TECHNICAL UNIVERSITY OF MUNICH

Chair of Electronic Design Automation

Faculty of Electrical Engineering and Information Technology

INTERNSHIP REPORT

Distributing Bit-Level Parallelism based on different communication requirements on Wavelength-Routed Optical Networks-on-Chips

Author: Subarnaduti Paul

Advisor: Dr.-Ing. Tsun-Ming Tseng

Supervisor: Prof. Dr.-Ing. Ulf Schlichtmann

Table of Contents

ACKNOWLEDGMENT

First of all, I wish to express my deep gratitude to Prof. Dr.-Ing. Ulf Schlichtmann, Head of Chair Electronic Design Automation, and to Dr.-Ing. Tsun-Ming Tseng, Research and teaching assistant at the Chair of Electronic Design Automation, for allowing me to do this internship.

I sincerely thank Dr.-Ing. Tsun-Ming Tseng, for his guidance and encouragement in carrying out this work. Without his kind direction and proper guidance, this internship would have been a little success. I choose this moment to acknowledge his contribution gratefully.

I am also grateful to my friends who have helped me while preparing this report by giving their suggestions and supply of information, which were valuable to me.

1. Introduction

1.1 Background

Network-on-Chip (NoC) is an alluring innovation to expand the data transmission of the present multi-core frameworks. With the shrinking size of device dimensions in nanoscale technologies, the performance of electrical interconnects of the global communication network is degrading relative to the gate performances. With technology scaling, the quantity of processor, memory, and equipment on a chip are expanding. This has resulted in increased computation and communication complexity of the design, and scalable approaches are needed to design the system. NoCs have risen as the most feasible aspect for structuring adaptable correspondence engineering for Systems-on-Chips (SoCs). Conventional on-chip networks for SoCs face a few issues, for example, poor versatility, restricted transmission capacity, and low usage. NoCs use modern communication and networking theories to relieve these issues.

In NoCs, unlike the conventional non-versatile buses, on-chip small scale systems are utilized to interconnect the different cores. NoCs have better productivity and design consistency when contrasted with bus-based frameworks. Some of the most important phases in designing the NoCs are the synthesis of the topology or structure of the network and the setting of various design parameters (such as frequency of operation, link-width, etc). The standard topologies (mesh, torus, etc.) that have been utilized in full-scale systems come up with bad performance and have enormous power and area overhead when utilized for SoCs.

1.2 Optical Network-on-Chip

Optical Network-on-Chip (ONoC) is becoming a promising technology to overcome the bottleneck of low bandwidth of electronic NoC. ONoCs are based on on-chip optical interconnects and routers. The usage of optical interconnections instead of metal wires can make unprecedented communication bandwidth available while minimizing dynamic power dissipation. Photonic integrated circuits are emerging as a key segment for network-on-chip with high processing power, just as other developing innovations, for example, lab-on-chip sensors. With these key innovations, they have predominantly developed themselves as a keynote in various fields of application. But these applications require novel features, such as reconfigurable light pathways, acquired by appropriately tackling on-chip optical radiation. They can also be a promising candidate to overcome the limitations of traditional metallic-interconnect based NoCs.

