Grenoble, February 13th 2015
OPTICS Workshop

Towards a Vertically Integrated Synthesis Flow for Predictable Design of Wavelength-Routed Optical NoCs

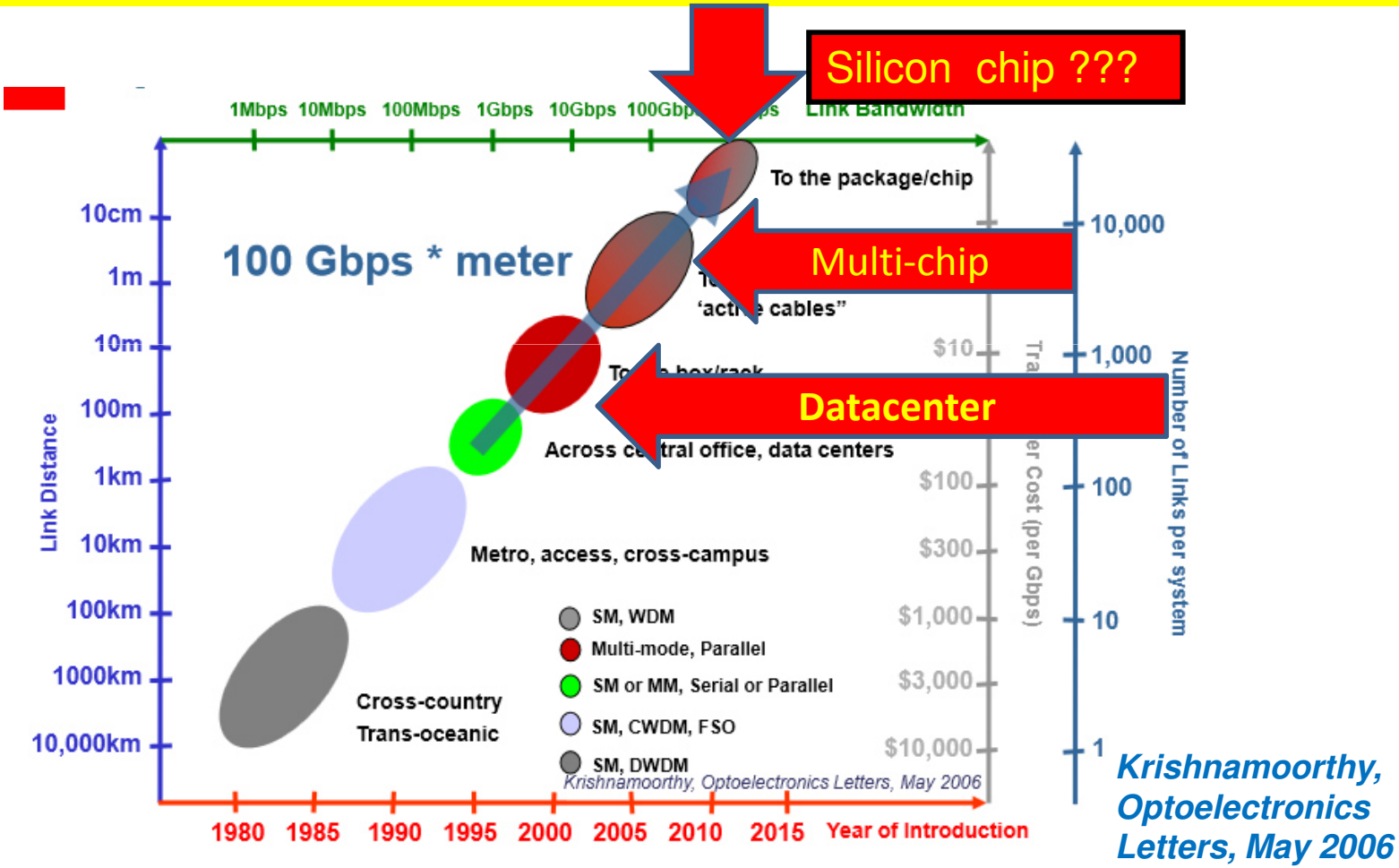
Davide Bertozzi
MPSoC Research Group

UNIVERSITA' DEGLI
STUDI DI FERRARA

Acknowledgement:
Luca Ramini, Marta Ortin, Anja Boos, Marco Balboni, Sandro Bartolini, Paolo Grani

Optics into Smaller-Scale Systems

Optics has made a long way from long-haul telecommunication networks to data centers and multi-chip systems.



Krishnamoorthy, Optoelectronics Letters, May 2006

- **The early days of ONoCs remind me of the early days of electrical interconnection networks!**

Behind the scene....

- Commercial exploitation of NoCs started from ***contradicting numbers***

Frequency (MHz)	Netlist	Floorplan	Post-P&R drop
AMBA Multilayer	480	400	16.7%
NoC/21 bits	910	885	2.7%
NoC/38 bits	910	885	2.7%

Higher clock speed, higher predictability

Area (mm^2)	Overall Floorplan	Fabric + slack
AMBA Multilayer	35	5
NoC/21 bits	45	15
NoC/38 bits	45	15

Higher area though!

Bandwidth (GB/s)	Overall bandwidth
AMBA Multilayer	26.5
NoC/21 bits	100
NoC/38 bits	180

Higher bandwidth, sometimes better latency!

Power (mW)	Sequential	Combinational	Overall	Seq. ratio
AMBA Multilayer	6	66	72	18.5%
NoC/21 bits	296	81	377	78.5%
NoC/38 bits	416	85	501	83.0%

Higher power consumption though!

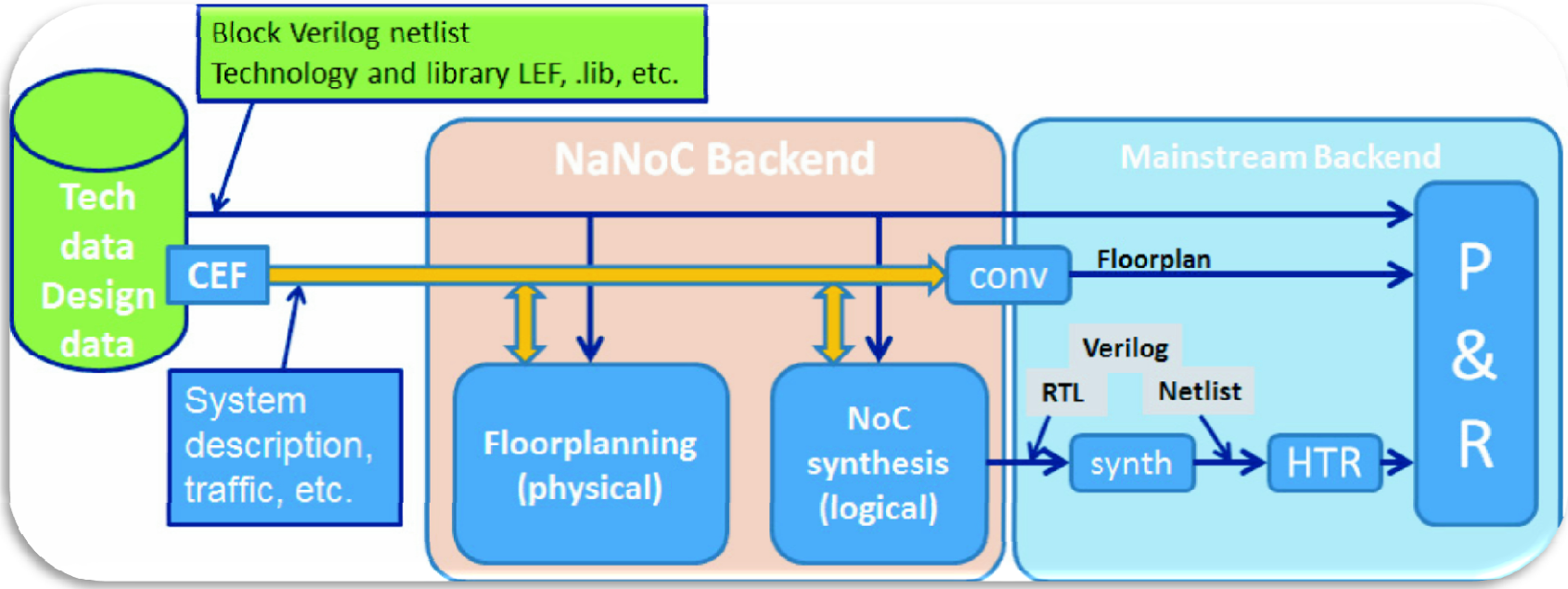
Energy (mJ)	Benchmark run time	Fabric only	1W system	5W system
AMBA Multilayer	1 ms	0.072	1.07	5.60
NoC/21 bits	0.9 ms	0.339	1.34	5.32
NoC/38 bits	0.85 ms	0.426	1.37	5.13

Behind the scene....

