3rd OPTICS Workshop, co-located with DATE 2017

Design Automation Beyond its Electronic Roots: Toward a Synthesis Methodology for Wavelength-Routed Optical Networks-on-Chip

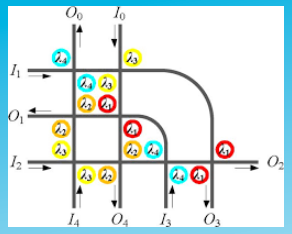
Davide Bertozzi

MPSoC Research Group,
University of Ferrara (Italy)

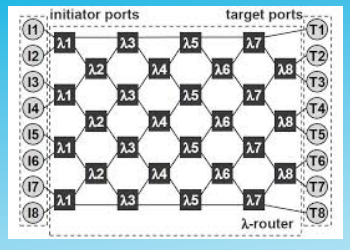
EARLY STAGE OF OPTICAL NETWORKS-ON-CHIP

Optical interconnection networks stand out as the most promising emerging technology for on-chip communication

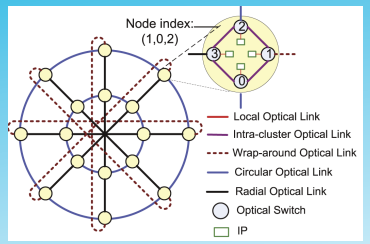
So far, lots of topologies, routing strategies, system architectures, control policites, etc.



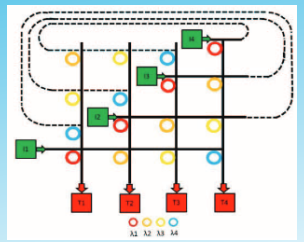
GWOR



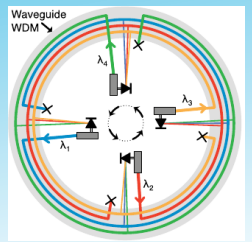
λ-Router



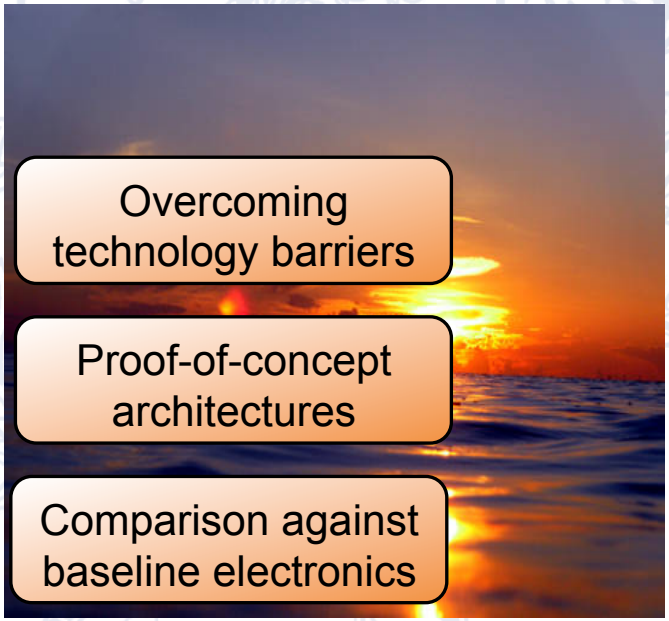
2D-HERT



Snake



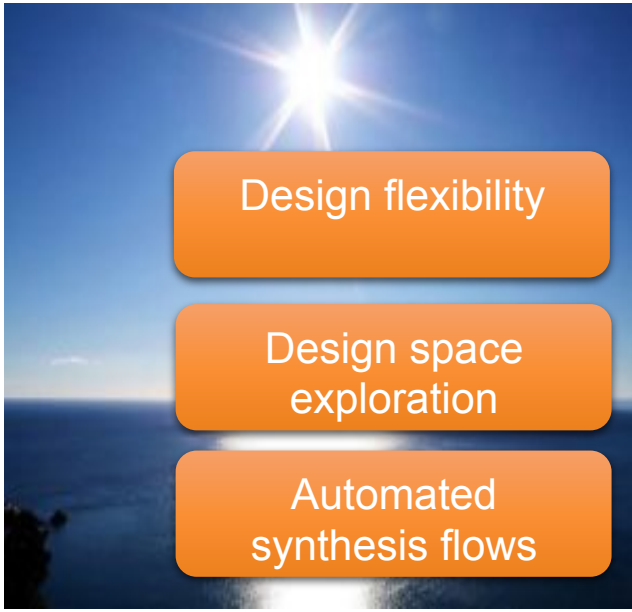
Ring



Today

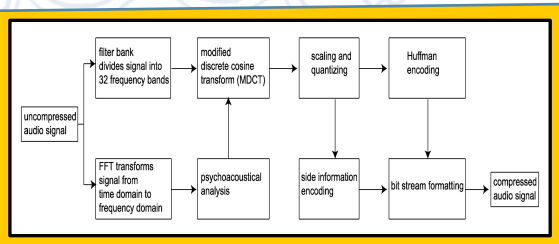
Small subset of the design space

Still trying to make the case



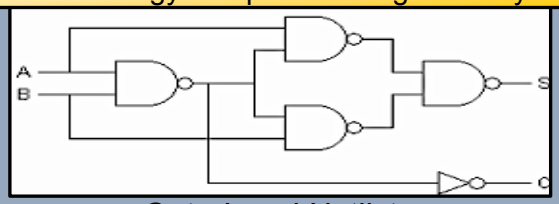
EDA BEYOND ITS «E-ROOTS»

Starting from a high-level description, EDA flows operate on abstractions and refine them into an actual implementation with components from a technology library.



