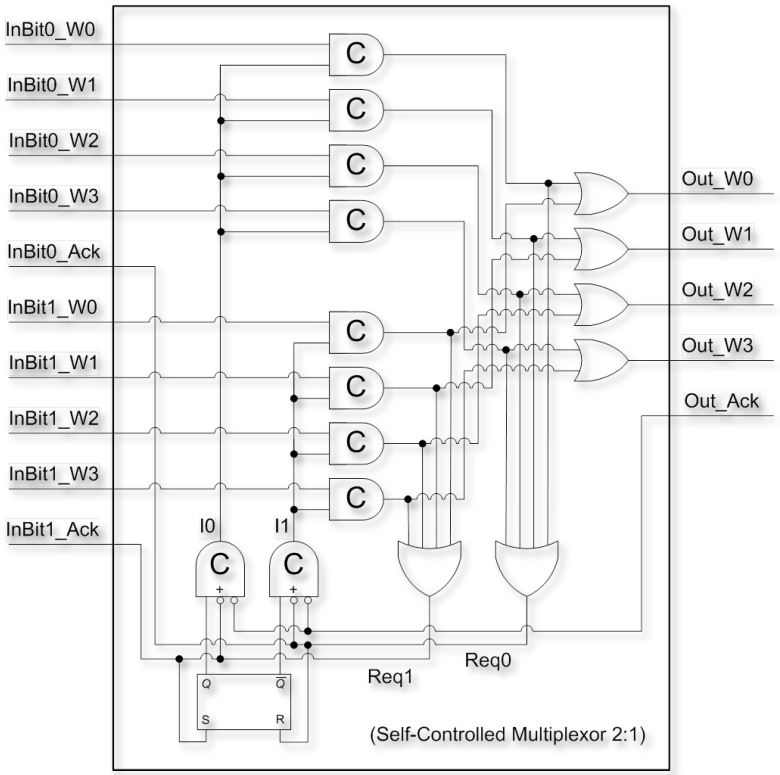
	Self-Controlled Multiplexer 2:1	Self-Controlled Demultiplexor 1:2	Serializer 4:1	Deserializer 1:4	Serializer 8:1	Deserializer 1:8
Transistor count	130	132	390	396	910	924
Latency	80 ps	70 ps	150 ps	130 ps	220 ps	190 ps
Throughput	2.9 Gflits/sec	3.2 Gflits/sec	2.5 Gflits/sec	2.8 Gflits/sec	2.5 Gflits/sec	2.8 Gflits/sec

