

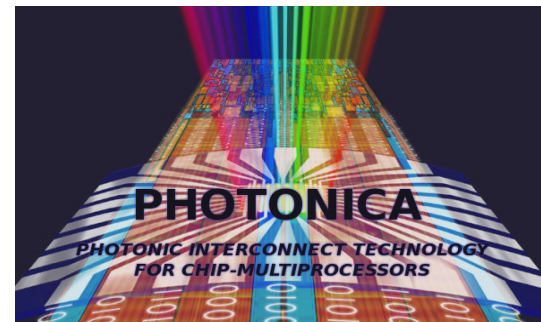
Meet in the Middle: Leveraging Optical Interconnection Opportunities in Chip Multi Processors

Sandro Bartolini*

Department of Information Engineering , University of Siena, Italy

bartolini@dii.unisi.it

OPTICS Workshop, Grenoble, 13/3/2015

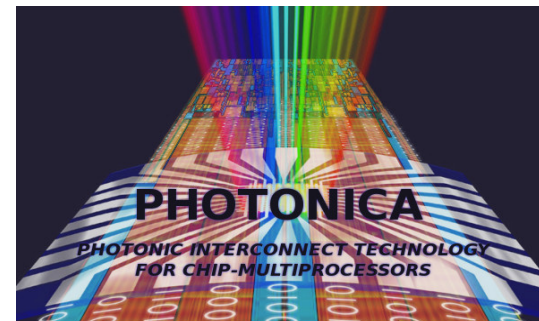


Ministero dell'Istruzione, dell'Università e della Ricerca

Acknowledgement: Paolo Grani, Davide Bertozzi, Luca Ramini, ...

Outline

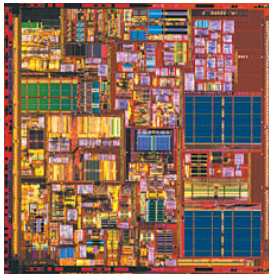
- Introduction
- Ring-based optical interconnection tradeoffs
- Fast path-setup for switched optical networks
- Software restructuring for matching ONoC features
- Conclusions



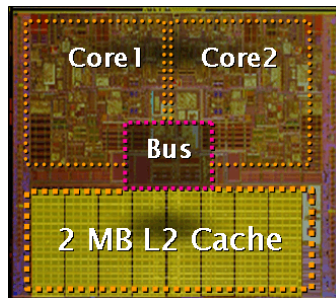
Ministero dell'Istruzione, dell'Università e della Ricerca

Introduction and motivation – processors (1)

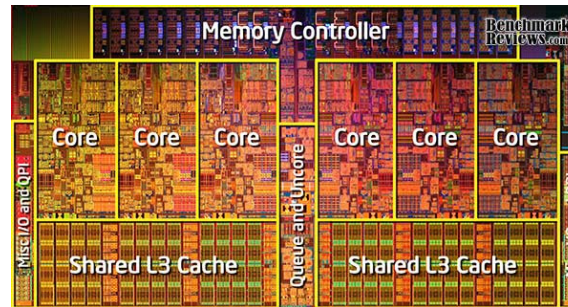
- Nowadays processors are parallel
 - Biggest reason was the emerging of wire-delay issues ... i.e. on-chip latency



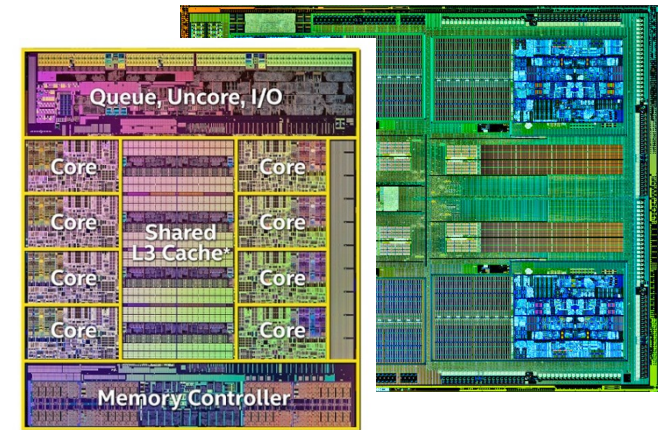
Pentium 4 (1)



CoreDuo (2)

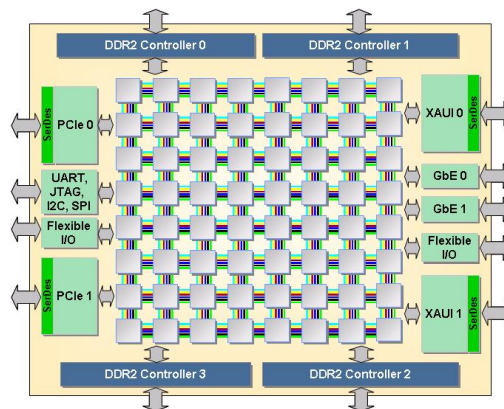


i7- 980X (6)



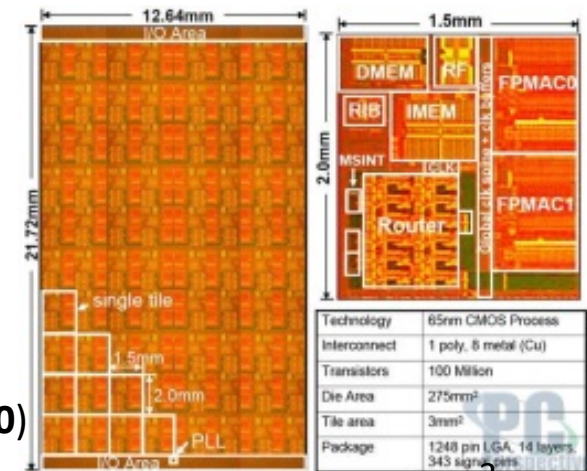
i7-5960X / AMD FX8370 (8)

- Beyond about 10 → tiled design



Tiler Tile64 (64)

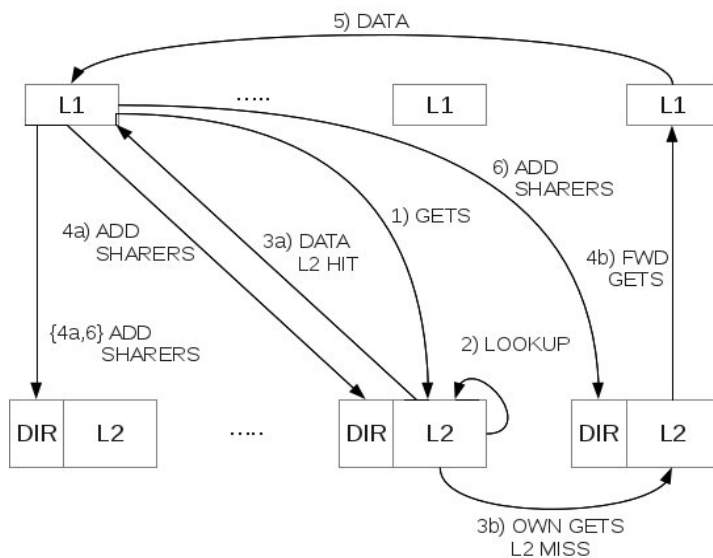
Intel Polaris (80)