- Commercial exploitation of NoCs started from ***contradicting numbers***

The Boosting Factor

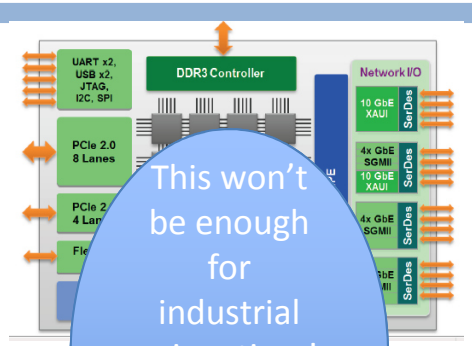
- Initially, the NoC IP portfolio was the “business card” of NoC vendors
- Very soon it became clear that the real business card was the availability of toolflows to bring designers’ productivity to a new level



NoC vendors then started to deliver what designers actually needed:

- *Fast and automated design space exploration*
- *Floorplanning constraints in the early design stages for faster and quicker convergence*
- *not just IP models, but also technology models*
- *NoC customization was the main goal*
- *IP portfolio: of course you should have it!!*

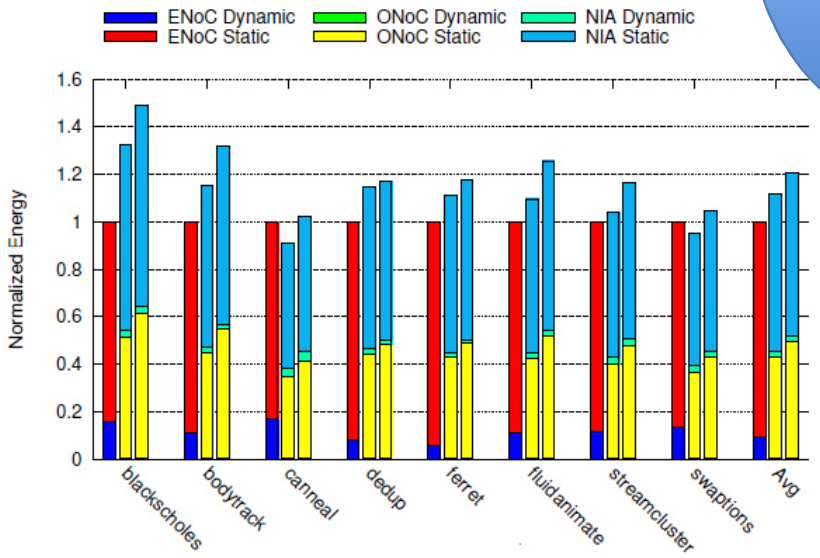
The ONOC Business Card



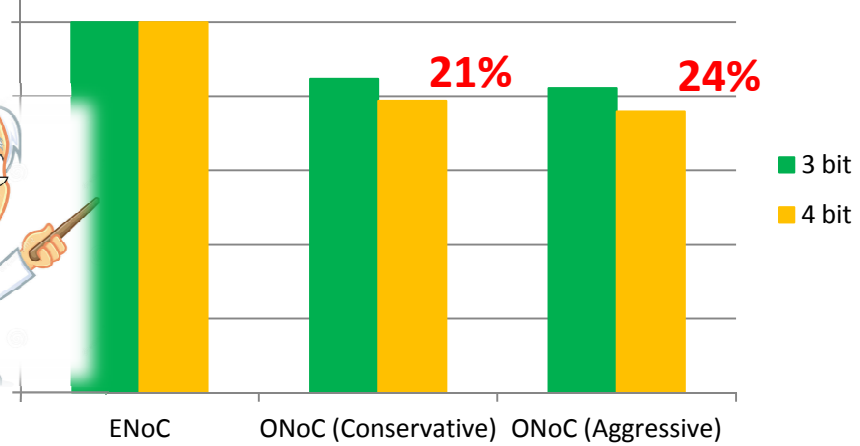
- The Electronic Tile-based CMP Architecture burns on average **15 Watts on target applicative loads**

This won't be enough for industrial prime time! What about the boosting factor?

Interconnect Energy (aggressive optical technology)



Normalized System Energy



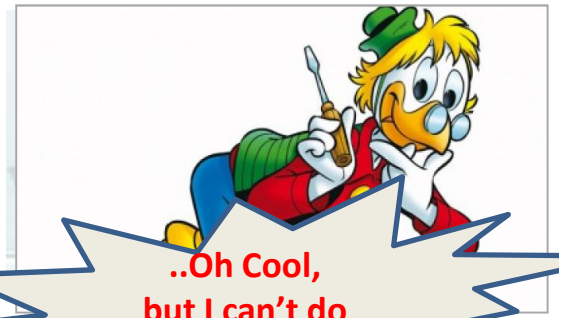
ONoC makes the system more energy efficient, although the interconnect itself does not achieve the energy break-even point with the electrical NoC

The trick: on average ONoC outperforms ENoC by about 18% @ 3bit and 23% @ 4bit.

Pathfinding

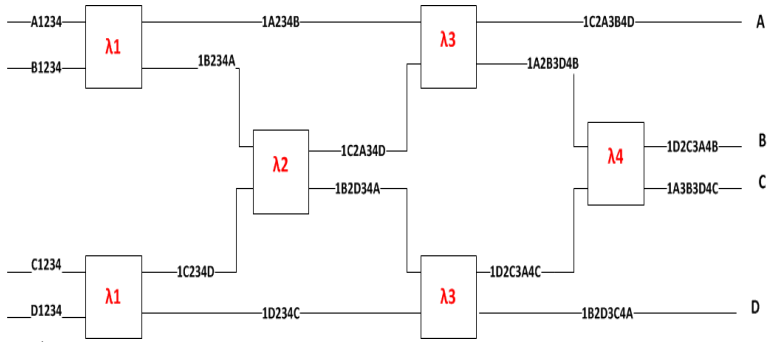
There is currently a huge gap between Technology Developers & System Level Designers

I have a great device; it works!

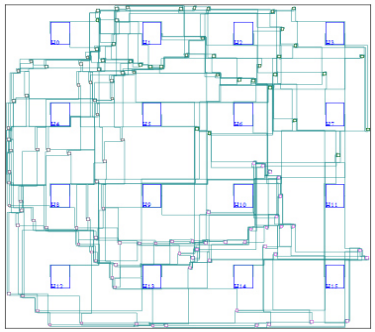


..Oh Cool, but I can't do design with it!

Are we ready to bridge this gap?



Physically speaking:
When we change P&R constraints, its layout becomes crossing-dominated

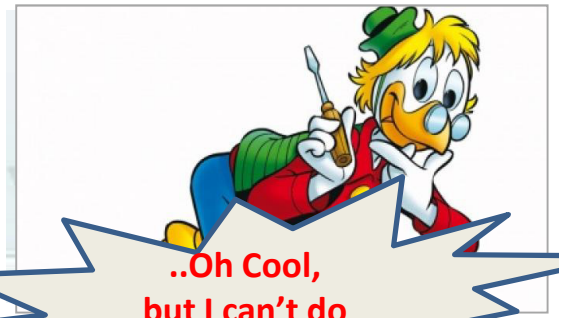


- Descriptive information at different abstraction layers **are mixed and hardwired in the same design description**.

Pathfinding

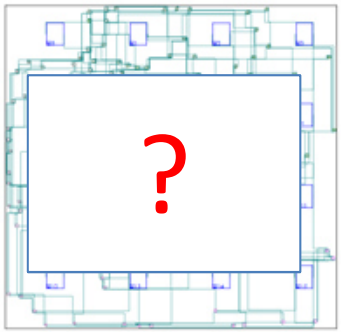
There is currently a huge gap between Technology Developers & System Level Designers

I have a great device; it works!

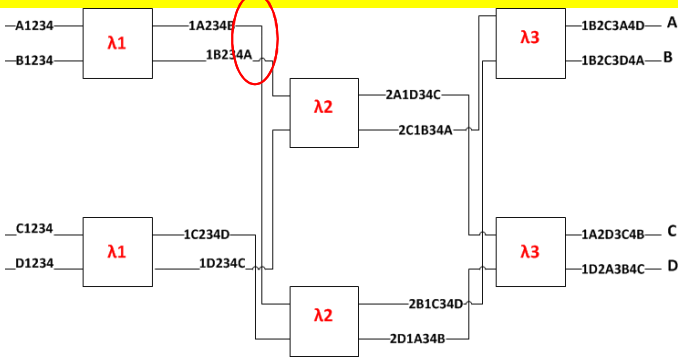


..Oh Cool, but I can't do design with it!

Are we ready to bridge this gap?



Logically speaking:
This topology does the same
Question:
Its inter-switch crossings should be viewed from a logical (irrelevant) or physical viewpoint?



- Descriptive information at different abstraction layers **are mixed and hardwired in the same design description**.

Pathfinding

There is currently a huge gap between Technology Developers & System Level Designers

I have a great device; it works!

..Oh Cool, but I can't do design with it!

Are we ready to bridge this gap?

	Crossbar	Reduced Crossbar	Hitless Router	WRON	λ-router	GWOR
MRRs	16	12	8	12	12	8

Claim: "GWOR uses a lower number of MRRs than lambda-router (e.g., in a 4x4 ONoC)
Yes, since GWOR does not support self-communication!