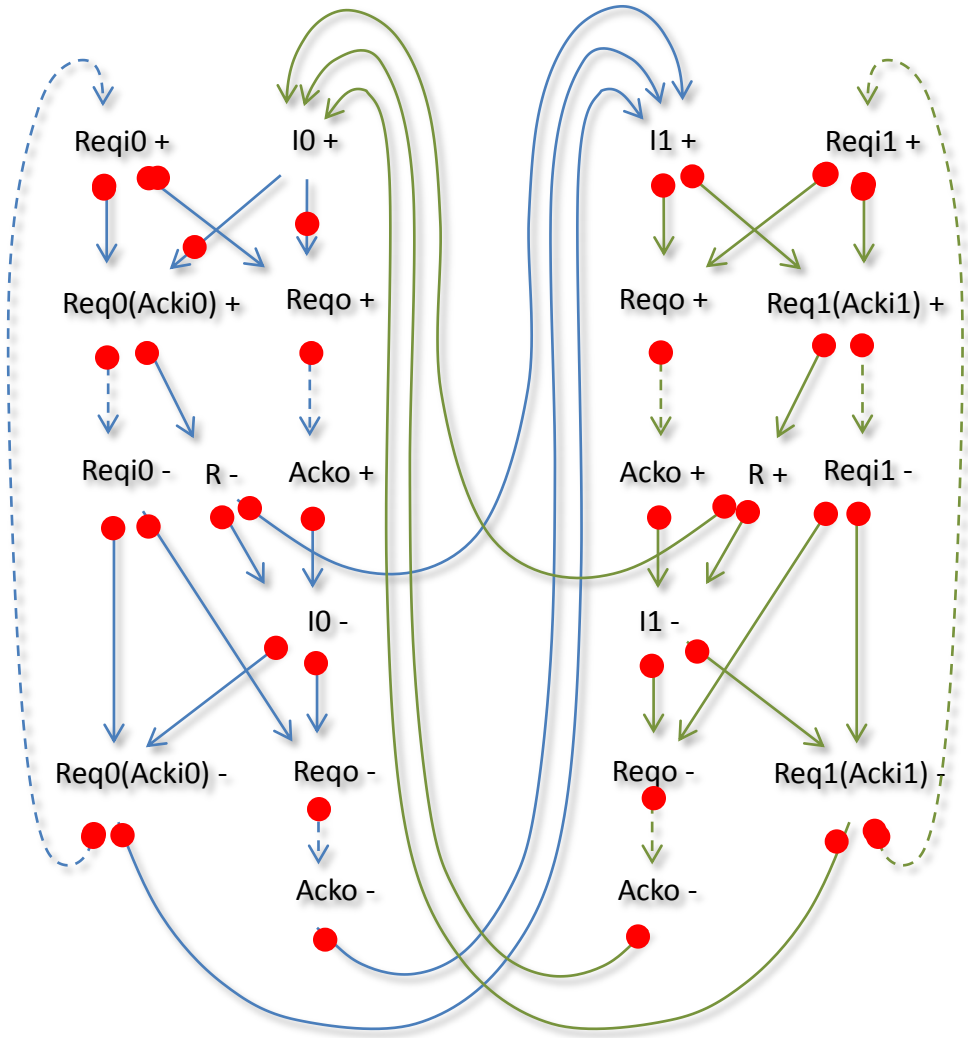
Self-Controlled Multiplexor



* French Patent 09/53637

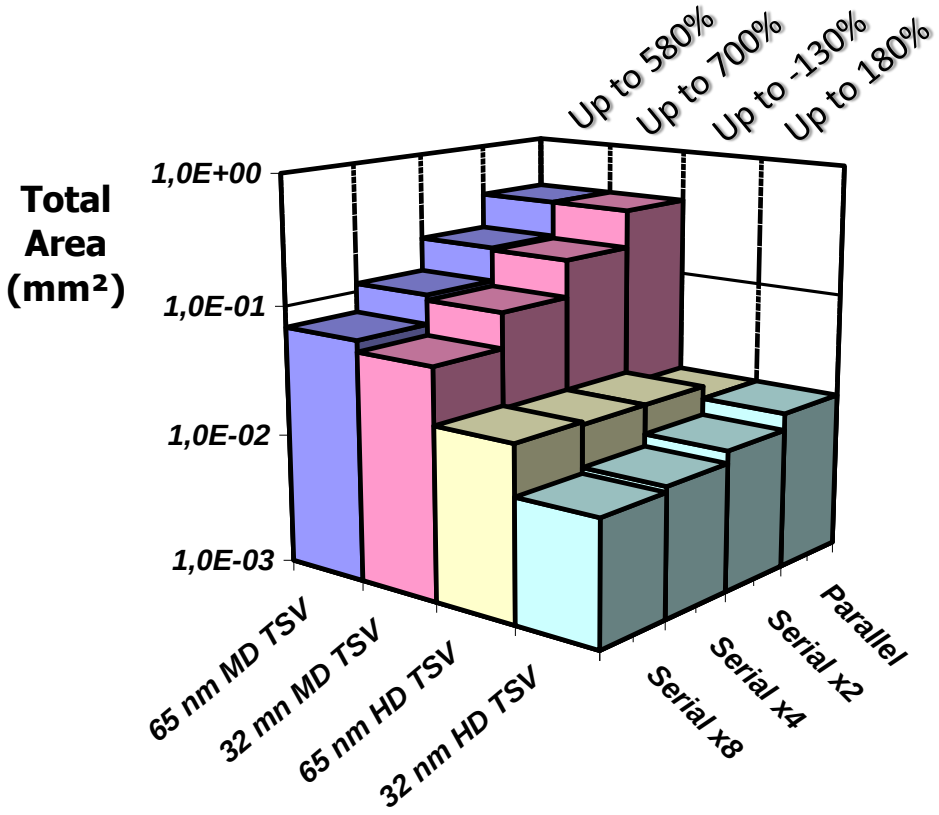
Signal Transitions





Serialization Area Cost Analysis

	MD TSV	HD TSV	65 nm	32 nm
Parallel	0.4 mm²	0.016 mm²	0 mm²	0 mm²
Serial x2	0.2 mm²	0.008 mm²	0.012 mm²	0.0039 mm²
Serial x4	0.1 mm²	0.004 mm²	0.016 mm²	0.0056 mm²
Serial x8	0.05 mm²	0.002 mm²	0.019 mm²	0.0067 mm²

