

# 3D Optical ManyCore

UMR CNRS 6082 – INSTITUT FOTON - Fonctions Optiques pour les Technologies de l'information

INRIA Rennes – Bretagne Atlantique – IRISA UMR CNRS 6074

UMR CNRS 5270 - Institut des Nanotechnologies de Lyon - École Centrale de Lyon

# Core of the project

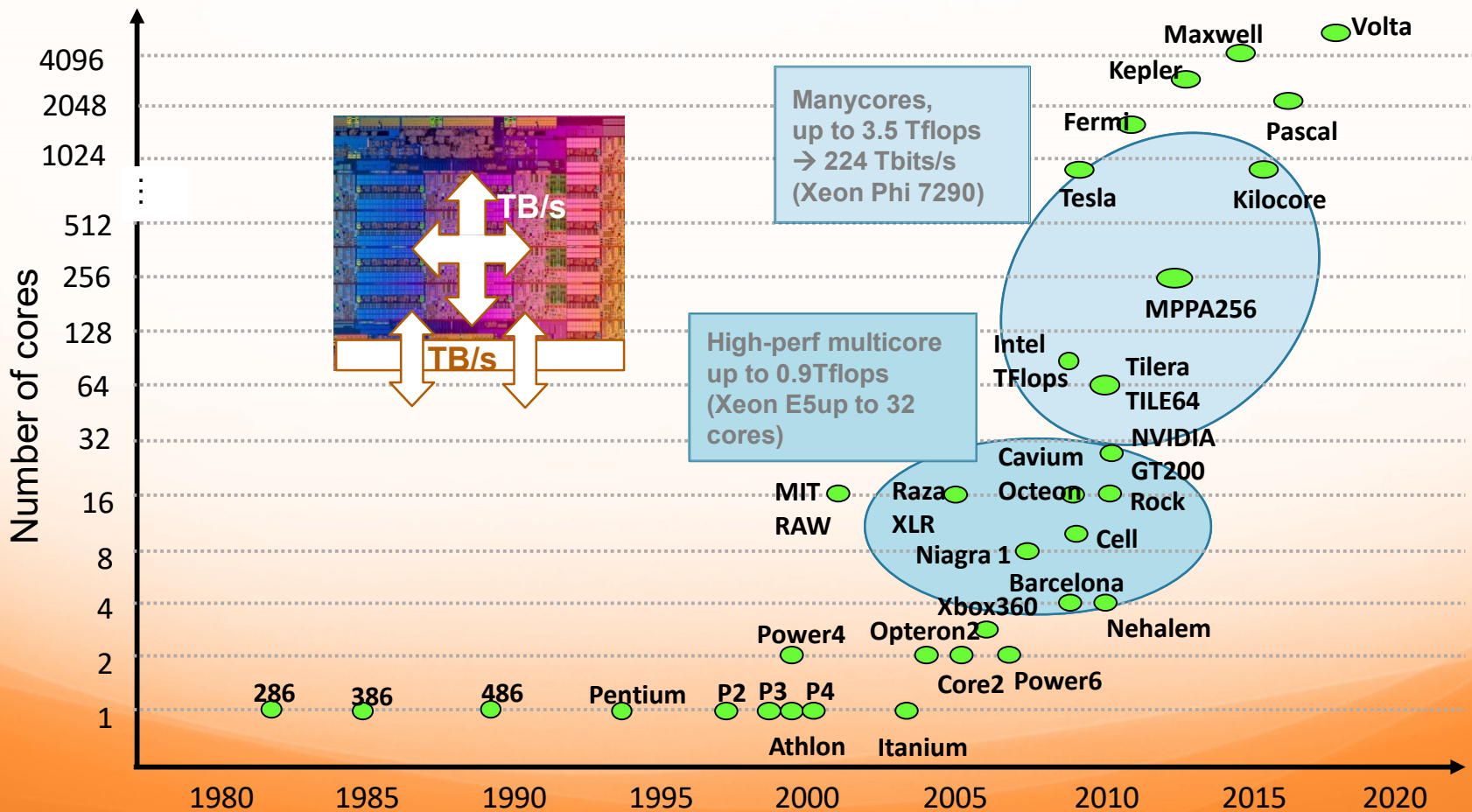
- Optical technologies could improve energy efficiency and data rate of on-chip interconnects used in many-core architectures for embedded and high-performance computing
- Designing a specific photonic layer suitable for a flexible and energy efficient high-speed optical network-on-chip (ONoC)

