THESIS

DESIGN AND ANALYSIS OF ENERGY-EFFICIENT HIERARCHICAL ELECTRO-PHOTONIC NETWORK-ON-CHIP ARCHITECTURES

Submitted by

Srinivas Desai

Department of Electrical and Computer Engineering

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Spring 2015

Master's Committee:

Advisor: Sudeep Pasricha

Sanjay Rajopadhye
Yashwant K. Malaiya

ABSTRACT

DESIGN AND ANALYSIS OF ENERGY-EFFICIENT HIERARCHICAL ELECTRO-PHOTONIC NETWORK-ON-CHIP ARCHITECTURES

Future applications running on chip multiprocessors (CMPs) with tens to hundreds of cores on a chip will require an efficient inter-core communication strategy to achieve high performance. With recent demonstrations of feasibility in fabricating photonic components for on-chip communication, researchers are now focusing on photonic communication based on-chip networks for future CMPs. Photonic interconnects offer several benefits over conventional electrical on-chip interconnects, such as (1) high-bandwidth support by making use of dense wavelength division multiplexing, (2) distance independent power consumption, (3) significantly lower latency, and (4) improved performance-per-watt. Owing to these advantages, photonic interconnects are being considered as worthy alternatives for existing electrical networks.

In this thesis, we design and explore a hierarchical electro-photonic network-on-chip (NoC) architecture called NOVA. NOVA aims to optimize several key design metrics such as throughput, latency, energy-delay-product, and power, which determine the overall system performance of a CMP. NOVA has three levels of communication hierarchy. The first level has a broadband-resonator based photonic switch. The second level consists of a low-loss, silicon-nitride arrayed waveguide grating based router. The last level of the hierarchy is made up of photonic ring waveguides. We have modeled and simulated multiple configurations of the proposed architecture with different designs of the photonic switch and several arbitration

techniques on the photonic rings. This comprehensive analysis of NOVA allows us to arrive at an optimal configuration of the network for a given set of input applications and CMP platform. Finally, experimental results are strong indicators for considering the proposed architecture, as the improvements achieved were up to 6.1×, 55%, 5×, and 5.9× in terms of throughput, latency, energy-delay-product, and power compared to other state-of-the-art photonic NoC architectures.

ACKNOWLEDGEMENTS

I would like to express my heartfelt thanks to all the magnificent people whose encouragement and continuous support made this thesis possible.

First and foremost, I would like to express my sincere appreciation and deepest gratitude to my advisor, Dr. Sudeep Pasricha who took me into his research group and provided me with great guidance and navigation throughout the course of this research. He constantly motivated me and helped me make progress through the work. Dr.Pasricha is a person with sheer dedication, conviction, and has been a great source of inspiration to me. Weekly meetings and regular brainstorming sessions with Dr.Pasricha helped me overcome many roadblocks through the journey. I am indebted to him forever for all the advice, inputs, and knowledge he has provided me.

I am extremely grateful to my committee members - Dr. Yashwant Malaiya and Dr. Sanjay Rajopadhye for agreeing to be on my thesis committee. I appreciate them for taking time out of their busy schedule to review my thesis and for providing pointers to improve my work further.

I would also like to thank my colleagues at Multi-core Embedded Computing Systems (MECS) lab – Nishit Kapadia, Yong Zou, Yi Xiang, Ishan Thakkar, and Sai Vineel for offering their valuable feedback on my work from time to time.

Finally, I would like to thank my family and friends without whom, my dream of obtaining a master's degree would not have turned into a reality. I am thankful to my father for being there for me always. My loving sister – Sushma, my brother in-law Raghavendra, and my wife Neela have been instrumental in this process with their unconditional support.

DEDICATION

To my parents

Sudha Desai and Prahlad Desai

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

Technology scaling has resulted in integrating an increasing number of cores on a chip. With the increase in the core count, computing platforms are evolving into highly parallel architectures with several applications running simultaneously on them. The bandwidth requirements of these applications are huge and the responsibility of meeting these requirements lies on the on-chip communication network. Electrical networks-on-chip (NoC) have been proposed as worthy alternatives to bus based architectures. Based on the projections of International Technology Roadmap for Semiconductors (ITRS), integration of as many as 256 cores on a single chip is possible in the near future. At high core counts, the network complexity grows and it becomes difficult for an electrical network to meet the bandwidth requirements within the power budget of a multi-core chip. Electrical networks have long communication latencies and suffer from scalability issues. As a result of these issues a viable alternative in the form of photonic NoC is being considered today by researchers. Photonic NoCs enjoy the benefits of reduced power consumption while supporting high bandwidth requirements. Photonic interconnects have bandwidths in the range of terabits per second, along with lower access latencies compared to electrical interconnects [1].

In this chapter we provide a brief introduction on the general trend in the design of processor architecture, explain the challenges associated with electrical NoCs for chip-multiprocessors (CMP). We then describe the motivation behind considering photonic NoC architectures for CMPs. Towards the end of this chapter, we list the contributions made in this work.

1.1 Why multi-core systems?

CMP architectures have become mainstream in CPU designs due to advances in integrated circuit technology and performance limitations of a single core, wide- issue, and super-speculative processors. The limited performance gain with single core processors is due to the fact that instruction level parallelism (ILP) exploited by superscalar architectures has begun to see a decline [2]. Another reason for moving towards multi-core system is because single-core processors consume large amounts of power to deliver good performance. With multi-core systems, huge bandwidth requirements are met while running every core on the system at lower frequencies compared to a faster single core processor. A modern CMP consists of many cores connected to each other, executing various workloads in parallel.

As the core count on a chip continues to increase, the inter-core communication traffic also increases. To achieve high performance with multi-core systems, it is absolutely necessary that on-chip communication network delivers a high bandwidth, low power, and reliable network. A brief discussion on several commercially available CMPs is given below.

TILE-Gx72 multi-core system

Figure 1.1 shows Tilera's TILE-Gx8072 CMP [4]. It consists of 72 identical processor cores interconnected with Tilera's iMesh on-chip communication network. Each tile consists of a 64-bit processor core as well as L1, L2 caches, and a non-blocking Terabit/sec switch. The non-blocking switch connects the tiles to the network providing full cache coherence among all the cores. The maximum core frequency is 1.2 GHz. On-chip processing elements communicate with the off-chip DDR3 memory through four integrated memory controllers.

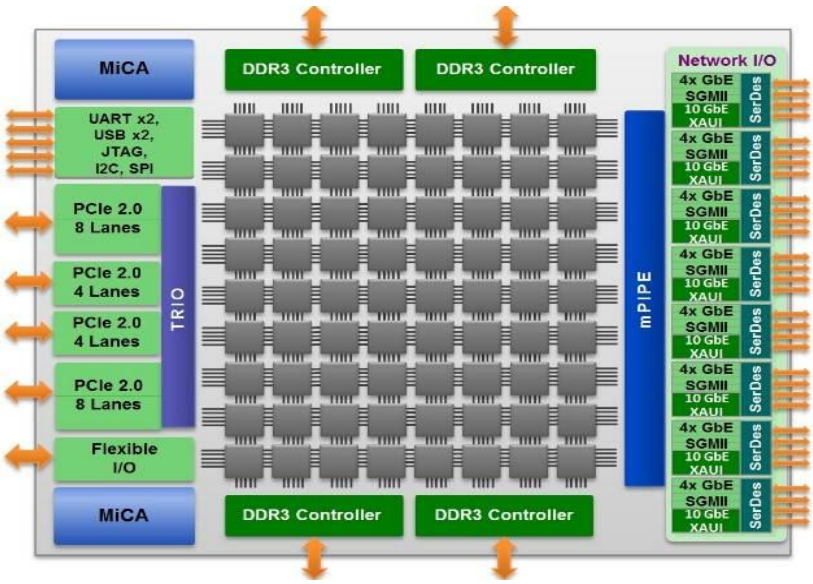


Figure 1.1 TILE-Gx8072 processor block diagram [4]

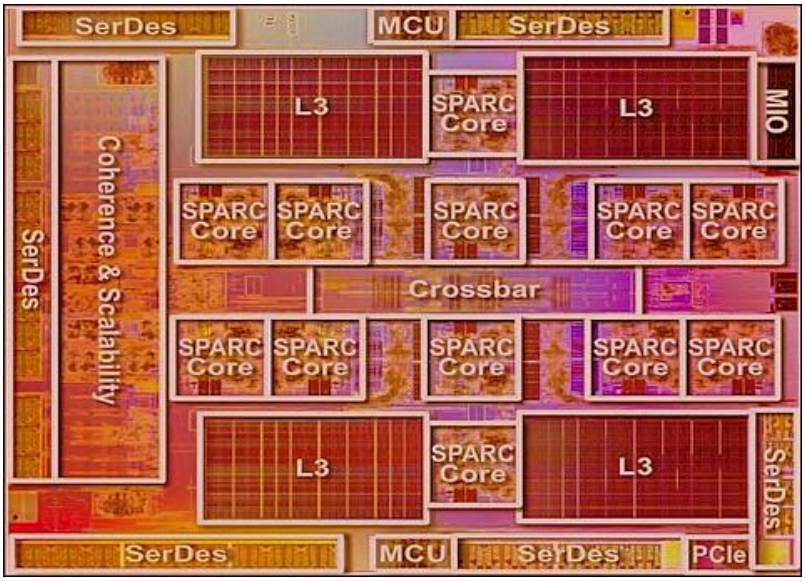


Figure 1.2: Oracle's M6 processor die [5]

Oracle's SPARC M6

Figure 1.2 shows Oracle's SPARC M6 CMP [5]. SPARC M6 contains 12 physical processor cores. Each processor core has two integer pipelines, one floating-point execution pipeline, and one memory pipeline. The processor core has a 16-stage integer pipeline to achieve

high operating frequencies. It also has an advanced branch prediction mechanism to mitigate the effect of a deep pipeline, and dynamic allocation of processor resources to threads. SPARC M6 can support up to 8 threads per core. It provides robust out-of-order, dual-issue processor cores that are heavily threaded among eight strands.

Intel's single-chip cloud computer (SCC)

Figure 1.3 shows Intel's SCC [6]. SCC is a 48-core experimental prototype and a research chip built to study many-core architectures. The tiles are connected in 6×4 two dimensional mesh fabric and each tile consists of two processing cores. The bisection bandwidth of the network is 2 TB/s. The SCC die has multiple frequency and voltage domains. It has 24 tile clock frequency dividers and seven voltage domains in total. It has four memory controllers providing a memory capacity of 64GB in total.

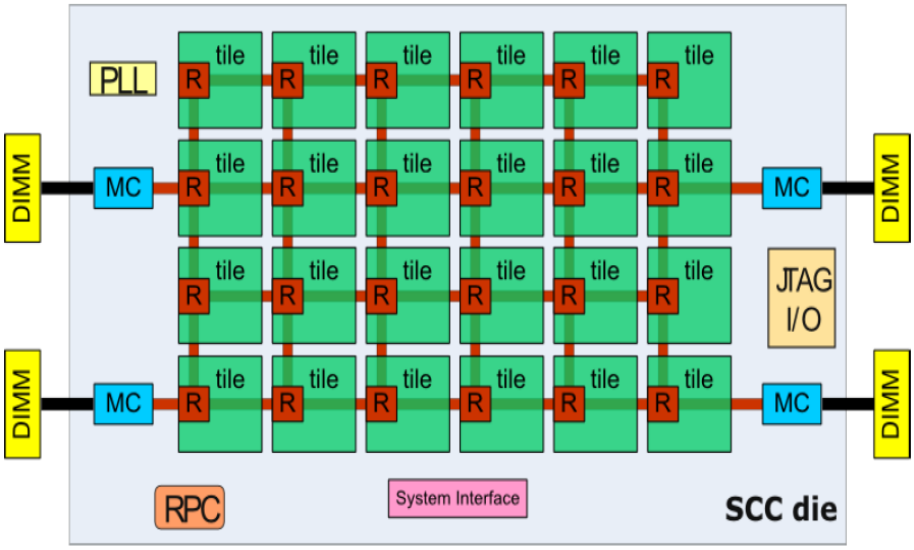


Figure 1.3: Intel's SCC block diagram [6]

AMD's FX-series

Figure 1.4 shows AMD's FX series multi-core system. It is comprised of eight cores with private L1 and L2 caches. There is a single memory controller integrated with DDR3 main memory. The system has four 16-bit HyperTransport links. Each 16-bit HyperTransport link can support up to 5600MT/s [7].