Introduction and motivation – processors (2)

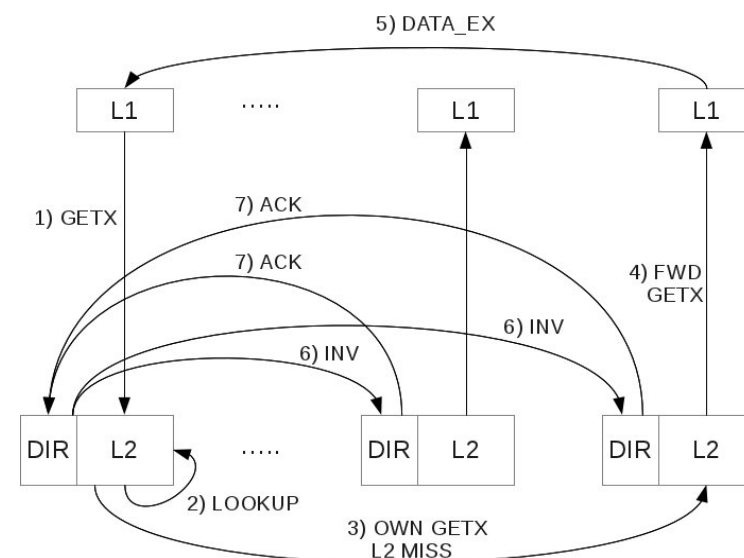
“... when processor A wants to talk to processor B ...”

- Shared memory model is here to stay ... some more
 - At least within *core clusters*
 - Ease of programmability
 - Scalable directory-based coherence
 - Numerous message exchanged for each load/store
 - E.g. 80 % *control* (8 byte) and only 20% *data* (64 byte)



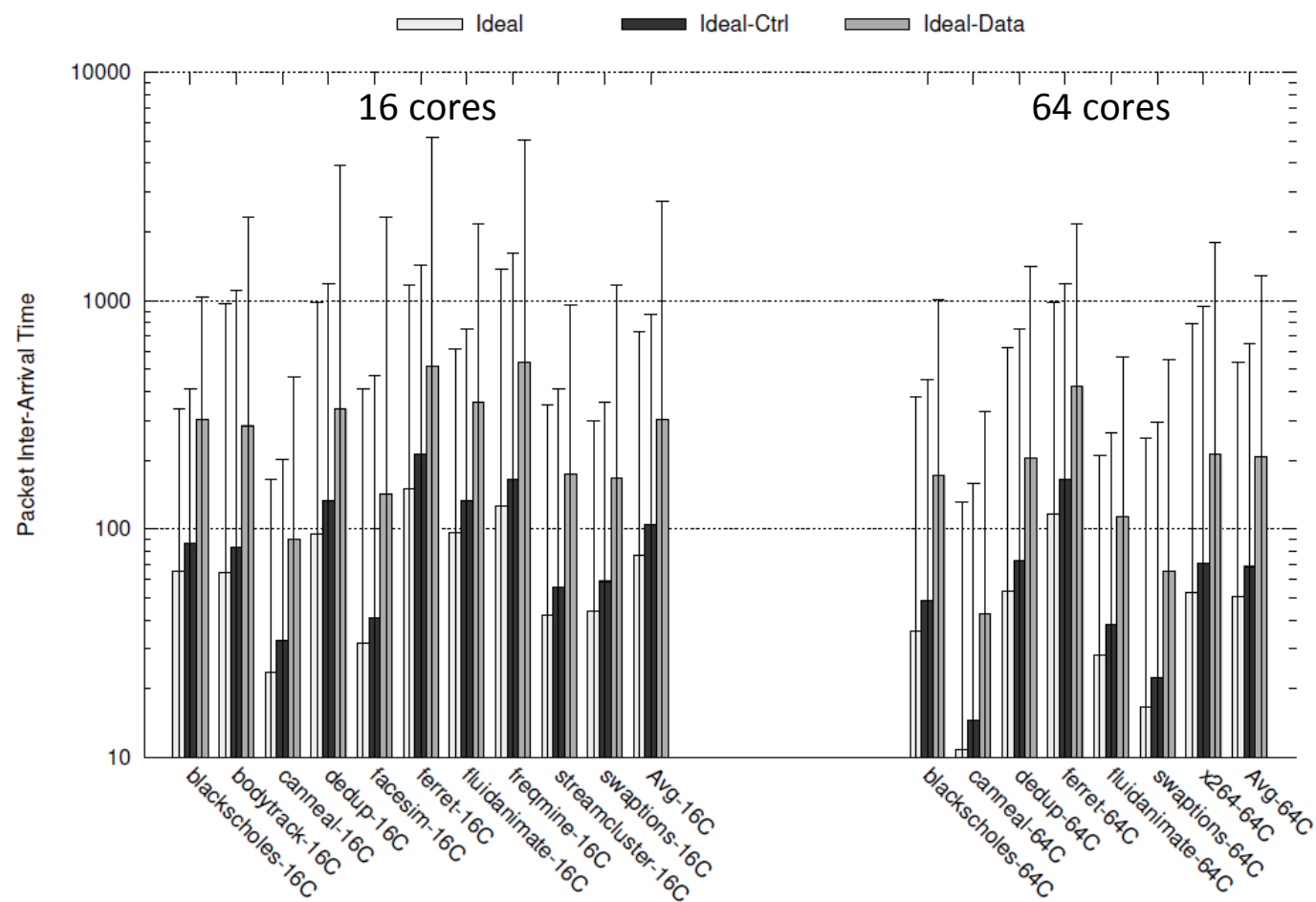
read miss

MOESI



write miss

Introduction and motivation – coherence traffic

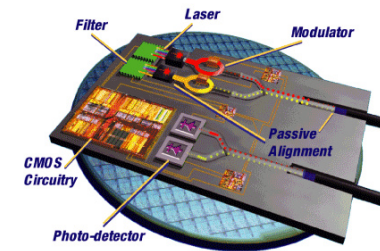


Total and control- (Ctrl) or data- (Data) only stats for ideal network

- Low average load, bursty → message latency is critical

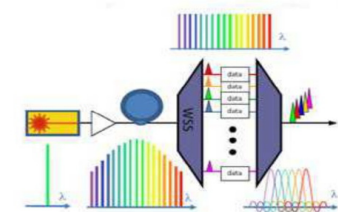
Introduction and motivation – Integrated photonics (1)

Optical communication technology can now be integrated in CMOS process



- Pros:

- Fast propagation (16 ps/mm)
 - 10x less latency compared to electronics (e.g. 22nm [1])
 - E.g. Couple of cycles @3 GHz to cross a 2cm chip corner to corner
- Compatible with CMOS fabrication
- High-bandwidth: 10-40 GHz frequency and WDM (*order of 1 Tbps*)
- End-to-end: energy consumption almost insensitive to distance

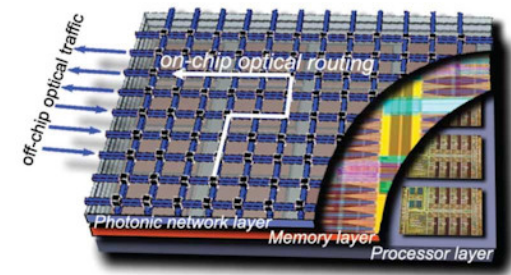


- Cons:

- End-to-end ... no store-and-forward → Throw away a lot of knowhow
- Active components (lasers, photodetectors) have some integration problems
 - Can have or induce significant static power consumption

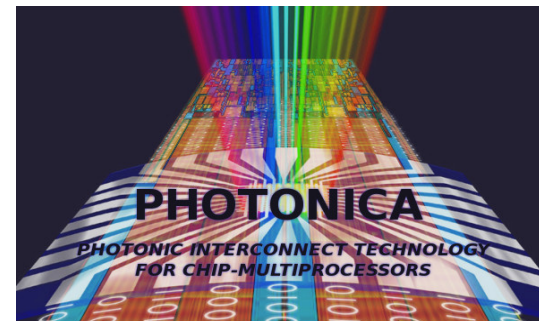
Introduction and motivation – Integrated photonics (2)

- Integrated photonics is still in its infancy in serving computer systems local requirements
 - Traffic close to cores is very different from the aggregated traffic at wider scale (e.g. blade/rack/datacenter)
- Layered design is not yet consolidated enough
 - Application-, architecture-, network-level requirements and choices are not orthogonal to optical design choices
 - Like: topology, access schemes, resource provisioning, DWDM, technological choices
 - Interactions can induce very different performance/consumption in the optical network
- Need for an integrated multi-layer approach
 - Effective designs
 - Exploration and consolidation of best practices



Outline

- Introduction
- Ring-based optical interconnection tradeoffs
- Fast path-setup for switched optical networks
- Software restructuring for matching ONoC features
- Conclusions

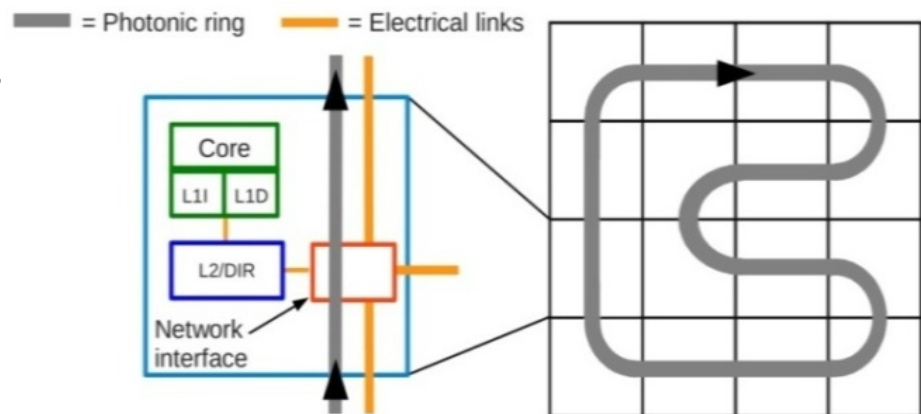


Ministero dell'Istruzione, dell'Università e della Ricerca

Ring-based ONoC tradeoffs

- Hybrid Electronic-Optical on-chip networks or all-optical
 - Optical network based on *ring* logical topology
 - Simplicity of ring topology can be a good reference design point
- We analyze the relationships between optical resource provisioning (waveguides) and core number, versus:
 - Traffic quote offloaded to the optical network
 - (One ring) Only read-requests, invalidations and invalidation acks
 - (Multiple rings) All traffic
 - MWSR, MWMR access schemes

Considering performance and power metrics



[Grani, Bartolini, "A Simple On-chip Optical Interconnection for Improving Performance of Coherency Traffic in CMPs ", DSD, 2012]

[Grani, Bartolini, "Design Options for Optical Ring Interconnect in Future Client Devices", ACM JETC, 2014]

Simulation Environment and Methodology

- Gem5 simulator in Full-System mode (Linux 2.6.27 booted in gem5), 8/16/32/64 core running Parsec 2.1 multi-threaded benchmark suite
 - Multi-threaded applications → We forced core affinity on the application execution to avoid non determinism due to OS scheduling

Architectural parameters

Cores	8/16/32 cores (64 bit), 4 GHz
L1 caches	16 kB (I) + 16 kB (D), 2-way, 1 cycle hit time
L2 cache	16 MB, 8-way, shared and distributed 8/16/32 banks, 3/12 cycles tag/tag+data
Directory	MOESI protocol, 8/16/32 slices, 3 cycles
ENoC	2D-Mesh/Torus, 4 GHz, 4/5 cycles/hop, 32 nm, 1 V, 64/128 bit/flit
Optical Ring	3D-stacked, 1-9 parallel waveguides, 30 mm length, 8/16/32 I/O ports, 10 GHz, 64/70 (16 and 32) wavelengths, 460 ps full round
Main memory	4 GB, 300 cycles

Results: multiple rings and all traffic

- No electrical NoC : searching for good design points
 - From low bandwidth up to 8/9 64-wavelength rings

Bandwidth per endpoint is too low (1/2/4 bits) and each message transmission suffers avalanche conflicts and long serialization

Best tradeoffs:

8-core

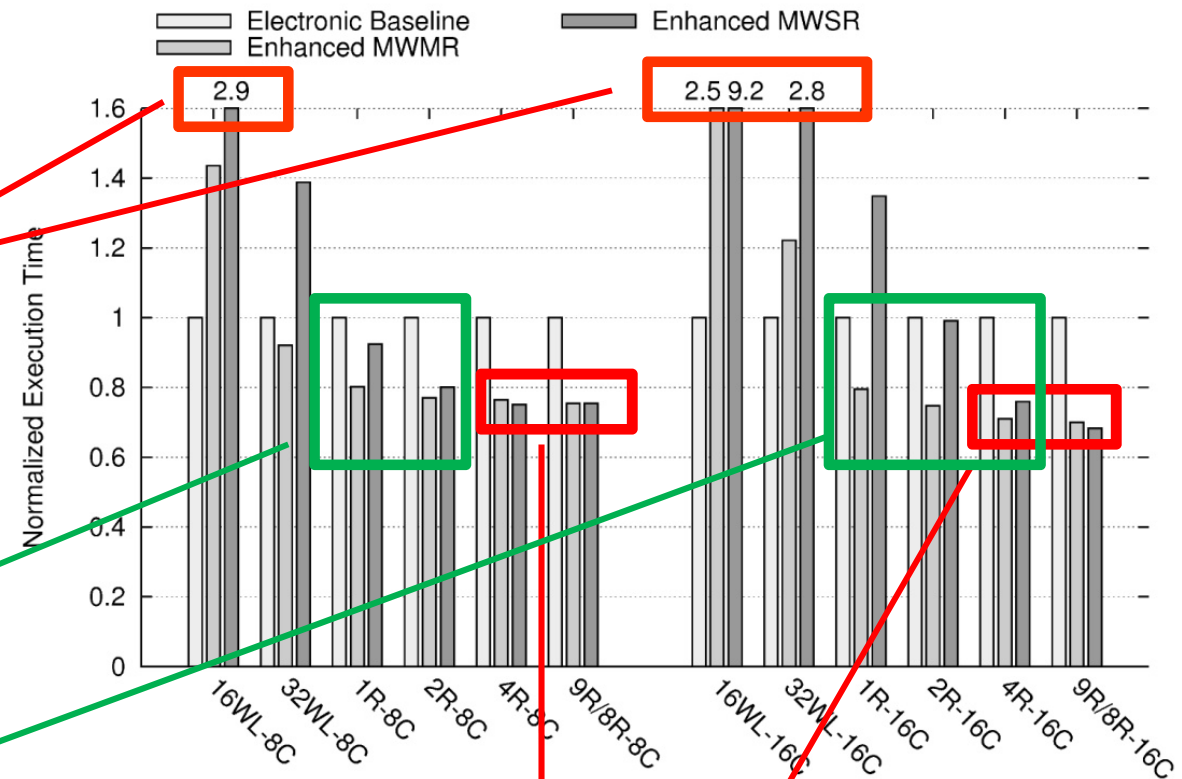
MWMM, 1-ring

MWSR, 2-ring

16-core

MWMM 1- or 2-ring

MWSR 4-ring



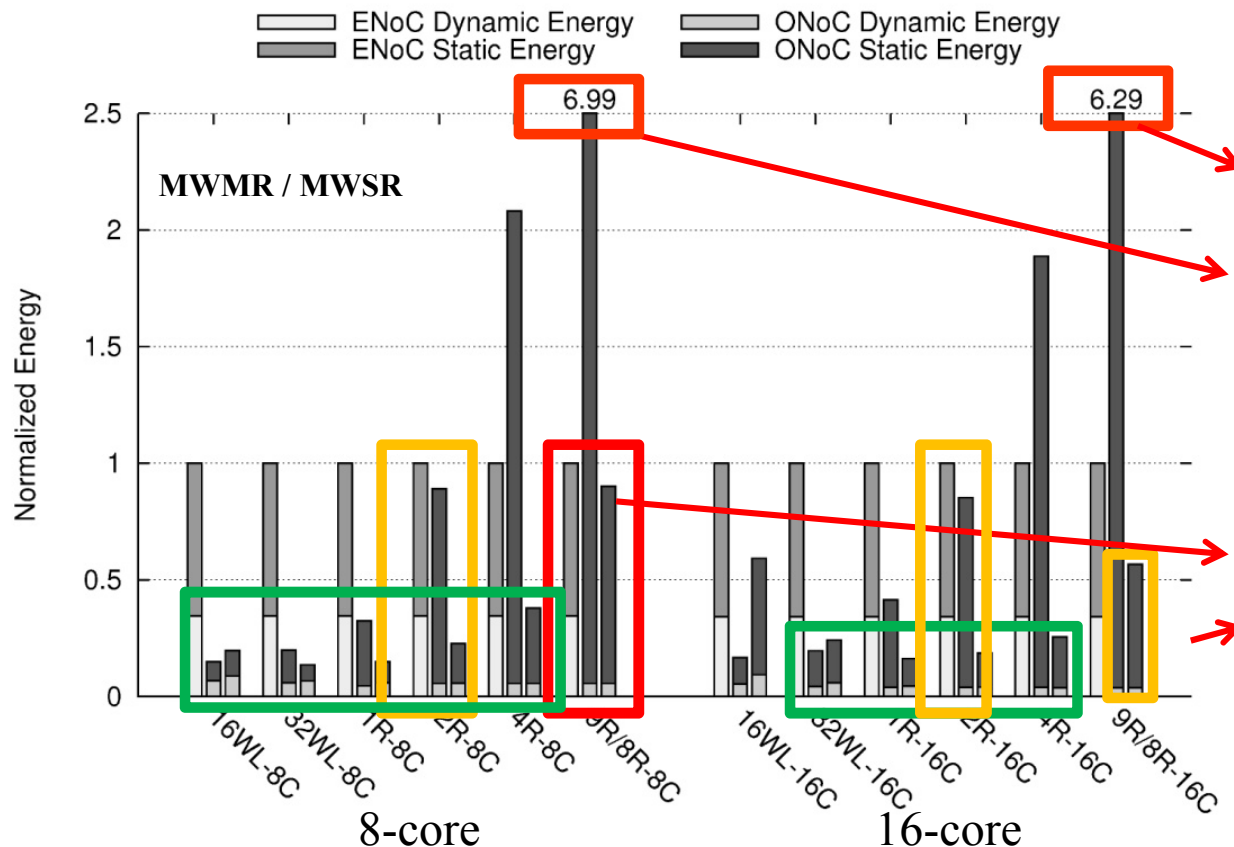
8-core

16-core

Saturation → Waste of optical resources

Results: energy

- Sensitivity to the number of photonics rings for MWMR and MWSR
 - **Ring number increase** → Overall NoC optical power increase, MRR increase, IL increase (crossing, splitting, ...), increased laser power
 - *But some topology assumptions can make the difference ...*

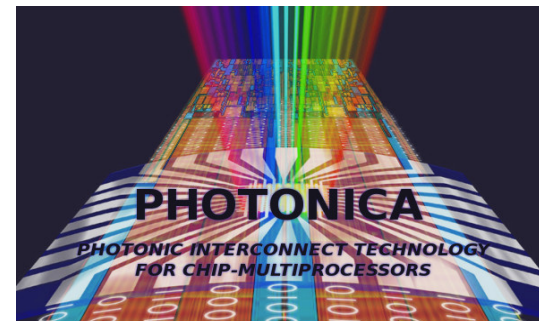


- Extremely high static power for MWMR with multiple rings.
- Static consumption of the MWSR ring increases for 8-ring also due to token waveguide