# State of mind

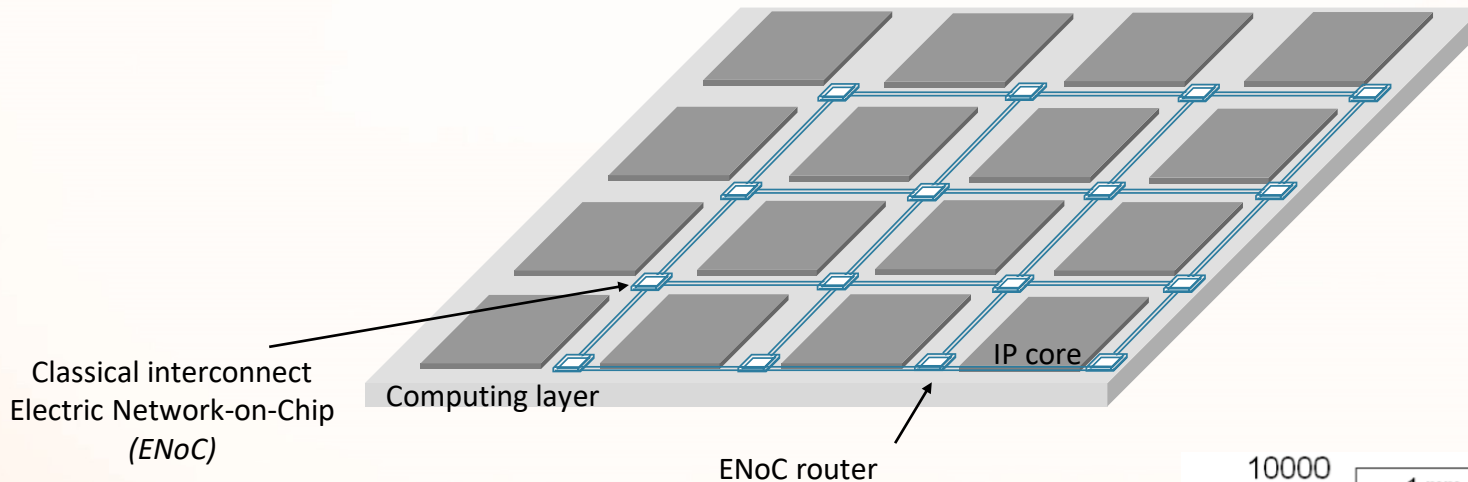
- Exploratory project
- Bring together Labs with different cultures
- Mixing ideas: Broadcast, free-space, spatial diversity, WDM, monolithic vs pseudomorphic photonic integration...

# Context

- Evolution of System-on-Chip architectures
  - Huge increase in communication between cores
  - More energy to move data (1pJ/bit/mm) than to compute (1aJ/bit)