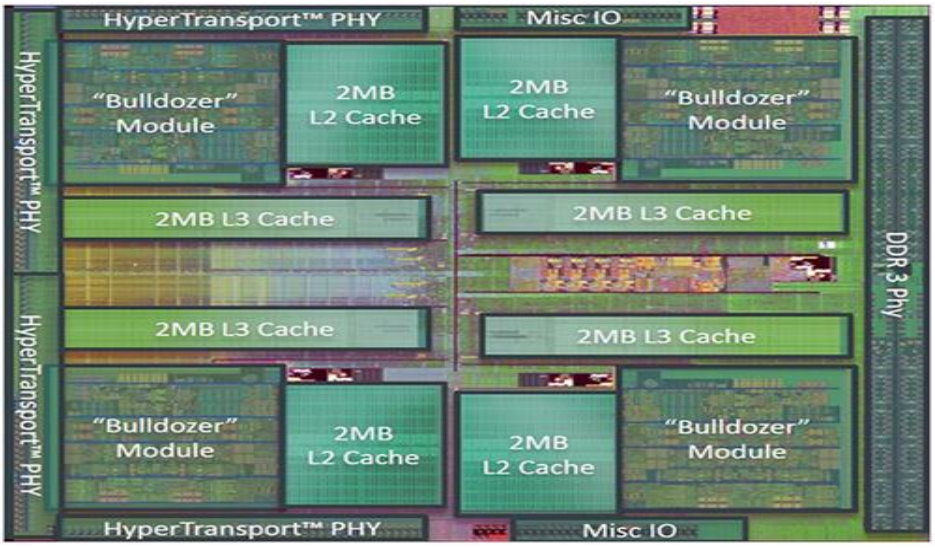


Figure 1.4: AMD FX-series processor die map [7]

From the above examples, we can observe that in a CMP architecture, there exists a multitude of shared resources. The system architecture of these CMPs, is progressively shifting from computation-centric towards communication-centric architectures. In fact, CMP performance will be increasingly determined by the ability of the communication infrastructure to efficiently accommodate the communication needs of the integrated/shared computation resources. To tackle the requirements of high throughput, low power and scalability, bus based architectures are evolving into network based architectures. Several advantages that NoC based architectures have over bus based architectures are:

- Better scalability at the architectural and physical levels.
- Better decoupling of protocol and transport level issues in the communication stack.
- Quicker design time and more freedom in the design due to decentralization.
- Modularity
- More customizability
- Streamlined design flows

For these reasons and several other, NoCs are emerging as the most promising interconnect medium for modern CMP systems. In the next section, we present a brief overview of electrical NoC interconnection architectures.

1.2 Overview of electrical NoC

A NoC is composed of three main building blocks. The first and the most important ones are the links that physically connect the nodes of a network and implement the communication. The second block is the router, which implements the communication protocol. And the final block is the network interface (NI). Each processing core may have a distinct interface protocol with respect to the network, and NI helps in making the logical connection between cores and the network. A detailed description of each of these blocks is presented in the following sections.

1.2.1 Links

A communication link connects two routers in the network and is composed of a set of wires. Links can be made of one or more logical channels and each channel is composed of a set of wires. Typically a NoC link has two physical channels forming a full-duplex connection between the routers. The source and target nodes agree over a synchronization protocol for transfer of data and this protocol can be implemented by a dedicated set of wires or through other

approaches such as FIFOs. The minimum amount of data sent over the links in one transaction is determined by the width of the channel and is defined as phit (Physical unit).

1.2.2 Routers

The basic functionality of the router is to receive a communication packet from shared links and forward it to the associated processing core or another shared link based on the address information in the packet. The router protocol consists of a set of policies defined during the design time that govern the packet handling mechanism. Typically, routers have five input and five output ports. A block diagram of a 5×5 router is as shown in the Figure 1.5. As shown in the figure, the router has a pair of input/output ports in each direction of North, East, West and South and has an additional port connected to the local core. Routing process is divided into five pipeline stages. Five stages of router pipeline are 1) buffer write and route compute 2) virtual channel allocation 3) switch allocation 4) switch traversal 5) link traversal. Figure 1.6 shows the micro-architecture of a 5 stage router.

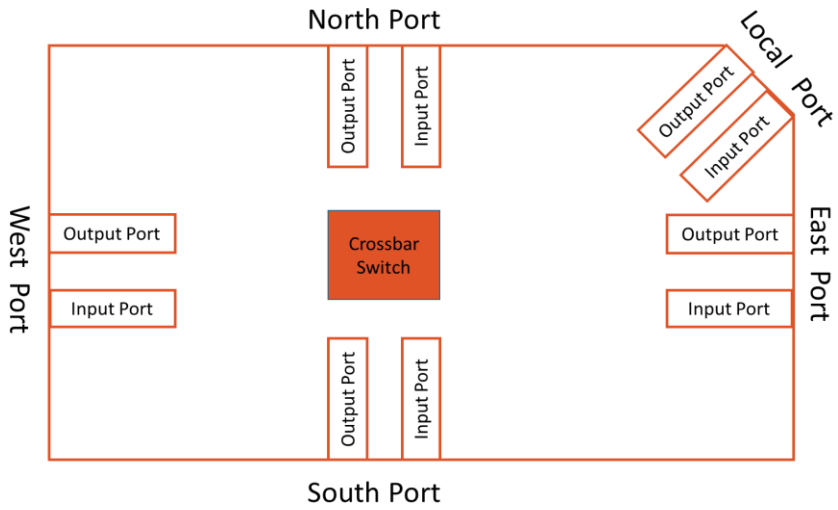


Figure 1.5: Block diagram of a 5×5 NoC router [8]

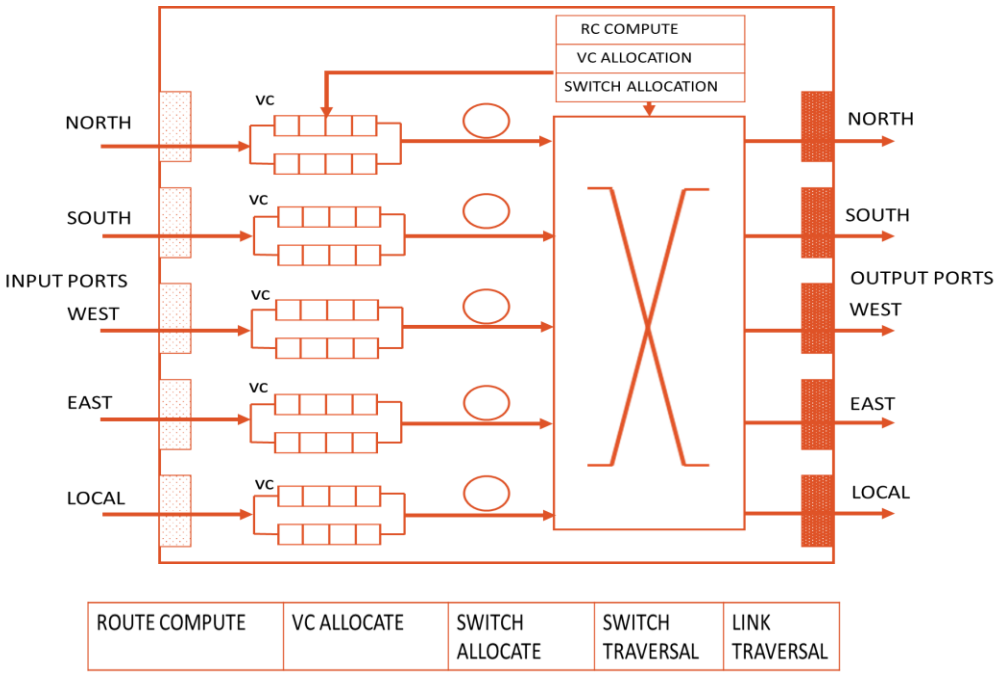


Figure 1.6: Micro-architecture of a 5 stage router [8]

A brief explanation of the 5 stages involved in routing a packet is as follows:

- **Buffer write and route compute:** An incoming flit is stored in the input buffer during this stage. Based on the information in the header flit, route computation for the packet is done.
- **VC allocation:** Virtual channel (VC) allocation stage follows the route compute stage. Only the header flits go through this stage. This stage is responsible for reserving a virtual channel for body and tail flits of the packet to which the header flit belongs.
- **Switch allocation:** Next stage is the switch allocation stage and in this stage, all the flits of a particular input port arrive at the crossbar of the switch and contend for output ports with other flits present in a different input port. The switch allocation logic determines the winner and allocates the output port accordingly.

- **Switch traversal:** In this stage, flits that won the contention in the switch allocation stage, advance to the next stage by traversing the crossbar and arriving at the respective output ports.
- **Link traversal:** After the switch traversal stage, flits traverse the link and go to the next router along their destination path.

1.2.3 Network interface (NI)

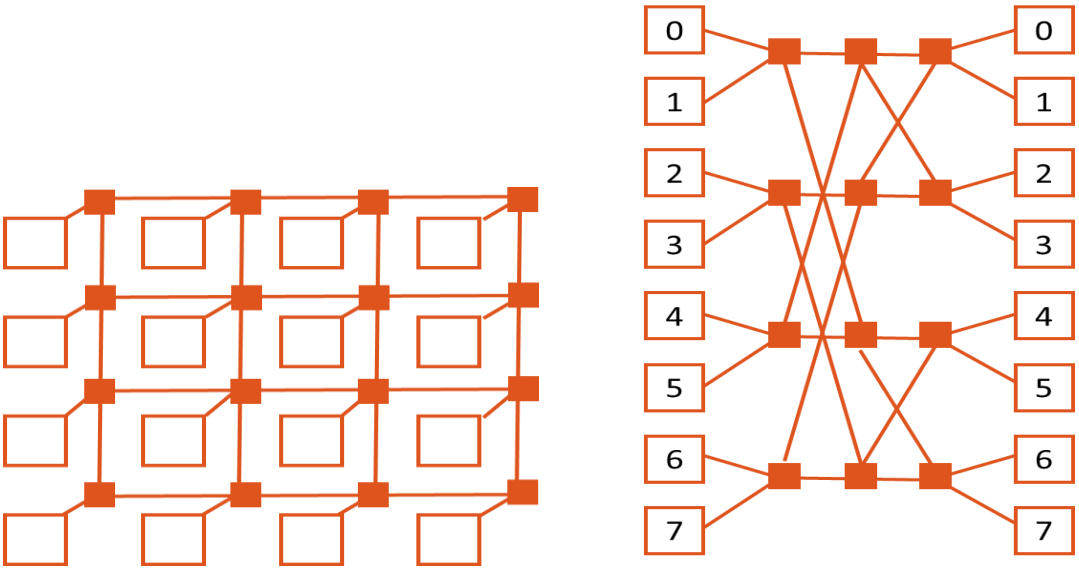
The third NoC building block is the network adapter (NA) or network interface (NI). This block makes the logical connection between processor cores and the network. NI is important because it allows the separation between computation and communication. A poorly designed NI can become a throughput bottleneck and greatly increase network latency. NI module can be reused, irrespective of the core and communication environment. It is divided into two parts: a front end and a back end. The front end handles all the requests from the core and is generally unaware of NoC, while the back end part handles the network protocol.