- **Designs are difficult to compare with one another**

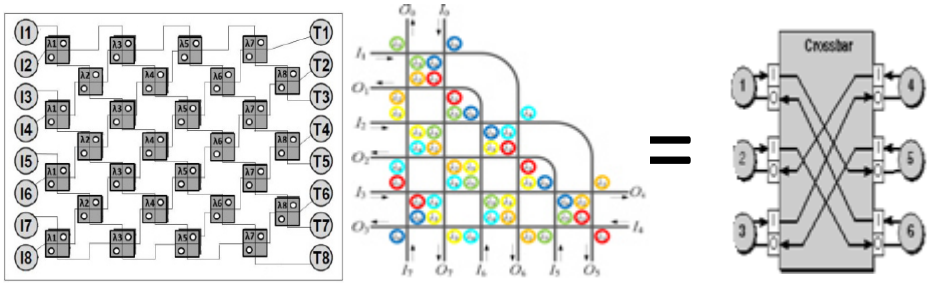
Pathfinding

There is currently a huge gap between Technology Developers & System Level Designers

I have a great device; it works!

..Oh Cool, but I can't do design with it!

Are we ready to bridge this gap?

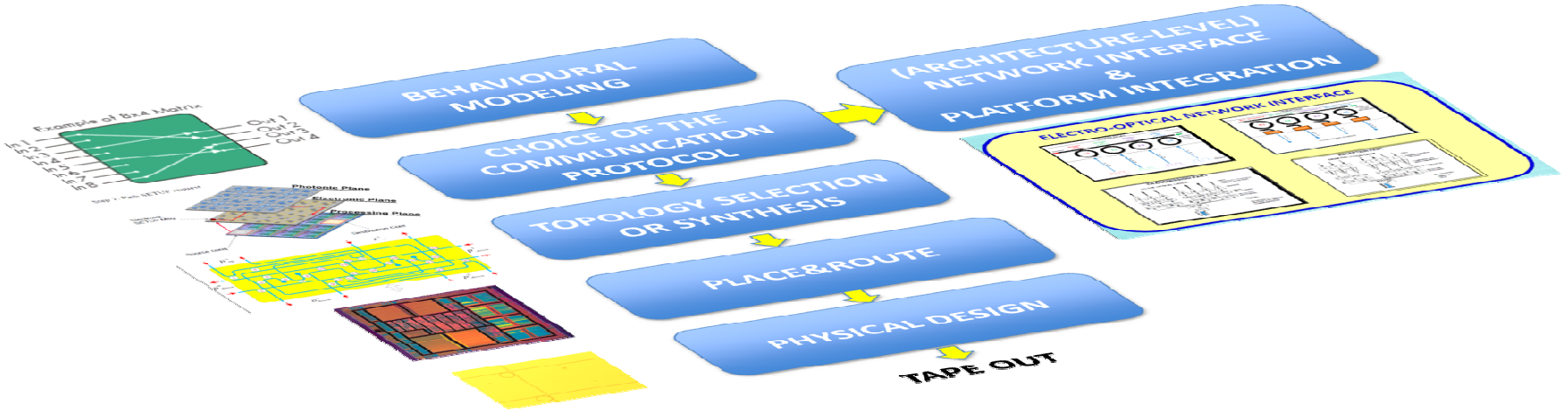


- *Flow control should be end-to-end and topology-independent*
- *Receiver buffer size should be topology-specific, since it should cover the round-trip latency for max. throughput operation*

◆ The **application of well-known** interconnection network techniques **is more difficult**.

Design Methodology and Synthesis Flow

Only a cross-layer design methodology and automated synthesis toolflow can accelerate or even determine the evolution of the ONoC concept into an industry-relevant and viable interconnect technology



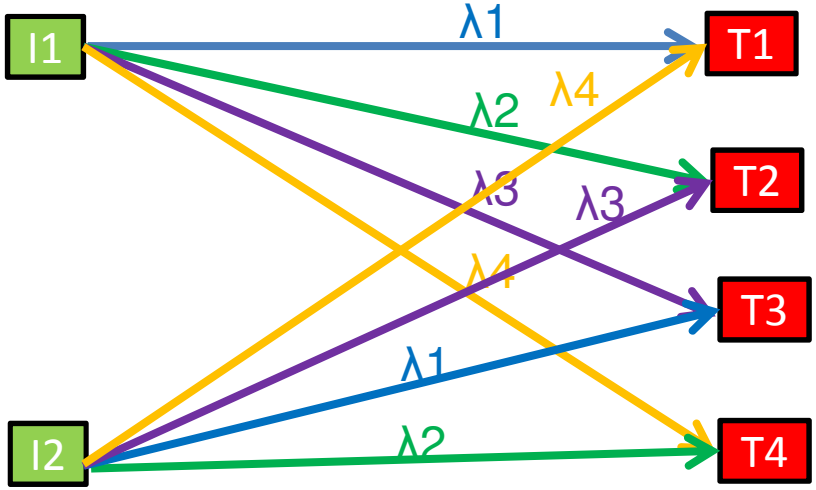
-***Specification of abstraction layers for ONoC design.***

Learning from the past:

- *Fast and automated design space exploration*
- *Floorplanning constraints in the early design stages for faster and quicker convergence*
- *not just IP models, but also technology models*
- *NoC customization should be the main goal (high-end embedded computing)*
- *IP portfolio: of course you should have it!!*

Early signs of a top-down design methodology

- Let us keep the focus on Wavelength-Routed Optical Networks-on-Chip
- Contention-free and performance-guaranteed communication

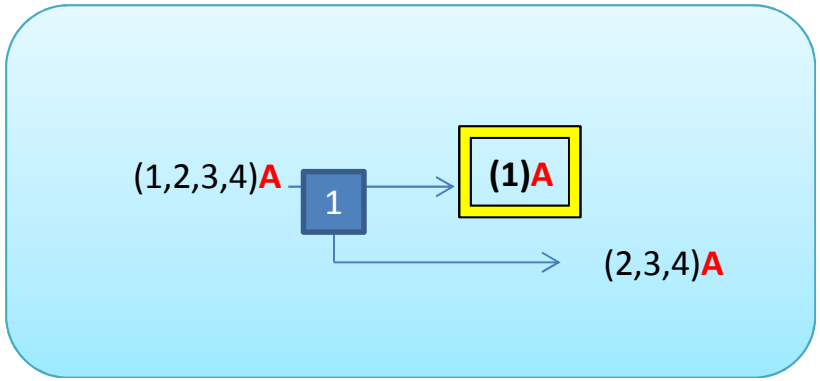


How to synthesize the topology?

Key abstract operator: add-drop filter

(1,2,3,4)A I1

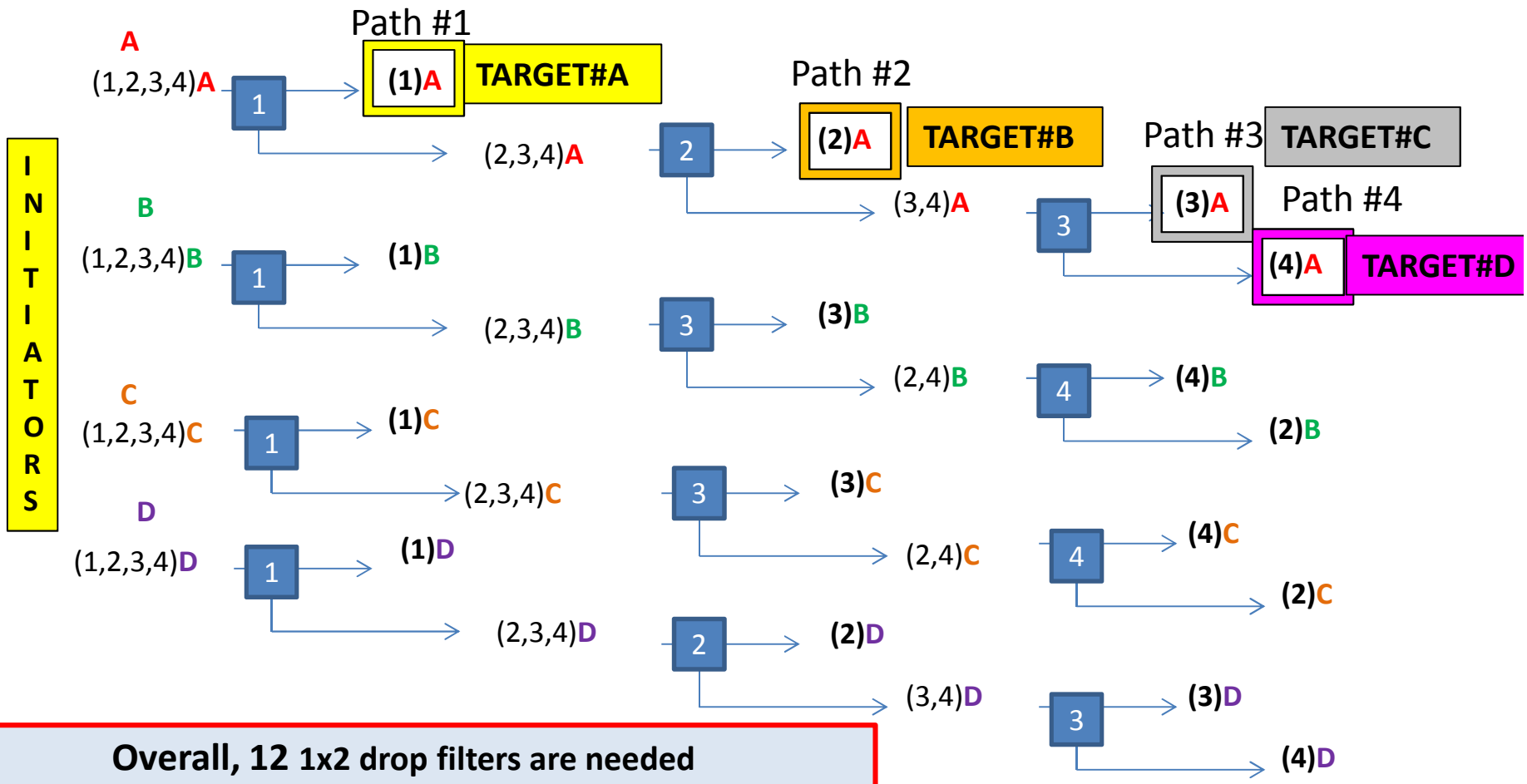
Separation of chromatic signals: each wavelength component should be separated to reach a different destination