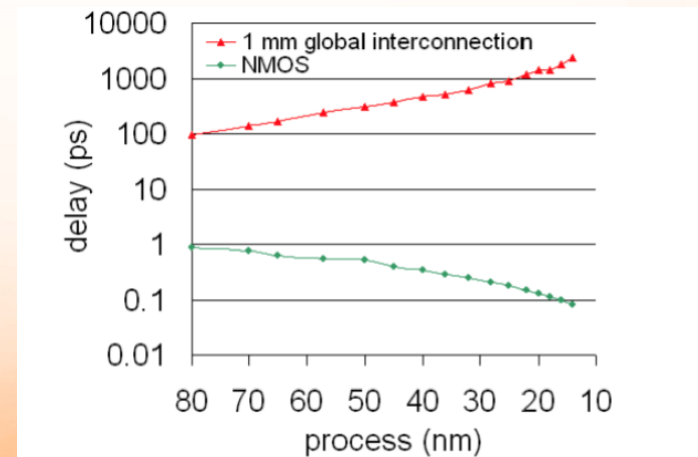


- Conventional architecture

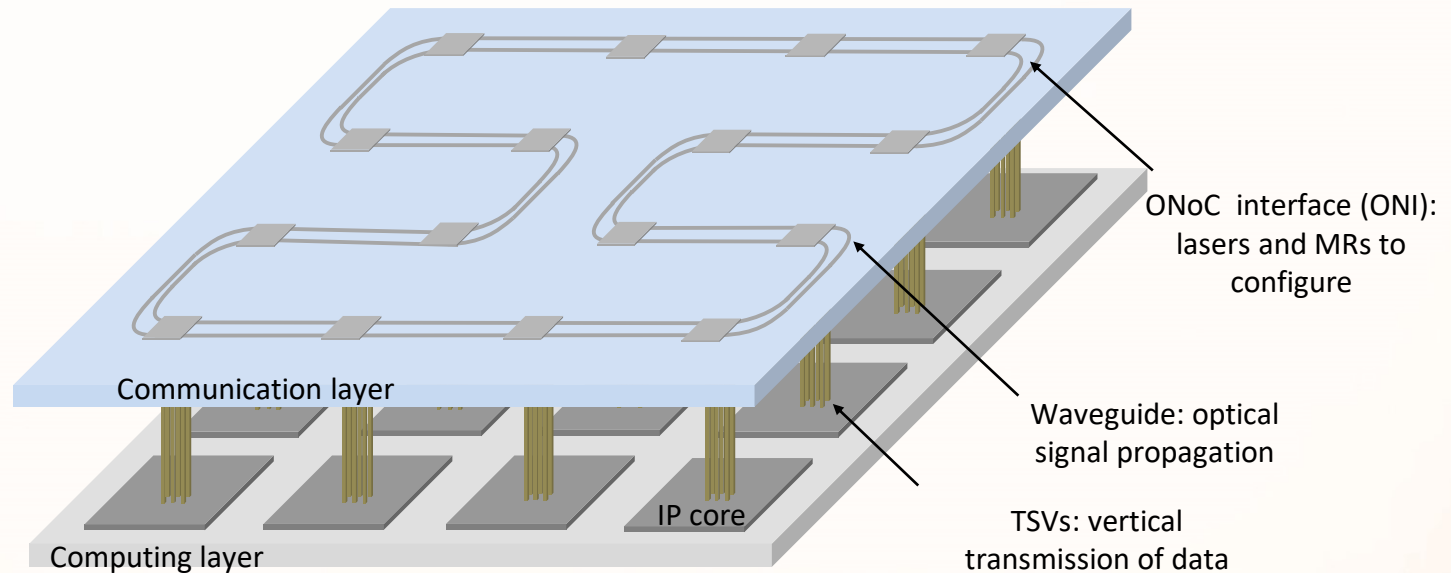


- ENoC limitations

- **Dynamic** power and **latency** are related to number of hops (router crossed)
- With transistor shrinking, gate **propagates** faster compared to wire
- On-chip interconnect can represent up to 30 to 80% of chip **power consumption**



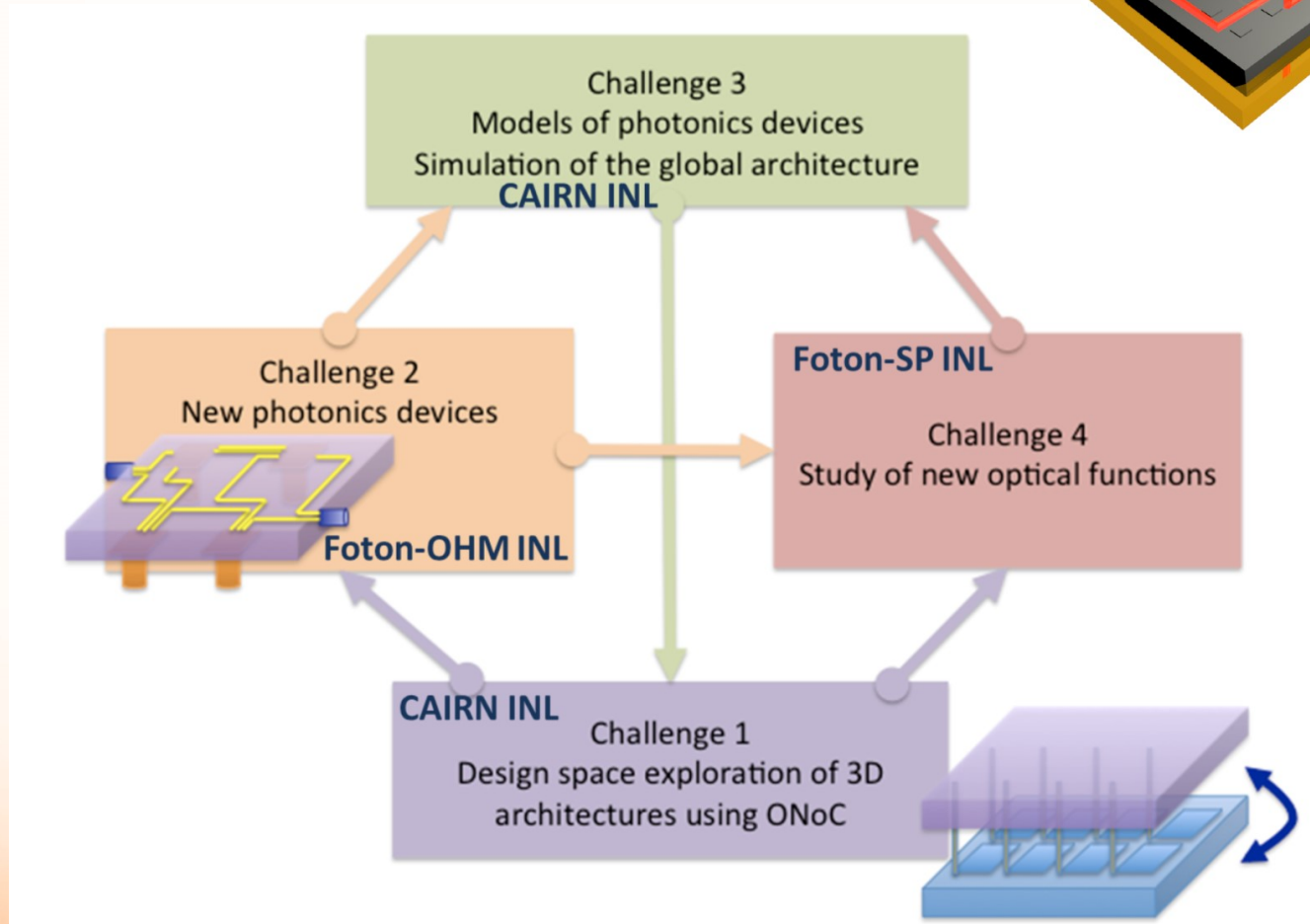
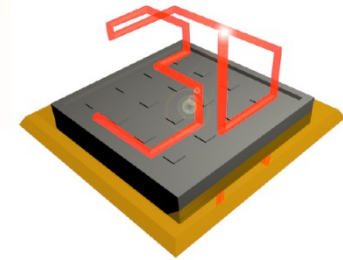
# Considered architecture