1.2.4 Characteristics of a basic NoC

A NoC can be characterized by several key concepts such as topology, routing algorithms, flow-control, and switching. In this section, we describe the fundamentals of these concepts that define a NoC.

Topology: A NoC can be categorized by the structure that connects all the routers in the network. This structure or organization of NoC elements is called topology. Topology refers to the static arrangement of channels and nodes in an interconnection network. The routers can be connected in two kinds of topologies viz., direct and indirect.

In direct topologies, each node has a fixed set of neighbors and is directly connected to all its neighbors. Communication is based on the routing algorithm implemented on routers. In an indirect topology, a set of routers are connected directly to the processing cores and another set of routers are used only to propagate messages through the network and do not share a direct connection with the processing cores. Figure 1.7 shows 2D mesh network, an example for direct topology and a butterfly network that forms an indirect topology. As shown in the figure, in 2D mesh all the routers are connected directly to the processing cores and in butterfly network, a set of routers form an intermediate stage and are not connected to any of the processing cores.



(a) **(b)**

Figure 1.7: NoC topologies: (a) 2D mesh, (b) k-ary n-fly butterfly [8]

Routing algorithms: Routing involves selecting a path from source node to a destination node in a particular topology. There are several routing algorithms that can be used in a NoC, each one leading to different trade-offs between performance and cost. There are two

main classifications of routing algorithms viz., deterministic and adaptive routing [8]. In a deterministic routing a packet always uses the same path between two specific nodes. Common deterministic routing schemes are source routing and XY routing. In the adaptive routing, alternative paths between two nodes may be used if the original path or a local link is congested. This involves a dynamic evaluation of the link load and implies a dynamic load balancing strategy.

Flow control: A flow control policy characterizes the packet movement along the NoC and it controls both global (network level) and local (router level) issues. Flow control determines how a network's resources, such as channel bandwidth, buffer capacity are allocated to packets traversing the network. Flow control policy can be used to ensure a deadlock free routing. Also, flow control techniques ensure optimal utilization of NoC resources.

Switching: Switching defines how communication data is transmitted from the source node to the destination node. Switching techniques can be of different types namely, circuit switching and packet switching. In the circuit switching approach the whole path from source to destination is established a priori and reserved for the transmission of the whole packet. The data is not sent until the whole path has been reserved. After the transfer of data is complete, the path that was setup will be freed up for other transfers. In packet based switching, flits of the packet are sent one by one, without reserving an end-end path in advance. Packet based switching is the most cost-effective way to enable inter-node communication in multi-core systems [8].

1.3 Challenges associated with electrical NoC architectures

As discussed in Section 1.1, all the commercially available CMPs use electrical networks for on-chip communication. Compared to bus based architectures, electrical networks have certain advantages such as modularity and higher bandwidth support. But as the technology scaling continues, electrical networks suffer from several bottlenecks. Traditional electrical based interconnects will have increasing difficulty in providing the required bandwidth support for future CMPs [9]. Adding to this is the problem of power dissipation in electrical networks. For example, power consumption of interconnect components in the Intel TeraFLOPS processor is more than 28% of the total power budget [10]. Figure 1.8 shows the distribution of power consumption among interconnect components, with technology scaling.

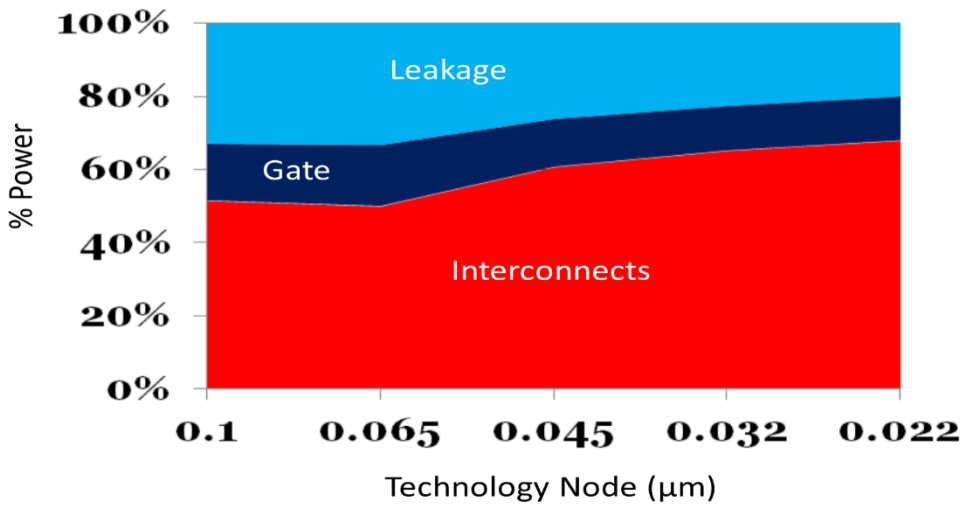


Figure 1.8: Interconnect power consumption with technology scaling [11]

Technology scaling has resulted in scaling down of the size of electrical wires that form the interconnection network [11]. As the resistance of electrical wires is inversely proportional to their size, reduction in size causes the resistance to increase. With the increase in the number of

components connected to a particular shared bus, the parasitic capacitance on the bus also increases proportionally. The combined effects of increased resistance and parasitic capacitance are 1) increase in the propagation delay and 2) higher power dissipation. To support the huge bandwidth requirements of future CMPs, multiple parallel electrical wires have to be laid down on the chip and this becomes increasingly complex and therefore affects the scalability of the system. Finally with electrical NoCs, I/O pin count problems persist and according to ITRS, "TILE64" - a commercially available CMP, has already reached the maximum pin count [12].

The challenges associated with the electrical networks have created an opportunity for exploring new technological solutions that can address the huge bandwidth requirements of future CMPs while keeping their power consumption under check. One such solution is integrated photonics which has the potential to mitigate the current challenges and issues associated with electrical networks [13].

1.4 Motivation

Advancement in silicon photonics technology has created a new opportunity in the area of on-chip communication networks [13]. Integrated photonics present an opportunity to reduce power, latency, and the area of interconnects on the chip compared to electrical counterparts. Photonics is ideal for global on-chip communication because the energy cost is only incurred at the endpoints of communication. The power consumption of photonic interconnects is largely independent of the distance [14]. Dense wavelength division multiplexing (DWDM) enables multiple wavelength communication channels to share a single communication medium – waveguide, thus providing a way to significantly increase the bandwidth support. Carefully designed photonic networks also lead to lower cross-talk and lower signal attenuation compared

to electrical networks [14]. Another reason why photonics is appealing is that it helps in reducing the I/O pin count drastically while maintaining high bandwidth support. As an example, TILE64's 1000 I/O pins could be replaced with just 50 waveguides, each supporting 20 wavelength channels [4] [10]. Finally, fabrication advances have made it possible to realize basic photonic communication elements such as micro-ring resonators on a silicon chip and this is a major motivating factor to consider photonic NoC as a viable alternative [15] [16].

Many photonic NoC topologies have been proposed in the recent past to turn the theoretical advantages over electrical networks into a reality [21] - [24]. Following subsections discuss some of the example photonic NoC topologies.

1.4.1 Circuit switched architectures

Photonic networks cannot buffer the data in intermediate stages and are unlike electrical networks which have store and forward mechanism. To get around the problem of data storage at intermediate stages, end-to-end arbitration is used in circuit switching. In circuit switching, a path setup stage precedes the high speed data transfer stage. Once the path is setup, data transfer from source to destination takes place in an uninterrupted fashion [24]. Broadband WDM is used to achieve high-bandwidths between communicating nodes, by multiplexing data onto many parallel wavelengths. The size of the communication data plays an important role in the performance of circuit switched networks. To mitigate the overhead of path setup, circuit switched networks transfer large chunks of data during a single transfer. An example of a path setup protocol is shown in Figure 1.9. As shown in the figure, path setup is accomplished using different control messages such as path-setup, path-acknowledge, path-blocked, and path-

teardown (after the data transfer is done). A source NI, a destination NI and intermediate routers are involved in the path setup/teardown process.

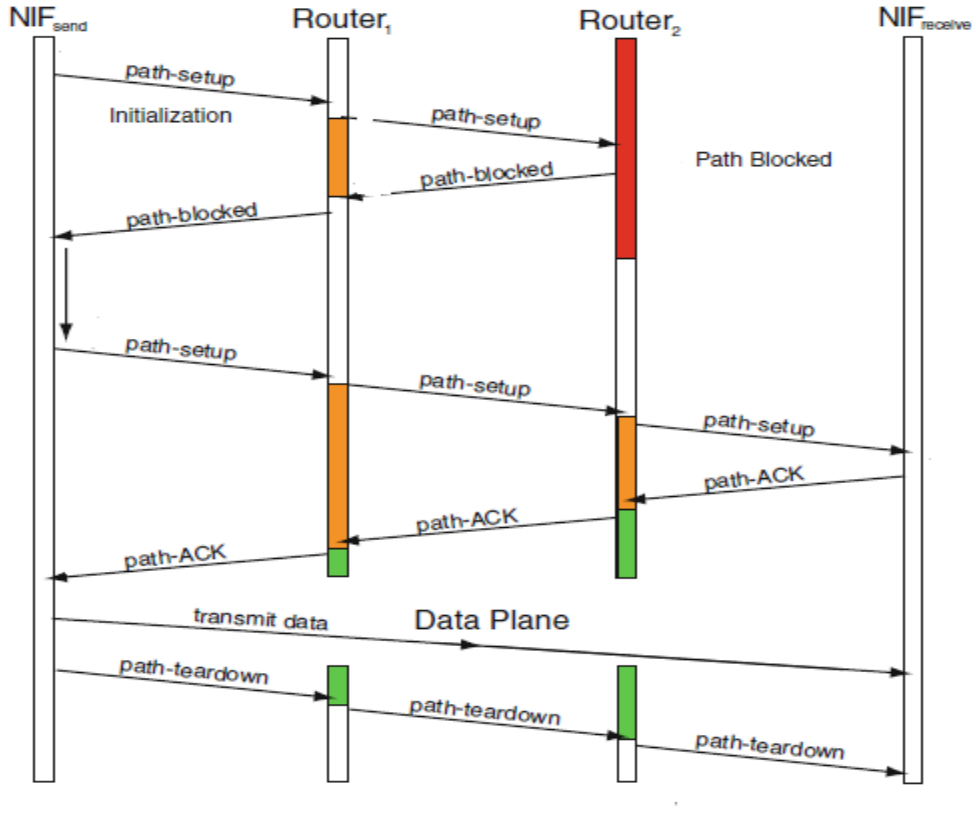
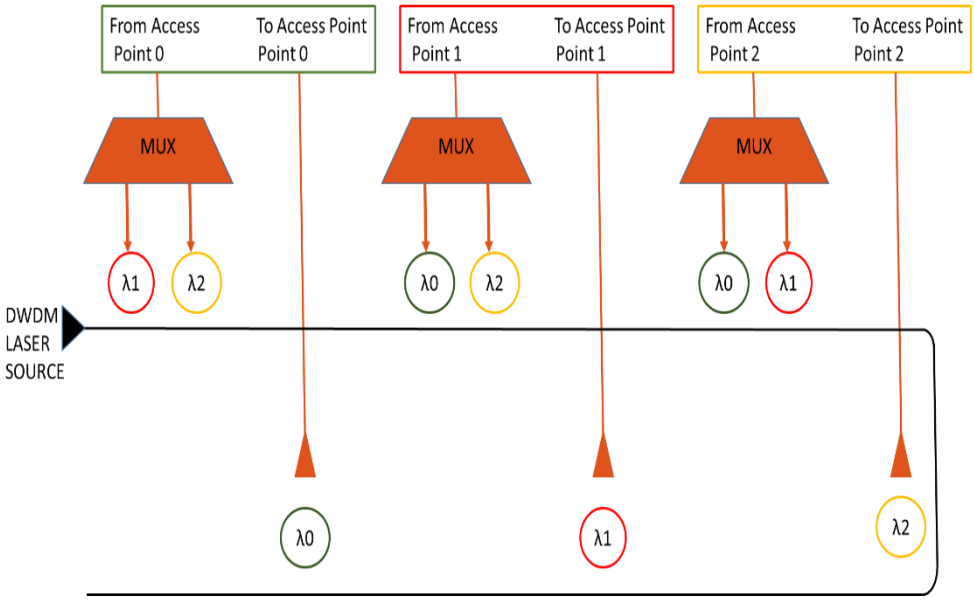