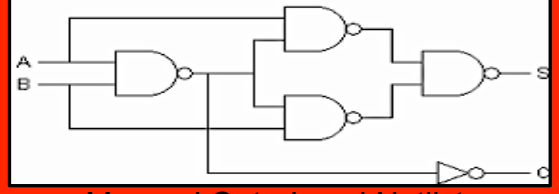
High-level specification

Technology-independent Logic Library

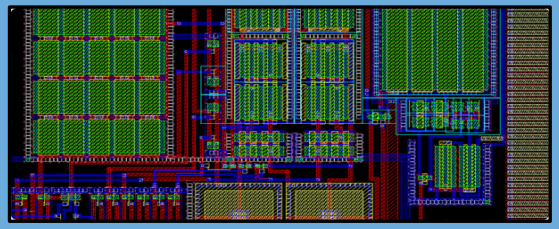


Gate-Level Netlist

Technology Library



Mapped Gate-Level Netlist



Planar geometric shapes

0 Routing Protocol Selection

I. Wavelength Resolution

II. Technology Mapping

III. Wavelength Assignment

IV. Topology Connection

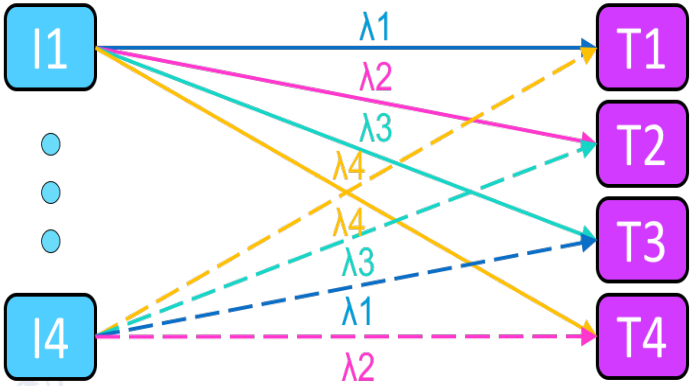
V. Device Parameter Selection

VI. Physical Mapping Flow

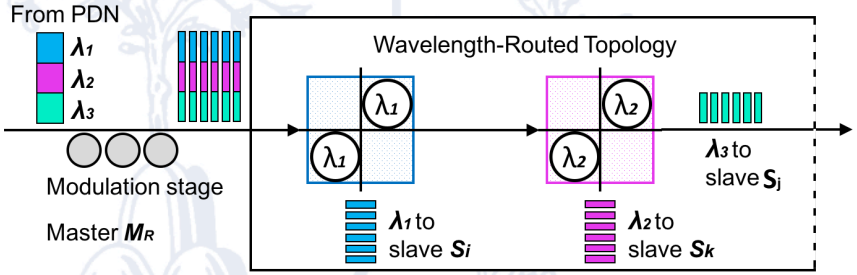
VII. Physical Design

WAVELENGTH-ROUTED OPTICAL NOCS (WRONOCS)

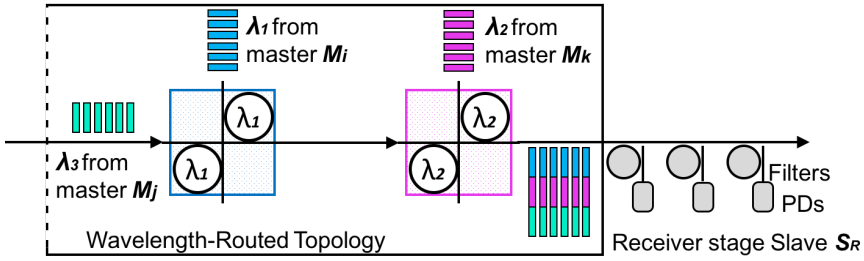
- **Wavelength-selective routing**:



- **The wavelength- selective routing function is ultimately fulfilled by means of two abstract primitives:**



Drop function at initiator's side.

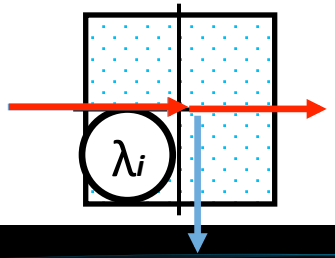


Add function at target's side.

Basic building block for the implementation of the add and of the drop function:
the **1x2 ADD/DROP FILTER**

On-resonance signal

Off-resonance signal



Proposed Synthesis Methodology

I. Wavelength Resolution

II. Technology Mapping

III. Wavelength Assignment

IV. Topology Connection

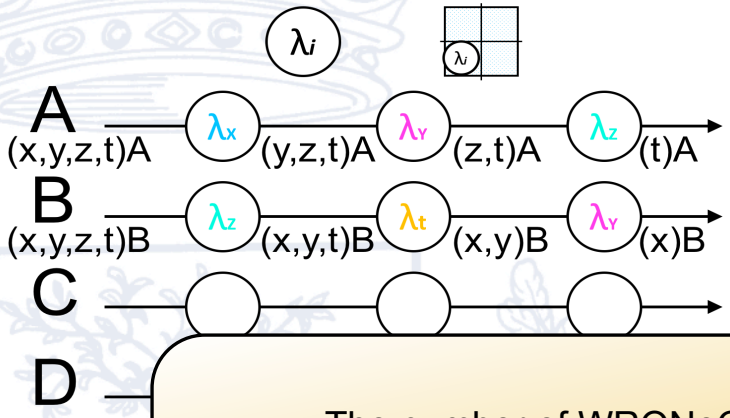
Can we understand all topology design points in the context of a unified design framework?

Can we populate the design space of wavelength-routed optical NoC topologies?

Synthesis Methodology

1. Wavelength Resolution

WDM input signal from masters should undergo the drop function by going through *n − 1* ADFs, when assuming the connectivity of *n* masters with *n* slaves.

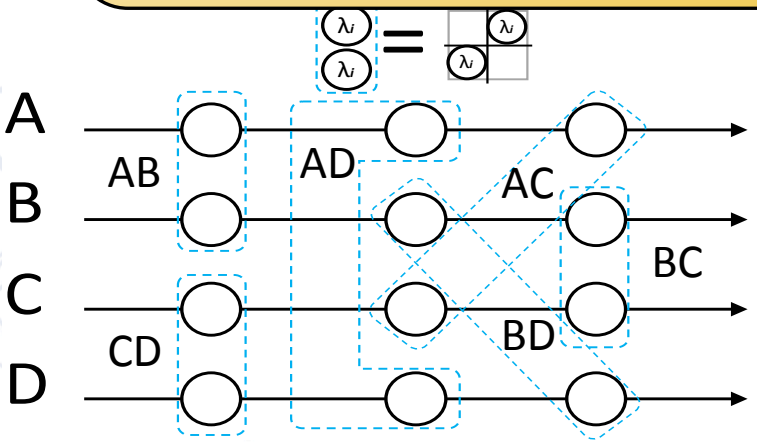


Fixing the parameters of the WRG indirectly prunes the design space and biases the methodology toward a specific

The number of WRONoC topologies in the design space amounts to
$[(n-1)*(n-2)*.....*(n-n+2)]^n$
A 4x4 WRONoC topology can be implemented in 1296 different ways
For n > 4, the full enumeration of all possible design points becomes computationally unaffordable.

2. Techno
Groupin

library!

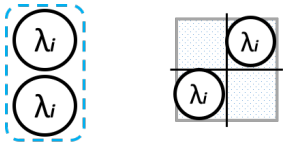


Mapping Constraint:

Legal mappings are the combinations without repetitions
C(n,2)
for picking 2 unordered outcomes from *n* possibilities.
(…assuming 4 masters and 4 slaves, C(4,2)=6 possible mappings AB,AC,AD,BC,BD,CD)

3. Wavelength Assignment

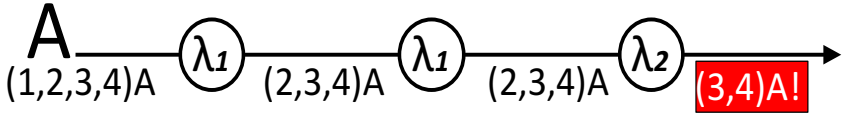
Assumption: we do not consider real wavelengths at this stage (e.g., 1550 nm) but symbolic ones, such as *λ1* to *λn* in an *n*×*n* WRONoC.