• **Overall MWSR is the best choice for all-traffic support**

Outline

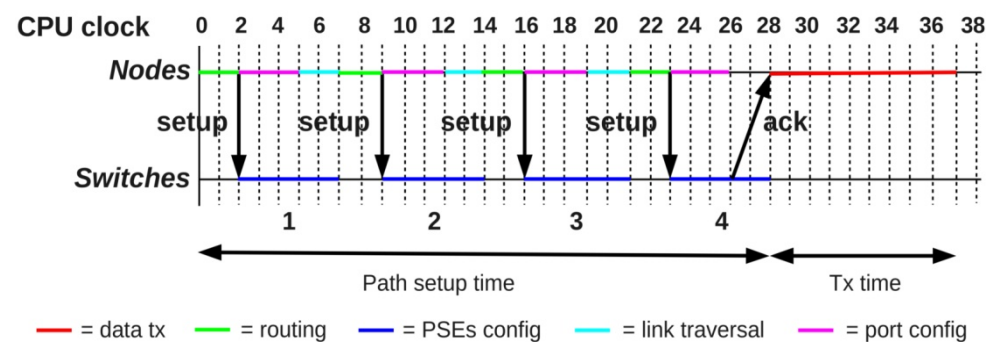
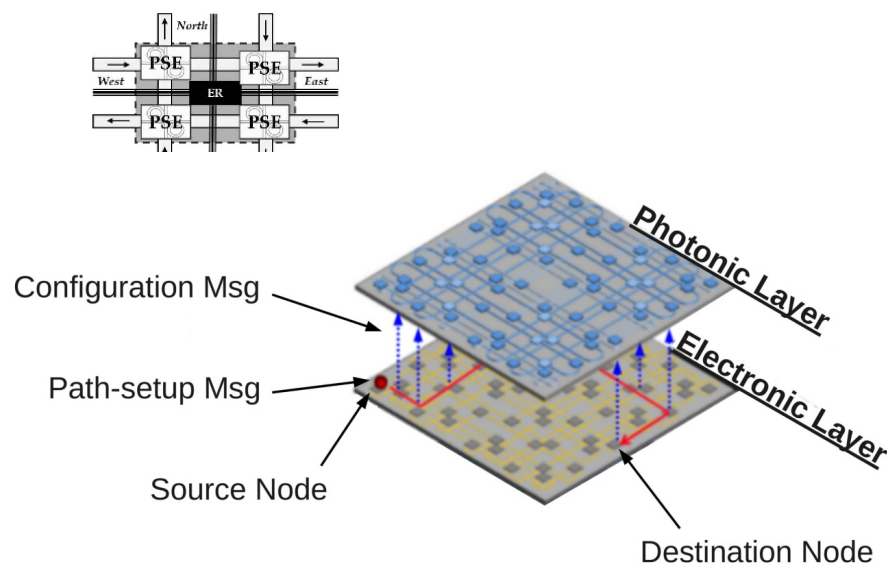
- Introduction
- Ring-based optical interconnection tradeoffs
- Fast path-setup for switched optical networks
- Software restructuring for matching ONoC features
- Conclusions



Ministero dell'Istruzione, dell'Università e della Ricerca

Fast path-setup for switched optical networks

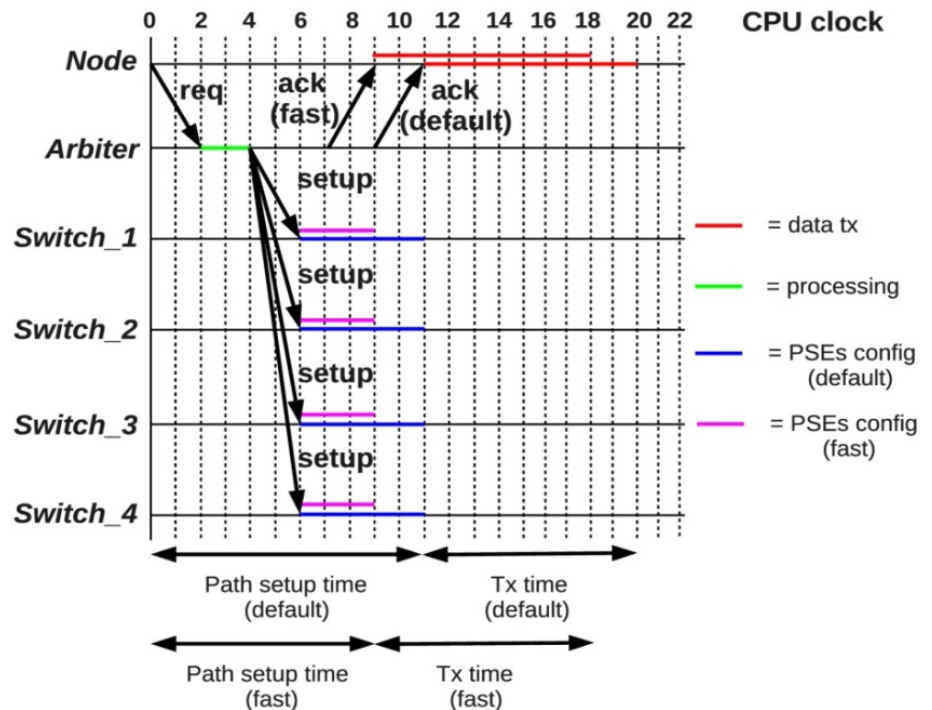
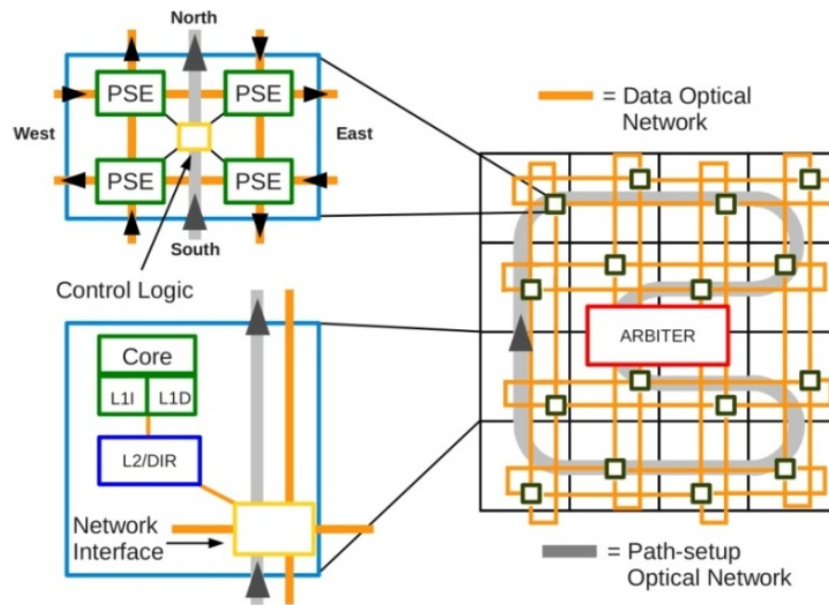
- Switched optical networks can provide
 - Potentially higher scalability than ring (crossbar)-based approaches
 - Require broadband optical switches
 - **But suffer from sequential path-setup time**
 - High overhead for high endpoint number and for “small” message sizes
 - “small” can mean less than 1000s bytes !
 - **Coherence traffic is out of game**