Wavelength Separation Graph

ASSUMPTIONS

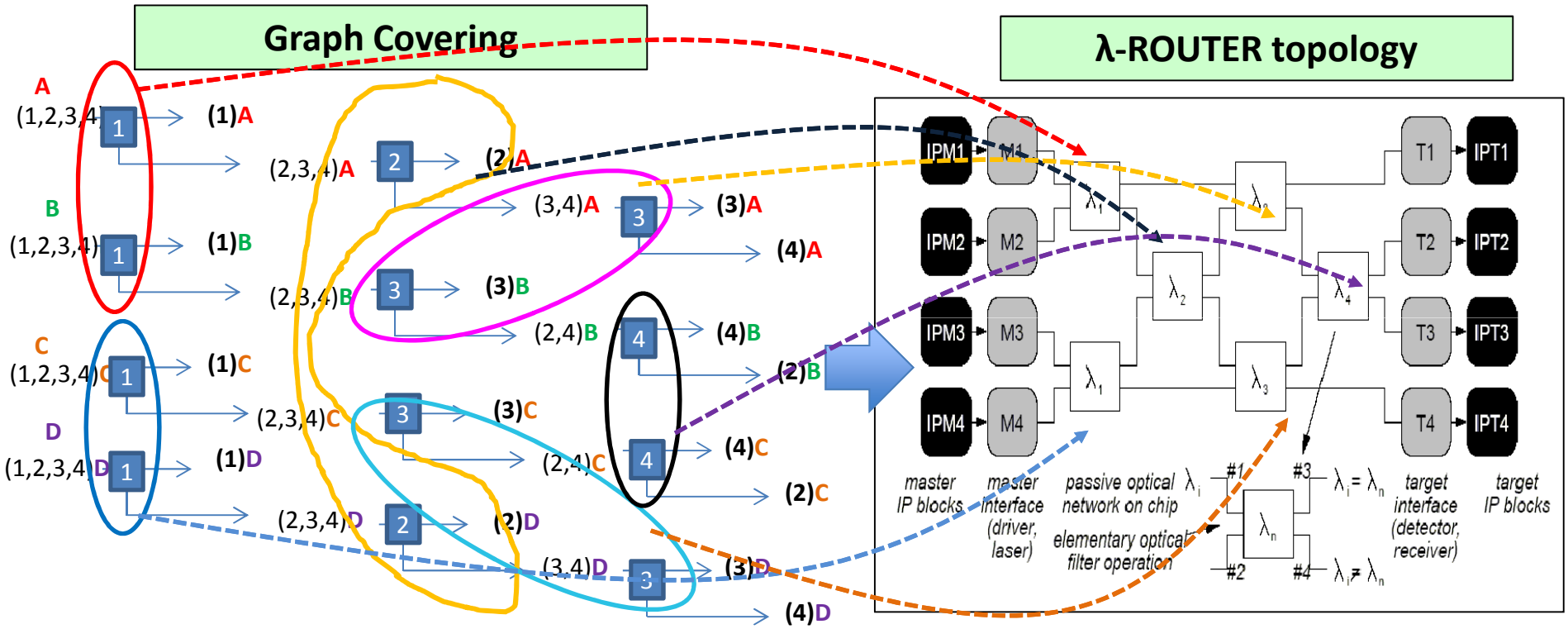
- 4 INITIATORS AND 4 TARGETS (i.e., A,B,C D).
- INITIATORS USE THE SAME 4 WAVELENGTHS (i.e., 1,2,3,4).
- UTILIZATION OF 1X2 LOGICAL FILTERING OPERATORS



Overall, 12 1x2 drop filters are needed to realize 16 contention-free optical paths!!!

Covering the Separation Graph

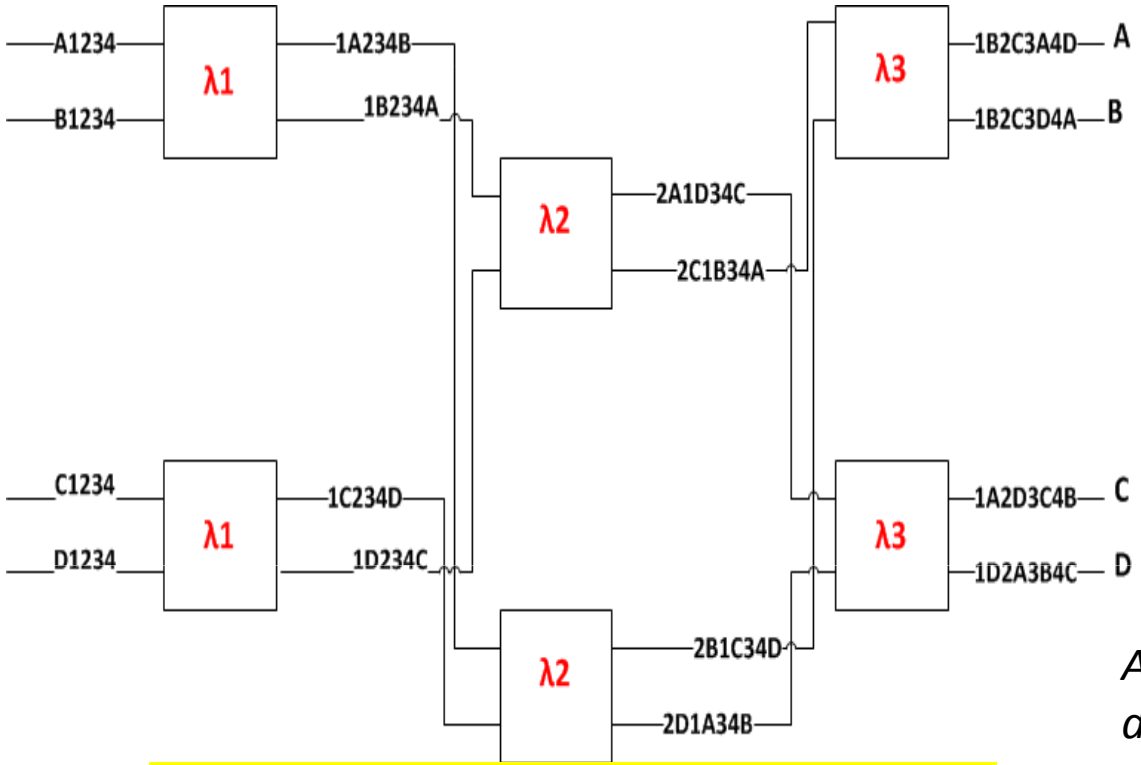
***Let us "cover"* the wavelength graph with higher-order logic filters (e.g., 2x2) in order to obtain logic topologies**

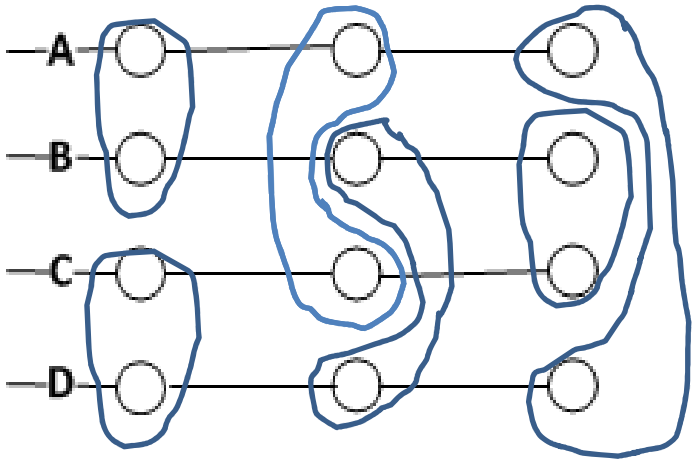


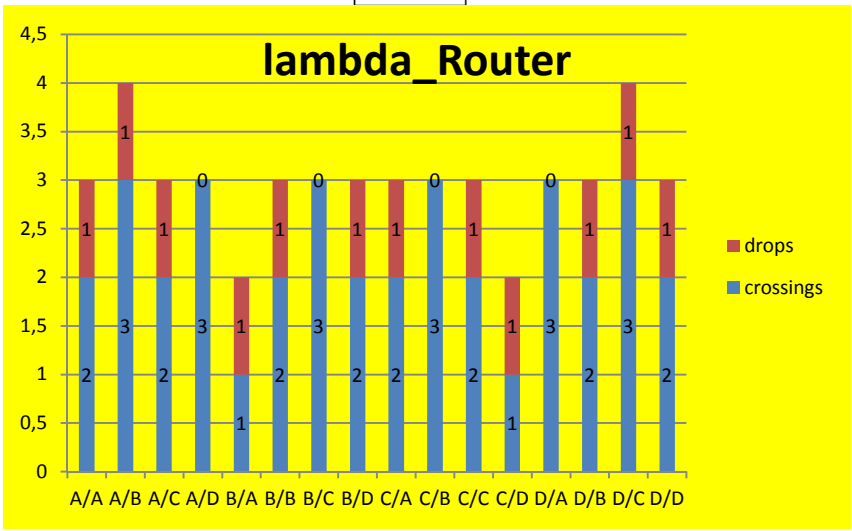
There are precise covering rules for the functional correctness of the topology
(e.g., never recombine split signals; never mix wavelength-homogeneous signals)

All known WRONoC topologies can be materialized this way!
What about unknown topologies?