Figure 1.9: Circuit switched network: A timing diagram showing an example sequence of control packets for constructing a path through the network [24]

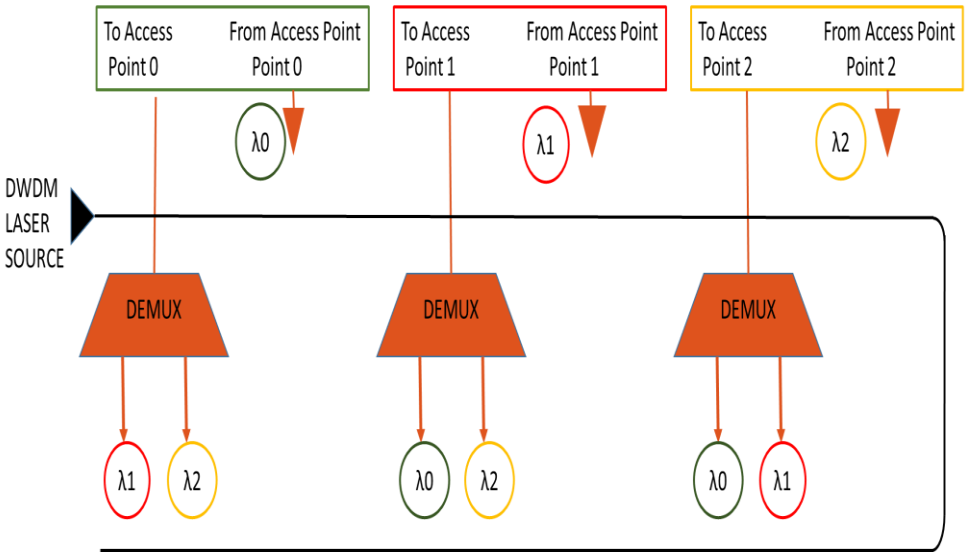
1.4.2 Wavelength arbited architectures

Wavelength arbited architectures make use of the fact that waveguides can support WDM and can selectively filter out some or all wavelengths to different destinations. Unlike circuit switched networks, wavelength arbited networks need not establish an end-end path, prior to data transfers. There are several variants in wavelength arbited architectures such as: 1) source-routed

bus, 2) destination-routed bus, and 3) crossbar-based bus architectures 4) single write multiple read 5) multiple read multiple read 6) multiple write single read [24] [25].



(a)



(b)

Figure 1.10: Wavelength arbited photonic NoC architectures (a) source-routed wavelength bus (b) destination-routed wavelength bus [25]

Figure 1.10 (a) and (b) show an example source routed and destination routed architectures respectively. In source routed bus, as shown in Figure 1.10 (a), each node reads from a single channel associated with it. All other nodes can write to this channel and conversely, each node can write to any node's channel [25]. Figure 1.10 (b) shows destination routed bus where each node can write to a single channel associated with it. All other nodes can read can read from this channel. Description about other wavelength arbited architectures is given in the subsequent chapters.

1.5 Contributions

To mitigate some of the issues such as high network contention, high static power dissipation of existing electrical and photonic NoCs, and to achieve a high-throughput, low latency network, we propose a novel photonic NoC architecture – NOVA. NOVA is hierarchical in nature and supports contention free wavelength routing scheme, implemented using photonic switching elements. NOVA has three levels of communication hierarchy. A broadband-resonator based photonic switch at the first level, an arrayed waveguide grating based router at the second level, and configurable photonic rings at the last level of hierarchy. Below is the summary of our contributions in this work:

- We propose an electro-photonic hierarchical NoC – NOVA that supports high bandwidths. We demonstrate that by employing a hierarchical topology, system performance parameters such as latency and contention can be notably improved.
- We use a silicon-nitride arrayed waveguide grating (AWG) based router for communication between two hierarchical levels. Usage of AWG based router helps in demultiplexing of photonic data and simplifies the process of communication between different levels of hierarchy.

- We do a comprehensive analysis of the proposed architecture by implementing (i) three variants of the photonic switch, each varying in their support of bandwidth and (ii) three different configurations on the photonic rings at the last level of hierarchy. This comprehensive analysis of NOVA helps in arriving at the optimized architectural configuration for a given set of input applications.
- We perform a qualitative comparison of NOVA with several state-of-the-art photonic NoC architectures for 64-core and 256-core CMP platforms. We demonstrate the improvements of our proposed architecture in terms of throughput, latency, power, and energy-delay-product over the prior works.

1.6 Outline

This thesis is organized as follows. In Chapter 2, we describe fundamental photonic communication elements such as waveguides, micro-ring resonators, lasers, and couplers. We also discuss the photonic communication mechanism and analyze the basics of photonic NoC architectures. In Chapter 3, we elaborate on the problem statement of this thesis. We then describe in detail about the various state-of-the-art photonic NoC architectures in Chapter 4 and explain different communication protocols used in photonic NoCs. We also present a discussion on dynamic reconfiguration support in photonic NoCs in Chapter 4. In Chapter 5, we discuss our proposed energy-efficient, high bandwidth novel photonic NoC architecture – NOVA. We present the micro-architectural details of the modules involved in different levels of communication hierarchy. We also discuss various configurations of the proposed architecture in this chapter. In Chapter 6, we provide a detailed description of the experimental setup, and discuss the performance impact of various configurations of the proposed architecture. The

chapter also presents comparison results of the proposed architecture with other state-of-the-art photonic NoC architectures in terms of throughput, latency, power, and energy-delay-product. Finally, in Chapter 7, we present the summary and conclusion of this thesis. Also, we present a brief discussion on possible future extensions of this work.

2 BACKGROUND

In this chapter, we present an overview of the fundamental elements that are used in building photonic networks-on-chip (NoC). We then discuss the aspects of modulation and demodulation of photonic data. We also explain the process of dense wavelength division multiplexing (DWDM), which helps in achieving high bandwidths. Finally, we discuss the working of photonic switches and analyze a few basic photonic NoC architectures.

2.1 Basic photonic elements

In this section, we will delve into the details of basic elements that are used for building a platform for on-chip photonic communication. Recent advances in fabrication have made it possible to realize basic elements on a silicon chip [15] [16]. In this chapter, we will focus mainly on waveguides, ring resonators, couplers, and lasers.

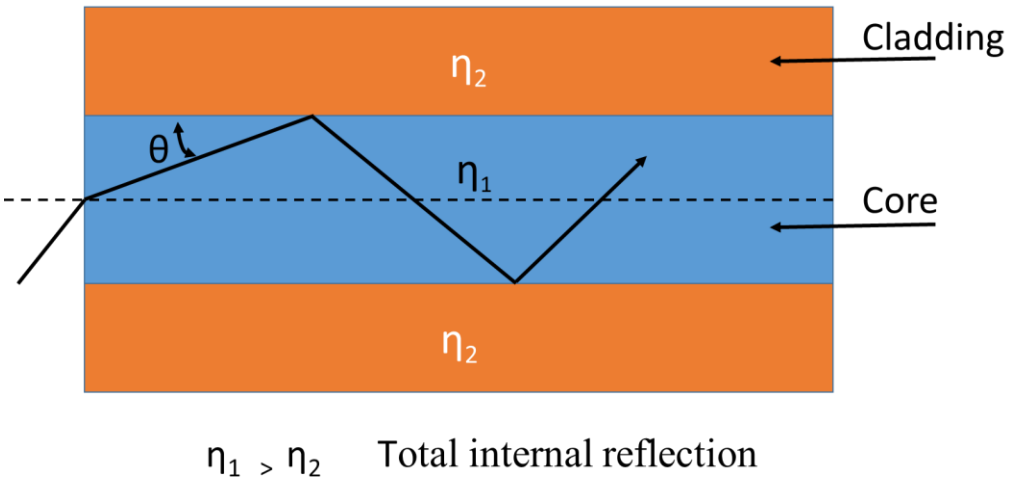


Figure 2.1: Total internal reflection in a waveguide

2.1.1 Waveguides

A waveguide can be considered as photonic equivalent of an electrical wire. It forms the basic medium for transfer of information from source to destination. Light is confined in the waveguide by a phenomenon called total internal reflection. Figure 2.1 shows the basic diagram of a waveguide with the light confined within the walls of waveguide. When a propagating wave is incident at an angle larger than the critical angle with respect to the normal of surface, light gets reflected internally. Waveguides are characterized by their effective indices, which define how the propagation of light takes place within the waveguide. To confine the light within the core, as indicated in the figure, the refractive index of the core has to be greater than the refractive index of the cladding. Based on several factors such as operating frequency, amount of power to be transferred, and the amount of transmission losses that can be tolerated, different structures of waveguide are used [25].

As the light passes through the waveguide, it experiences attenuation due to insertion losses associated with the waveguide. Factors such as light scattering at sidewalls and substrate leakage lead to insertion losses in waveguides [24]. Along with insertion loss, light also suffers from losses due to bends of waveguides and cross-talk losses. For on-chip communication silicon oxide waveguides with a typical cross-section of 500 nm are used. Waveguide losses are in the range of 1 ~ 2 dB per cm [17] [24]. All these losses have to be considered while building photonic interconnect models to accurately depict a practical photonic communication network.

Waveguides are laid down on a silicon chip in both straight and bent fashions to realize proper routing of photonic data. Since signal attenuation or the power loss in a waveguide is inversely proportional to the bending radius, sharp bends have to be avoided, as excessive bends can cripple the scalability of the network.

Finally, waveguides help in realizing large bandwidths by supporting many wavelength transfers simultaneously, with DWDM mechanism. Significant research efforts are being carried out to improve the waveguide performance in terms of power loss and bandwidth support [26] [27] [28].

2.1.2 Ring resonators

Ring resonator is a versatile element in the realm of integrated photonics. Because of the high refractive index contrast, silicon ring resonators can be manufactured in extreme small sizes [25]. A ring resonator is made from a photonic waveguide which is looped back on itself such that resonance takes place when the photonic path length of the resonator is exactly an integral number of the incident wavelength. Ring resonators can be used to create modulators, detectors, and switches. A detailed explanation of these building blocks is given in the next section of this chapter. A ring resonator has resonant modes during which the interaction of light with the resonator takes place. Figure 2.2 (a) shows a ring resonator in the 'off – resonance' mode. As shown in the figure, in this mode, the light waves injected into the communication waveguide are rejected by the resonator and light passes through the waveguide without being absorbed. Figure 2.2 (b) shows a ring resonator in the 'on-resonance' mode where light waves travelling in the communication waveguide are absorbed by the closely positioned resonator.

Ring resonators can be used as All-Pass filters by feeding an output of a directional coupler back into its input. They can also be used as Add-Drop filters where the ring resonators are coupled to two waveguides and the incident field is partly transmitted to the drop port [29]. Another application of ring resonator is data modulation. Modulation can be achieved by electrically manipulating ring resonators through free carrier injection. Electrical manipulation

can be accomplished by creating a P-I-N structure on the ring with the waveguide acting as the intrinsic region. Using this technique ring resonators operating at frequencies as high as ~25Gb/s have been demonstrated successfully [29]. This modulation of light using resonators provides a mechanism for converting an electrical signal to photonic signal.