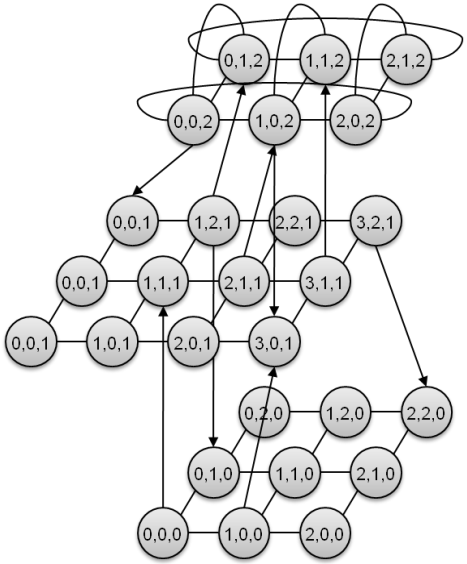


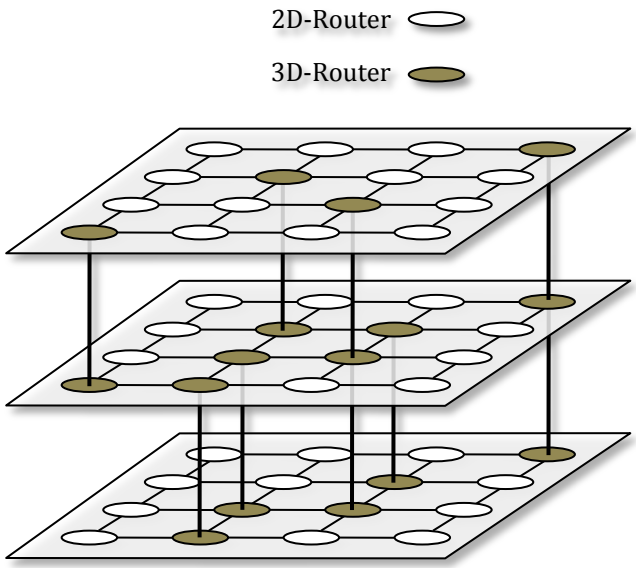
Outline

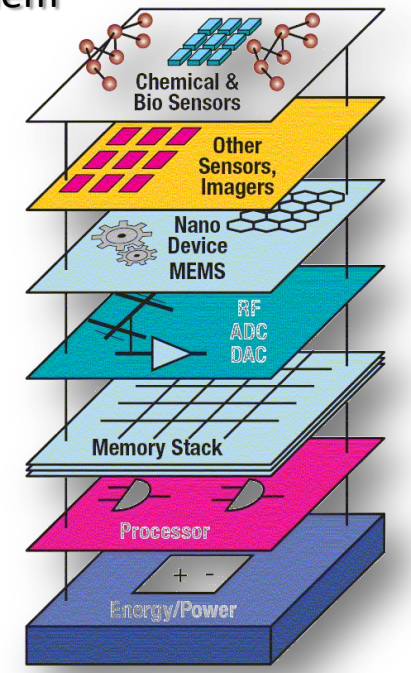
- 3D-Integration and Asynchronous NoCs
- Vertical Link Serialization
- Vertically-Partially-Connected 3D-NoC
- Conclusion

Vertically-Partially-Connected 3D-NoC !

- Limited number of vertical connections (TSVs)
- Network with different dies fabricated with different technoilogies
  - Heterogeneity
  - Irregularity
- Vertically-Partially-Connected Topology as an efficient solution
  - Routing strategy in such an irregular topology as the major problem