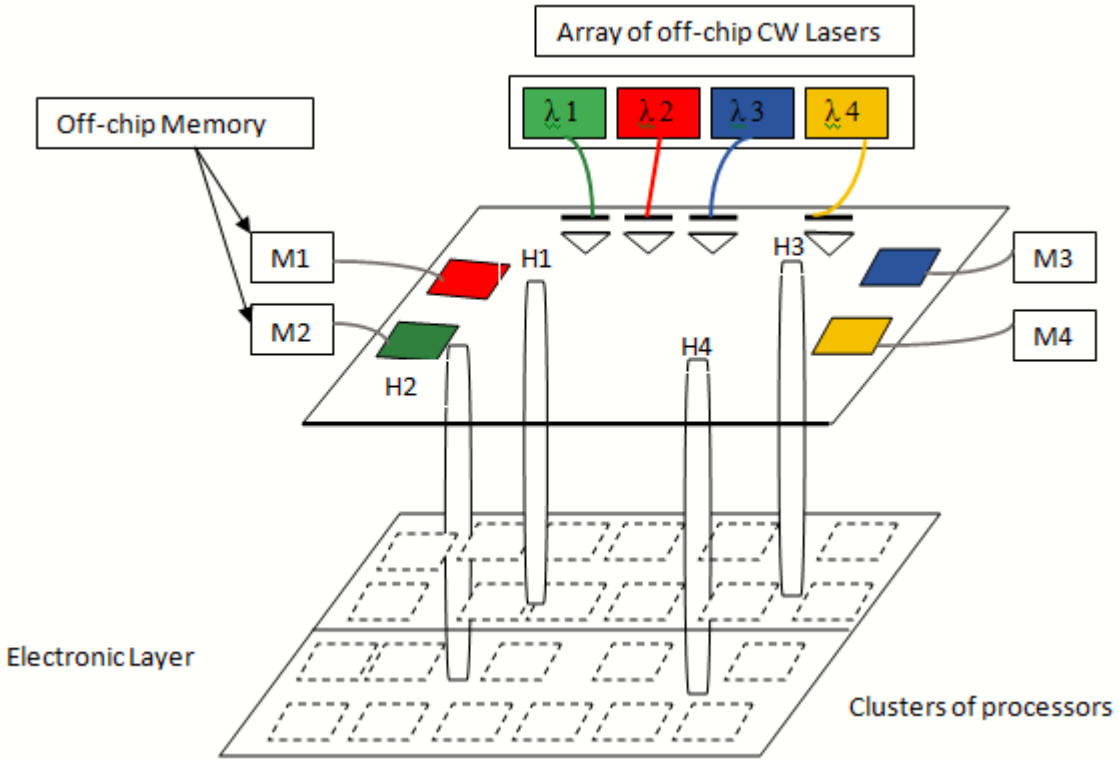


Fig.1 Optical Network-on-chip Architecture

The optical router forms the basis of an optical NoC. The progress in photonic technologies, especially the development of microresonators, makes optical on-chip routers feasible. Microresonators can be fabricated on silicon-on-insulator (SOI) [3] substrates, which have been used for CMOS-based high-performance low-leakage SoCs. They are also smaller in diameter. Microresonators are a good candidate for very large scale integrated optoelectronic circuits and can be used as an optical switch, optical add-drop multiplexer (OADM) [3], modulator, and optical sensor. Optical routers implement the routing and flow control functions, and switch packets from an input port to an output port using an optical switching fabric. The optical switching fabric is usually composed of multiple basic optical switching elements, which implement the basic 1x2, 2x1, or 2x2 switching functions. A control unit commands the optical switching fabric based on routing requests and a routing algorithm. The control unit can be built from CMOS transistors to process routing requests. For microresonator-based routers, the control unit uses electrical signals to control basic optical switching elements.

2. WRONOC

WRONOC stands for Wavelength-Routed Optical Network on Chip.

This work focuses on the wavelength-routed optical network on chips (WRONoCs), which belong to a promising family of ONoCs. As their name suggests, WRONoCs route optical signals based on their wavelengths. A dedicated signal path is predominantly reserved for each pair of communicating nodes in the network at the design time to make the initiator node transmit at any time, regardless of other on-chip communications in progress [9]. This is how WRONoCs tend to improve the latency and energy overhead for arbitration. WRONoCs can be a potential candidate for the applications based on future high-capacity wide-area networks. WRONoC's are particularly advantageous when there are no implementations of wavelength translation or switching in the intermediate wavelength routing networks which further improves the management overheads and thus establishing a single-hop logical topology. All optical channel paths are determined by several factors such as the position of the transmitting and receiving nodes, the wavelength transmitted and the configuration of the routers. It routes simple wavelength based optical paths on the channels, easing management and processing of the network in comparison to the digital cross-connect routing systems. The feasibility of WRONoC's is based on certain factors such as the number of wavelengths used to determine a local connection path and the demand of the traffic. Therefore, the primary focus of this report will be aimed at minimizing the network wavelength requirements as the wavelengths being a limited resource.

2.1 Challenges

To reduce the number of resources, there have been quite a few solutions concerning WRONoC topologies where the MRRs and wavelengths can be productively shared among various routing paths. In any case, two important challenges still exist before the adaptability concerns can be mitigated.

- First, the topology should be custom-fit to suit different correspondence situations. Standard WRONoC topologies are intended to ensure full availability, for example, every initiator is associated with each target. This over-traditionalist reservation prompts noteworthy assets to squander in application-specific frameworks where full connectivity isn't required.
- Second, while topologies just spotlight on the rational schemes, organizing components, as such, MRRs and waveguides should be set and steered in physical plane thinking about the constraints of the layer. Specifically, shorter waveguides with fewer bendings and intersections are ideal for limiting the laser power. Because of the enormous design space, handcrafting the format is exceptionally wasteful and the nature of the outcomes is unusual.

Design automation for WRONoCs has emerged as an aid to diminish design trouble and to upgrade the quality of the design. During the last decade, wonderful progress has been made in both topological and physical synthesis. By systematically investigating the design space, networks can be consequently customized, and a significant decrease in usage of resources and laser force can be accomplished. Despite the inspiring advances, design automation for WRONoCs is still in its earliest stage. There are significant execution factors to be investigated, and possible drawbacks in the design flow to be crossed over. In this project, we will quickly review the design aspects of WRONoCs, and propose a design combining both of the active and passive types of routers. I will be presenting a router design that will be capable of supporting bit-level parallelism. In accordance with this, I also need to address the potential barriers concerning this design proposal and will try to find suitable outcomes in order to achieve the goal.

3. Description of Internship

3.1 Motivation

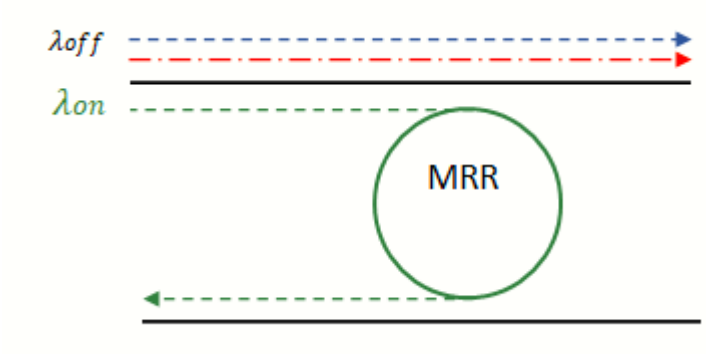
So far most of the WRONoC topology designs assume that each initiator/target pair communicates with one bit at a time using one wavelength. This means that each initiator/target pair is connected through a single communication path. However, MRRs have transmission characteristic that allows them to support multiple resonant wavelengths. Thus, the communication parallelism of a signal path can be increased by assigning multiple wavelengths between each pair. This report proposes different router architectures that support multiple signal paths for a single initiator/target pair. The idea behind it is that by assigning multiple signal paths for a single initiator/target pair, we can increase the amount of information being transmitted between a specific pair whenever it is needed. However, the achievable parallelism in the communication path not only relies on the device parameters but also depends on the topology. As many NoC systems have various requirements in terms of bandwidth, customization of the topology will significantly enhance the application spectrum of WRONoCs. An extensive study of the trade-off needs has to be done besides enhancing the bit-level parallelism (using more wavelengths for a single/initiator pair) which in turn increases power overhead and crosstalk noise, to formulate the feasibility of the designs.