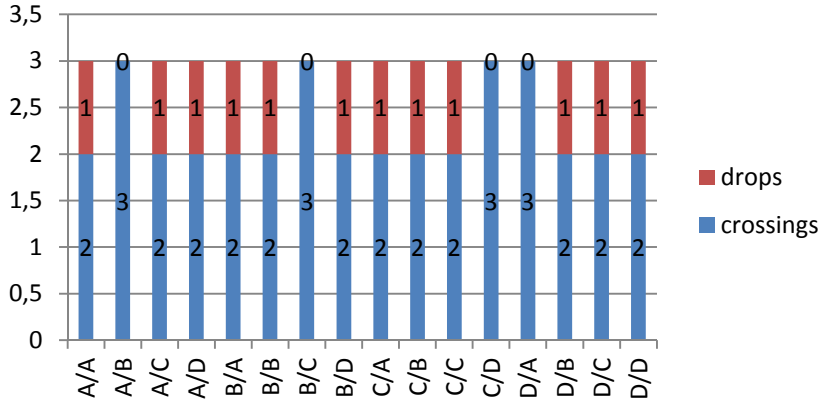
New Topology: Equalized Lambda-Router



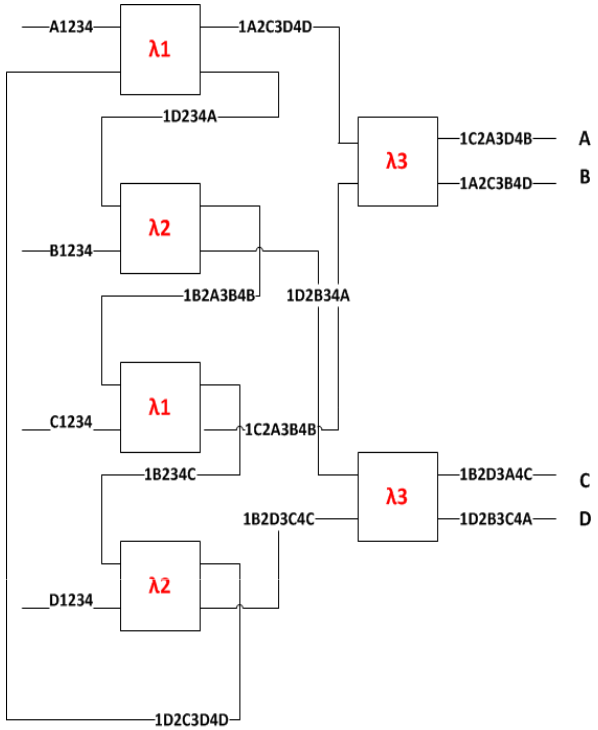


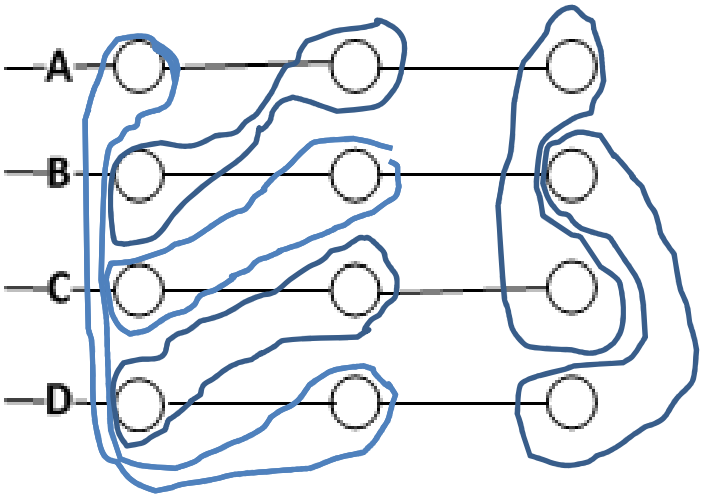


Assumption:
do not count inter-switch crossings
- No drops on the critical path,
- more balanced optical paths

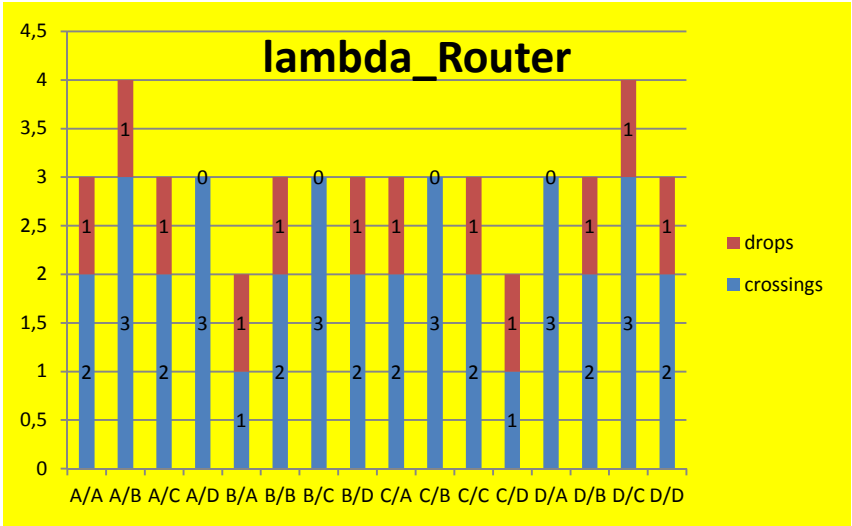


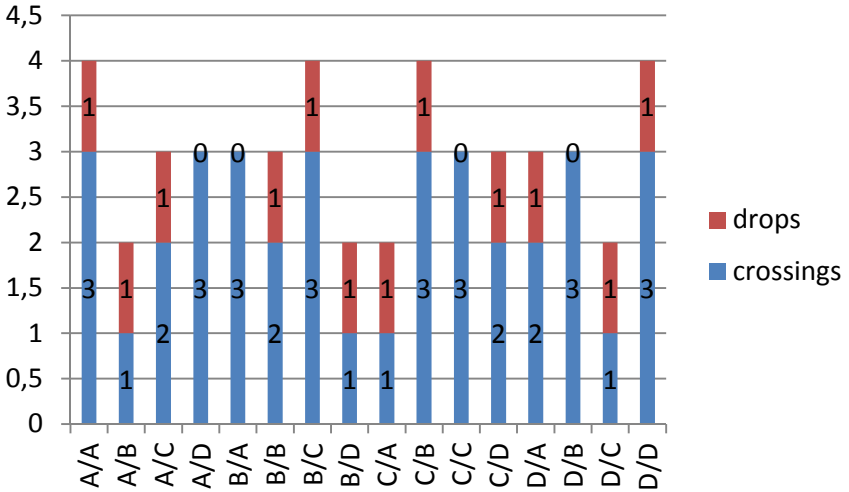
New Topology: GWOR With Self-Communication



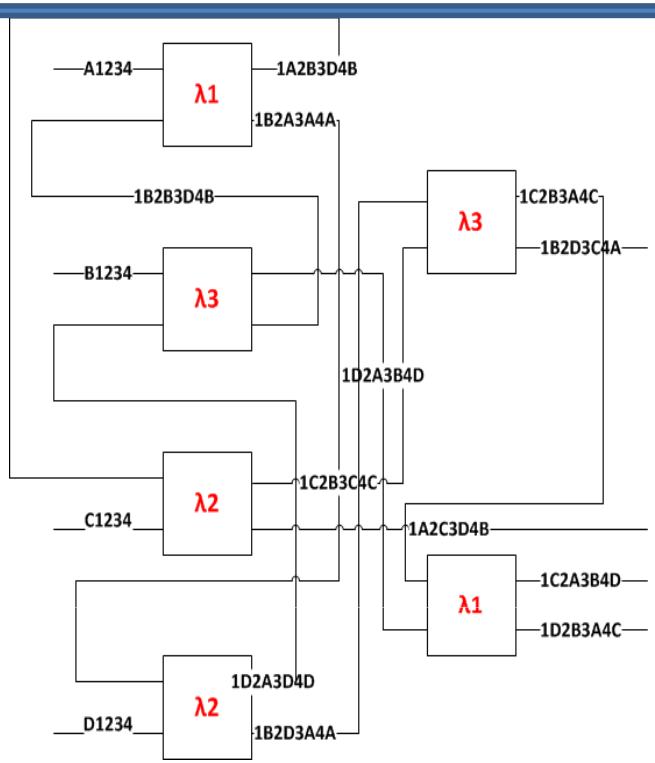


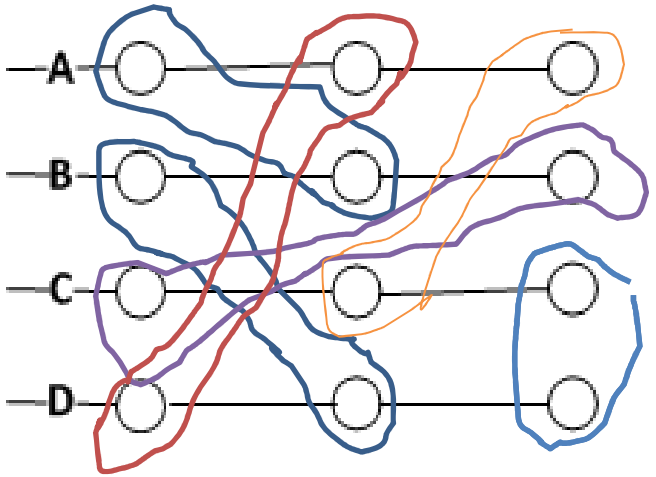
- This covering increased the number of overly short and overly long optical paths