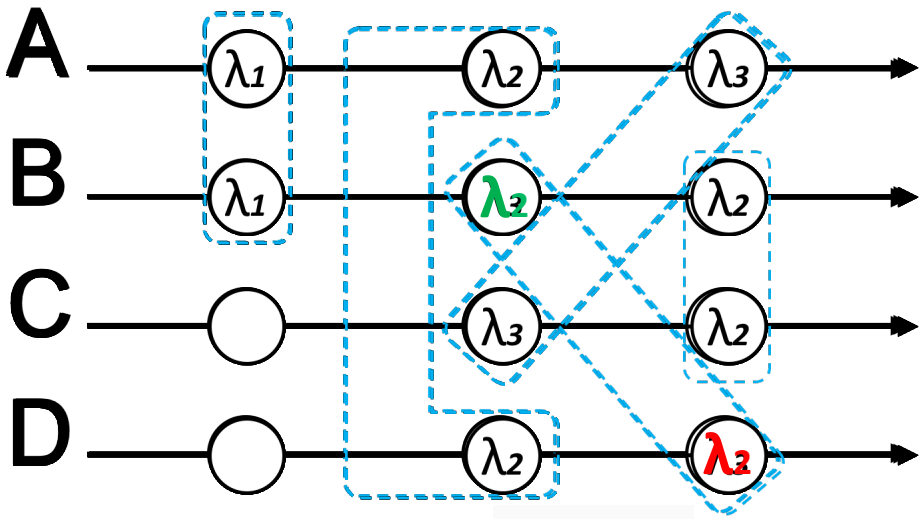
Greedy algorithm to perform complete wavelength assignment:

Constraint:

- ✓ **the resonant wavelengths of the add-drop filters along a single row should be different from each other!**

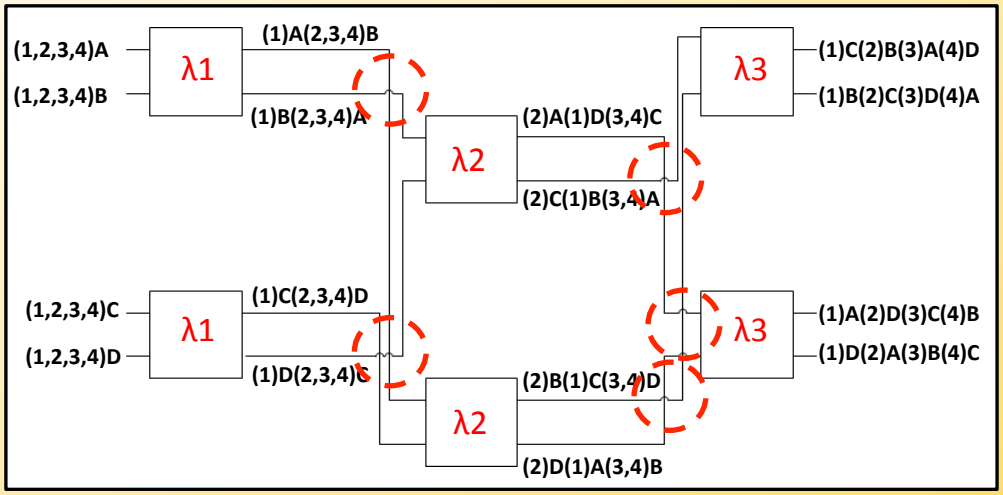


- ✓ **the mutual exclusion property should be checked on coupled rows at each wavelength assignment**



When multiple wavelength options are available, the algorithm selects that with the lowest ID. This way the algorithm typically optimizes the number of microring resonator types.

Topology Connection



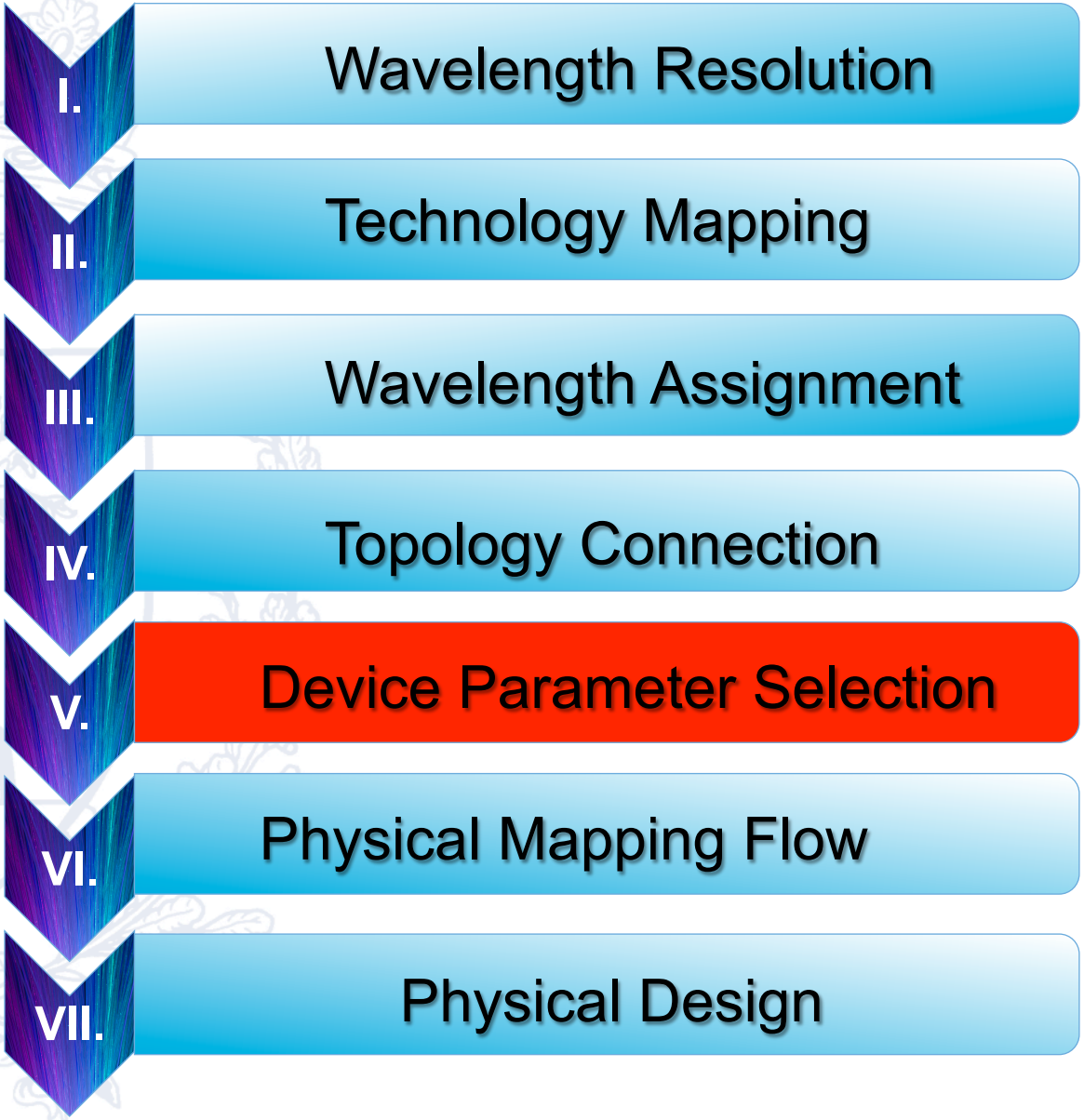
Generic Topology from the Front-End Flow

- Generic topologies from our flow exhibit a more intricate drawing on a 2D-plane,
  - ✓ there are crossings even outside PSEs.