3.2 Working Mechanism

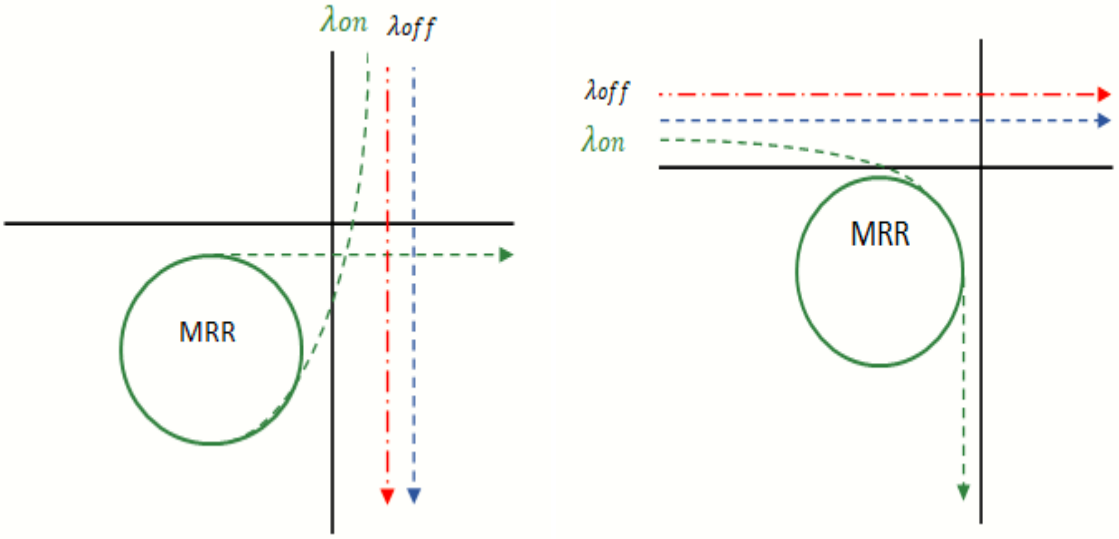
A typical setting for WRONoC application is a 3D-integrated chip consisting of both the electronic layer and photonic layer (in a vertically stacked position) [7], [8], such as the processor-memory network as shown in Figure 1. On the electronic layer, there consists of clusters of processors, all of which are attached to a hub in the photonic layer by through-silicon via (TSVs). With the aid of wavelengths provided by laser sources, electrical signals are transformed into optical signals and vice versa at electrical/optical (E/O) and optical/electrical (O/E) interfaces. WRONoC router manages the transmission of the data on the photonic layer among/between hubs and memory controllers (MCs) of off-chip memories.

WRONoC routes optical signals inactively. For every initiator-target pair, the desired routing path is decided at configuration time. The message directing system in WRONoCs depends on microring resonators (MRRs) [1]. An MRR comprises a micro ring and a coupling mechanism to direct the wavelengths of light as. Each MRR is designated to a particular wavelength of light. Two MRRs may have the equivalent wavelengths yet can bolster diverse exchanging systems. There can be two types of switching components such as the parallel switching component which comprises a micro ring between two equal waveguides, and the cross switching component which comprises a micro ring close by a waveguide crossing. When the signals with a wavelength equivalent to that of the MRR approach the switching components through a close-by waveguide, they will be coupled to the micro ring and afterward changed to an inverse/symmetrical heading on the other waveguide. In these cases, we call the signals as being "on-resonance" with the MRRs. On the other hand, when signals are routed on wavelengths different than the resonating wavelength, they will pass the micro

rings and simply keep their previous designated directions. In this condition, we term these signals as “off-resonance” with the MRRs. In WRONoCs, multiple MRRs of various resonant wavelengths are provided to form several paths for signals on different wavelengths.



(a) Parallel Switching Element



(b) Cross-switching Element

Fig. 2: *Switching Elements*

3.3. Building Blocks of WRONoC

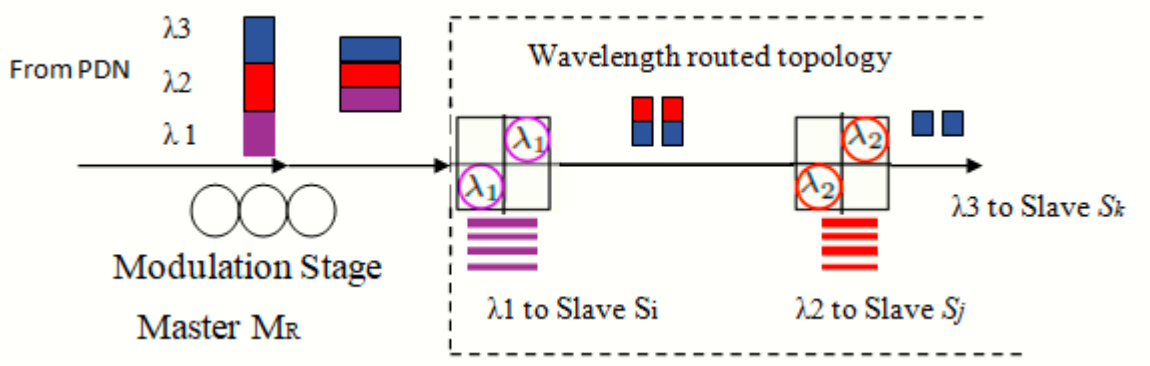
The building blocks of the most widely recognized WRONoC topologies consist of add-drop filters (ADFs) [1]. They include a narrow band wavelength of light coming from a more extensive optical signal being conveyed along a bus waveguide. They utilize microring resonators (MRR) [1] for this reason. When the input optical signal is on resonance with the MRR, the signal changes the course of its path, else, it continues to transmit unaffected. Every ADF's are assigned to a specific wavelength which ultimately determines the route of the optical signals.