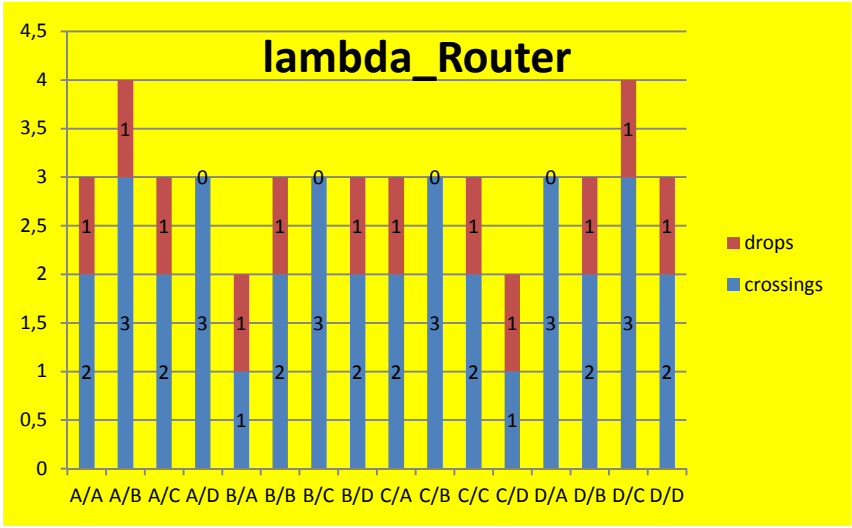




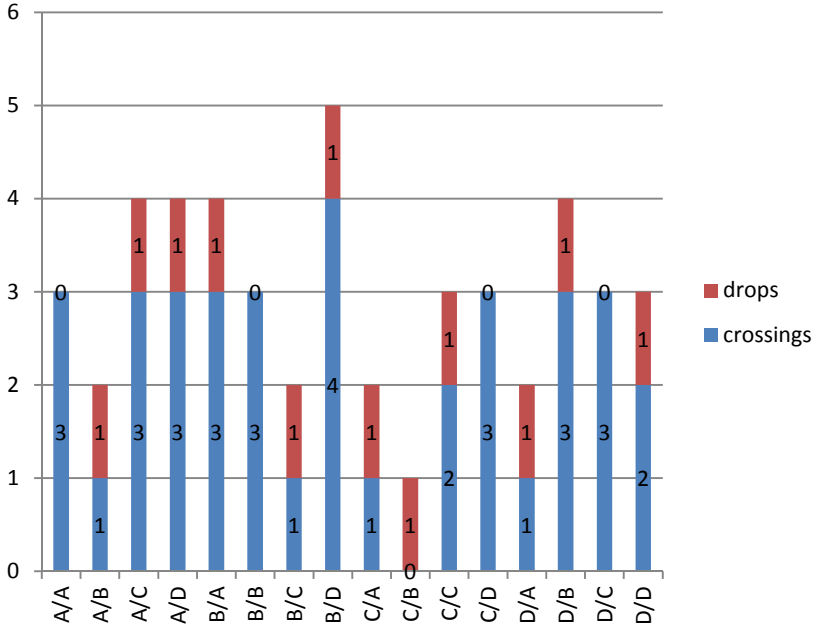
New Topology: Random







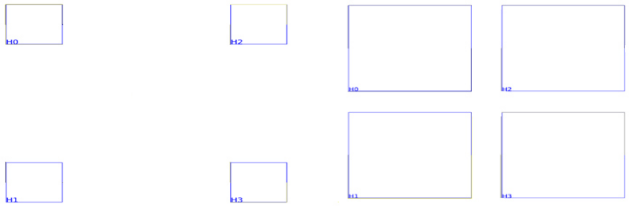
- Extreme path differentitation



Wavelength Separation Graph

❑ **First case assumptions:**

- Floorplan area: 8mmx8mm
- Hub size: 1mmx1mm



❑ **Second case assumptions:**

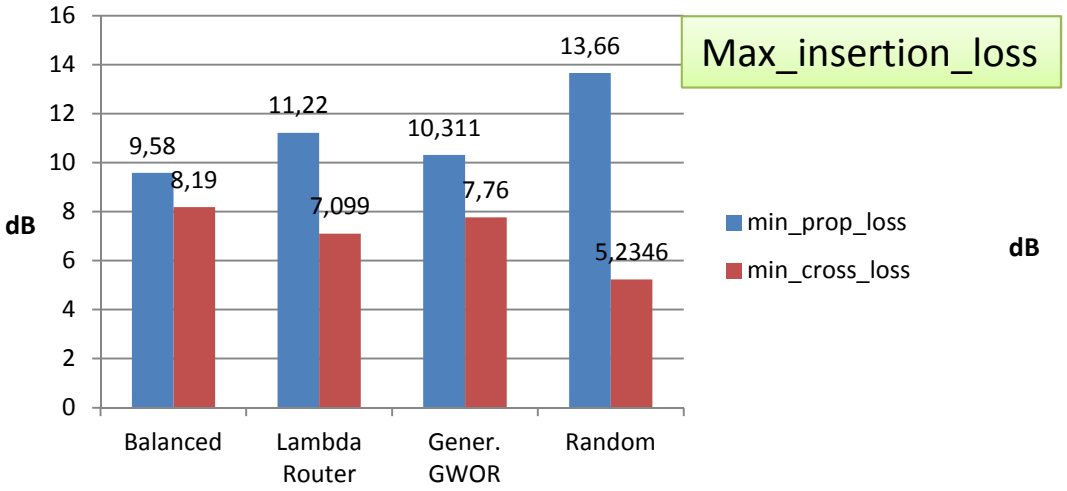
- Floorplan area: 2.95mmx2.95mm
- Hub size: 1mmx1mm

❖ **The Proton P&R tool for ONoCs is used to obtain the Physical layout and the maximum insertion loss.**

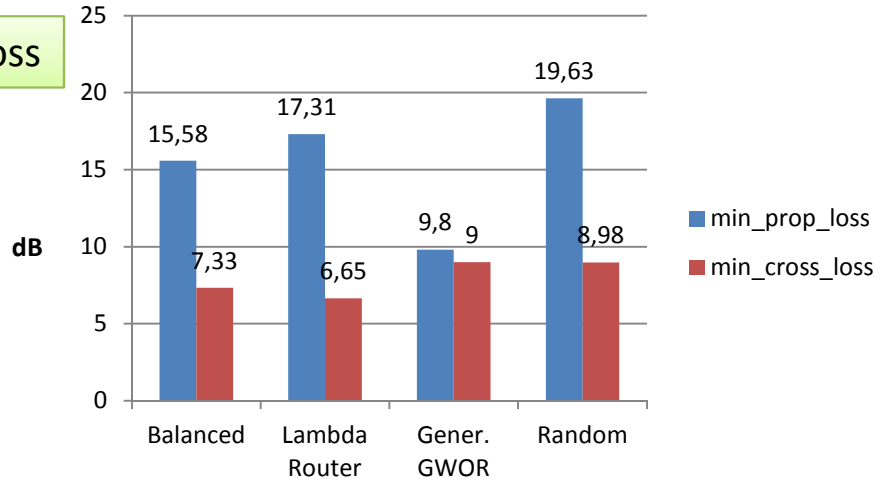
❖ **Proton can be instructed to pursue different primary design goals (or a mix thereof):**

✓ Minimize_propagation_loss.

✓ Minimize_crossing_loss.