Figure 2.2: Ring resonators (a) off resonance (b) on resonance

The final aspect of ring resonator is its spectral characteristic – Free Spectral Range (FSR) which defines the spectral distance between wavelengths that couple and resonate with the ring. The FSR of a ring is inversely proportional to the radius of the loop that forms the ring. FSR has to be very high to allow modulators and filters to operate at a single frequency, rejecting all other frequencies. Broadband resonators have low FSR values, thus supporting multiple wavelengths per resonator. Techniques such as inferometric combining and vernier effect help in creating resonators with different FSR values [25].

2.1.3 Couplers

Couplers allow on-chip elements to physically interface with off-chip components and are therefore an important class of photonic elements. For single mode operation, most silicon photonic applications have extremely small photonic mode sizes, attributed mainly to high-index

contrast between silicon core and silicon dioxide cladding. This makes efficient coupling between photonic modes of the waveguides and single-mode fibers (SMF) challenging [30].

There are two commonly employed methods for efficient coupling between fibers and on-chip waveguides: 1) vertical coupling and 2) edge coupling. Both techniques support numerical aperture conversion for efficient coupling to standard SMF. Vertical coupling makes use of diffractive grating to provide a photonic mode comparable to SMF [31] [32]. Edge coupling uses inverse-tapered silicon waveguides for coupling and frees both top and bottom chip surfaces for electrical and thermal connections [33] [34].

2.1.4 Lasers

Laser is the most important element required for any photonic communication. Lasers are used to generate or amplify light, producing a coherent beam of radiation, making use of stimulated emission of photons in the laser medium. The light produced by laser is used for encoding the communication data for transmission over the network. This section first describes the basics of lasers and then dives into the explanation of different types of lasers that can be employed for on-chip communication. Lasers have several unique properties such as:

- **Monochromaticity**: Conventional light sources emit light in a broad range of wavelengths whereas a laser has the ability to emit a very narrow range of wavelengths.
- **Directionality**: Light emitted by laser does not diverge with distance, compared to other light sources.
- **Coherence**: All the electromagnetic waves emitted by laser are in phase.

Lasers emit light through a process of photonic amplification based on the stimulated emission of photons. A laser is made of three fundamental parts 1) a resonant photonic cavity 2) a laser gain medium and 3) a pump source to excite the particles in the gain medium. The resonating photonic cavity consists of two mirrors between which the light bounces back and forth. One of the mirror is partially reflective allowing a portion of light to be transmitted. The gain medium is required to amplify the light and is hence placed inside the resonating photonic cavity for stimulated emission. Pump source supplies the required energy for active particles in the gain medium to be in a state of inversion. Pumping process can be achieved using either electrical current or photonic pumping in a solid-state or dye laser. Characteristics such as stability, compatibility with CMOS, and durability determine the feasibility of laser usage for on-chip communication. Another key parameter of a laser is its wall-plug efficiency (WPE). WPE of a laser system is its total electrical-to-photonic power efficiency [35].

Using silicon based lasers for photonic NoC presents unique challenges. Silicon by its nature is an indirect-bandgap material and is not naturally capable of achieving efficient stimulated radiation. Many solutions to overcome the bandgap problem of silicon have been demonstrated. One such solution is epitaxial growth of germanium on silicon resulting in a direct bandgap material which can then be used as feasible gain medium with either optical-pumping or electrical-pumping [35]. Intel has demonstrated a silicon photonic amplifier using "Raman Effect" which overcomes the inefficiency of silicon material to act as a feasible gain medium [36].

Another practical solution for light source is leveraging III-V compound semiconductors to produce efficient external off-chip lasers which can be coupled to chip. Since most of the proposed photonic NoC architectures make use of WDM, a laser that supports WDM is highly

desired. Based on the network design requirements, a laser comb source is preferred sometimes to generate WDM signals [35]. While other laser sources such as quantum well based lasers can be used, they suffer from stability issues for multiple wavelength lasing. Mode locked lasers are another variant that generate multiple wavelengths [21].

2.2 WDM and data conversion mechanism

In this section, we first present the details of WDM - a technique used to realize high bandwidths in photonic communication. We then describe the methods in modulation (electrical - photonic) and detection (photonic - electrical) of photonic signals using the elements described in the previous section.

2.2.1 WDM

To achieve unprecedented bandwidth support, photonic communication implements WDM, where parallel data streams can be sent on a single waveguide by utilizing a unique wavelength for each independent data stream. WDM is one of the main reasons for huge performance improvements of photonic communication over electrical communication. The extent to which a photonic device can make use of WDM is defined by its spectral bandwidth property [35]. As an example, a waveguide that supports a spectral bandwidth of 100 nanometers can allow the transfer of a digital signal at any center frequency within the range. Since a single data stream exhibits a spectral bandwidth of less than one nanometer, multiple data streams can be multiplexed and transmitted as simultaneous data stream. Data streams do not interfere with each other because they occupy different center frequencies and hence are mutually exclusive. At the receiving side, the multiplexed data streams are de-multiplexed and detected separately.

WDM can be generated using a mode-locked on-stack laser that can support up to 64 unique wavelengths [21]. Hybrid WDM laser supporting huge bandwidths with good WPE value has been demonstrated in [35]. Figure 2.3 shows an end-end path of photonic data, from *generation side* where a WDM laser is used as a source of light, to *modulation side* where electrical signals are modulated and converted to photonic domain, and to *reception side* where the signals are demultiplexed and converted back to electrical domain.

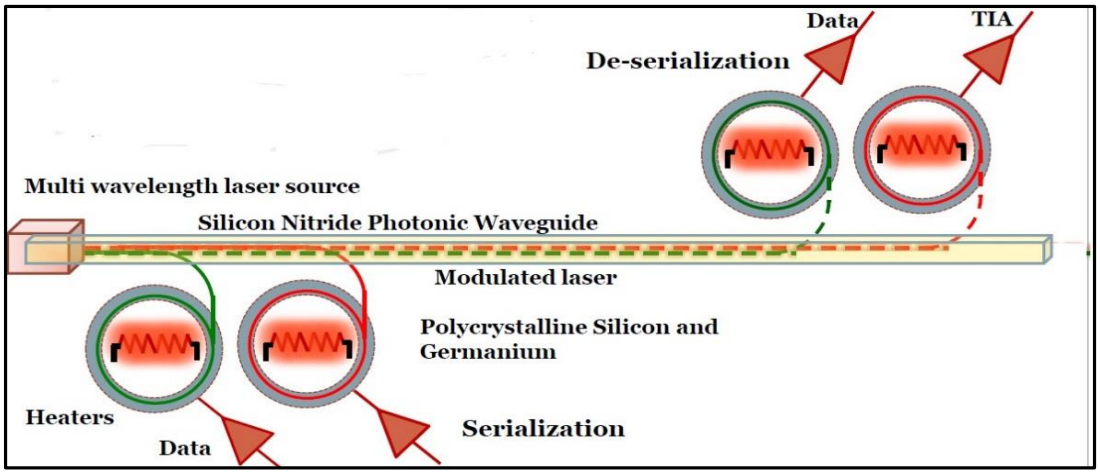


Figure 2.3: Photonic link showing WDM laser source and ring resonators for modulation, demodulation [11]

2.2.2 Modulators

Modulators are photonic elements that help in modulation of digital electrical data to photonic data. On-chip modulation can be achieved using CMOS compatible elements such as micro-ring resonators that are compact in size and are extremely efficient devices. Ring resonators can be designed to manipulate the flow of data by enabling modulation.

One of the ways to achieve modulation using micro-ring resonators is to use on-off keying (OOK) technique [25]. Figure 2.4 shows a basic ring resonator that can be used as a modulator. The fundamental idea in OOK is to change the resonant wavelength of the ring by injecting

carriers into it and by varying voltage across p+ and n+ regions shown in the figure. Usage of OOK has an inherent disadvantage during transmission of logical 0s. Transmission of logical 0s is done by absorbing the light completely inside the ring that gets dissipated eventually. As a result, if the power injected into the ring is high, it leads to two-photon absorption and free-carrier absorption problem. This problem can be overcome by using a drop port to let the photonic power out of the micro-ring [29].

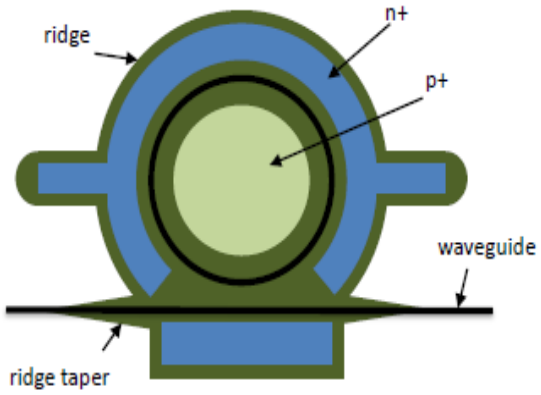


Figure 2.4: Ring resonator as a modulator [25]

2.2.3 Detectors

Detection of light is the last phase of photonic communication. Figure 2.5 shows a basic ring resonator that can be used as a detector. While the laser produces the light, modulator modulates the digital information to light, a detector which is at the receiving side helps in decoding the data sent over the photonic link. A ring resonator can be used to construct a detector and is done by introducing defects into the silicon, thereby enabling absorption of photons. Another way to construct a detector is to use deposited poly-crystalline for the ring as shown in the figure [29]. The detector converts incoming streams of photons into an electrical

current which is then picked by a trans-amplifier, responsible for producing a voltage based on the input current.

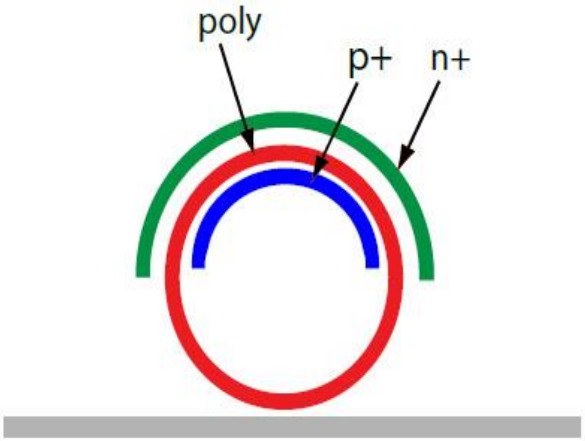


Figure 2.5: Ring resonator as a detector [25]

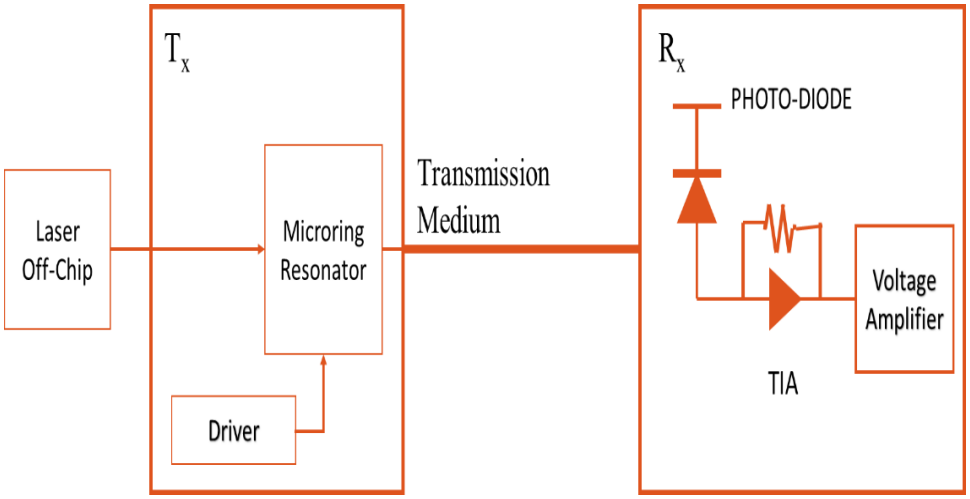


Figure 2.6: End-end transmission in a photonic link

Figure 2.6 shows a block diagram of end-end process of photonic communication involving generation, modulation, transmission, and detection. The figure gives a pictorial description of the photonic communication process described in this section. Along with lasers, modulators,

and detectors, a photonic communication process also uses other elements such as 1) a serializer, whose task is to up-convert an input data stream to a higher serialized data rate by combining multiple incoming wires 2) a driver circuitry, which converts the incoming electrical signal to higher voltage to actuate the modulator device 3) trans-impedance amplifier that is responsible for producing an equivalent voltage based on the input current 4) a de-serializer, which converts the high-speed photonic data to a clock rate suitable for electrical data bus.