Wavelength Routing Functions

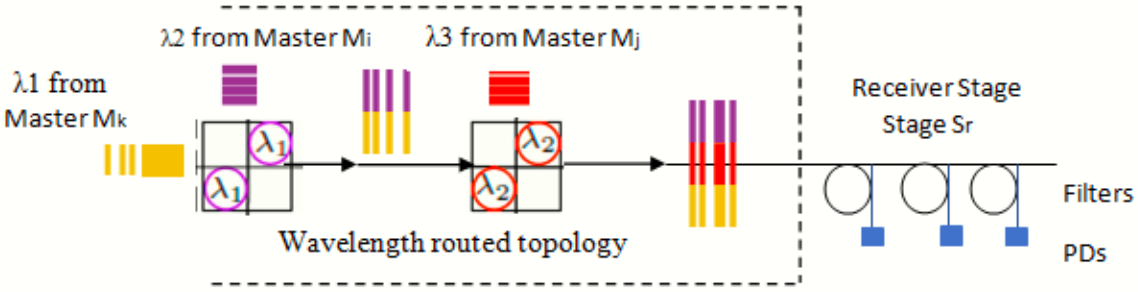
The wavelength routing functions are satisfied by each WRONoC router and can be categorized logically into two sub-functions:

•Drop function
(Fig.3a). Power distribution network (PDN) [1] provides each master with multiple carriers with wavelengths 1 to n merged into a wavelength division multiplexed (WDM) [1] optical signal. Each carrier is modulated and transmitted to the routing plane for transmitting to a specific and distinct slave. Then, the network resolves these multiplexed compound signals into individual wavelength-channels so that every component can be transmitted to a unique slave. In reality, this specific function can be performed by using add/drop optical filters, which are made into resonance to a particular wavelength, and therefore separate the optical channels from the compound signal.

•Add function
(Fig.3b). Wavelength channels from the different masters heading to the same slave should be recombined together into a WDM optical signal propagating onto the output waveguide of that slave. This way, a selective filtering stage can eject the desired wavelength channel and feed it to a photodetector (PD) [1] stage. In practice, this task can be accomplished by using different inputs of add/drop optical filters.



(a) Drop Function of a WRONoC



(b) Add Function of a WRONoC

***Fig. 3**: Functionalities of WRONoC*

4. Implementation

4.1. Design

The λ-router [1] is a passive optical network [4] composed of 4- port optical switches (based on add-drop filters designed to route information data through System on Chip components). The below figure represents a structural representation of an N × N λ-router architecture (each grey box indicates an add-drop filter where a physical architecture is embedded into it).

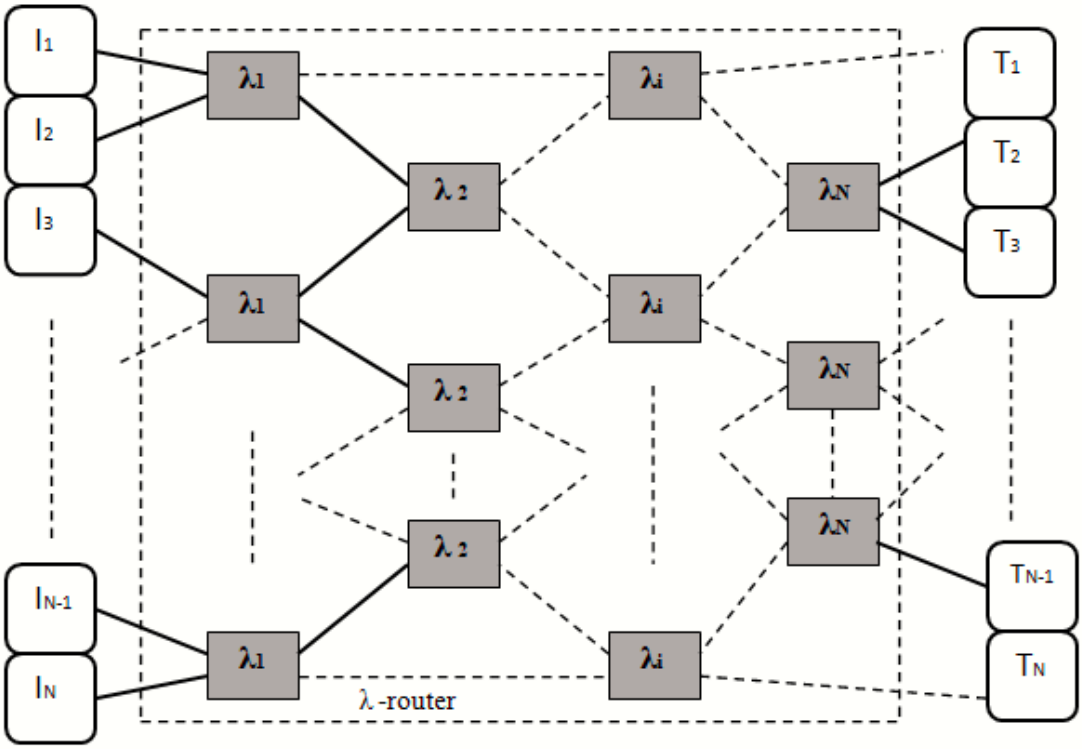


Fig. 4: *λ-Router Architecture*

Three possible switching states are depending on the input signal, which are:

• Straight state happens when a particular wavelength, termed as resonant wavelengths (λ_i) are embedded in the filter and is routed through the micro-resonator.

• The diagonal state happens when other wavelengths (λ_j) are embedded in the filter and are not routed through the micro-resonator.

• Cumulative state happens when signals of both resonant and non-resonant wavelengths (λ_i and λ_j) are inflicted into the architecture and are either routed or not routed through the micro-resonator.