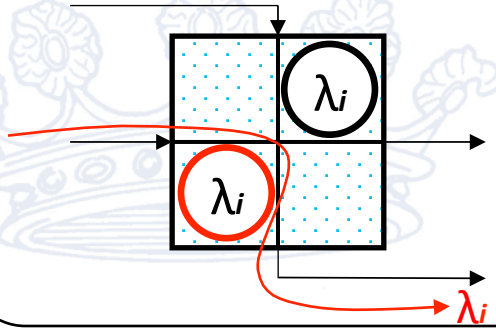
We would like to stress that these crossings should be considered as *apparent*, since at this stage we are drawing the *logic* topology, not the physical one.

it very important to differentiate between abstraction layers!

Proposed Synthesis Methodology



DEVICE PARAMETER SELECTION



- What is the exact radius length of the MRR, typically in the range 5-20um?
- What is the exact value of the n wavelengths used by each initiator in an *n x n* wavelength-routed optical NoC?

This is not just a refinement step! It has implications on the quality of the solution!

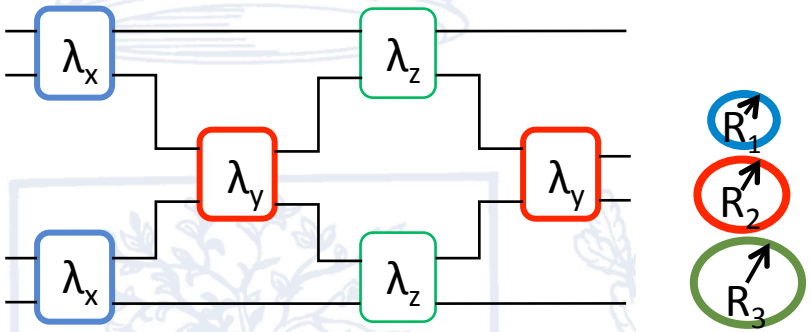
→ Scalability: do **legal assignments exist for any topology size? Routing faults should be prevented!**

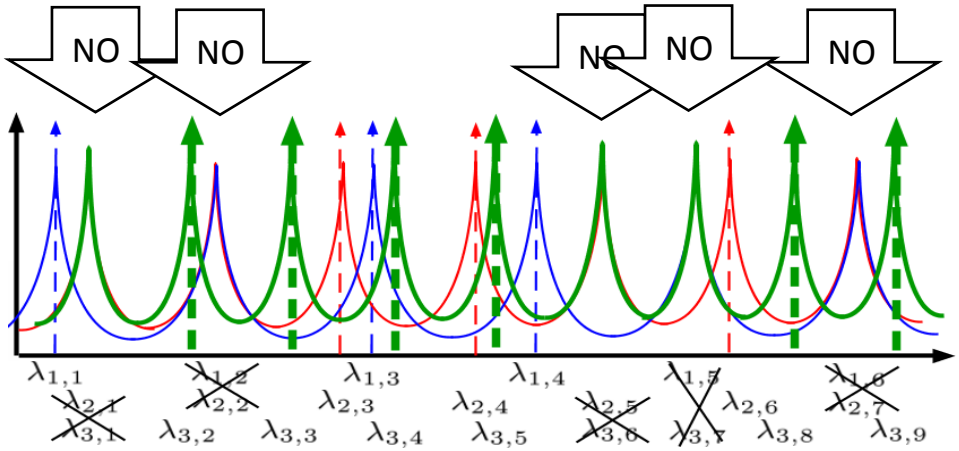
→ Parallelism: **what is the maximum supported communication parallelism on each wavelength channel?**

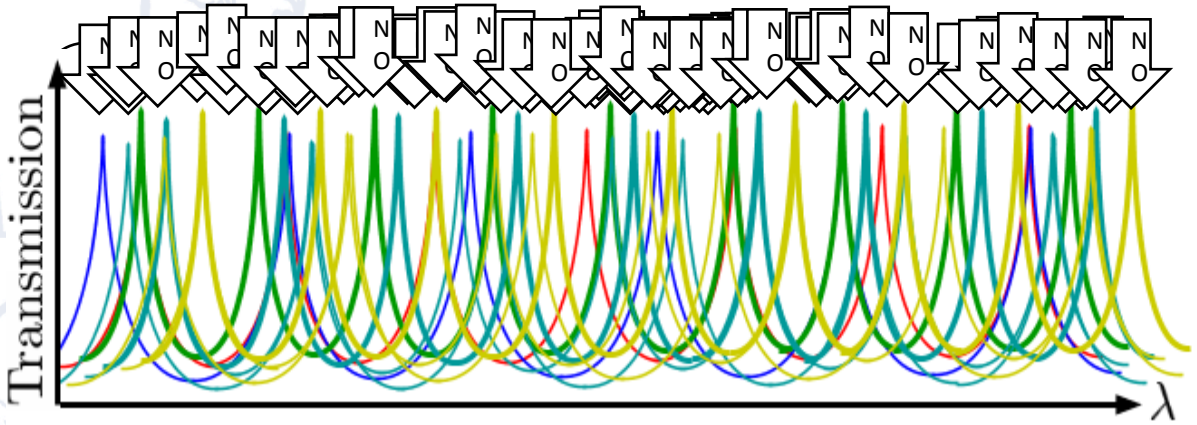
→ Technology: does **the uncertainty of the manufacturing process affect achievable scalability and parallelism**?

ROUTING FAULTS

→ Available parallelism is ~~6~~ ~~4~~ **3** in PSE_x
→ Available parallelism is ~~7~~ ~~5~~ **3** in PSE_y
→ Available parallelism is ~~9~~ **7** in PSE_z