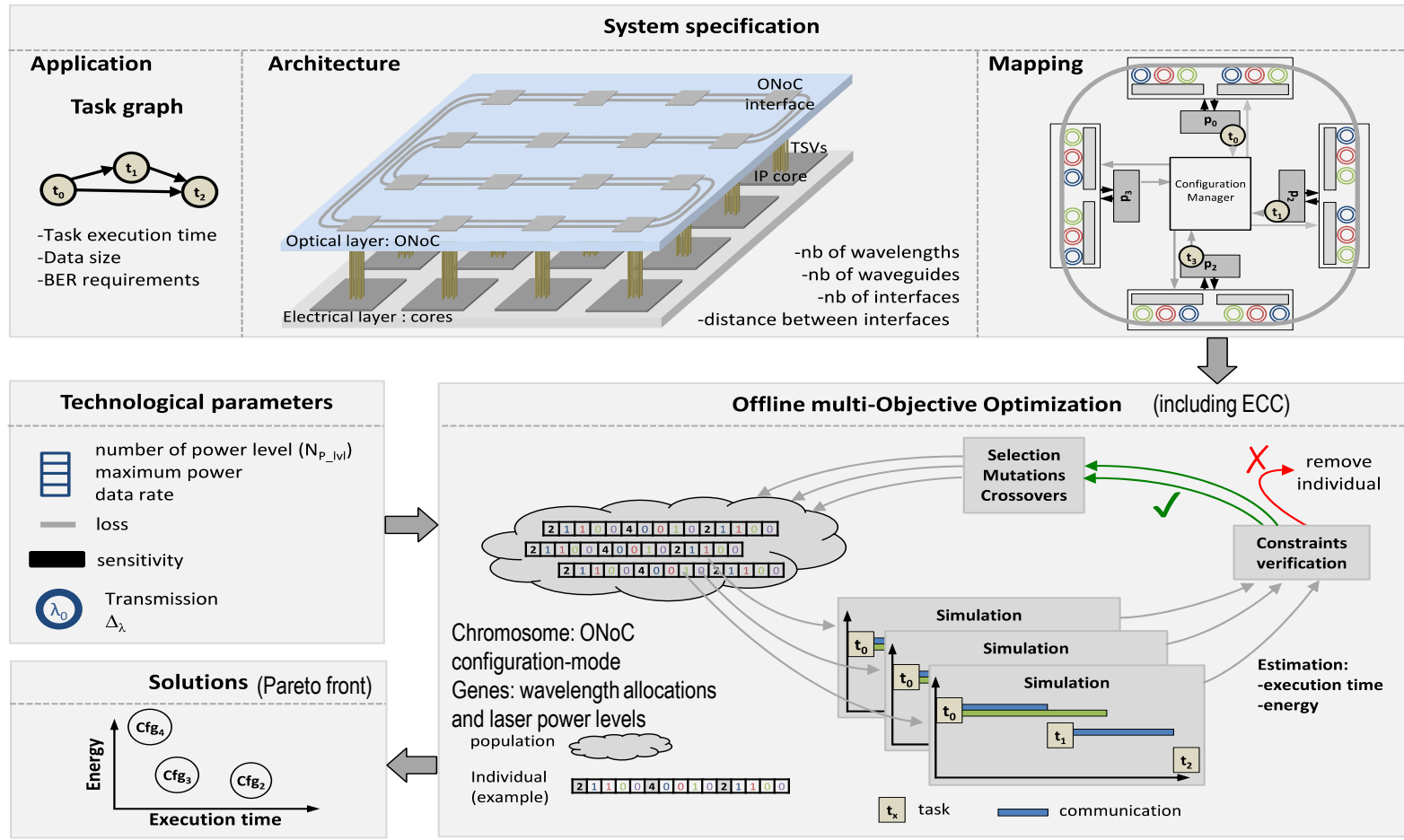
*Architecture Chameleon : Le Beux (INL) et al, DATE 2014*

- Optical NOC
  - Dedicated communication layer
  - On-chip lasers
  - Link between core and ONoC through ONI