4.2. Inputs & Outputs

This work mainly focuses on a 4*4 λ-router with two Hub Coordinators and two Memory ports. I am proposing a router design that will support both the characteristics of full connectivity and the bit-level parallelism. The two Hub Coordinators should be connected with every other port except itself while the memory ports will only be connected with the Hub-Coordinators. The connection between the two memory ports is considered redundant. To establish full connectivity between all the ports, we have used optical signals of different wavelengths. The optical signals of 5 different wavelengths are listed below separated by their specific colours. The different MRR's are assigned with specific wavelengths of light and to support bit-level parallelism as well as the flexibility to change the destination ports at any point in time, some MRR's are kept redundant and are assigned with no wavelength of light.

Types of Wavelengths

Red → λ1

Blue → λ2

Yellow → λ3

Green → λ4

Purple → λ5

Distinguished signal patterns for every port

HUB 1: ——————————

HUB 2: — · — · — · — · — · —

MEMORY 1: ·····

MEMORY 2: — — — — — — — —

4.3. Procedural Steps for the Final Router Design

In the following section, I will be describing different router architectures in detail. First, I will be starting with a basic router design and will slowly develop into a final router architecture that supports multiple signal paths for a single initiator/target pair and full connectivity among its ports. Each Router design is a modification and improvement of the previous Router design overcoming some of the disadvantages and slowly advancing into an ideal design.

Step 1: I will start with a Basic All-1 Router Architecture. Here, the term "All-1" indicates that every MRR in this design will always be turned on.

Issue: This Router neither supports full connectivity among its ports nor the communication parallelism. So we modified the present design into a new one which we will see in the next step.

Step 2: I then decided to select a specific topology for our design. So I ultimately choose λ-router to be our preferred architecture.

Issue: Still we couldn't achieve full connectivity among the ports with this specified router architecture. So we further improved it into modified λ-router architecture.

Step 3: A conventional λ-router follows the 2-1-2-1 rule for MRR placement. To fulfill our objectives, I have modified it into a 2-2-2-2 MRR based architecture. With this architecture, we are finally closing into our full connectivity goal for our router design.

Issue: A large number of waveguide dividers and couplers. With the increasing number of waveguide dividers and couplers, our design is prone to too high power consumption and latency.

Step 4: In the following design, we will be solving the problem of high power consumption and latency. This architecture is capable of supporting full-connectivity among its ports. Next, we need to shift our focus on how we can achieve a design that is capable of supporting parallelism among its ports.

Issue: Incapable of supporting communication parallelism criteria. Although it supports some level of parallelism, which in turn also leads to a few redundant connections in our design which are undesirable.

Step 5: Final Router Design- In this following design, we did some significant modifications to our architecture. First, I have used 5 different wavelengths of light. Then I have added some more MRR's which will ultimately lead to our proposed architecture. Instead of giving a specific wavelength to every MRR's, I decided to keep some MRR's as redundant which can, in turn, be turned on by the user according to his need.

In the following pages, I will describe this specified λ-router topology with a more detailed view.