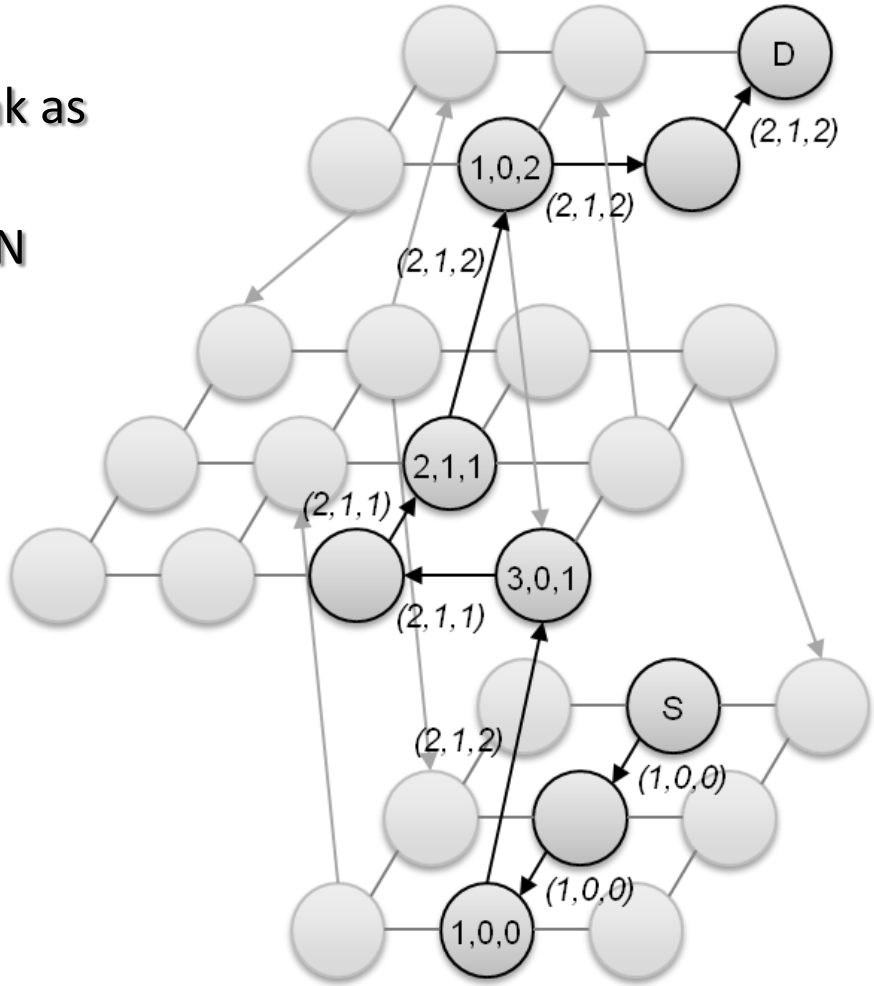






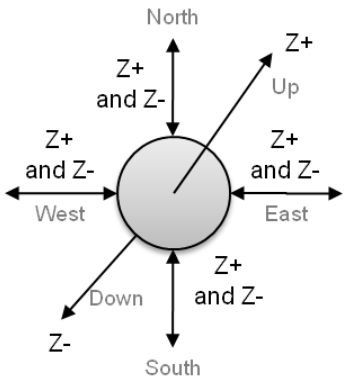
Elevator-First Routing Algorithm

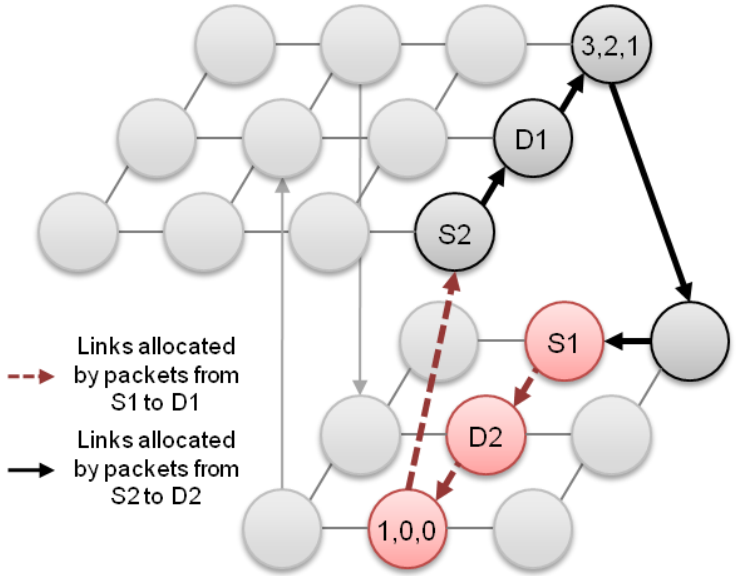
- Each router registers
  - A router in its layer with UP link as ascending elevator
  - A router in its layer with DOWN link as descending elevator



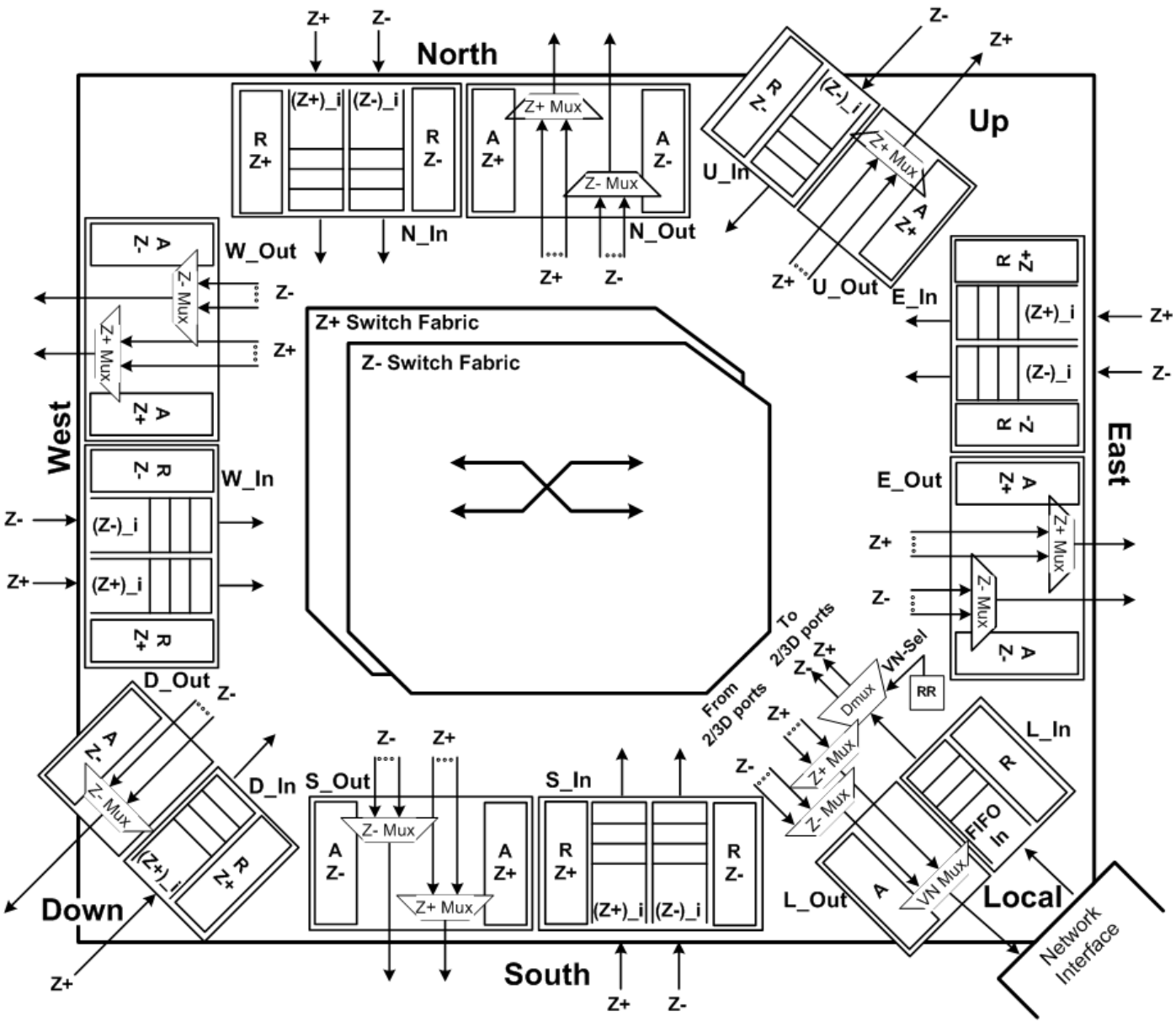
… and … Deadlock !