# General organisation



# Global simulation of the computing architecture





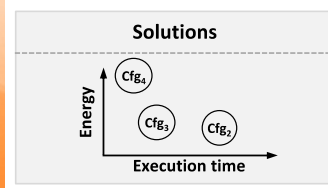
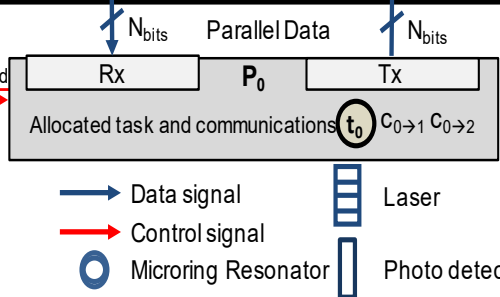
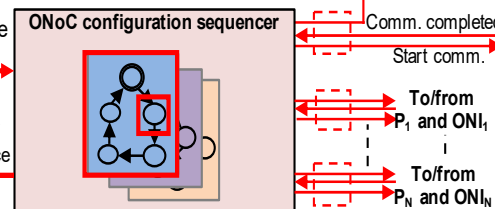
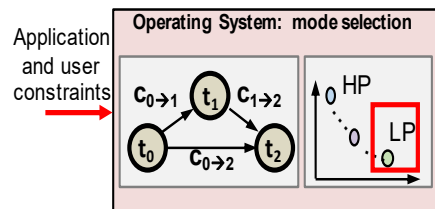
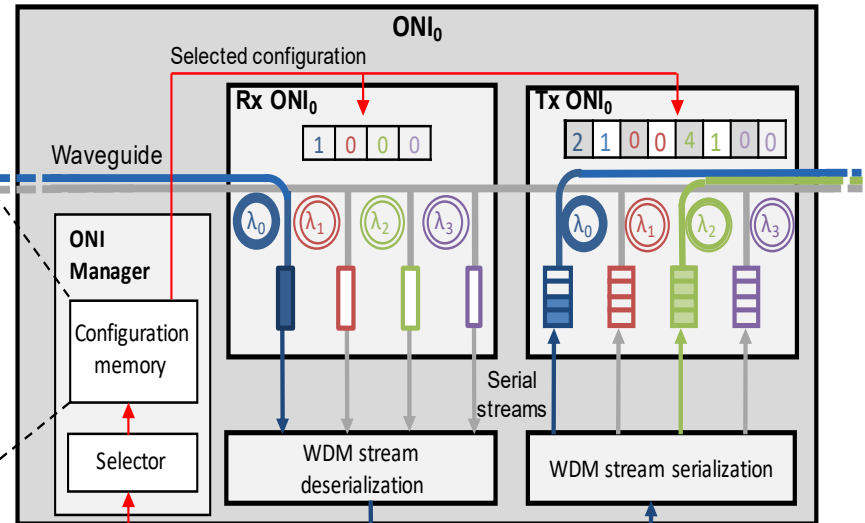
# Design space exploration and Offline analysis

allows the reuse of wavelengths

Optimized optical communication-configuration

MR state is '1' for ON or '0' for OFF ;  $P_{W_i}$  is the selected laser power level

		ONI configurations memory											
Mode	Step	Rx				Tx							
		$\lambda_0$	$\lambda_1$	$\lambda_2$	$\lambda_3$	$\lambda_0$	$\lambda_1$	$\lambda_2$	$\lambda_3$	$P_{W_0}$	$P_{W_1}$	$P_{W_2}$	$P_{W_3}$
High Performance	$Q_{init}$	-	-	-	-	-	-	-	-	-	-	-	-
	$Q_A$	1	1	-	-	4	1	4	1	4	1	4	1
	$Q_B$	-	-	-	-	-	-	-	4	1	4	1	-
	$Q_C$	-	-	-	-	-	-	-	-	4	1	4	1
	$Q_D$	-	-	-	-	-	-	-	-	-	-	-	-
Balance	$Q_{init}$	-	-	-	-	-	-	-	-	-	-	-	-
	$Q_A$	1	1	-	-	2	1	2	1	4	1	-	-
	$Q_B$	-	-	-	-	-	-	-	-	4	1	-	-
	$Q_C$	-	-	-	-	-	-	-	-	4	1	-	-
	$Q_D$	-	-	-	-	-	-	-	-	-	-	-	-
Low Power	$Q_{init}$	-	-	-	-	-	-	-	-	-	-	-	-
	$Q_A$	1	-	-	-	2	1	-	-	4	1	-	-
	$Q_B$	-	-	-	-	-	-	-	-	4	1	-	-
	$Q_C$	-	-	-	-	-	-	-	-	4	1	-	-
	$Q_D$	-	-	-	-	-	-	-	-	-	-	-	-