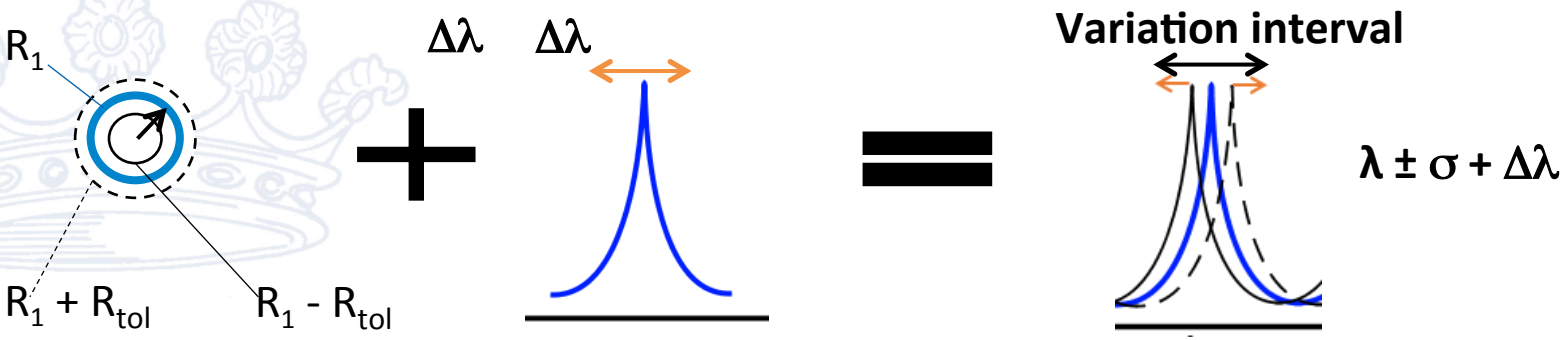




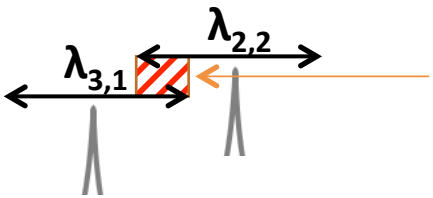


As topology size increases, the proliferation of filter types and wavelength channels may limit the availability of non-overlapped transmission peaks, which may cause the topology to be practically infeasible

PARAMETER UNCERTAINTY



For the time being, we made the most conservative assumption: avoiding routing faults in any variability scenario.



There exists a variation scenario that yields a routing fault!!

Routing fault avoidance constraints:

→ A selected on-resonance of a PSE type must be off-resonance of all the other PSE types

→ **Two on-resonances conflict if their variation intervals overlap**

Future work: statistical analysis of routing faults

DEVICE PARAMETER SELECTION: PROCEDURE

We set up a table where the r-th row contains the resonances of radius Rr

r	R_r [μm]	$\lvert\{\lambda_{r,j}\}\rvert$	$\lambda_{i,1}\langle\sigma_{i,1}+\Delta\lambda\rangle$, $\lambda_{i,2}\langle\sigma_{i,2}+\Delta\lambda\rangle$, $\cdots$ [nm]
1	5	5	$1496.4\langle3.5\rangle$, $1521.3\langle3.6\rangle$, $1547.1\langle3.6\rangle$, $1573.8\langle3.6\rangle$, $1601.4\langle3.6\rangle$
2	6	6	$1500.5\langle3.0\rangle$, $1521.3\langle3.0\rangle$, $1542.7\langle3.1\rangle$, $1564.8\langle3.1\rangle$, $1587.5\langle3.3\rangle$, $1610.8\langle3.3\rangle$
3	7	6	$1503.4\langle2.6\rangle$, $1521.3\langle2.6\rangle$, $1539.6\langle2.7\rangle$, $1558.4\langle2.7\rangle$, $1577.7\langle2.7\rangle$, $1597.4\langle2.8\rangle$
4	8	7	$1505.6\langle2.2\rangle$, $1521.3\langle2.2\rangle$, $1537.3\langle2.2\rangle$, $1553.7\langle2.2\rangle$, $1570.4\langle2.0\rangle$, $1587.5\langle2.0\rangle$, $1604.9\langle2.0\rangle$

We have to select

- **As many radius lengths (i.e., table rows) as the number «s» of microring resonator types in the topology**
- **As many resonant wavelengths (i.e., table columns) as possible for each resonator type**

The possible decisions:

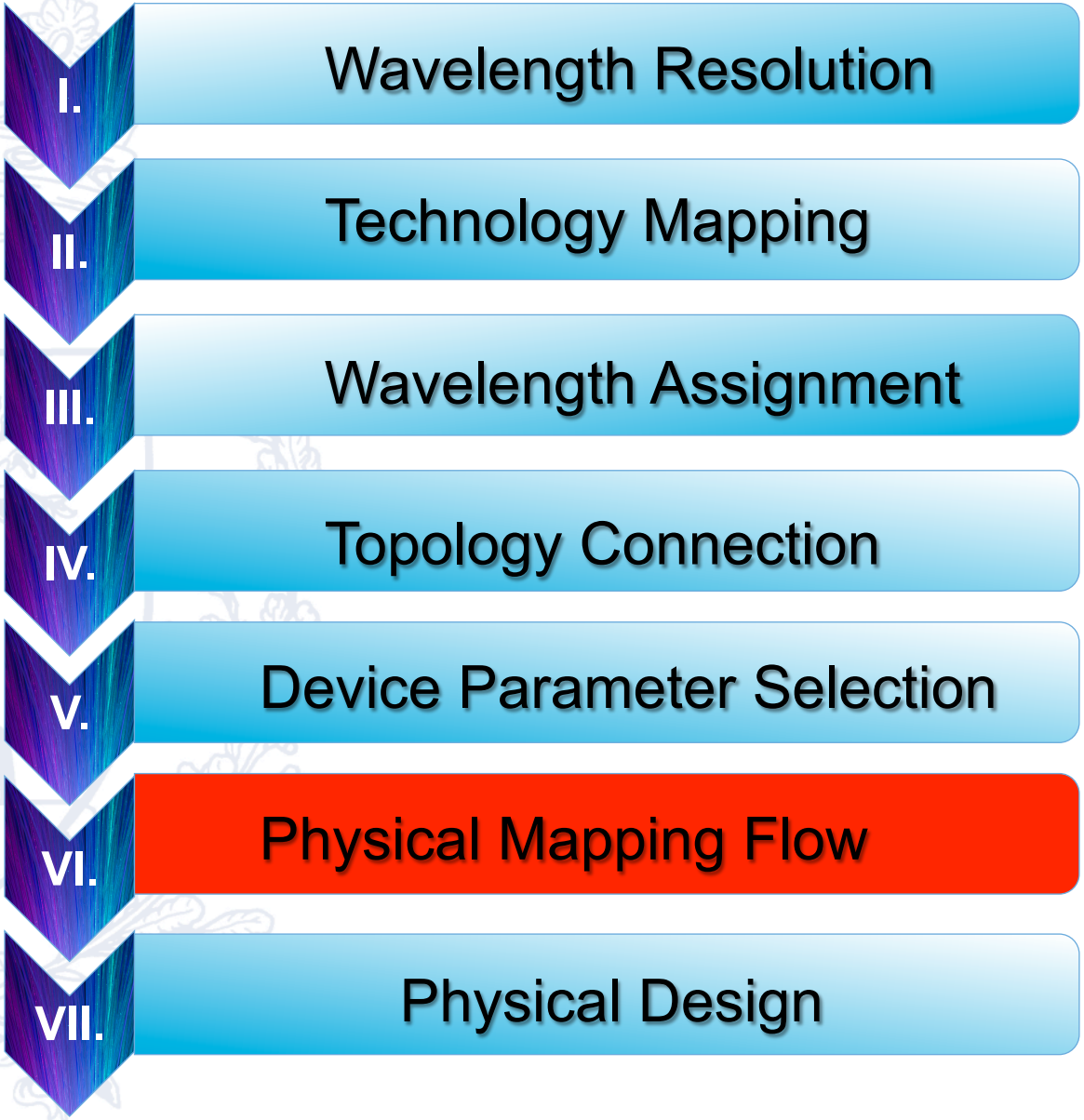
1. **must avoid routing faults**
2. **have to maximize the minimum parallelism P(*s*) across the «s» selected rows**

This problem can be modelled as a **Constrained Optimization Problem**:

Decision variables: *which resonances to be selected*
Constraints: *routing fault avoidance, select s radii*
Objective: *maximize the bottom parallelism*

We have chosen Answer Set Programming as declarative technology and Clasp as COP solver.