5. RESULTS

5.1. All-1 Router Architecture

Wavelengths used: λ1, λ2, λ3, λ4

Transmitter Ports: Master1, Master2, Master3, Master4

Receiver Ports: Slave1, Slave2, Slave3, Slave4

No. of MRR used: 8

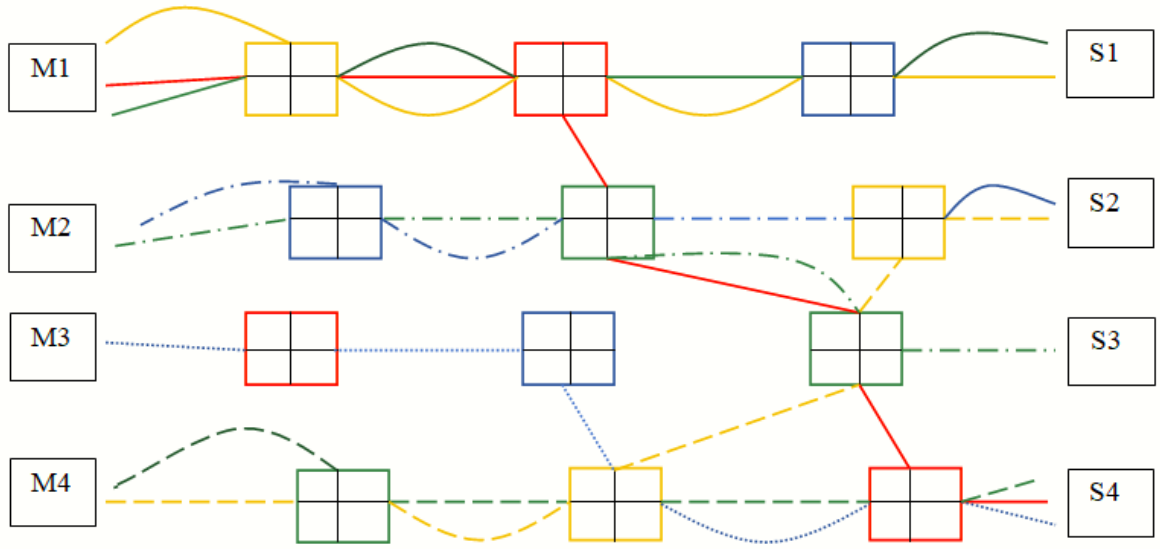


Fig.5: All-1 Router

Transmissions Matrix:-

	S1	S2	S3	S4
M1	λ3, λ4	0	0	λ1
M2	0	λ2	λ4	0
M3	0	0	0	λ2
M4	0	λ3	0	λ4

5.2. λ-Router

Wavelengths used: λ1, λ2, λ3, λ4

Transmitter Ports: Hub-Coordinator 1, Hub-Coordinator 2, Memory 1, Memory 2

Receiver Ports: Hub-Coordinator 1, Hub-Coordinator 2, Memory 1, Memory 2

No. of MRR used: 6

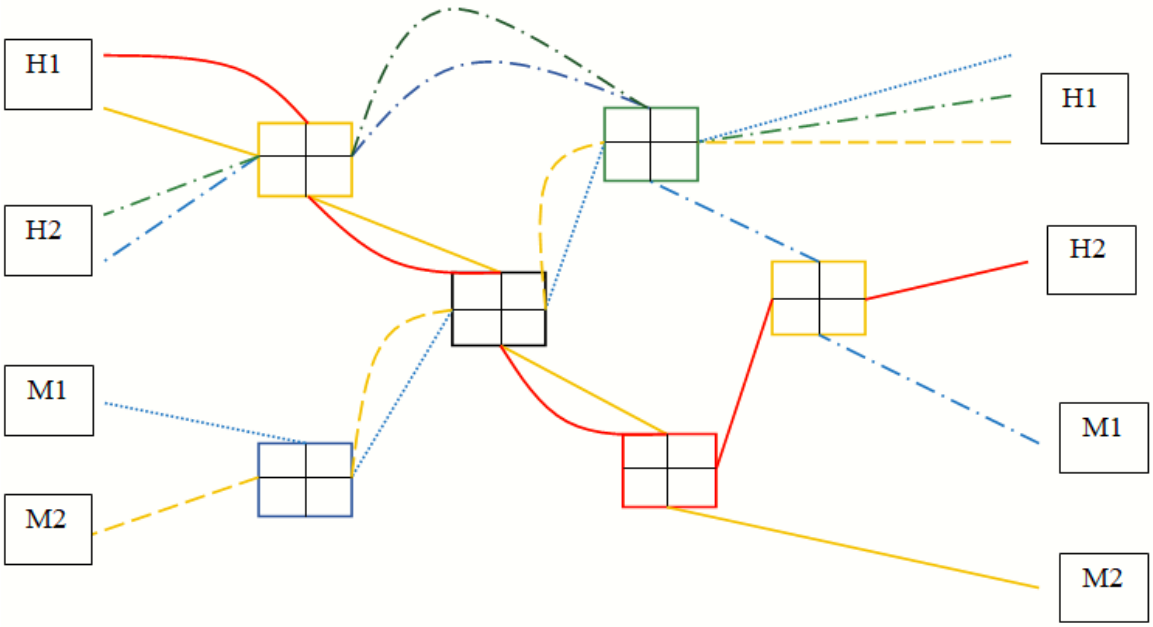


Fig.6*: *λ-Router

Transmission Routes:

H1	H2	M1	M2
H1→ H2::λ1	H2→ H1::λ4	M1 →H1::λ2	M2→H1::λ3
H1→ M2::λ3	H2 →M1::λ2		

5.3. MODIFIED λ-Router

Wavelengths used: λ1, λ2, λ3, λ4

Transmitter Ports: Hub-Coordinator 1, Hub-Coordinator 2, Memory 1, Memory 2

Receiver Ports: Hub-Coordinator 1, Hub-Coordinator 2, Memory 1, Memory 2

No. of MRR used: 8

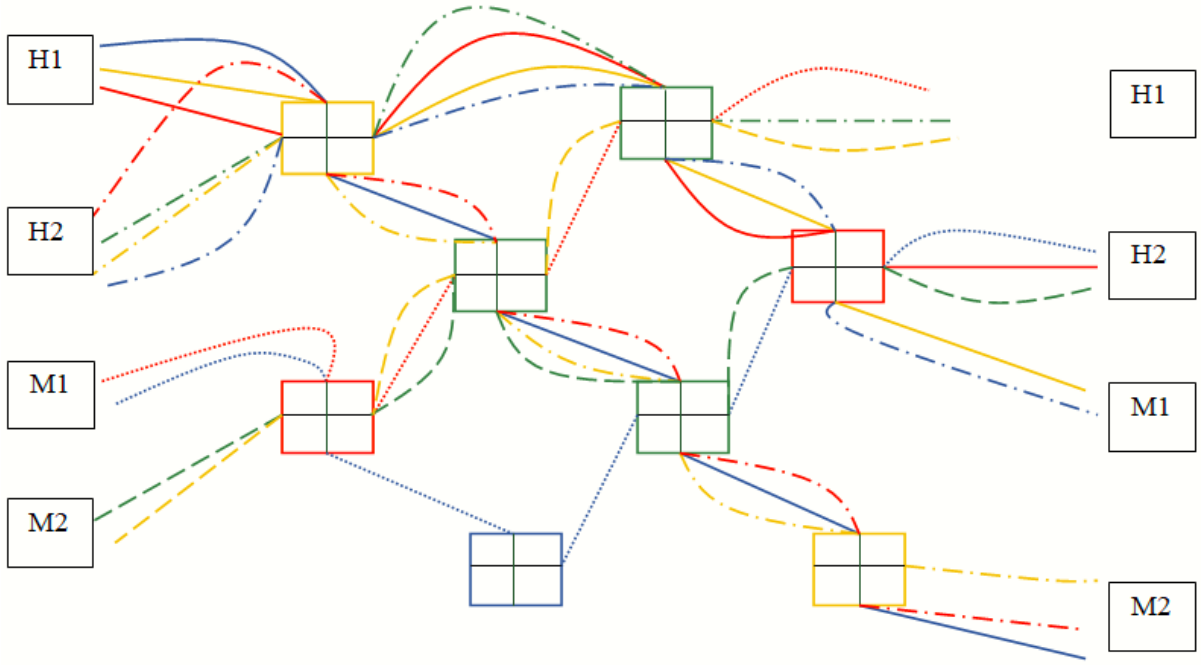


Fig.7: *Modified λ-Router*

Transmission Routes:

H1	H2	M1	M2
H1 →H2::λ1	H2 → H1::λ4	M1 →H1::λ1	M2 →H1::λ3
H1→M2::λ2	H2 →M1::λ2	M1 → H2::λ2	M2 →H2::λ4
H1→M1::λ3	H2 →M2::λ1		

5.4. Router with minimum waveguide dividers and couplers

Wavelengths used: λ1, λ2, λ3, λ4

Transmitter Ports: Hub-Coordinator 1, Hub-Coordinator 2, Memory 1, Memory 2

Receiver Ports: Hub-Coordinator 1, Hub-Coordinator 2, Memory 1, Memory 2

No. of MRR used: 8

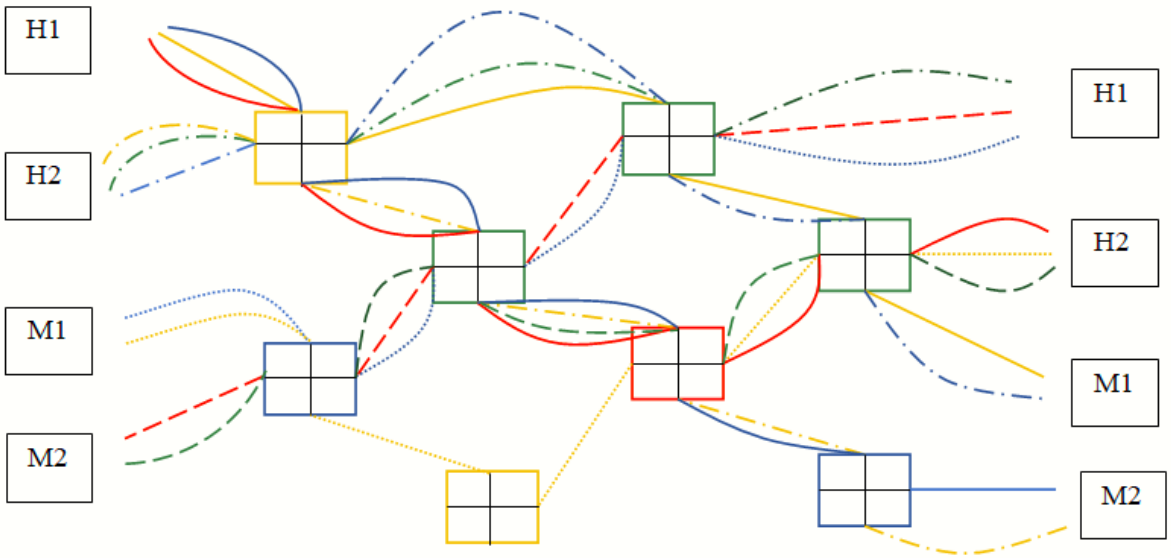


Fig.8: *λ-Router with less no of components*

Transmission Routes :

H1	H2	M1	M2
H1 →H2 :: λ1	H2 → H1 :: λ4	M1 →H1 :: λ2	M2 →H1 :: λ1
H1 →M2 :: λ2	H2 →M1 :: λ2	M1 → H2 :: λ3	M2 →H2 :: λ4
H1 →M1 :: λ3	H2 →M2 :: λ3		

5.5. Final Router Architecture

Wavelengths used: λ1, λ2, λ3, λ4, λ5

Transmitter Ports: Hub-Coordinator 1, Hub-Coordinator 2, Memory 1, Memory 2

Receiver Ports: Hub-Coordinator 1, Hub-Coordinator 2, Memory 1, Memory 2

No. of MRR used: 9

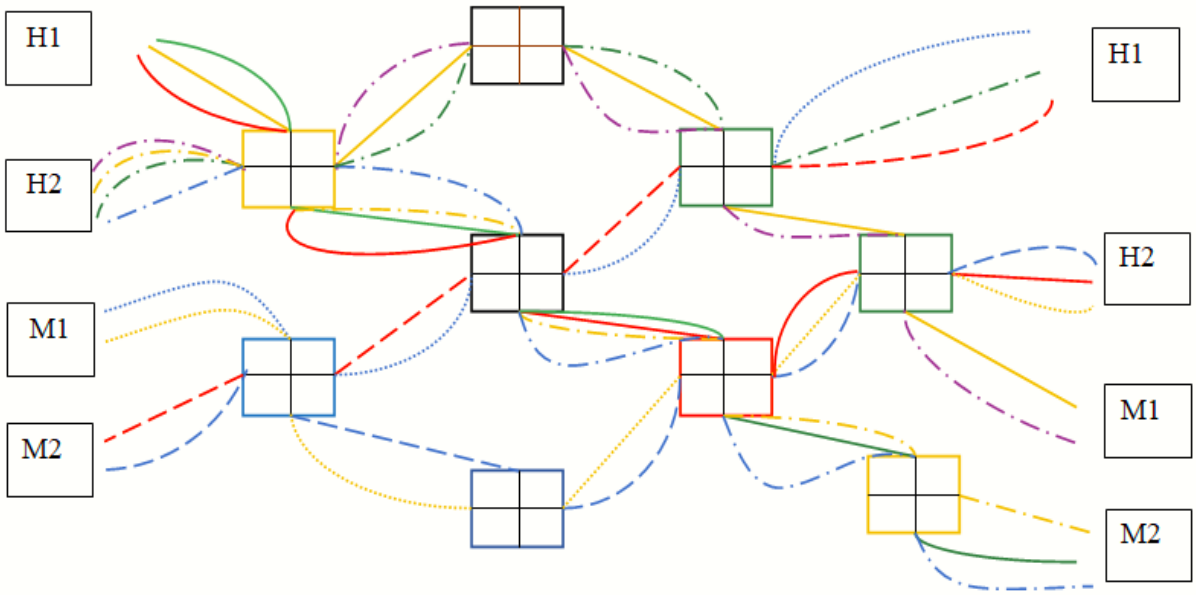


Fig.9: *Final Router Architecture*

Transmission Routes:

H1	H2	M1	M2
H1 →H2 :: λ1	H2 → H1 :: λ4	M1 →H1 :: λ2	M2 →H1 :: λ1
H1 →M2 :: λ4	H2 →M1 :: λ5	M1 → H2 :: λ3	M2 →H2 :: λ2
H1 →M1 :: λ3	H2 →M2 :: λ3, λ2		

5.6. Supported Parallelism Techniques

In the final router architecture as stated before, we have kept some MRR's as redundant which can later be fed with some specific wavelength of light based on the application. We have also inherited the on-off feature of the passive routers [4]. Turning one MRR 'OFF' means it becomes completely redundant allowing all the wavelength of light to pass through without being deviated from its original path. I can in turn switch 'ON' some redundant MRR's with a specific wavelength of light to support full connectivity among all ports. The user can also navigate more than one bit to an output port by using more than one distinct wavelengths of light. Below I will list the supported transitions that are possible between the input and the output ports.

The MRR's are numbered in a top-down manner and column-wise. So, in the 1st column we will have MRR 1, 2, then in the second MRR 3, 4, 5, and so on. Now, we will be switching some MRR's every time and show how the communication network between the ports gets switched up to support bit-level parallelism.

1. MRR-1 closed

Hub-Coordinator 1

λ3: H1→M1 changes to H1 →M2

Hub-Coordinator 2

λ3: H1 →M2 changes to H1→M1

2. MRR-2 closed

Memory 1

λ2: M1→H1 changes to M1→H2

Memory 2

λ2: M2 → H2 changes to M2→H1

3. MRR-3 on with wavelength λ5

Hub Coordinator 2

λ5: H2→ M1 changes to H2→ H1

4. MRR-4 on with wavelength λ4

Hub Coordinator 1

λ4: H1→ M2 changes to H1→ H2

5. MRR-6, MRR-8 closed

Hub Coordinator 2

λ4: H2→ H1 changes to H2→ M1

6. MRR-7 closed

Hub Coordinator 1

λ1: H1→ H2 changes to H1→ M2

Full Transmission Matrix

	H1	H2	M1	M2
H1	0	λ1	λ3	λ4
H2	λ4	0	λ4	λ3, λ2
M1	λ2	λ3	0	λ2
M2	λ1	λ2	0	0

Fig.10: *Full Communication matrix*

6. Conclusion

In this report, we have made a comparative analysis of WRONoC router architectures with the focus on λ-Router topology. We gradually went from a simple Router into a more advanced architecture that supports the communication parallelism concerning 4*4 λ-Router. Our proposed router architecture significantly improves many of the challenges faced during designing an application-specific WRONoC router. Although there had been a significant amount of work done based solely on the active and passive routers previously [4], in this report, we have tried to combine the functionalities of both types of routers and make it a more user concentric and application-oriented design. Still, there are a lot of factors we need to take into consideration before realizing it practically. In our architecture, we have closely analyzed factors like power consumption, latency, etc, and tried to minimize them as much as possible. I will further continue my research on this topic and will hopefully come up with more advanced designs and topologies.

7. References

1. A. Peano, L. Ramini, M. Gavanelli, M. Nonato, and D. Bertozzi, “Design technology for fault-free and maximally-parallel wavelength-routed optical networks-on-chip,” in Proc. Int. Conf. Comput.-Aided Des.,2016, pp. 3:1–3:8.
2. A. V. Beuningen, L. Ramini, D. Bertozzi, and U. Schlichtmann, “PROTON+: A placement and routing tool for 3d optical networks-on-chip with a single optical layer,” J. Emerg.Technol. Comput. Syst., vol. 12, no. 4, pp. 44:1–44:28, 2016.
3. H. Gu, K. H. Mo, J. Xu, and W. Zhang, “A low-power low-cost optical router for optical networks-on-chip in multiprocessor systems-on-chip,” in VLSI, 2009. ISVLSI ’09. IEEE Computer Society Annual Symposium on, 2009, pp. 19–24.
4. Mengchu Li, Tsun-Ming Tseng, Davide Bertozzi, Mahdi Tala, Ulf Schlichtmann, "CustomTopo: A Topology Generation Method for Application-Specific Wavelength-Routed Optical NoCs", *Computer-Aided Design (ICCAD) 2018 IEEE/ACM International Conference on*, pp. 1-8, 2018
5. A. Peano, L. Ramini, M. Gavanelli, M. Nonato and D. Bertozzi, "Design technology for fault-free and maximally-parallel wavelength-routed optical networks-on-chip," *2016 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, Austin, TX, 2016, pp. 1-8.
6. L. Ramini, P. Grani, S. Bartolini, and D. Bertozzi, “Contrasting wavelength-routed optical noc topologies for power-efficient 3D-stacked multicore processors using physical-layer analysis,” in Proc. Design, Automation, and Test Europe Conf., 2013, pp. 1589–1594.
7. S. L. Beux, I. O’Connor, G. Nicolescu, G. Bois, and P. G. Paulin, “Reduction methods for adapting optical network on chip topologies to 3D architectures,” Microprocessors and Microsystems: Embedded Hardware Design, vol. 37, no. 1, pp. 87–98, 2013.
8. T. Tseng, A. Truppel, M. Li, M. Nikdast and U. Schlichtmann, "Wavelength-Routed Optical NoCs: Design and EDA — State of the Art and Future Directions: Invited Paper," *2019 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, Westminster, CO, USA, 2019, pp. 1-6.
9. M. Li, T. Tseng, M. Tala and U. Schlichtmann, "Maximizing the Communication Parallelism for Wavelength-Routed Optical Networks-On-Chips," *2020 25th Asia and South Pacific Design Automation Conference (ASP-DAC)*, Beijing, China, 2020, pp. 109-114.