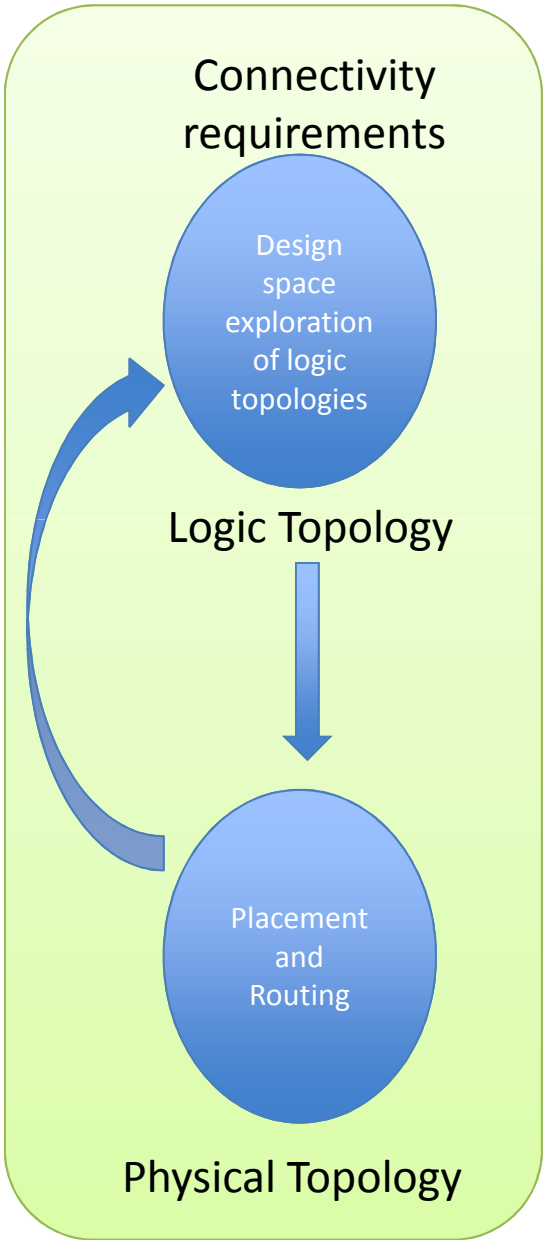
The random topology is the most power-efficient one



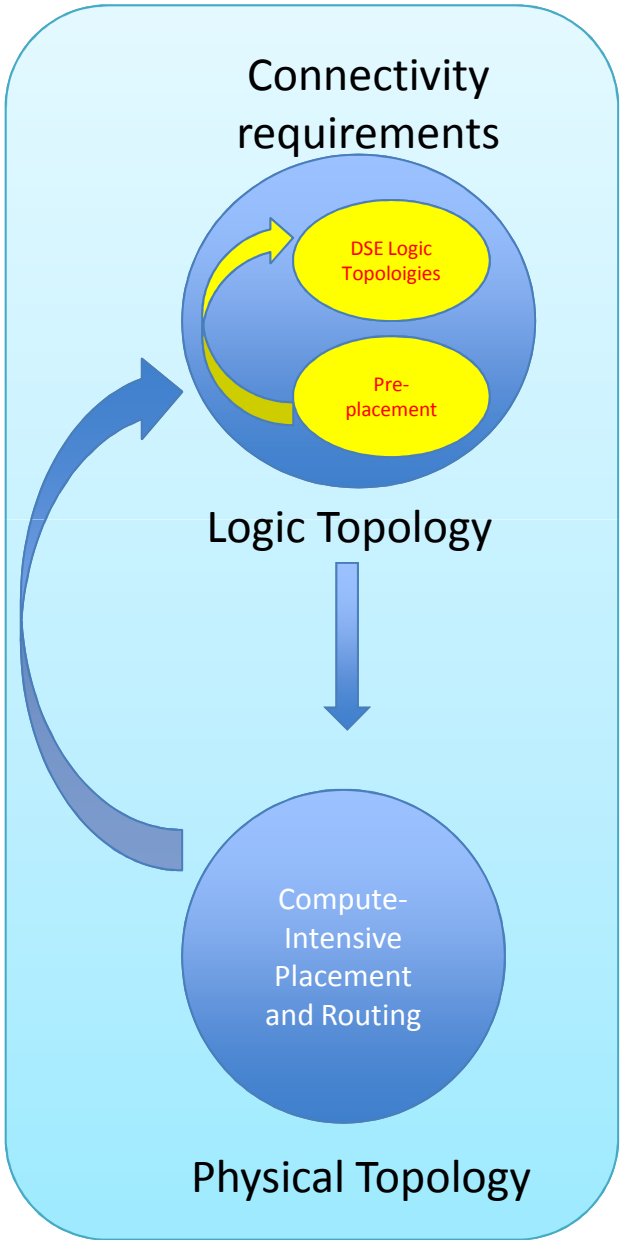
The lambda router is the most power-efficient topology

We start to have a design space to explore here! The common design abstraction is there! The pruning method to some extent depends on th P&R algorithm!

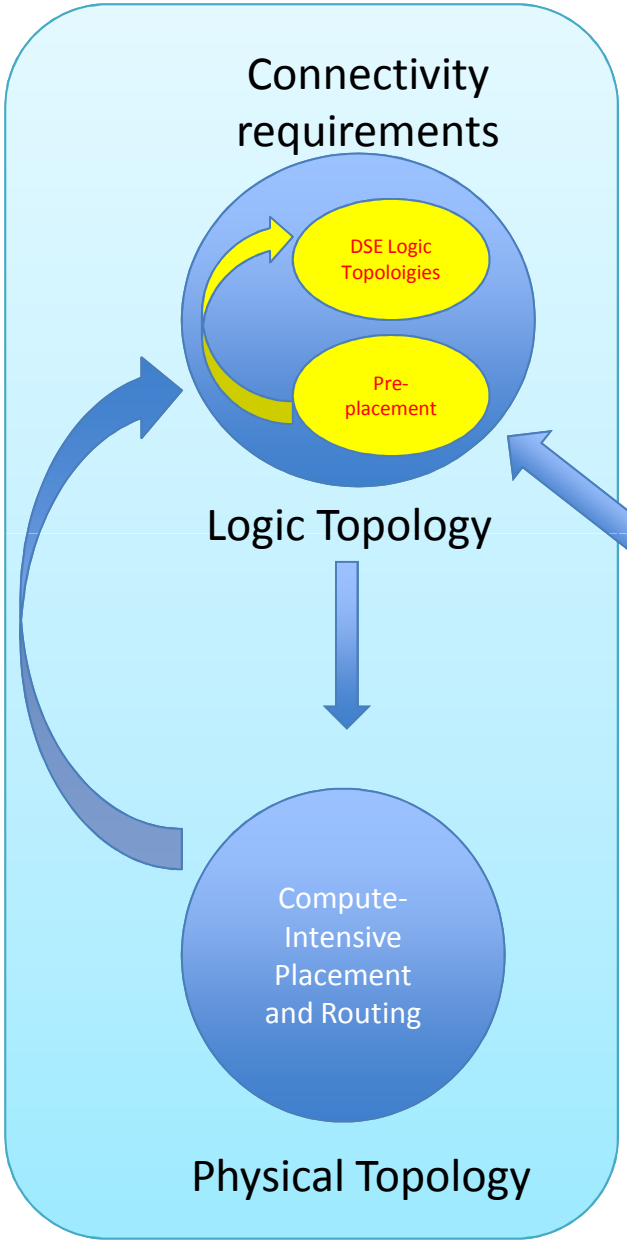
Lesson learned



Searching for more design predictability?