Proposed Synthesis Methodology



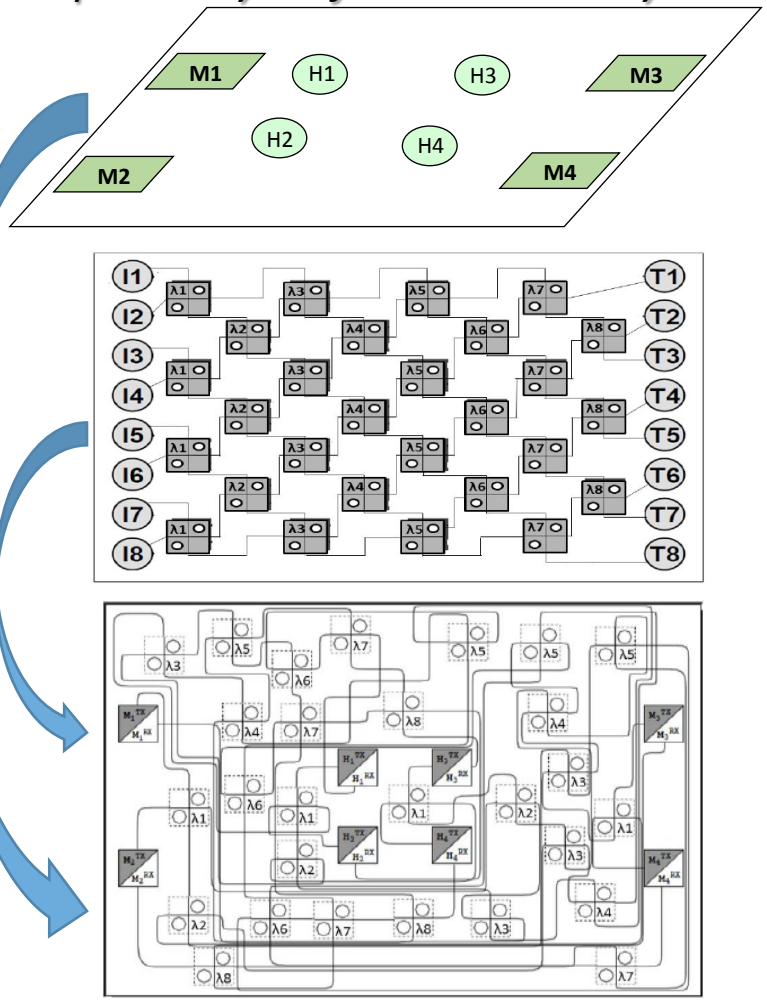
The Place&Route Problem

Path list, chip area, positions of hubs and memory controllers

- Minimize **insertion loss**, e.g. minimize
  - Waveguide length
  - Number of crossings between waveguides
  - Number of bends
- Constraints:
  - Place all PSEs and waveguides inside chip area
  - No overlap

Valid and optimal layout

Optical Layer of a 3D-stacked system



Lots of unexpected waveguide crossings (which burden on the static power budget)

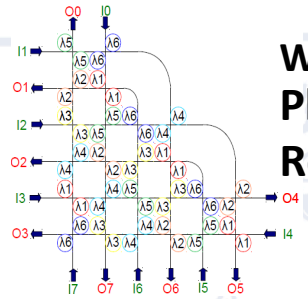
P&R engines end up optimizing insertion loss across one or multiple optical paths

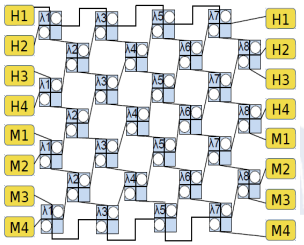
The tool has to strike a good balance between crossing losses and propagation losses, which might be conflicting objectives

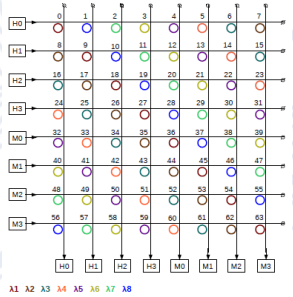
$$\min_{\mathbf{x},\mathbf{r}} \max_{p \in P} \alpha \cdot L_p(\mathbf{x},\mathbf{r}) + \beta \cdot C_p(\mathbf{x},\mathbf{r})$$

Our Work: the **Proton+** Tool

A collaboration with TU Munich (Prof. Ulf Schlitchmann)