- Two Virtual Networks to avoid Deadlock
  - One for ascending packets (Z+)
  - One for descending packets (Z-)





Elevator-First Router



The Algorithm

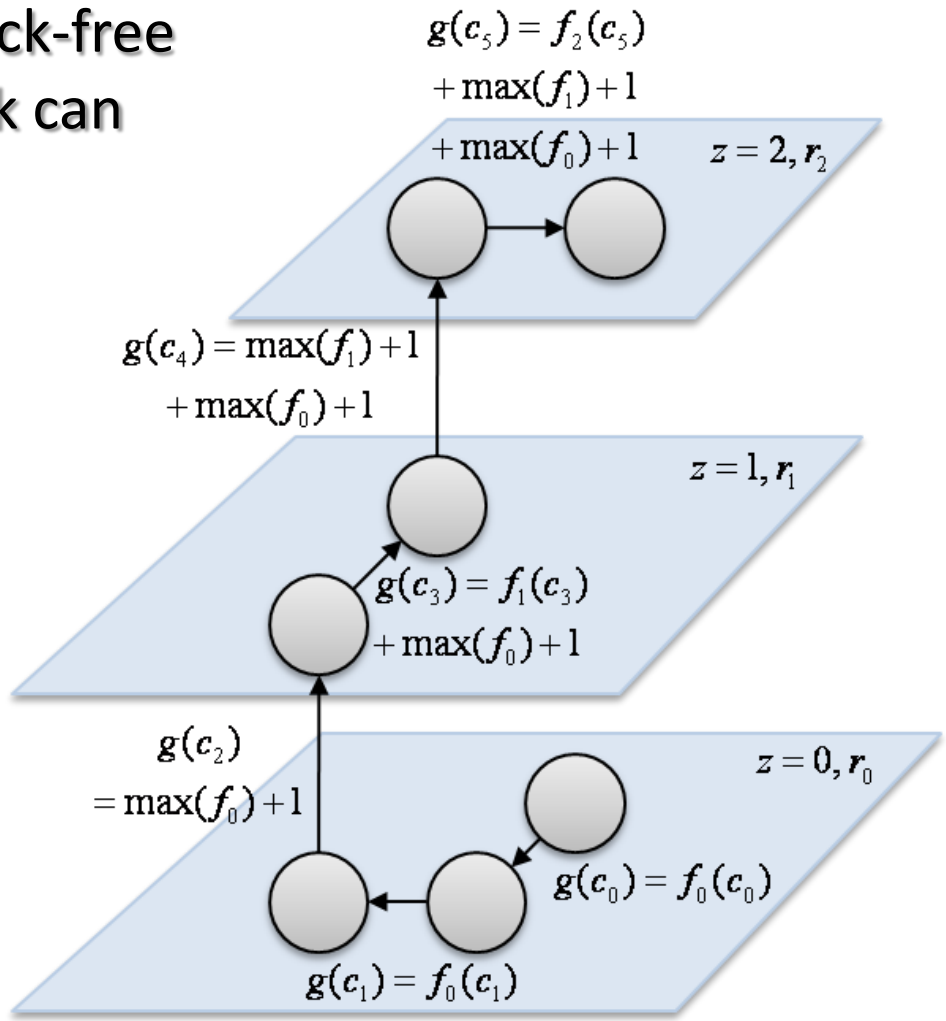
Algorithm 1 - Elevator-First Routing using two Virtual Channels

```
@c : current router address
@s : source router address
@d : destination router address
if ( @s == @c) then
  // The current router is the source
  if ( the destination is on a lower tier ) then
    - Assign the packet to the virtual network (channel) Z-
  else if ( the destination is on an upper tier ) then
    - Assign the packet to the virtual network (channel) Z+
  else
    // The destination is on the current tier
    - Randomly assign the packet to either the virtual network
      (channel) Z- or Z+
  end if
end if
```

```
if ( @d == @c) then
  if ( the elevator flag is set ) then
    // The current router is an elevator node
    - Remove the packet header
    - Get the original header (the next flit)
    - Send the packet to the ascending (if the assigned virtual channel
      is Z+) or descending (if the assigned virtual channel is Z-) vertical
      link
  else
    // The current router is the final destination
    - Consume the packet
  end if
else
  if ( The packet destination is in the current tier ) then
    - Send the packet to the port determined by the given planar
      routing algorithm (e.g. X-First in 2D-meshes)
  else
    // The packet destination is not in the current tier
    - Add a new header with the elevator flag set and an address of a
      vertical link (i.e. elevator) on the current tier (given from local
      registers) as an intermediate destination
    - Send the packet to the port determined by the given planar
      routing algorithm (e.g. X-First in 2D-meshes) toward the
      intermediate destination (i.e. the elevator)
  end if
end if
```