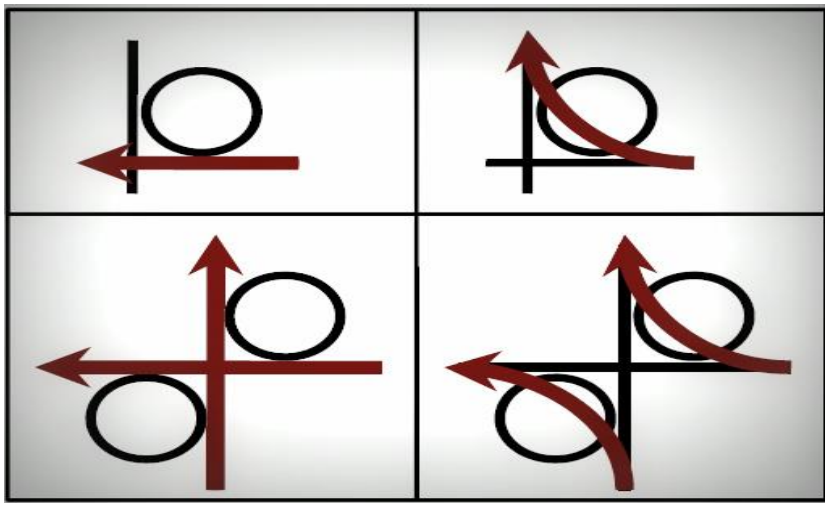


Figure 2.7: Ring resonator based 1×1 and 2×2 photonic switches [37]

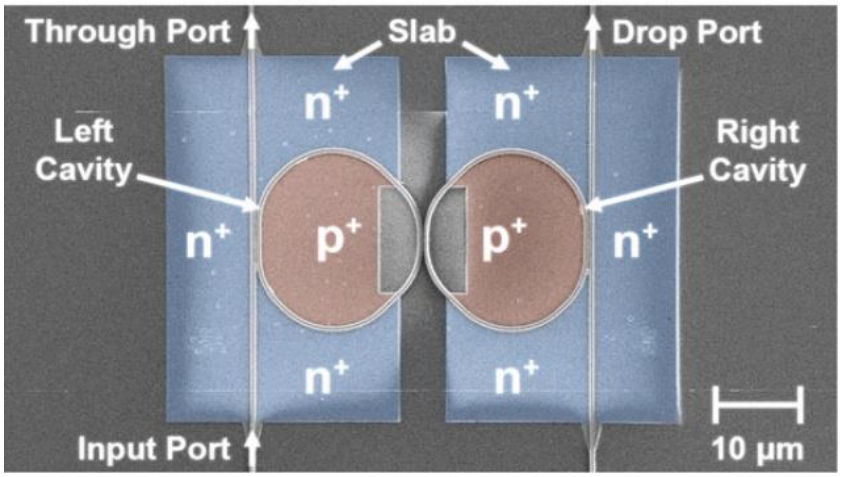


Figure 2.8: Scanning electron microscope image of 1×2 ring resonator based electro-photonic switch [24]

2.3 Photonic switches

Photonic switches help in routing of high bandwidth data. Micro-ring resonators are used to construct photonic switches owing to their high-bandwidth support and low energy dissipation [37]. Figure 2.7 depicts basic 1×1 and 2×2 photonic switches constructed using ring resonators.

Switching can be either all-photonic or electro-photonic switching. Photonic switch has a ring resonator coupled to two waveguides, one acting as the through port and the other as drop port. As shown in Figure 2.7, in an all-photonic switch during on-resonance state, light is coupled to be sent on the drop port and during off-resonance state, light is virtually let to the through port. Figure 2.8 shows an example electro-photonic switch. In electro-photonic switches, switching between drop and through port is accomplished by detuning of the right cavity's resonance, using the free-carrier dispersion effect arising from injecting and extracting electrical carriers through the P-I-N diode [37].

2.4 Photonic routing

Routing of data is required to transfer data from source to destination. Routing of photonic data can be achieved in several ways such as wavelength routing, spatial routing, time-division-multiplexed (TDM) routing, and wavelength selective spatial routing (WSSR) [38].

Photonic NoCs that employ wavelength routing use ring resonator based filters that route light waves based on their wavelength. Latency in wavelength routing based NoCs is considerably less compared to other techniques, as propagation delay is just the time of flight. Spatial routing based architectures use electro-photonic broadband resonators to guide large set of parallel data. In spatial routing, a path is established before the transfer of data. Here entire spectral bandwidth can be utilized for data transfers, unlike wavelength routing [37] [38].

TDM routing based photonic NoCs are an improvement over spatial routing architectures. TDM routing eliminates the circuit switching overhead that is associated with spatial routing. In this routing technique, transmission medium is divided into frames and theses frames are in turn divided into unique time slots, which represent different configurations of the network. Combination of all the unique time slots connects all nodes in the network. In WSSR routing based photonic NoCs, spectral multiplexing is utilized to create several simultaneous communication links with a single waveguide. WSSR technique is a combination of wavelength routing and spatial routing techniques.

In this chapter, we have so far discussed about various photonic elements, data conversion mechanism, different photonic switches, and routing of photonic data. In the following section, we will build on these basic elements and explain two generic photonic NoC architectures.

2.5 Analysis of generic photonic NoC architectures

One of the major drawbacks of photonic NoC architectures is their lack of ability to store communication data at intermediate stages. This limits photonic NoCs from using packet switched network's store and forward mechanism. In this section, we discuss two variants of photonic NoC architectures – circuit switched and wavelength arbited architectures.

Circuit switched photonic NoCs: To overcome the inability of data storage at intermediate stages, some of the photonic NoC architectures are constructed using circuit switching technique. In circuit switching, a path setup stage precedes the high speed data transfer stage. Once the path is setup, data transfer from source to destination takes place in an uninterrupted fashion [24]. Broadband WDM is used to achieve high-bandwidths between communicating nodes, by multiplexing data onto many parallel wavelengths.

Figure 2.9 shows a custom router used in circuit switched photonic NoC. The custom router has two main components - control router and data switch. The control router is electrical and it assists in setting up of end-end path for photonic data. The second component – data switch, carries the communication data from source to destination. As explained in Section 2.3, the data switch can be an all-photonic switch or an electro-photonic switch.

Circuit switching is achieved using different control messages such as 1) path-setup 2) path-acknowledge 3) path-blocked and 4) path-teardown. Source node sends a request message (path setup) and the destination node responds to this request. If the destination node is busy or if the intermediate stages are busy, then the source node receives a path-blocked message and cannot proceed with the transfer. Source nodes sends the data only upon receiving path-acknowledge message. When the data transfer is done, source node sends a path-teardown message to the destination node. This message assists in freeing up of the path.

The performance of photonic circuit switched networks in terms of network latency is better compared to electrical mesh and torus networks. Even the energy savings are high when the network size scales to more than 16 processing cores [39]. Some of the disadvantages of the circuit switched networks are – 1) high latencies due to path setup stage 2) achieving system fairness is difficult in circuit switched networks if the path conflict resolution schemes are not implemented correctly.

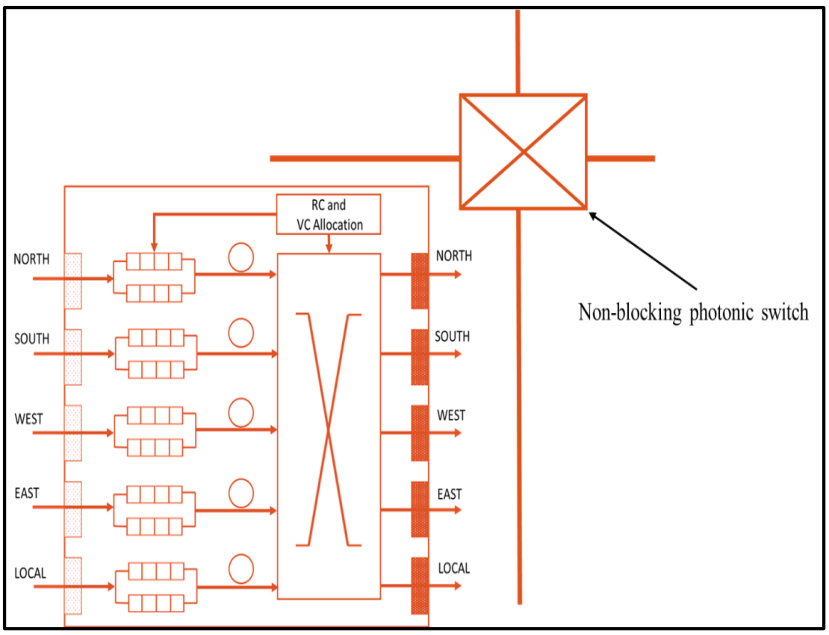


Figure 2.9: Custom router in a circuit switched network showing an electrical control router and a photonic data switch [24]

Wavelength arbited architectures: Wavelength arbited architectures make use of the fact that waveguides can support multiple wavelengths (WDM) and can selectively filter out some or all wavelengths to different destinations. Unlike circuit switched networks, wavelength arbited networks need not establish an end-end path prior to data transfers. An example wavelength arbited photonic NoC is as shown in Figure 2.10. Figure 2.10 shows a crossbar based photonic NoC architecture. As shown in the figure, in a crossbar based architecture, each source node modulates data on a different wavelength channel based on the destination node. Every destination on the receiving end utilizes different wavelength for receiving the data based on the source. Therefore each source-destination pair utilizes a dedicated wavelength channel for communication. Crossbar based architectures support huge bandwidths and have lower latencies compared to circuit switched photonic NoCs [21] - [23]. Compared to circuit switched architectures, crossbar based architectures will

need many wavelength support. For an N-node network, N × (N - 1) wavelengths are required for communication [24]. A detailed description of some of the crossbar based architectures is given in Chapter 4.

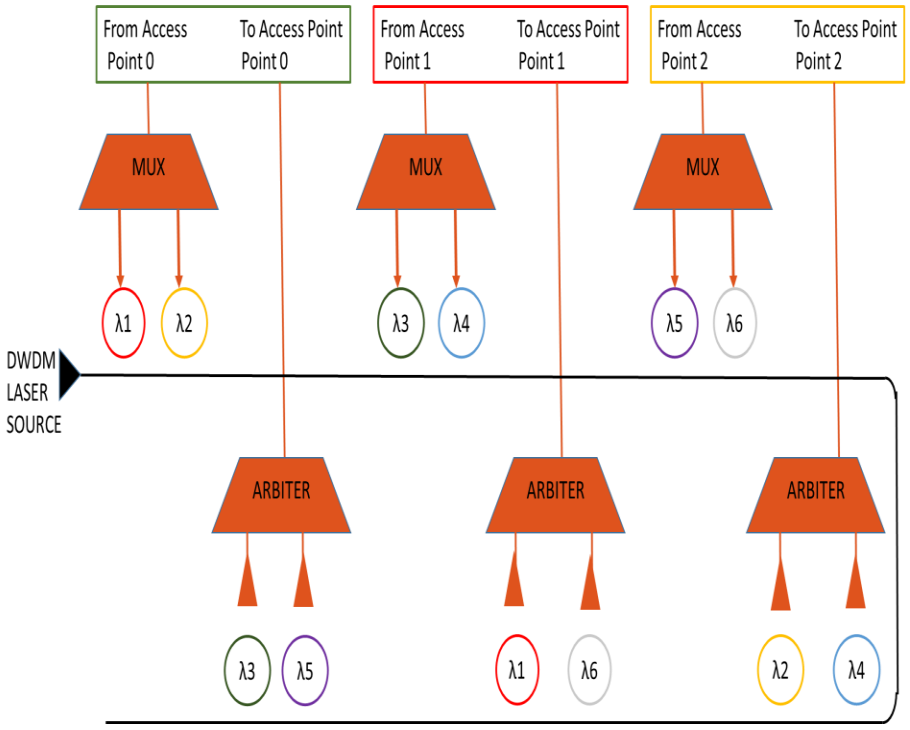


Figure 2.10: Wavelength arbited crossbar architecture connecting three nodes [39]

In this chapter, we described the basic elements used in photonic NoC communication. We provided a detailed explanation on the photonic data conversion mechanism and presented the basics of WDM. Also, we analyzed different flavors of photonic NoC architectures. In the following chapter, we will provide the problem statement of this thesis.