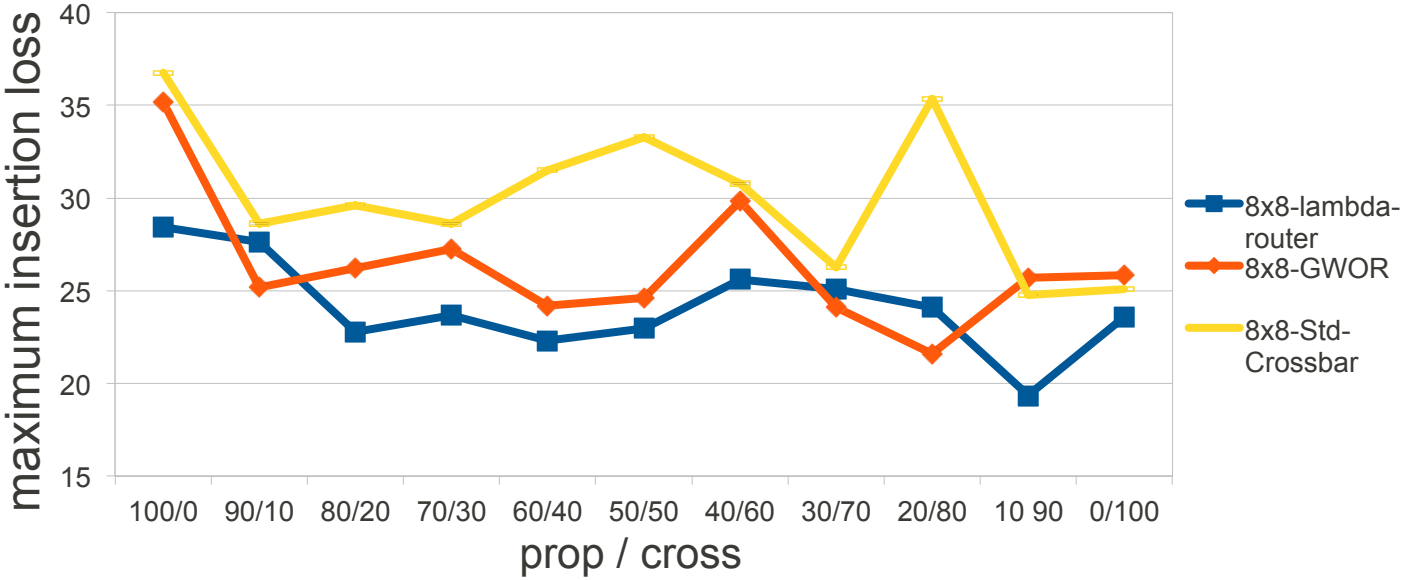




Where Lp and Cp are approximate functions of path lengths and crossings

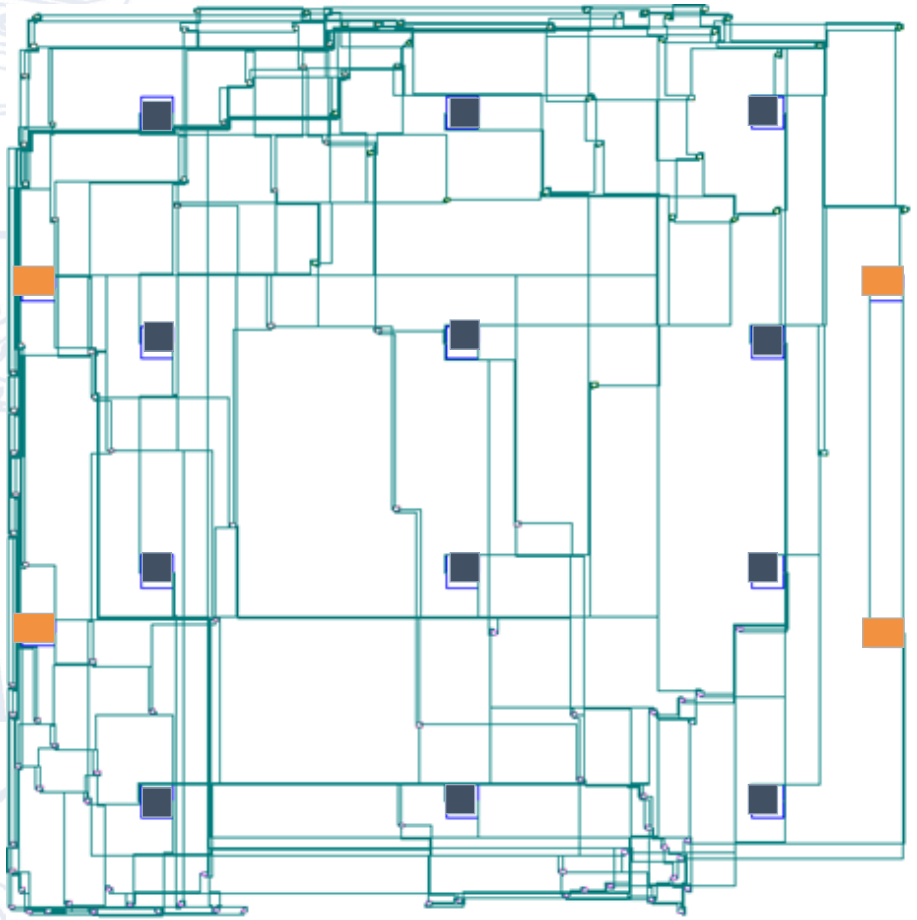
Placement: Non-linear optimization problem solved with an IPM

Routing: adaptation of the Lee's algorithm «Maze Router»



By setting the weights of the objective function, the best physical mapping for the technology at hand can be achieved

Resulting Layout of 16x16 λ-Router



Hubs

Memory controller

- Ins. Loss max = 44dB
- 255 crossings on the critical path
- 28636um waveguide length on the critical path
- 24425 sec of CPU time

(Intel Core 2 Quad CPU with 8GB RAM running at 2.33GHz)

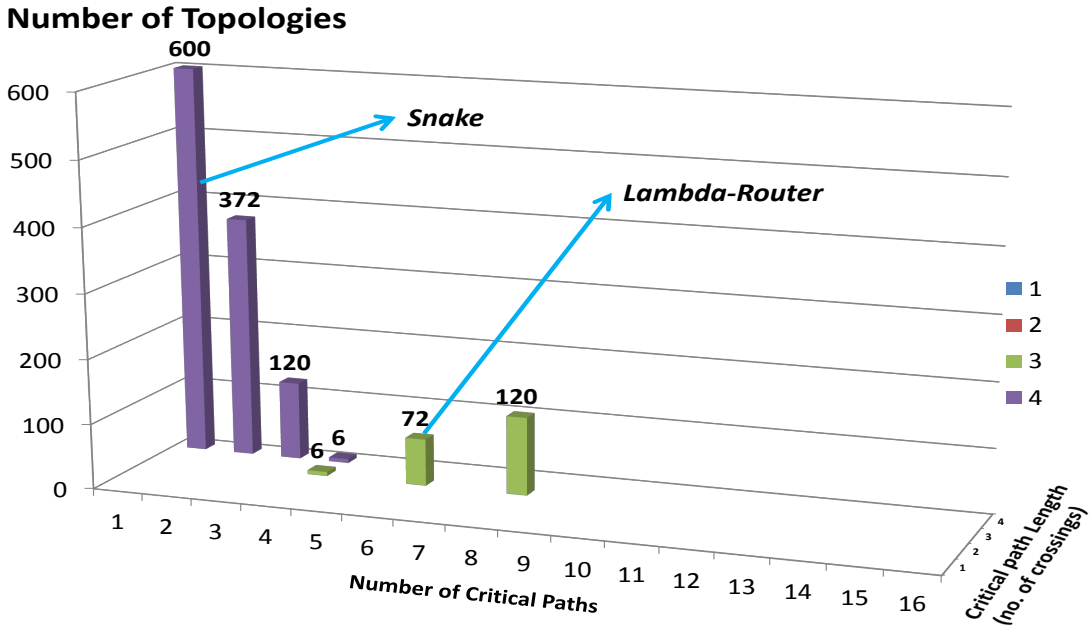
The new force-directed placement algorithm from TU Munich (PLATON) can bring the loss down to 22.5dB in 64 sec