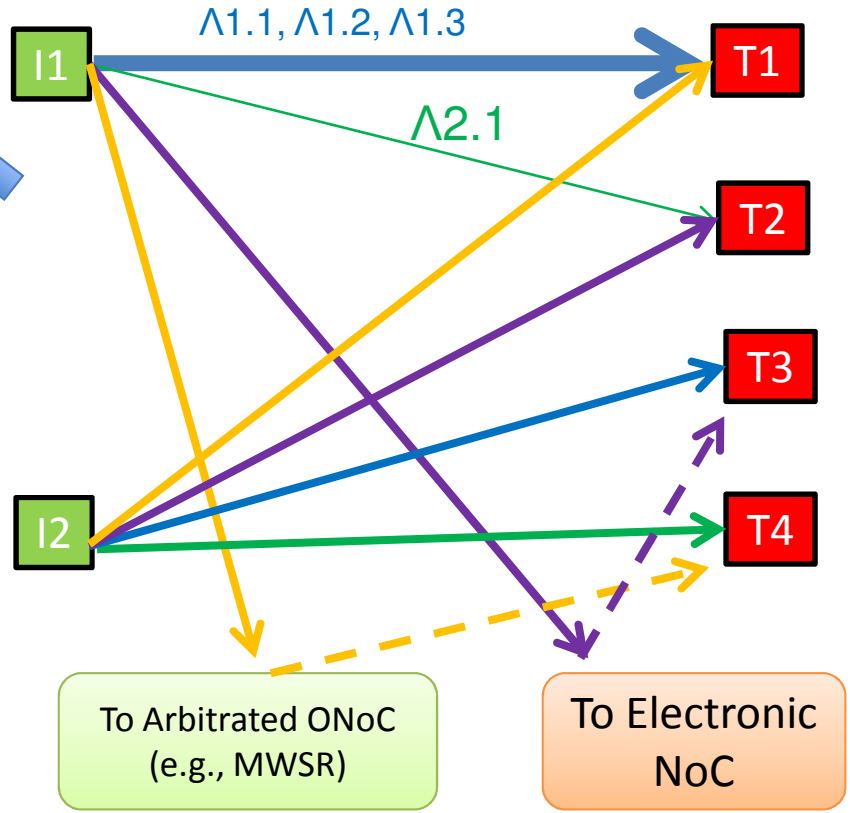


Augmenting the Flow



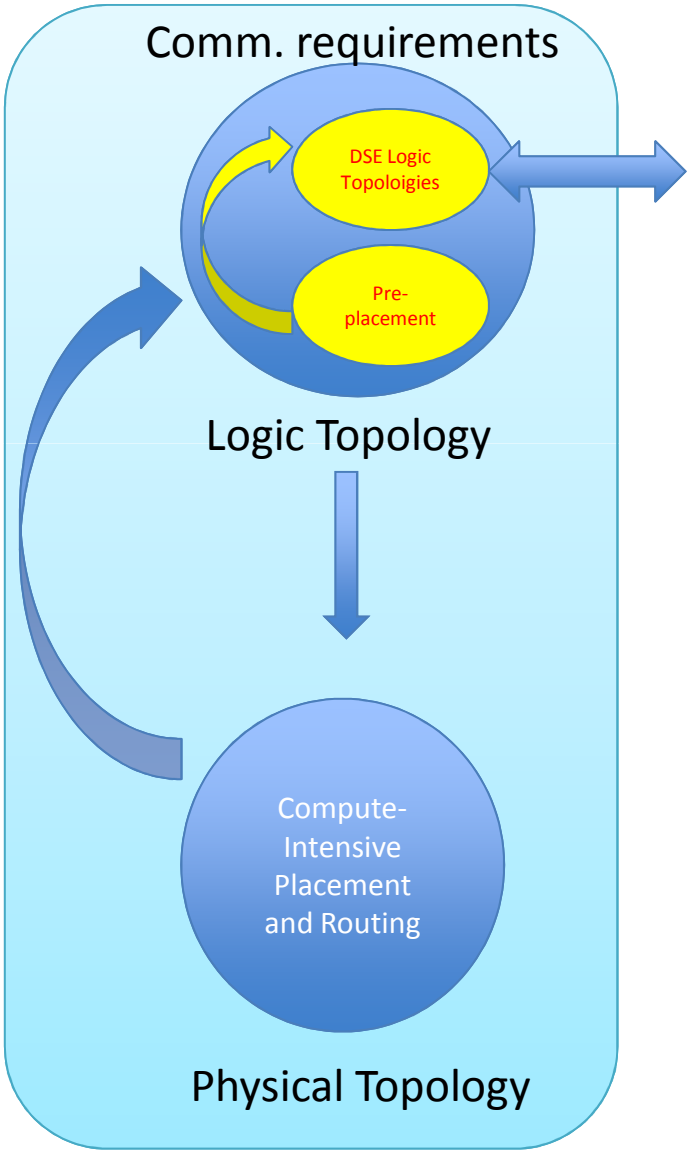
Specification of communication Requirements per flow (bandwidth, latency)

Customization and/or hybridization



Augmenting the Flow

Design iterations may be motivated by early-phase analysis of metrics pertinent to the physical layout

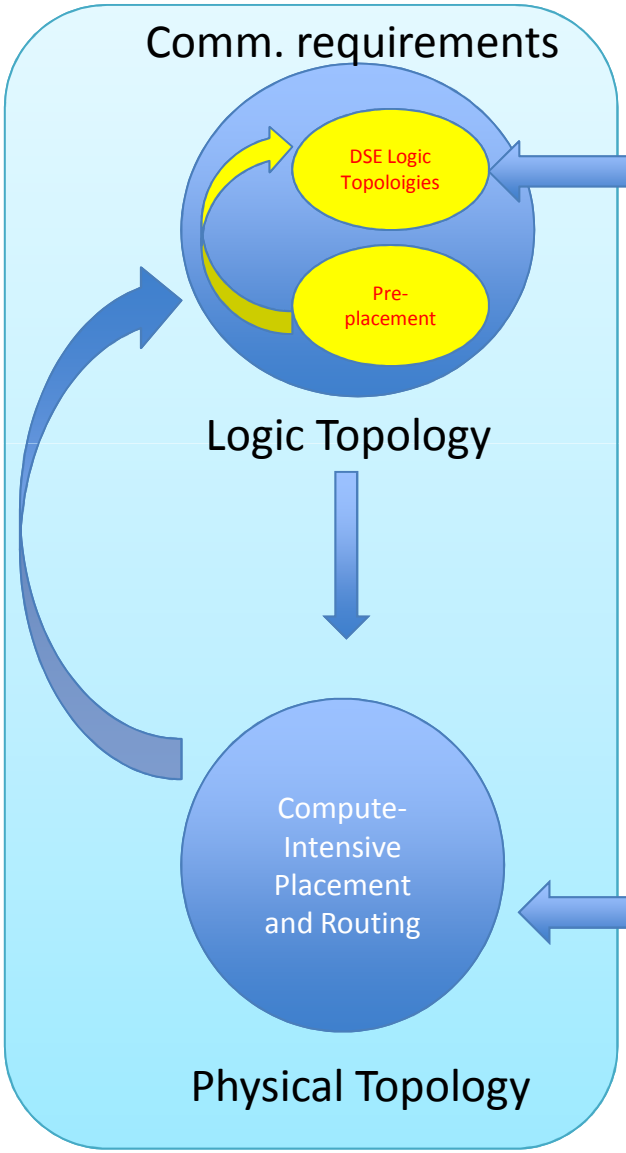


There is no "clean" (inverse) correlation between maximum insertion loss and worst OSNR in a topology

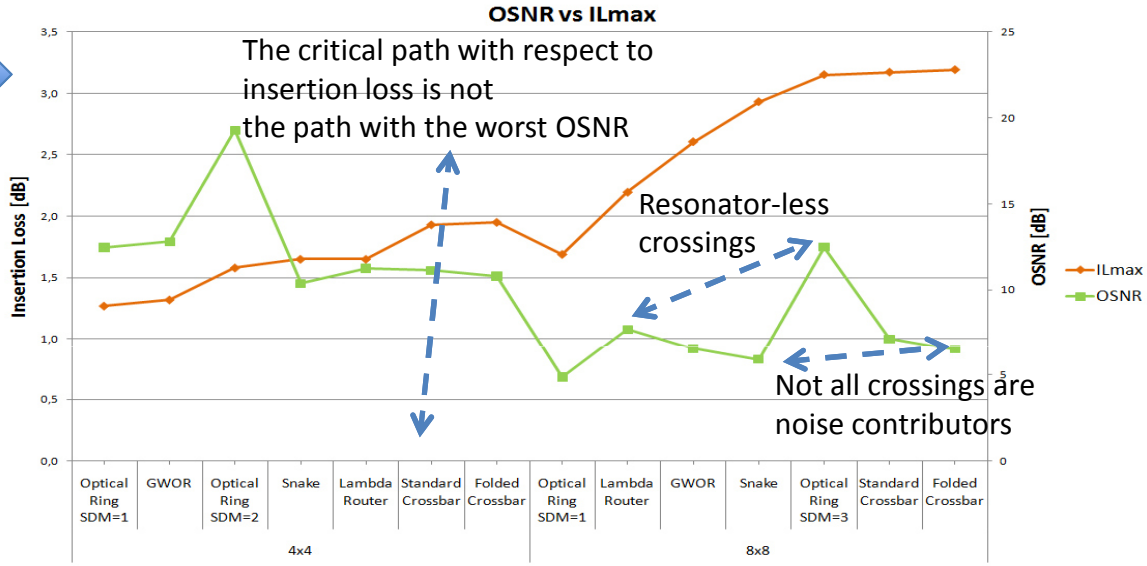
Automation of the flow will help designers capture subtle effects

Augmenting the Flow

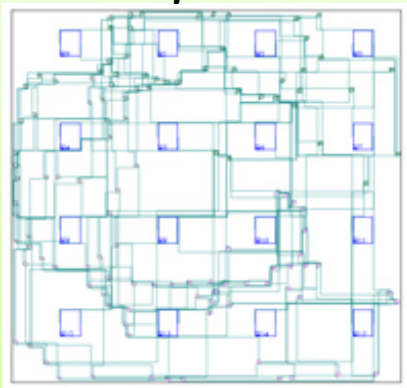
Design iterations may be motivated by early-phase analysis of metrics pertinent to the physical layout

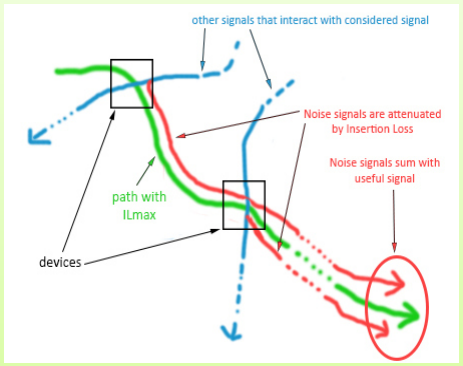


There is no "clean" (inverse) correlation between maximum insertion loss and worst OSNR in a topology



OSNR analysis should be performed on the layout as well, which requires extensive CAD tool support





Conclusions

- Optical NoCs have been demonstrated to enable system-level performance speedups and energy savings in academia.
- However, it is the availability of design methodologies and synthesis toolflows that makes the real difference when it comes to industrial exploitation.
- Clearly identifying abstraction layers in ONoC design is the ideal stepping stone to kick-off this process. Nonetheless, cross-layer optimizations are fundamental for predictable design.
- Customization should again drive ONoC design, especially in the embedded computing domain
- **In a sense, the history of electronic NoCs is repeating itself. However, there will be a fundamental difference: the cross-layer integration issue of optics with electronics.**
- It’s time to start bridging the EDA gap!