3 PROBLEM STATEMENT

Electrical networks-on-chip (NoC) provide significant improvements in terms of bandwidth, scalability, and power consumption compared to bus based architectures. However, future chip multiprocessors (CMP) will require a high-performance, low-latency, energy-efficient NoC to provide a base for communication substrate. Photonic NoCs are being considered as viable alternatives for electrical NoCs owing to several inherent advantages such as 1) high-bandwidth support due to wavelength division multiplexing (WDM) technique, 2) significantly low network latencies, as data flight time is essentially, the speed of light, 3) low dynamic power consumption, and 4) improved performance-per-watt [21]-[23]. Along with these advantages, recent advances in the silicon fabrication technology is another major driving factor to consider photonic NoCs for future CMPs [15] [16].

Several key metrics that determine the overall performance of a CMP system are:

- Throughput
- Latency
- Energy delay product
- Power consumption

Research in photonic NoC based CMP systems has mainly focused on optimizing one or more of these metrics to achieve enhanced performance. This thesis aims to optimize the aforementioned metrics to improve the overall system performance of future CMPs and it addresses the following challenges to realize high-bandwidth, low-latency, energy-efficient photonic NoC architectures:

- **Network topology design:** Some of the problems associated with existing photonic NoC architectures are 1) heavy interference of local and global traffic, 2) complex arbitration schemes for shared resources, 3) lack of hybrid (electro-photonic) architectures leading to costly data conversions for short distance transfers. This thesis addresses these issues by implementing a novel electro-photonic hierarchical topology – NOVA that 1) uses electrical wires for short distance transfers, 2) avoids interference among local and global traffic by taking advantage of its hierarchical nature 3) allows many simultaneous data transfers across the network.
- **Network latency and throughput:** Network latency is determined by the communication latency of packets traversing the network. This thesis aims at reducing the network latency by employing congestion free wavelength routing schemes for data communication. Wavelength routing schemes use static allocation of wavelengths to communicate with different nodes on the network and therefore do not require complex arbitration mechanisms to resolve contention issues. Another metric - throughput, which is a function of network latency, is defined as the total amount of work done over a defined number of cycles. Our work improves this metric by reducing the network latency and by employing 1) efficient switching and 2) hierarchical topology.
- **Static power dissipation in photonic network:** Photonic NoCs consume less dynamic power compared to electrical NoCs. Communication using photonic interconnects is subject to various losses such as 1) waveguide propagation loss, 2) ring-resonator through loss, 3) bending loss 4) coupler/splitter loss. These losses lead to high static power dissipation in photonic NoCs. This work aims at reducing the static power

dissipation of photonic NoC by decreasing the overall photonic hardware in the network.

- **Heterogeneity in system applications**: In modern systems, several applications will be running simultaneously, each with unique requirements of bandwidth. Some applications are compute intensive and are latency intolerant, spending most of their time in computing. Other applications are memory intensive that are latency tolerant, spending most of their execution time in communicating with memory. In this thesis, we perform experiments with heterogeneous workloads to mimic the requirements of modern systems.

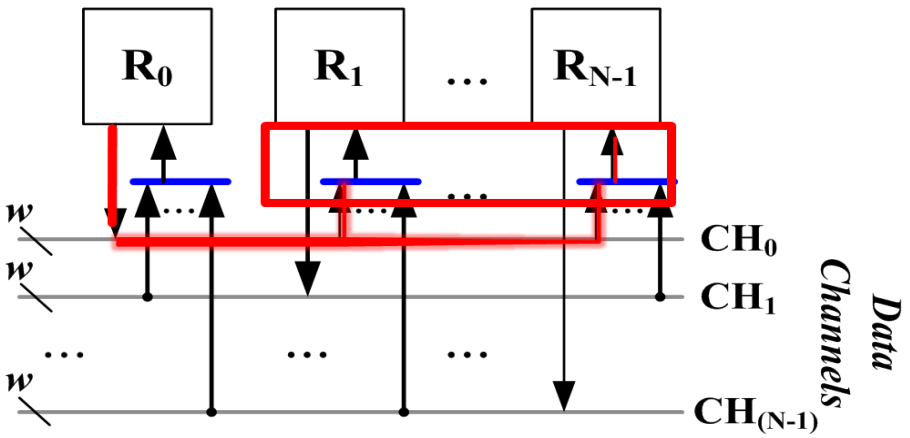
This thesis proposes a hybrid hierarchical photonic NoC – NOVA that addresses the abovementioned challenges leading to performance improvement in terms of system throughput, latency, energy-delay-product, and power.

4 RELATED WORK

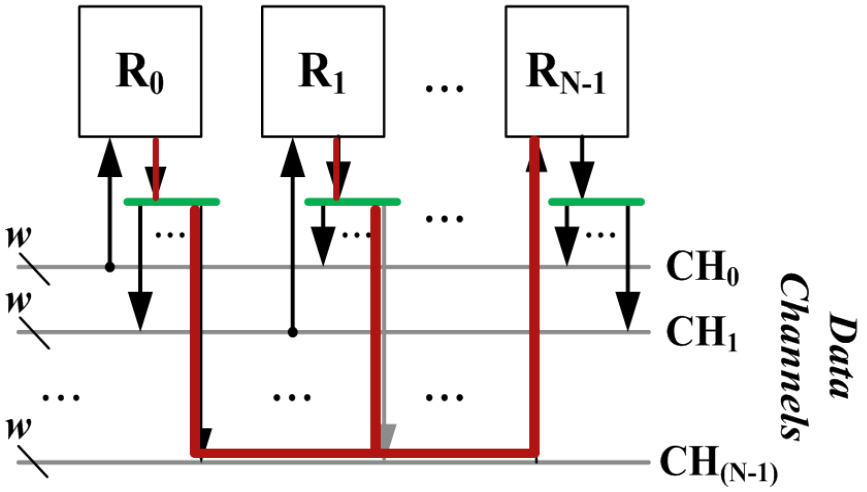
In this chapter, we first explain various configurations of a communication waveguide. We then describe several state-of-the-art architectures in the area of photonic networks-on-chip (NoC). We also make an effort to evaluate the pros and cons of these existing architectures. We describe various photonic communication protocols and end the chapter with a discussion on dynamic and reconfigurable architectures in photonic NoCs.

4.1 Various configurations of a communication waveguide

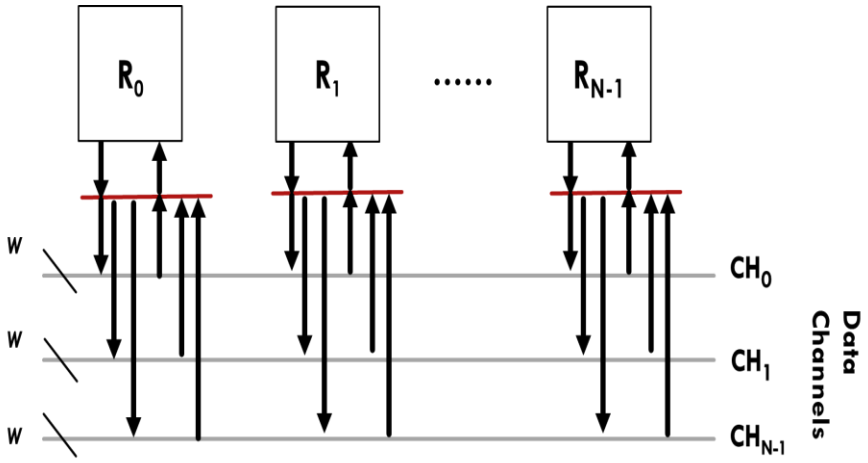
The choice of waveguide configurations in the network, define the communication protocols and arbitration schemes for crossbar based photonic NoCs. In this section, we discuss three different waveguide configurations. Figure 4.1 (a), (b), and (c) show single-write-multiple-read (SWMR), multiple-write-single-read (MWSR), and multiple-write-multiple-read (MWMR) configurations respectively.



(a)



(b)



(c)

Figure 4.1: N nodes connected using different waveguide configurations (a) SWMR (b) MWSR (c) MWMR

As shown in Figure 4.1 (a), in an SWMR configuration, a single node has write access over a particular data channel (waveguide) and all the other nodes have only read access on that channel. Each source node has its own waveguide and one set of modulators to write the data on the waveguide. Global arbitration is not required in such a configuration as there is no contention during the write stage. Figure 4.1 (b) shows MWSR configured data channels. In MWSR

configuration, multiple nodes of the network have write access over a particular data channel and a single node has read access over that channel. Each destination node has its own waveguide, on which every other node can transmit data using same wavelengths. Global arbitration is required to resolve any contentions during the write stage. In Figure 4.1 (c), MWMR configured data channels are shown. In MWMR configuration, a single data channel is shared across many nodes and all the nodes have both read and write access over this channel.

In the next section, we will describe three state-of-the-art photonic NoC architectures that are based on these configurations. We will also explain in detail about the arbitration schemes employed in these architectures.

4.2 Overview of photonic NoC architectures

The concept of photonic interconnects for on-chip communication was first introduced by Goodman et al. [41]. With the recent advances in fabrication and inherent advantages such as high-bandwidth support owing to dense wavelength division multiplexing (DWDM), several works in recent years have explored photonic NoC architectures for multi-core chip multi-processors (CMP) [21] [22] [23] [40] [42] [43] [44]. In this section, we summarize a representative subset of these works.

4.2.1 Corona [21]

Corona is an all-photonic, crossbar based CMP architecture, comprising of 256 general purpose cores, organized in 64 clusters (four-cores per cluster). Crossbar in Corona is configured in MWSR fashion. The crossbar enables a cache coherent design with uniform on-stack and off-chip memory communication latencies. Because it uses MWSR buses, all modulators must have

their OFF state (not injecting carriers) as OFF resonance, so that modulated signals from other access points can pass by on their way to the respective photodetector bank.

Clusters communicate through a photonic crossbar and the crossbar is managed using photonic tokens. Every cluster has a designated channel that is shared among address, data and coherence messages. The crossbars are designed such that any cluster can write to a given channel, but only a fixed cluster can read from it. By replicating such a MWSR channel 64 times, a fully connected 64 × 64 crossbar is formed in Corona. Each channel supports 256 wavelengths, using a bundle of four waveguides, with each waveguide supporting 64 DWDM. Light is sourced at a channel's home by a splitter that provides all wavelengths of light from a power waveguide.

A cluster that needs to send data to another cluster does so by modulating the light on the destination cluster's data channel. Modulation of data is done on both edges of the clock. Cluster bandwidth in Corona is 2.56 Tb/s and total crossbar bandwidth is 20 TB/s. Based on the geometric calculations, the propagation time of a photonic signal over the crossbar can be anywhere between one and eight cycles. To transfer a cache line of 64 bytes, the time taken can be a maximum of eight cycles. Global distribution of clock helps in avoiding the need for signal retiming at the destination.