Serial path-setup

Fast path-setup for switched optical networks

- We propose a centralized arbiter that can simultaneously configure the required optical switches through a wavelength-routed optical-ring
 - Network from cores to arbiter is optical ring-based (wavelength-routed)
 - Decoupling topologies of path-setup network and data network



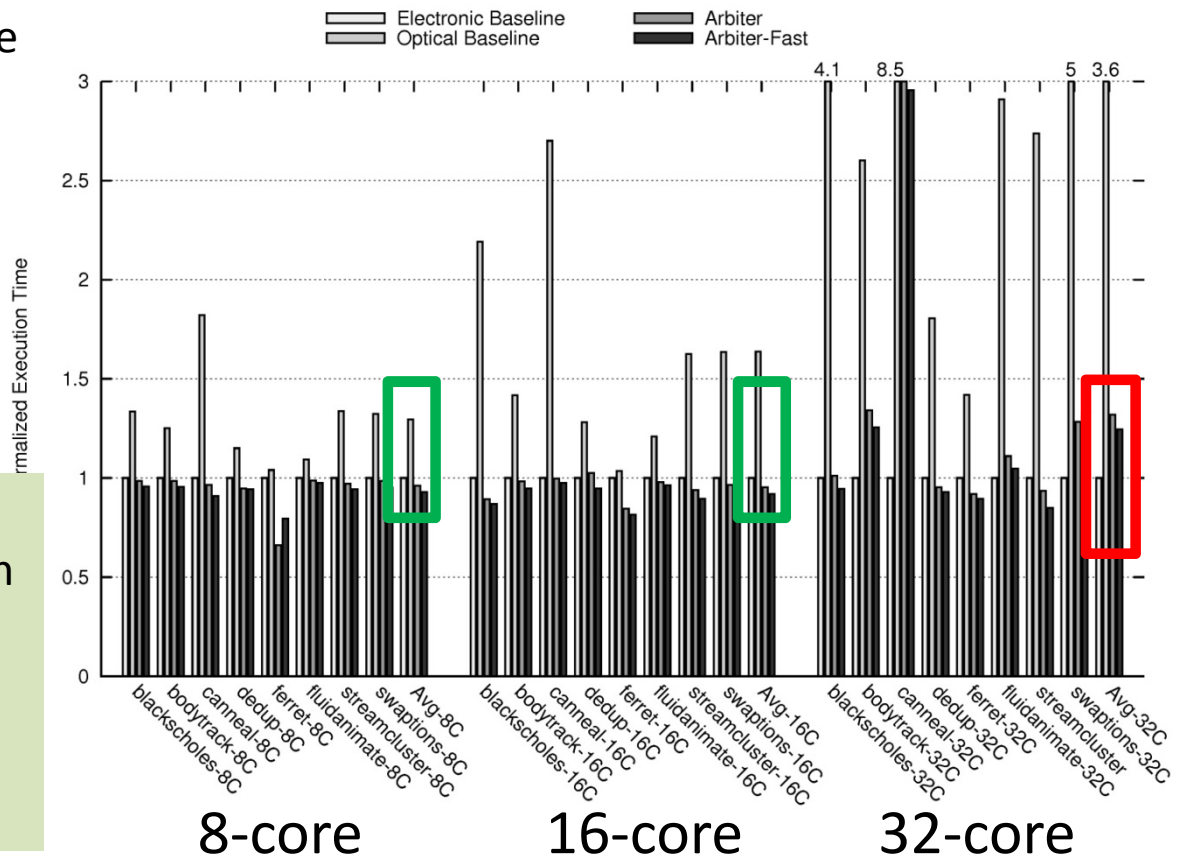
Results single arbiter

Setup	Serial PS [cycles]	Simultaneous PS [cycles]
8-core-AVG	51.26	25.94
16-core-AVG	70.37	27.45
32-core-AVG	156.18	65.60

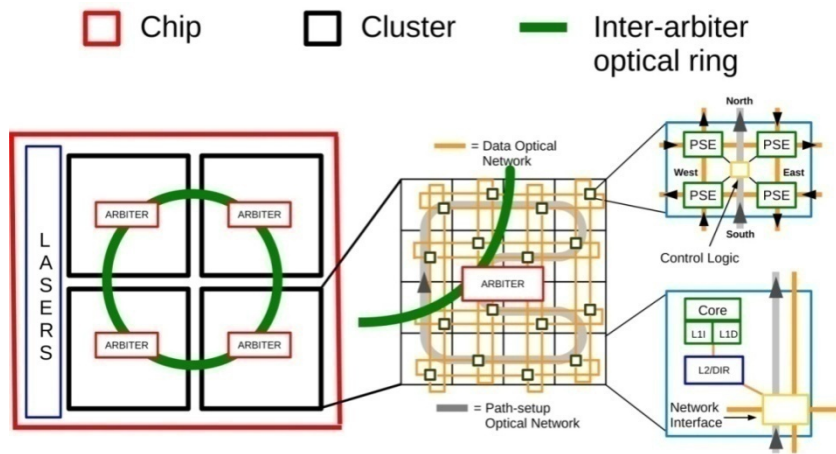
Path-setup latency: dramatic average reduction even if conceptually the arbiter serializes path-setup requests

- Arbiter works well 8-and 16-core setups
- For 32-core case arbiter induce 25% slowdown due to path-setup serialization

- In all cases arbiter performs much better than the serial path setup
 - **Serial PS prevents** cache coherent traffic to work
 - **Arbiter support this traffic**