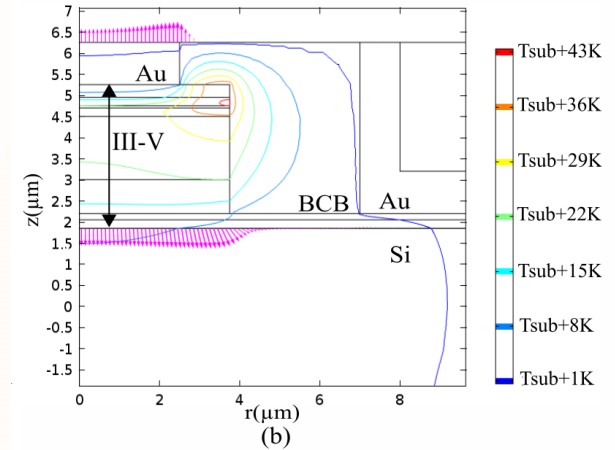


# Physical laser model

## Thermal management in lasers on-chip

Comparison of pseudomorphic and monolithic integration of lasers

Improved thermal dissipation for monolithic integration

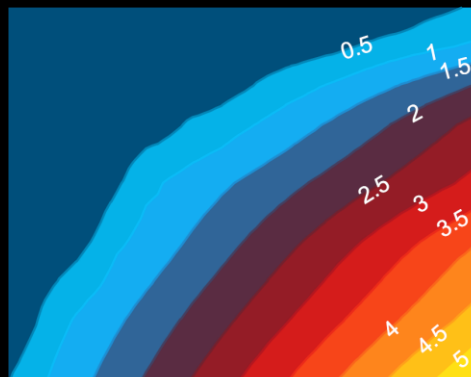


## Scalable laser model for output power and WPE

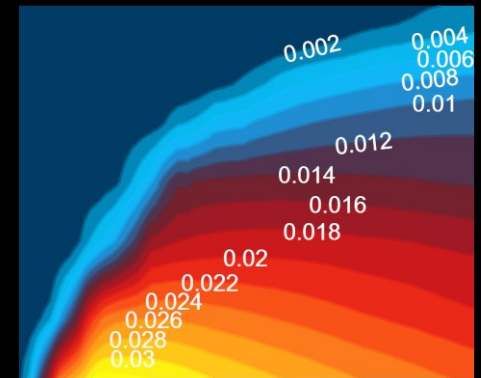
Highly sensitive to environment temperature

WPE still low for lasers on chip

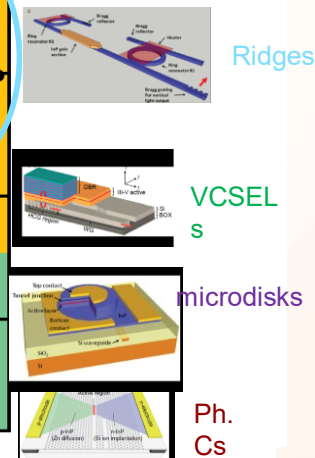
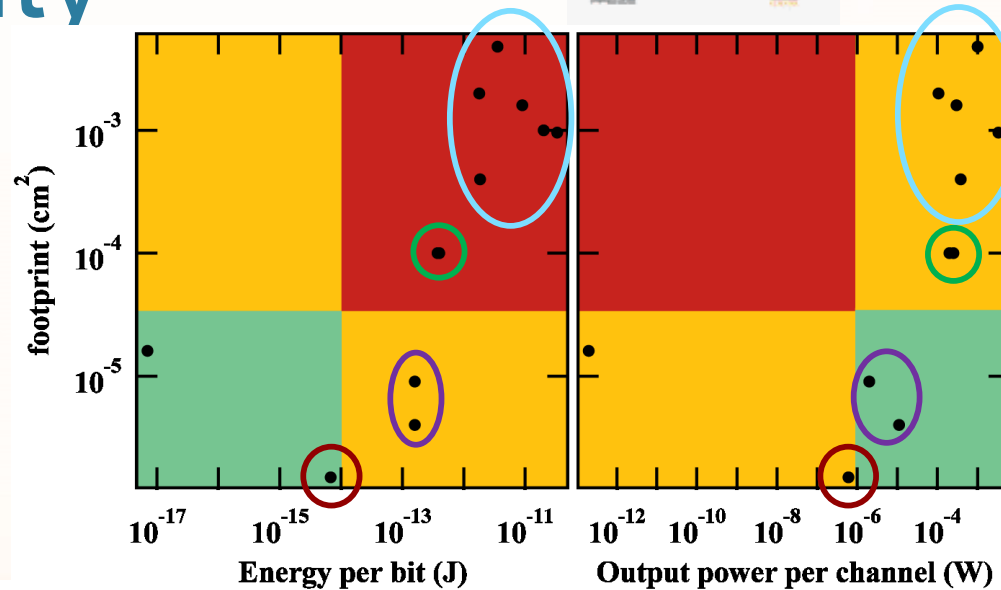
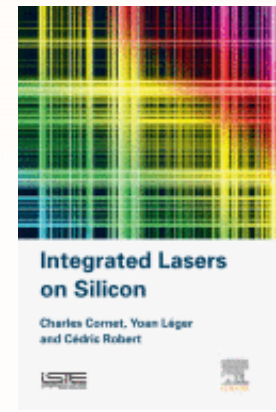
Output power (mW)



Wall plug efficiency



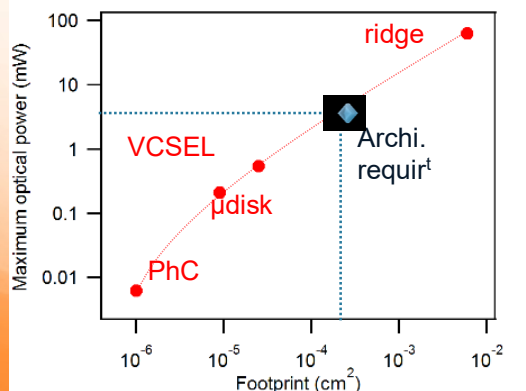
# Optical Sources for ONOC and their sustainability



## Integration criteria:

Grading laser solutions for 3D integration

Footprint and OPC calculated from Chameleon data.  
S. Le Beux *et al.*, ATD Europe (2014)