Formal Proof of Deadlock-Freedom

- A routing algorithm is deadlock-free if the channels in the network can be numbered such as every routing path uses strictly increasing (decreasing) channels



… an example …

b = max (n, m)

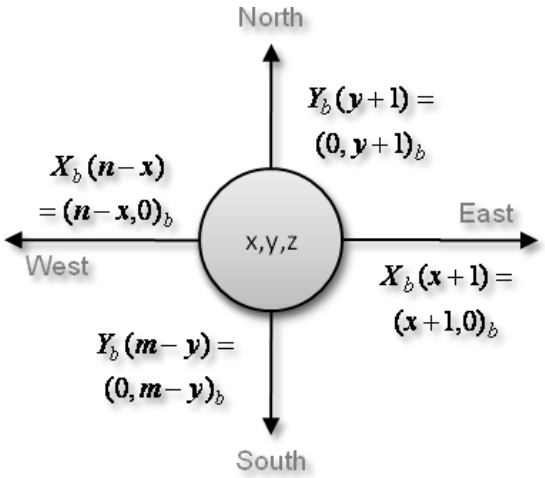
$$X_b : x \rightarrow (x,0)_b = x$$

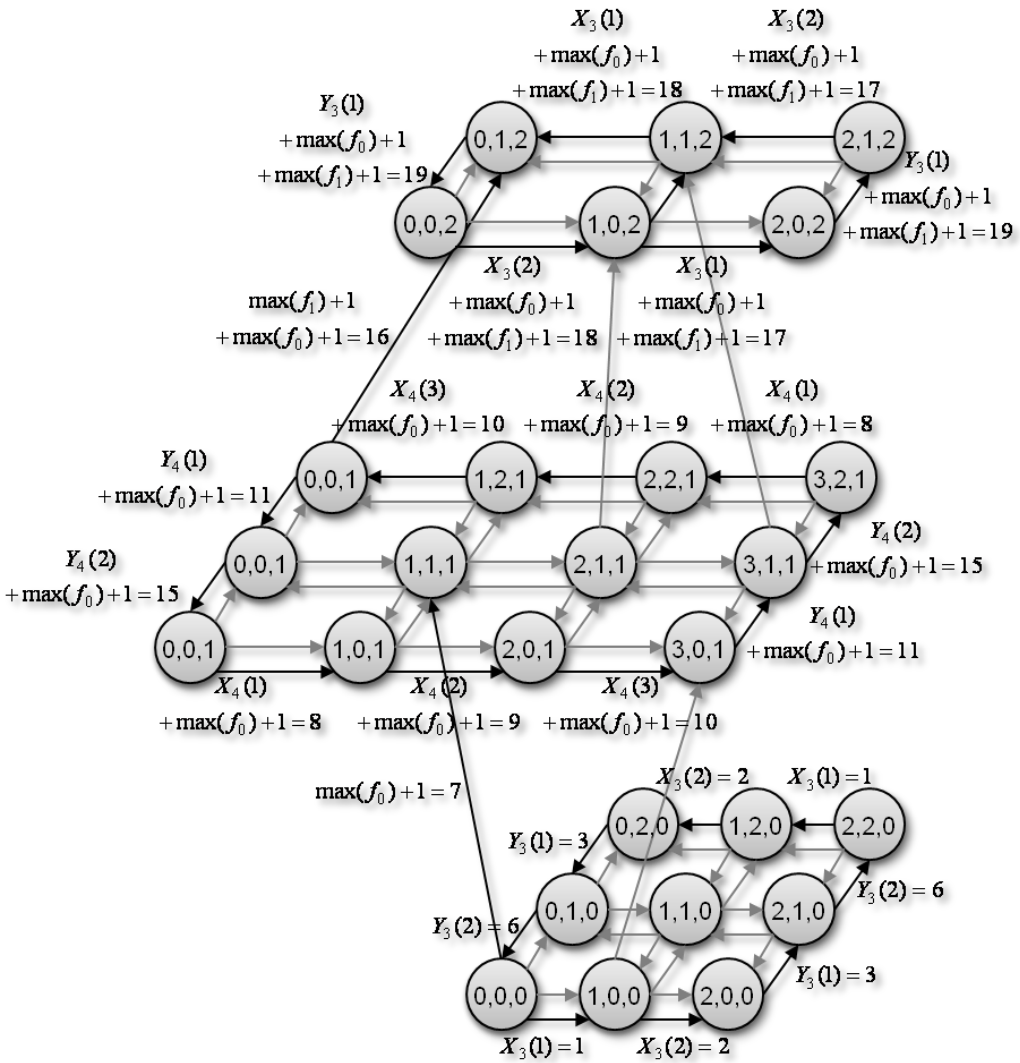
$$Y_b : y \rightarrow (0,y)_b = yb$$

$$f_z(c) \rightarrow \begin{cases} X_b(n-x) & if\ c = West \\ Y_b(y+1) & if\ c = North \\ X_b(x+1) & if\ c = East \\ Y_b(m-y) & if\ c = South \end{cases}$$