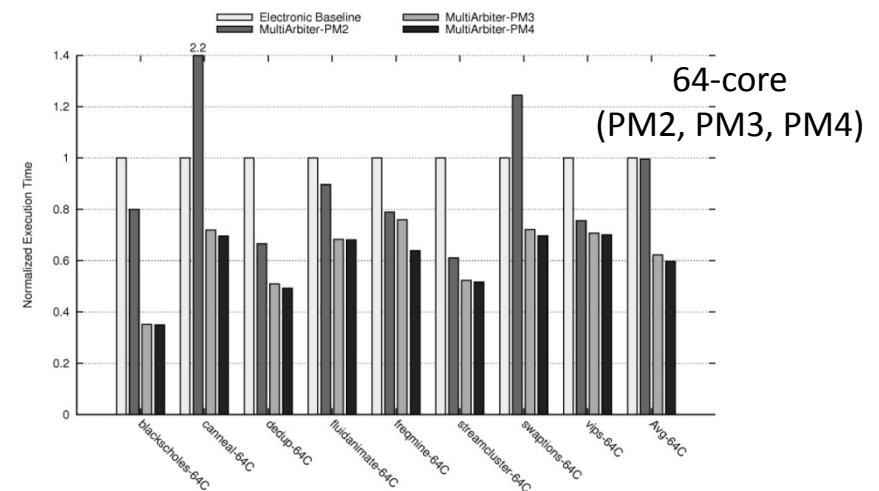
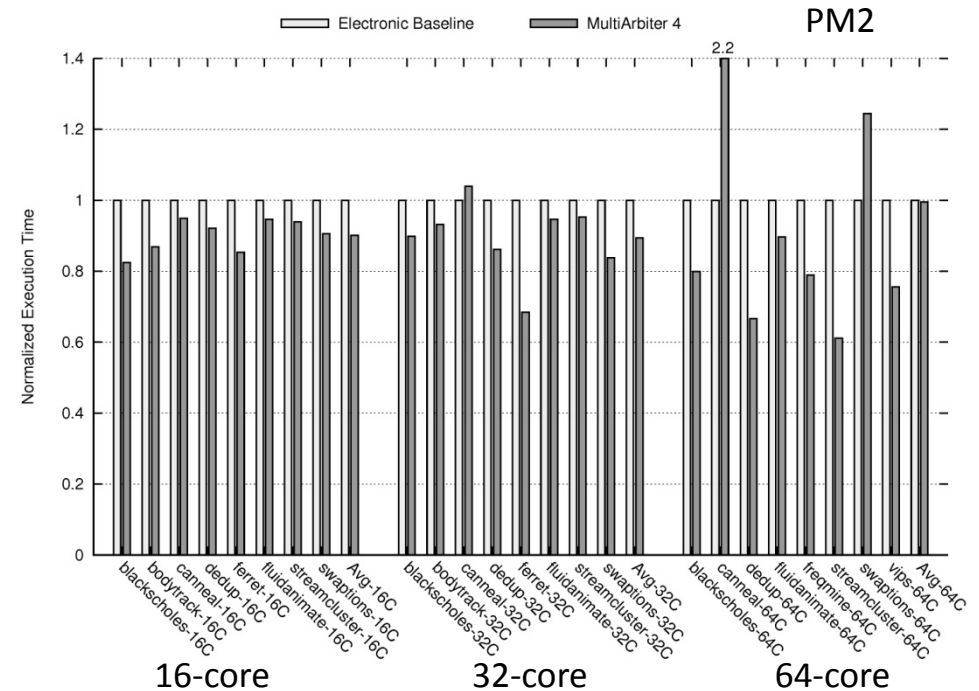


Scalable fast-setup: multiple arbiter



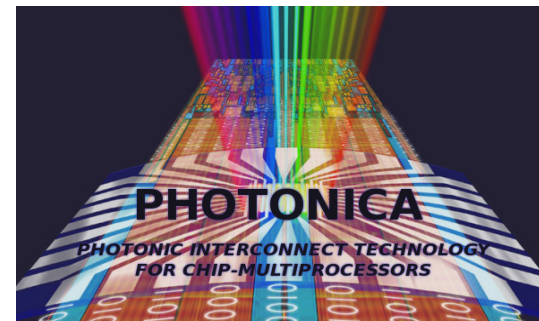
Multi-Arbiter 4-Clusters

- One arbiter per cluster, optical ring between clusters
- Independent path-setups within clusters
- Coordinated inter-cluster path-setup
- 2-level MOESI to increase in-cluster paths



Outline

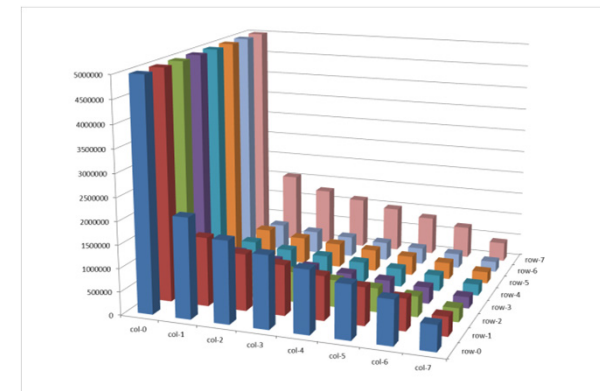
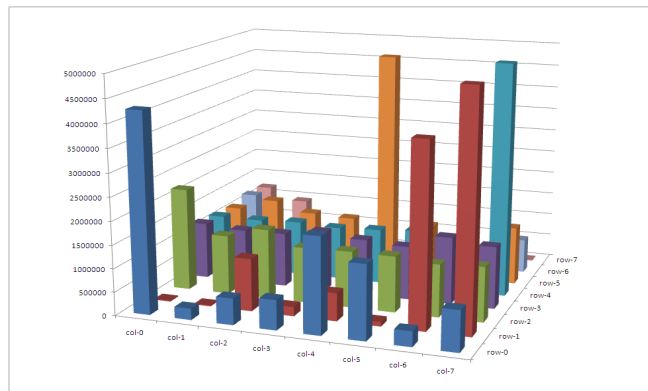
- Introduction
- Ring-based optical interconnection tradeoffs
- Fast path-setup for switched optical networks
- Software restructuring for matching ONoC features
- Conclusions



Ministero dell'Istruzione, dell'Università e della Ricerca

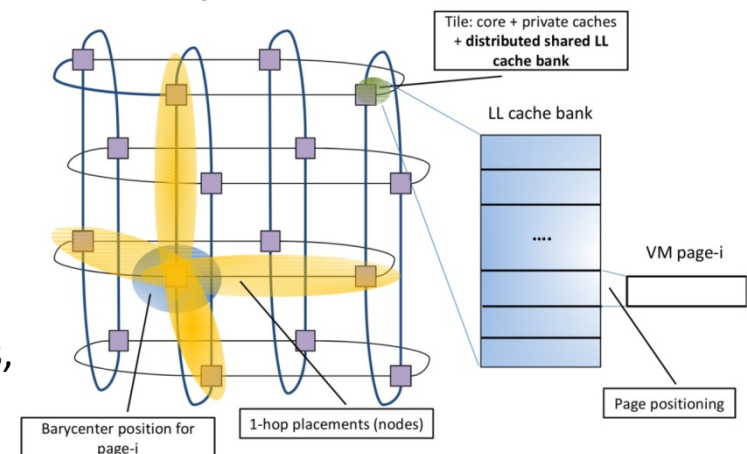
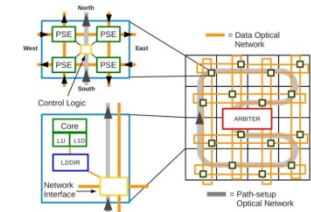
Software restructuring for matching ONoC features

- On-chip optical networks can potentially improve wire-delay issues in tiled CMPs
 - Distant cores (i.e. many “hops” away) can be actually reached in a few cycles
- Access time to distributed cache resources is **more-uniform**
 - Not Uniform Cache Access (NUCA) architectures (e.g. in tiled CMPs)
- Software restructuring techniques for locality typically aim at putting data close to usage (cores) as to reduce access time in NUCA caches
 - Delicate balance between ideal access barycenter from source cores and conflicts (misses) penalties in zones that need to be congested for low NoC access overhead