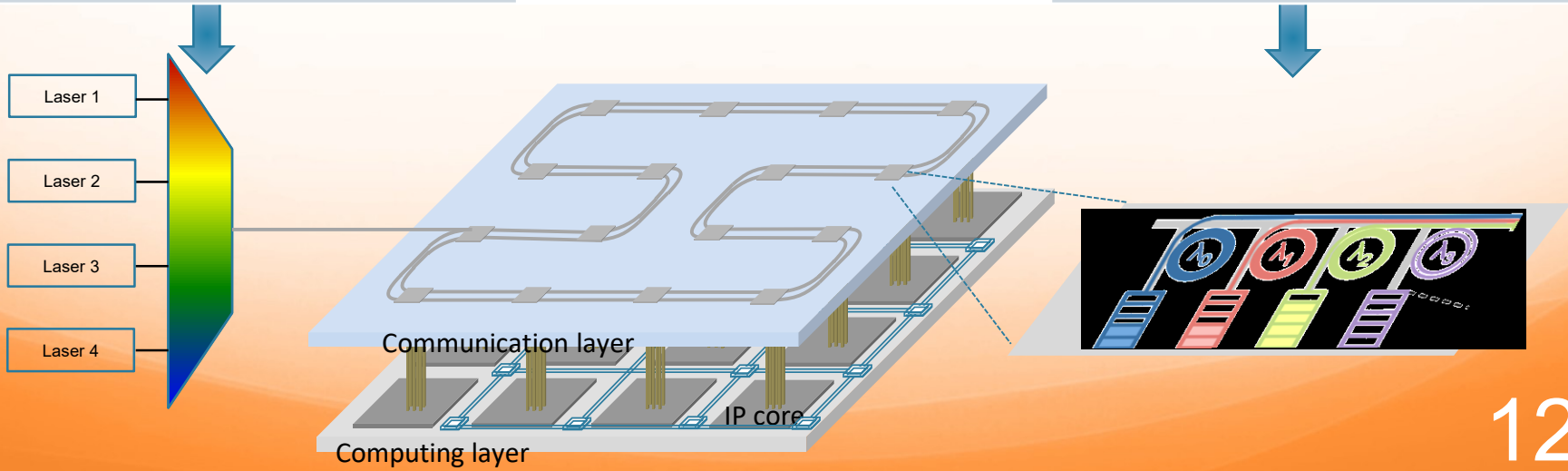


## Scaling law:

- design/performance empirical law
- Required output power defines the architecture and footprint of the device
- Application specific design

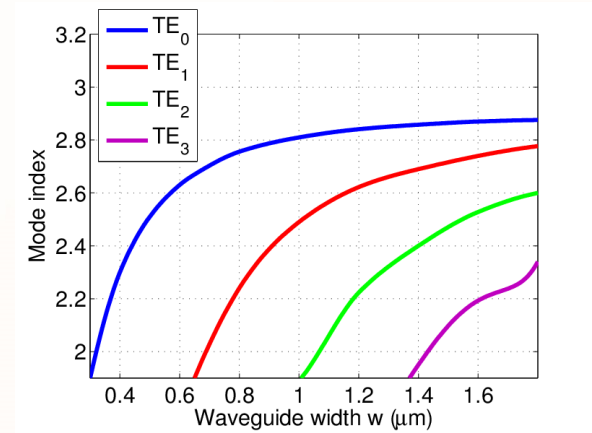
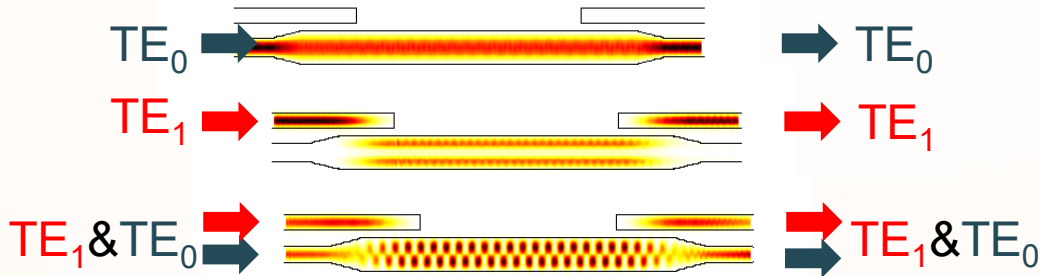
# Off-chip versus on-chip lasers

Off-chip		On-Chip
Laser always on	← Power management →	On-off management to save power
Always full power, higher power than on-chip lasers		Power level on demand
Must take into account of worst case	← Quality of services for → communication	On demande BER - SNR
Coarse grain management	← Level of management →	Fine grain management
Mature components/devices	← Availability of components →	Several improvement over the years
More efficient (Output power vs injected current)	← Efficiency →	Less efficient (10~15%)
Control is easier	← Temperature →	Sensitive

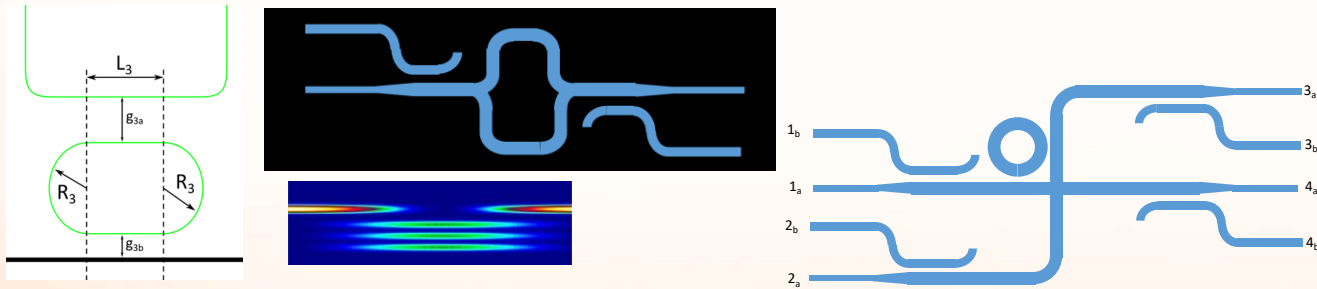


# Optical functionalities

- On-chip mode division multiplexing (MDM)