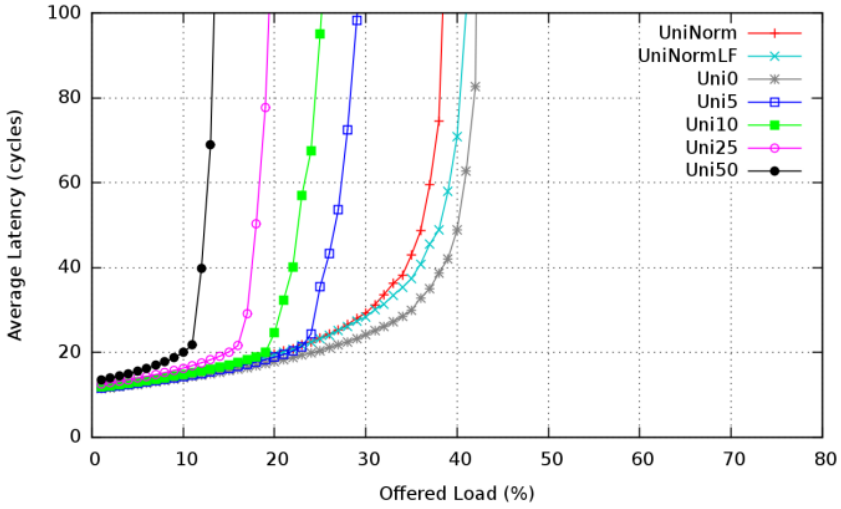
$$\forall (x,y,z),$$

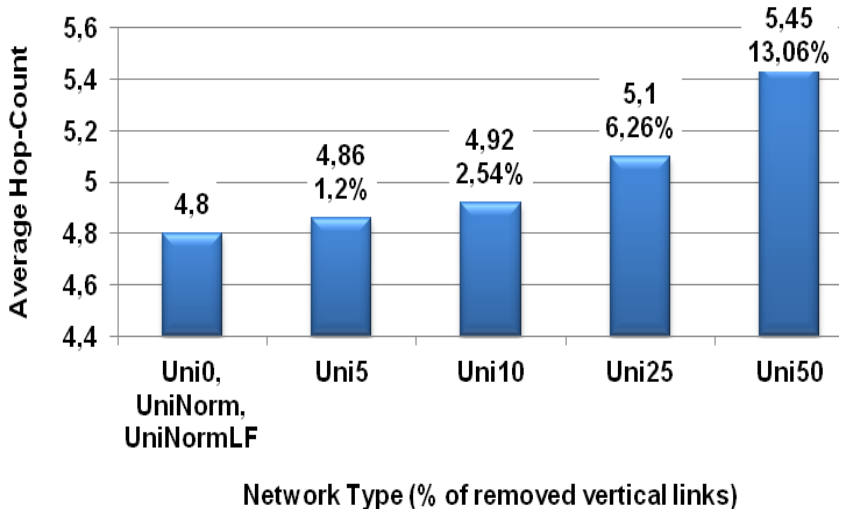
$$X_b(x) < X_b(x+1) < Y_b(y) < Y_b(y+1)$$