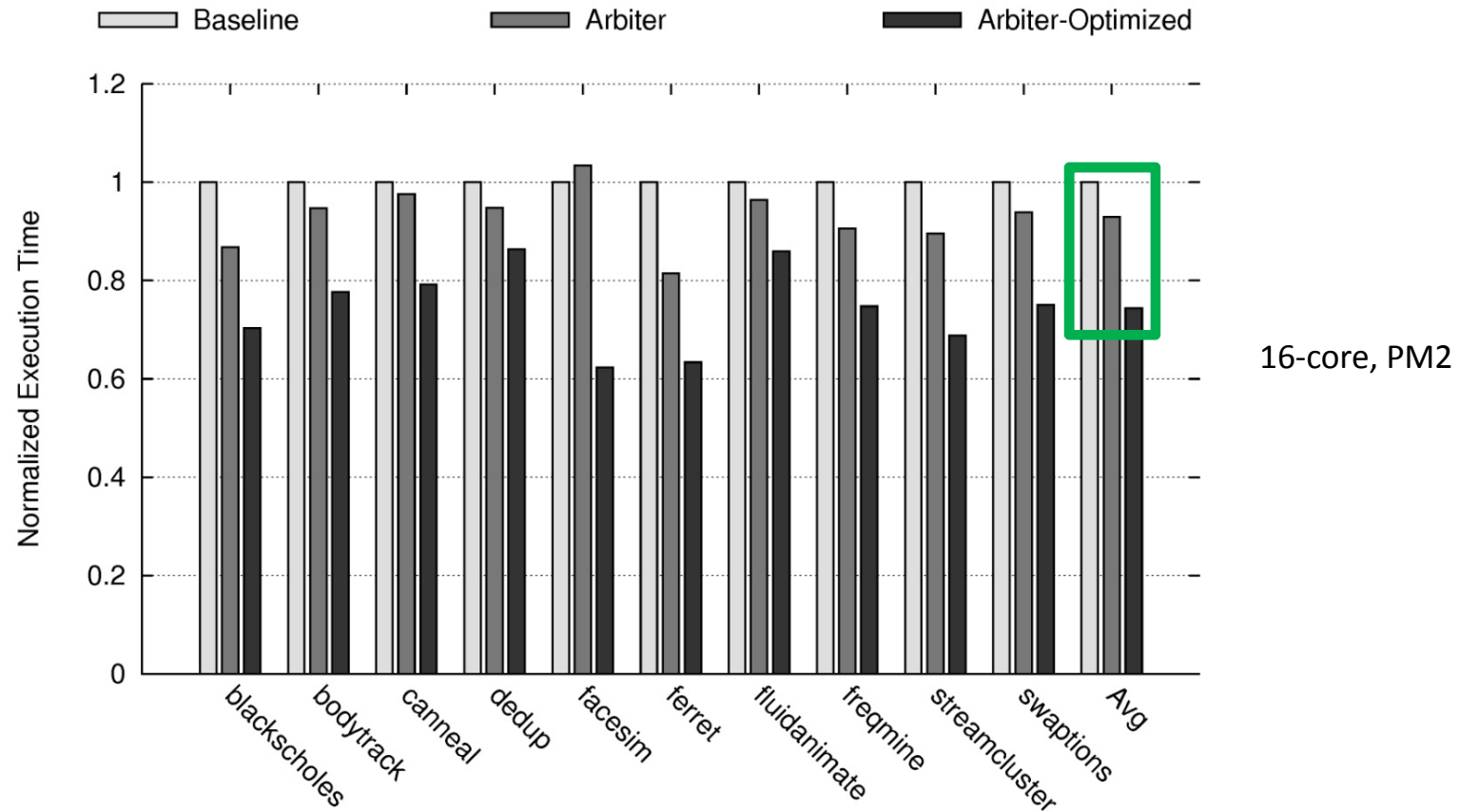


Software restructuring for matching ONoC features (2)

- With ONoCs on average we can afford to use far more cache
 - Spread data much more in the chip to reduce conflicts (misses) as distance overhead is almost constant
 - *Hyp: using our arbitrated optical switched network*
- Not exactly straightforward though:
 - Actually with switched networks, conflicts arise not only for data placement but for message paths ... and, specifically, sub-paths !
 - Need to
 - **Spread data to gain from reduced conflict misses**
 - **But (!) not too far from barycenter otherwise average path-length increases and path-setup conflict probability increases (big overhead !)**
 - Each VM page is positioned in the access barycenter first
 - Then it is tried around a “radius” of H hops, looking for the minimum of a cost-access function
 - Considering: misses, path-length and conflicts, cache, memory and ONoC access times



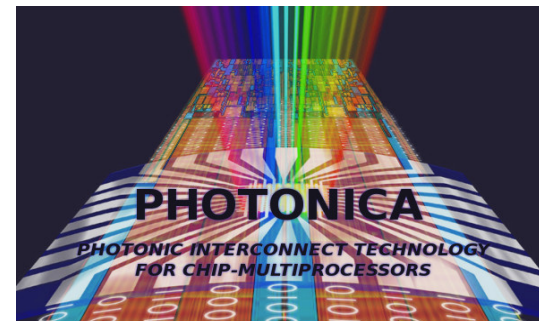
Results: Software restructuring for switched ONoC



- Software restructuring allows gaining 19% more speedup over the standard arbiter solution (7% over the Baseline)
 - Electronic baseline cannot benefit from this software restructuring due to heavy Non-uniform access time to tiles

Outline

- Introduction
- Ring-based optical interconnection tradeoffs
- Fast path-setup for switched optical networks
- Software restructuring for matching ONoC features
- Conclusions



Ministero dell'Istruzione, dell'Università e della Ricerca

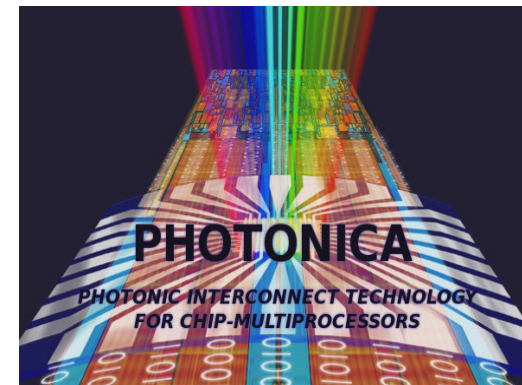
Conclusions

- Integrated optics is a breakthrough technology that brings a number of positive facets
 - Improvements in devices will make it even better perspective
- Its technological discontinuity needs to be integrated with patience and with a thick vertical design approach in modern processor and computer design
 - To master and exploit the various two-way inter-related effects that are nowadays present between layers
 - significant risk of sub-optimal designs ... if not even worse than some well-tuned electronic solutions
 - For reaching effective designs
 - Opportunities and constraints of different layers must ... meet in the middle ☺

**Meet in the Middle:
Leveraging Optical Interconnection Opportunities
in Chip Multi Processors**

Thanks for your attention!

Q & A



Ministero dell'Istruzione, dell'Università e della Ricerca