- Study of building blocks for MDM

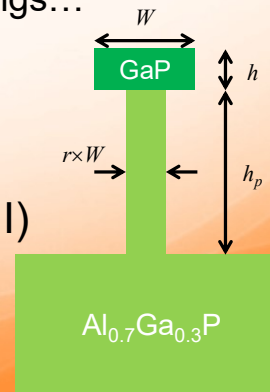


- Waveguide crossings
- Mode converters
- Micro-rings...

- MDM demonstrators

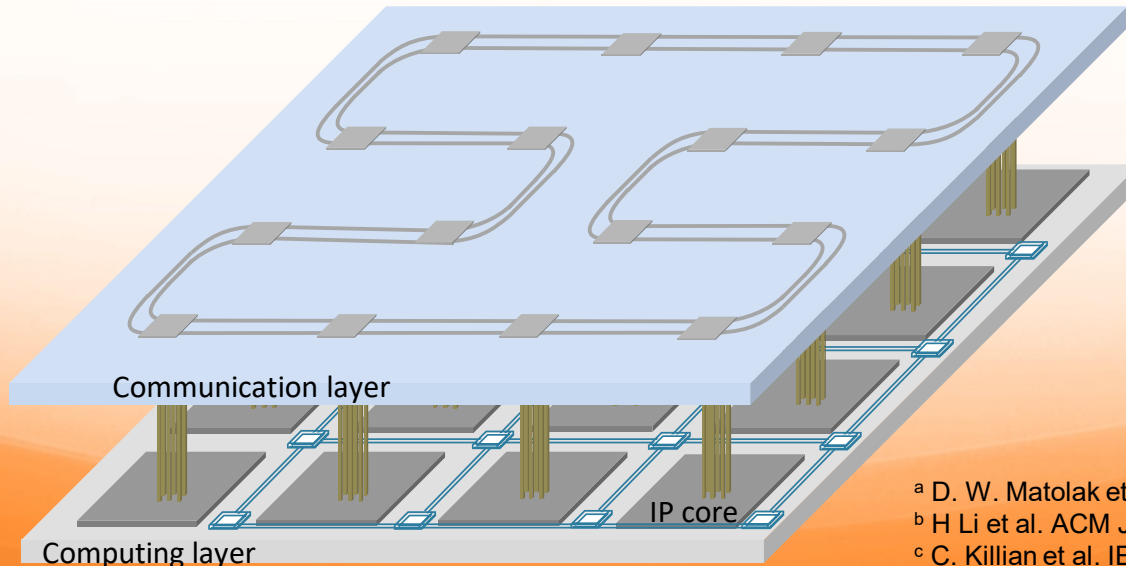
- Polymer waveguide building blocks (collaboration FOTON OGC, Univ. Ferrara, I)
- SOI ring ONoC (collaboration DTU Fotonik, DK)

- GaP nonlinear waveguide design for signal processing applications



# Electrical versus Optical NoC

Classic Electrical NoC		Optical NoC
Packet switching	← Switching →	Circuit switching
Linked to nb of hops Sensitive to network utilization	← Latency →	Independent from distance between core
64 Gbit/s (64 bits @ 1Ghz) per direction Up to 128 Gbit/s	← Bandwidth →	80 Gbits/s (8 wavelengths @ 10 Gbit/s) Up to 160 Gbit/s
0.12 pJ/bit/router + 0.18 pJ/bit/mm of wire <sup>a</sup>	← Energy →	Worst case is 1.2 pJ/bit for laser <sup>b</sup> + 0.7 pJ/bit TxRx <sup>c</sup>
Router = ~245*245μm <sup>2</sup> on 28nm FDSOI	← Area →	Optical component interface = ~22*192μm <sup>2</sup> Tx/Rx = ~50*50μm <sup>2</sup> each



<sup>a</sup> D. W. Matolak et al. IEEE Wireless Communication 2012

<sup>b</sup> H Li et al. ACM JETC 2017

<sup>c</sup> C. Killian et al. IEEE/ACM DAC 2017

# Conclusion and perspective

- Strategy for the physical layer (Model of communications losses integrated realistic photonic devices, ONI, optical framework for design space exploration)
- 3 Ph.D.
- 6 reviewed papers, 25 international conferences, 5 national events
- 1 book, 1 book chapter

# Conclusion and perspective

- 1 Ph.D. started in Dec. 2018
- 1 ANR
- 2 ANR JCJC
- collaboration with Ghent university, ULB, DTU, Univ.Ferrara
- 1 Ph.D. mixing Multimode waveguide and quantum key distribution
- Common project Institut Foton INRIA (CPER)