Experimental Results

Experimental Results

Populating the Design Space with the Front-End Flow

- We exhaustively populated the design space of *4×4* WRONoC topologies

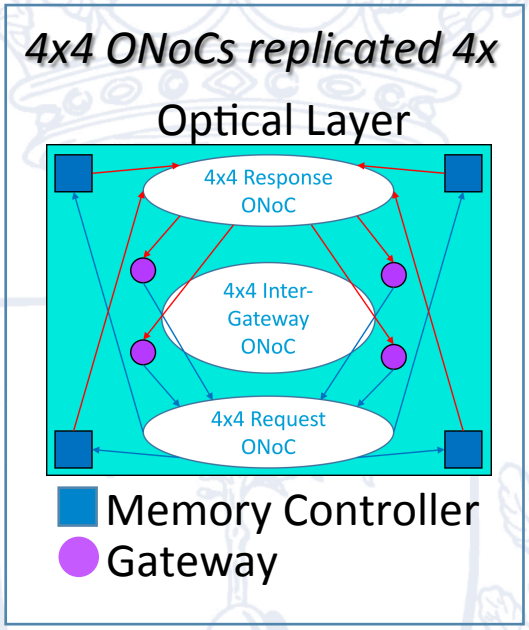


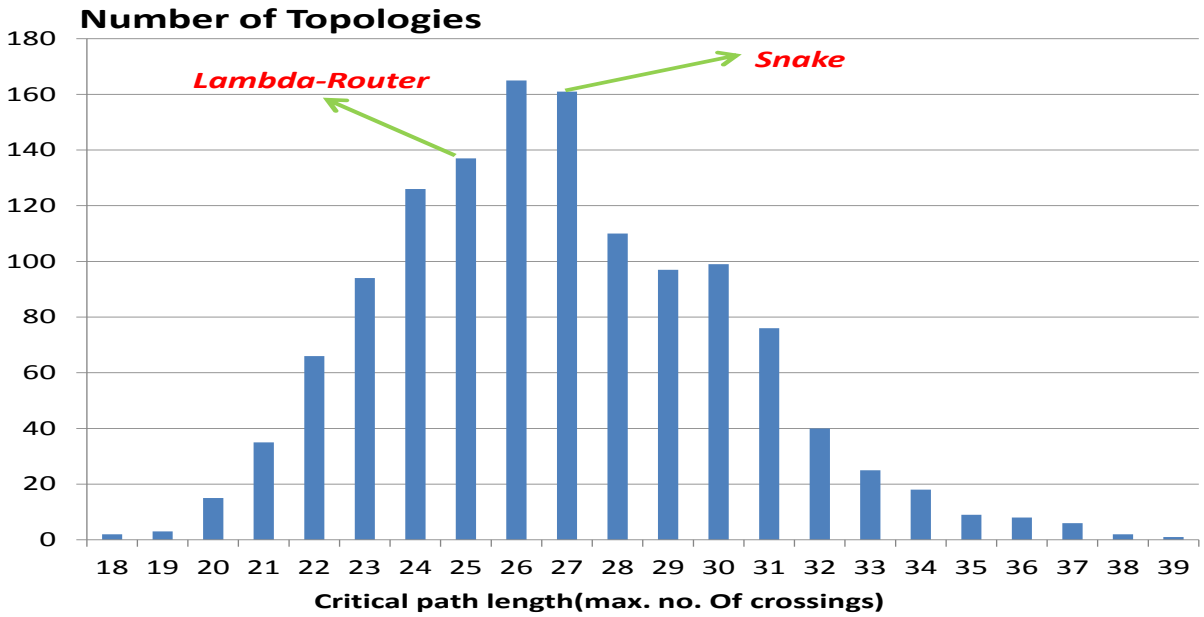
Complete design space for 4x4 WRONoC topologies built out of 2x2 PSEs.

- Most logic topologies have 4 crossings on the critical path, while 15% of the population has only 3 (the λ-router belongs to this latter category).
- When the critical path is the longest, then it is likely to be unique (46% probability).
- When the critical path is shorter (3 crossings), the topology is likely to have more than 1

Experimental Results

Physical Design with Proton+





Distribution of the critical path after physical mapping.

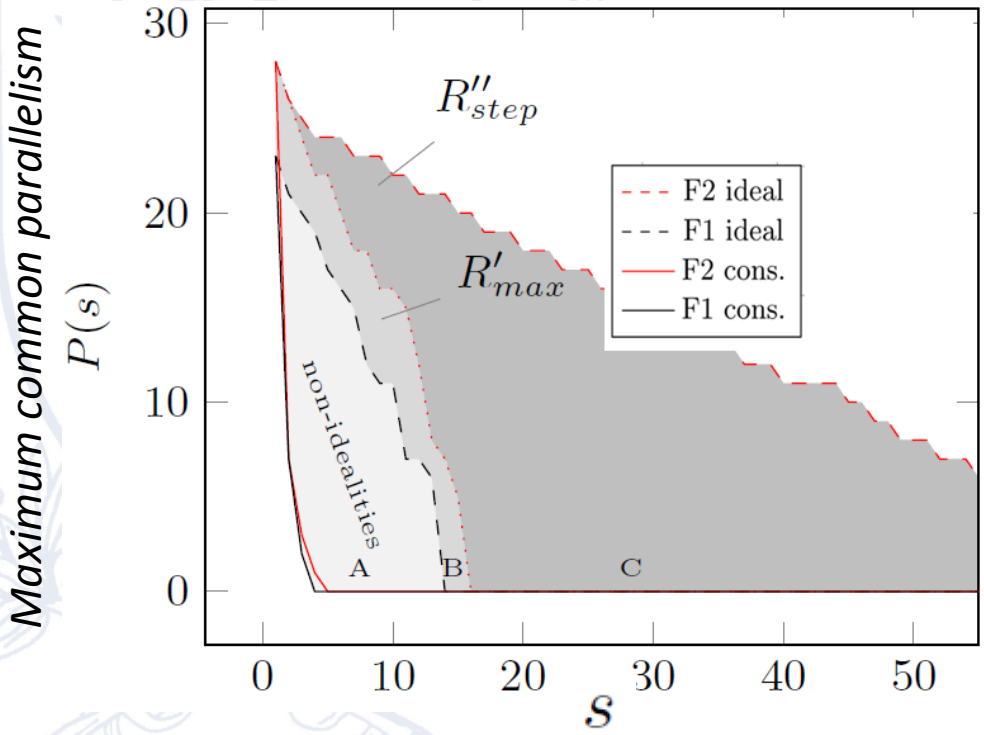
- Interestingly, now the critical path length ranges from 18 to 39 crossings, thus featuring a much larger variability than in logic schemes.
  - This raises the issue of placement-aware logic topology synthesis, completely new discipline for optical NoCs.
- λ-Router and snake proposed in literature are not the best topologies from the critical path length viewpoint!
  - Design automation helps to get the most out of a technology

Experimental Results: Scalability

	Fabrication options			Decision table statistics
Options	R_{min}	R_{step}	R_{max}	#Radii
F1	5µm	1µm	25µm	21
F2	5µm	0.25µm	30µm	104

Conservative case: R_{tol}=0.01µm and $\Delta\lambda$=0.5nm (from the literature)

Ideal case: R_{tol}=0.0µm and $\Delta\lambda$=0.0nm (no variations)



Conservative

Wavelength routing performance is highly sensitive to the uncertainty of the manufacturing process

Ideal case

The resolution step is a better driver to increase parallelism with respect to making available more resonating wavelengths through larger ings

In the future we can hope to connect topologies with «s» up to 60, with limited parallelism tough!

CONCLUSIONS

- This work proposes an early-stage methodology to refine an abstract specification of a wavelength-routed optical NoC topology into a mapped layout intent
- The methodology bridges the gap between system designers and technology developers
- Milestones:
  - ✓ Wavelength filtering abstractions and legal combination criteria to infer any point of the logic design space
  - ✓ Formulation of a constrained optimization problem to select device parameters without incurring any routing fault
  - ✓ A place&route algorithm capable of mapping a logic topology onto the layout of an optical layer of a 3D-stacked system, meeting placement and routing constraints

Future work targeting automatic synthesis of WRONoC topologies that meet the connectivity requirements and the floorplanning constraints of the system at hand