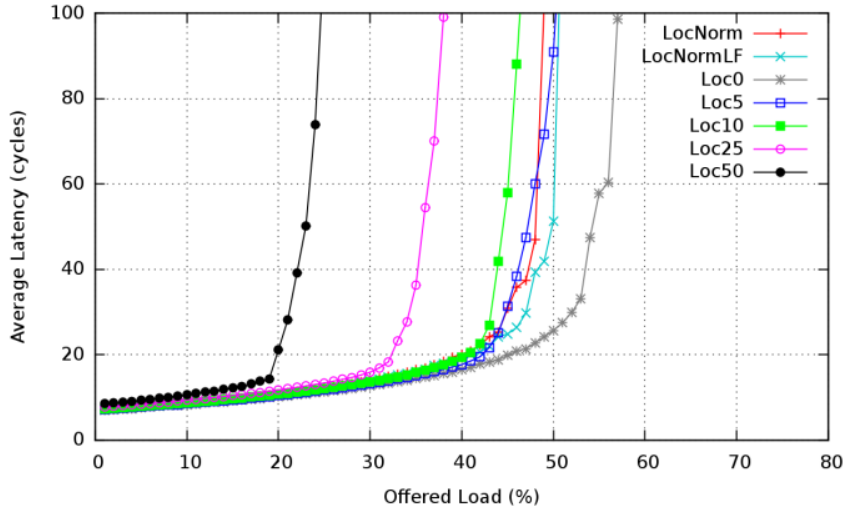


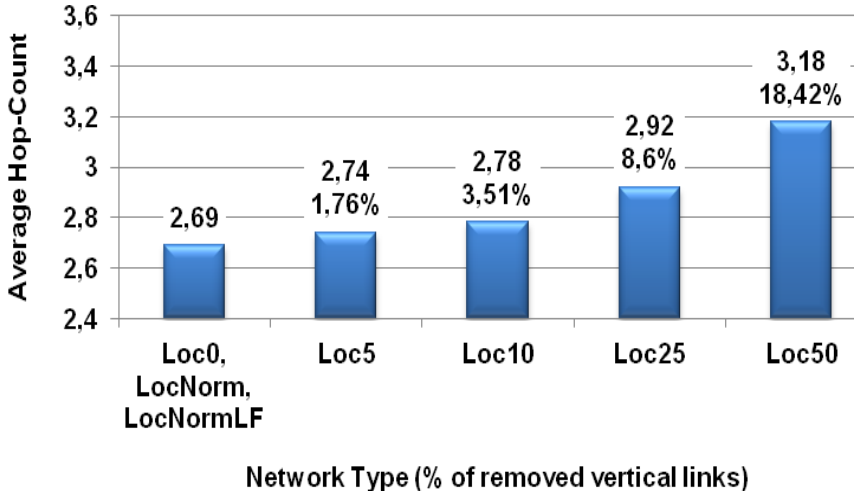


Performance …









Conclusion

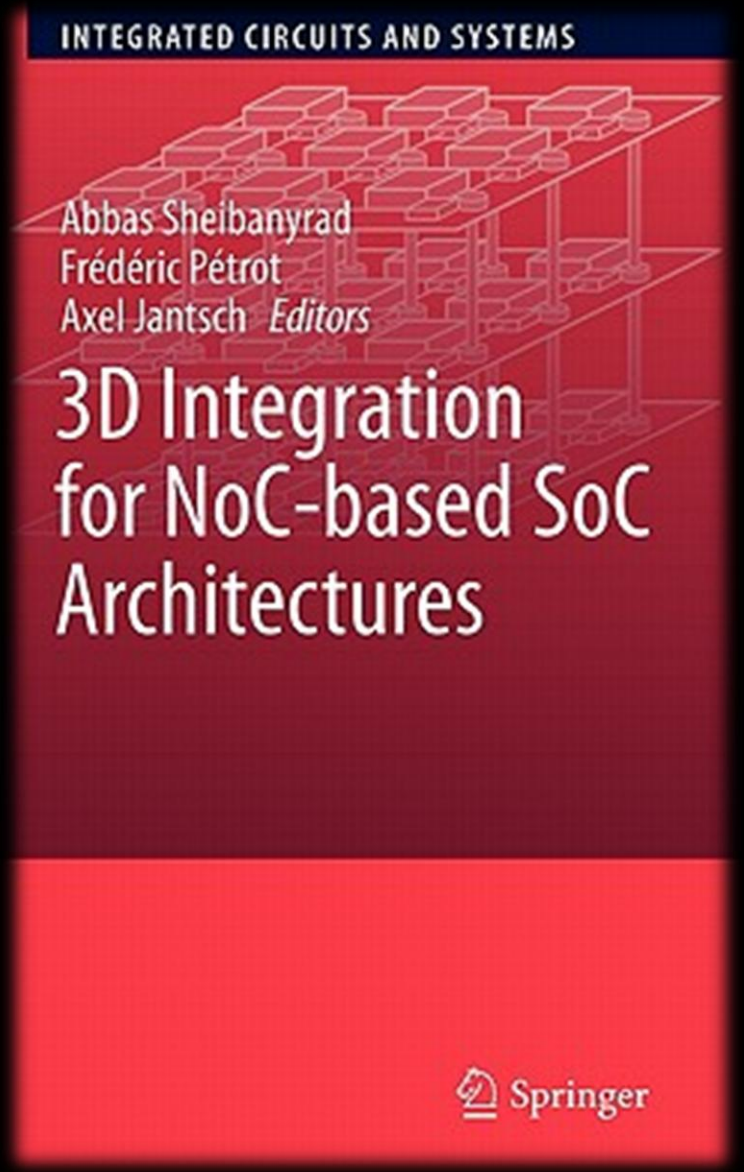
- The new technology of 3D-Integration opens a new windows to more and more integration of components
- Contribution of the third dimension in network architecture helps to improve the system performance
- TSVs are the most promising technology of vertical connection with a high bandwidth and a low power consumption
- Due to the cost and yield reduction, 3D-Integrated Circuits are limited on the number of TSVs to be exploited
- The GALS/DVFS paradigm is demanded as clock distribution in three dimensions is almost impossible
- Asynchronous Networks-on-Chip help to exploit the whole high bandwidth of vertical links (TSVs)
- Serialization of data of vertical links (TSVs) is a trade-off between cost and performance
- a Vertically-Partially-Connected Topology is an efficient solution for reducing the number of TSVs and adapting to the heterogeneity and irregularity of 3D-Integrated Systems

Conclusion

"Vertically-Partially-Connected Asynchronous 3D-NoC making use of Serialized Vertical Links is a viable technology and undeniably will be used as the communication infrastructure of the future Many-Core Systems"

Hamed S.

Merci…


INTEGRATED CIRCUITS AND SYSTEMS
Abbas Sheibanyrad
Frédéric Pétrot
Axel Jantsch Editors
3D Integration for NoC-based SoC Architectures
Springer