Since a channel is shared among multiple clusters, an arbitration scheme that prevents two or more clusters from simultaneously sending data to the same destination, is required in Corona. A distributed, all-photonic, token-ring based arbitration scheme that allocates available channels to clusters is implemented on the crossbars in Corona. Its asynchronous acquire-and-release nature tolerates variability in request arrival time, message modulation time, and message propagation time. Only the cluster that acquires a token has the right of transfer. After the transfer is complete, the cluster then injects the token back into the channel allowing other

contenders to acquire the right of transfer. Each cluster will absorb and regenerate the grabbed token to ensure that the token remains optically sound even after many trips around the ring without any "takers".

A summary of the photonic hardware used in Corona is given in Table 4.1. The total photonic interconnect power is estimated to be 39W [21]. Corona provides huge bandwidth support, but suffers from high static power dissipation. The high static power dissipation is due to the presence of overwhelmingly large photonic hardware. The photonic layer complexity in Corona is high compared to other architectures [22]. Also, token-ring arbitration employed in this work can limit the throughput of the network and result in poor utilization of the channel [23].

Table 4.1: Photonic hardware in Corona [21]

Photonic Subsystem	**Waveguides**	**Ring Resonators**
Memory	128	16K
Crossbar	512	2048K
Broadcast	1	8K
Arbitration	2	8K
Clock	1	64
Total	644	~2080K

4.2.2 Firefly [22]

Firefly is another crossbar based, hierarchical electro-photonic NoC architecture. Local, small distance transfers in Firefly is done using electrical networks and global long-distance transfer using photonic links. Four processing cores are connected to a single electrical router to

form a concentration. Four routers are connected electrically, to form a *cluster*. All *intra-cluster* communication is done using electrical interconnects. All *inter-cluster* communication is done using global photonic crossbars. Crossbars in Firefly are configured in a custom SWMR fashion. Routers belonging to different *clusters* are connected via photonic crossbars, forming an *assembly*.

Crossbars in Firefly are implemented using reservation assisted SWMR (R-SWMR) configuration. Use of R-SWMR, helps in reducing the power consumption compared to basic SWMR configuration [22]. In basic SWMR, all the micro-ring resonators (MRRs) are "on" by default. Unlike SWMR, in R-SWMR, all the MRRs on the crossbar are "off" by default. In Firefly, a router first broadcasts a reservation flit, which contains the destination and packet-length information, to all the other routers in the *assembly*. When the destination router sees this broadcast flit, it switches on the MRRs on the corresponding data channel to receive the packet. As the MRRs are "off" on all the other routers in the *assembly*, they will not be coupling the laser energy – resulting in a point-to-point or unicast communication instead of expensive broadcast on the wider data channels.

The router micro-architecture in Firefly is slightly modified to support *inter-cluster* communication. The router has an electrical-photonic and photonic-electrical conversion module. If an *inter-cluster* packet is detected, it is first converted to photonic domain before transferring the data. Once the reservation phase is complete, the source router puts the data on the global photonic data channels. On the receiving side, the detectors are switched "on" for the duration of transfer and help in converting the photonic data, back to electrical domain.

A summary of the photonic hardware used in Firefly is given in Table 4.2. Firefly reduces the power consumption compared to Corona, however the architecture has high implementation

overhead and has no support for controlling the distribution of traffic among the electrical and photonic paths. In the next subsection, we will describe the Flexishare architecture [23].

Table 4.2: Photonic hardware in Firefly [22]

Photonic Subsystem	Waveguides	Ring Resonators
Routing	128	32768
Broadcast	8	1024
Arbitration	-	1024
Clock	1	16
Total	137	34832

4.2.3 Flexishare [23]

Flexishare is a MWMR based photonic NoC architecture. Flexishare decouples the allocation of buffers from channels by introducing an efficient token-stream technique for channel arbitration and credit distribution.

In the previously described architectures, Firefly and Corona – the number of channels required in the design is proportional to the radix of the crossbar to provide full connectivity. If the network traffic is unbalanced or if the network bandwidth usage is low, communication channels in Firefly and Corona will be under-utilized. To overcome the problem of under-utilization, Flexishare employs MWMR configuration on its data channels, where on the transmitting side, a source can transmit (write) its packet on any of the available channels and on the receiver side, all the receivers (destinations) are equipped to listen (read) on all the channels. The data channels are either upstream or downstream and both are single round channels. Single

round channels are photonic waveguides that do not coil around and have reduced length compared to double round channels.

On the MWMR data channels, Flexishare implements a token-stream arbitration. Similar to token-ring arbitration of Corona, photonic tokens are used to arbitrate the data channels of Flexishare. In this architecture multiple tokens are injected serially to create a "stream" of tokens. Unlike, token-ring arbitration of Corona, a token in the token-stream represents the right to modulate on the corresponding data channel. Based on a static priority, the nodes of the network are allowed to grab the token from the channel and modulate the data in the adjacent data slot. The default token-stream arbitration in Flexishare is single-pass token-stream where each token is dedicated to a unique node of the network. To overcome the starvation problem in the single-pass token-stream, Flexishare also proposes a two-pass token stream arbitration technique. In the first pass, each token is dedicated to a unique node and in the second pass, all the nodes are free to grab any available token. Flexishare assumes a 3-D integration support such that a separate photonic die is stacked on top of a processor die. Such stacking enables specialized fabrication process for the photonic die and more freedom for waveguide routing. The photonic hardware requirement for the implementation is as shown in Table 4.3.

Table 4.3: Photonic hardware in Flexishare [23]

Photonic Subsystem	**Waveguides**	**Ring Resonators**
Routing	256	512K
Clock	1	64
Total	257	~512K

We have implemented all the three works described above for comparison with our proposed architecture. Unlike the works described above, we use a wavelength based routing scheme which requires no arbitration, supports extremely low latencies, and does not incur the power losses due to additional arbitration circuitry. We reduce the static power dissipation compared to above architectures by reducing the overall photonic hardware in our architecture. As broadcasting is a power hungry mechanism compared to unicast messages, we minimize the usage of broadcasting the in our architecture [22].

There are several other works that have proposed novel photonic NoC architectures, each of them aiming at optimizing one or more of the following parameters – power consumption, latency, bandwidth, energy dissipation [39] [43] [44] [45]. In [46], latency response and power consumption parameters have been improved using photonic ring interconnects. In PROPEL, electrical switching is used for routing and flow control and photonic interconnects are used only for long distance communication [47]. PROPEL implements a photonic crossbar based on wavelength routing where the number of distinct wavelengths is equal to the maximum number of nodes in the network.

4.3 Protocols for photonic NoCs

While significant efforts have been carried out in proposing new photonic NoC topologies, some of the recent works have focused on other optimizations in a photonic NoC such as efficient arbitration techniques, flow control techniques, and dynamic reconfiguration techniques [48] [49] [50]. In this section, we will explain some of the research efforts carried out in these areas.

4.3.1 Arbitration techniques

Time division multiplexing (TDM) based arbitration and distributed control of photonic switches has been proposed to overcome the problems such as lack of fairness, high contention, and performance degradation in circuit switched and token-ring based architectures [38]. In this architecture, switches in the network are configured to allow communication between one or more pairs of access points (nodes on the network) during a specified amount of time. Using control registers and global clock each switch can be made aware of its correct current configuration for any given time slot. This allows the control of switches to be completely distributed. The scheduling of node accesses to network resources is done statically at design time or at the application startup time.

In this work, two different techniques are proposed. In the first technique, nodes repeatedly cycle through every time slot and if a node has data to be sent, it waits for the correct time slot. If a node has multiple data to different destinations queued up, it can send them out of order. In the second technique, the total number of unique time slots required for arbitration is decreased by dividing the end-end transmission of first technique into two stages. In the first stage, X-dimension transmission is completed and the data packet is converted to electrical and stored in buffers at intermediate routers. In the second stage the Y-dimension transmission is completed. This improved arbitration scheme helps in conserving energy consumption [38].

Another interesting arbitration technique has been proposed in [51]. The work uses a dynamic scheduling scheme to coordinate the access of the shared photonic channels. In this work, efficient channel utilization without the use of electrical arbitration schemes is achieved. On the receiving (destination) side, once the arbitration flag is detected, one of the following three actions is taken by the receiver. If the forthcoming packet is destined for itself, it will

enable all the rings for the indicated duration of the packet. If the forthcoming packet is not destined for this receiver, then it will switch off all its rings. If there is collision detected, (multiple senders trying to send to the same receiver) then the receiver will enter into dynamic scheduling phase. On the transmitting side, the sender waits for any on-going packet transmissions to complete. After this, the sender modulates a copy of the arbitration flag to the appropriate arbitration wavelengths for each of the destination nodes in the network.

4.3.2 Flow control techniques

Flow control techniques for photonic NoCs have been proposed in several works [52] [53]. While arbitration among the shared channels helps decide the winner among the contenders, flow control information is required by the winning node to make a decision on sending the data over to the destination node. Flow control techniques avoid costly retransmissions and dropping of packets due to unavailability of buffer space at the destination side.

In [52], the destination node emits credit-filled tokens to communicate the available number of buffers to the source nodes. A source node that removes these credit filled tokens is guaranteed of equivalent credits (buffer space) at the destination. The source node marks its reservation by decrementing the credits and re-injecting the decremented token. The destination node upon receiving the token from source node updates the value of credits based on the available entries since the token's last visit. The destination node re-sends this updated token and the process continues with other source nodes. In [53], a circuit switch based flow control is used to indicate the availability of the buffer space at the destination side.

4.4 Dynamic reconfiguration of photonic NoCs

In this section, we will highlight some of the works on the effects of temperature and process variation in photonic NoCs [54] [55]. The absolute temperature and temperature fluctuations across the chip have been major concerns in chip design and packaging because high on-chip temperatures can cause performance degradation and even functional failures in CMOS circuits. The basic photonic elements such as waveguides, switching elements, and transmitters are all sensitive to thermal variations and show degradation in their performance with high thermal variations. In [54] several techniques have been proposed to reduce the adverse effects of temperature variations in photonic elements. The work suggests methods such as 1) reducing the temperature dependence of ring resonators, 2) increasing the 3-dB bandwidth of photonic switching elements by parallel coupling of MRRs, and 3) reducing the number of switching stages in photonic NoC architectures.

In [55], several techniques to tolerate the impact of process variation have been proposed. One such technique is to overprovision the MRRs, creating more opportunities for selectable and correctable rings. Another technique is to provide flexible wavelength assignment for network nodes by which nodes that are capable of using certain wavelengths are assigned only those wavelengths even in presence of process variations.

In summary, research in photonic NoC is spread across different domains. Every work in photonic NoC research aims at optimizing several metrics and characteristics of photonic on-chip communication. In this chapter, we have described several state-of-the-art architectures in detail along with their advantages and disadvantages. We also presented the details of various communication protocols implemented in photonic NoCs. In the following chapter, we will present the details of the proposed architecture.

5 DESIGN AND ANALYSIS OF ENERGY-EFFICIENT HIERARCHICAL ELECTRO-PHOTONIC NETWORK-ON-CHIP ARCHITECTURES

In Chapter 1, we discussed some of the key issues and challenges associated with electrical networks-on-chip (NoC) for chip-multiprocessor (CMP) architectures. In Chapter 2, we explained the fundamentals of photonic interconnects and described the basics of switching and routing in photonic communication. We presented several state-of-the-art photonic NoC architectures in Chapter 4. In this chapter, we will illustrate and analyze our proposed hierarchical electro-photonic architectures – NOVA. Our proposed architecture is inspired by the following challenges in a CMP network design:

- High interference among local and global traffic leading to an inefficient, low throughput, and high latency network.
- High contention among shared resources and inefficient arbitration schemes leading to starvation in the network.
- High static power consumption in existing photonic architectures.

To overcome the problem of high interference among local and global traffic we implement a hierarchical architecture that has three communication levels. This hierarchical nature of the architecture not only avoids the interference among different network requests, it also reduces the hop count of a communication packet and increases the overall network efficiency. High communication locality exhibited by applications is also an encouraging factor to consider

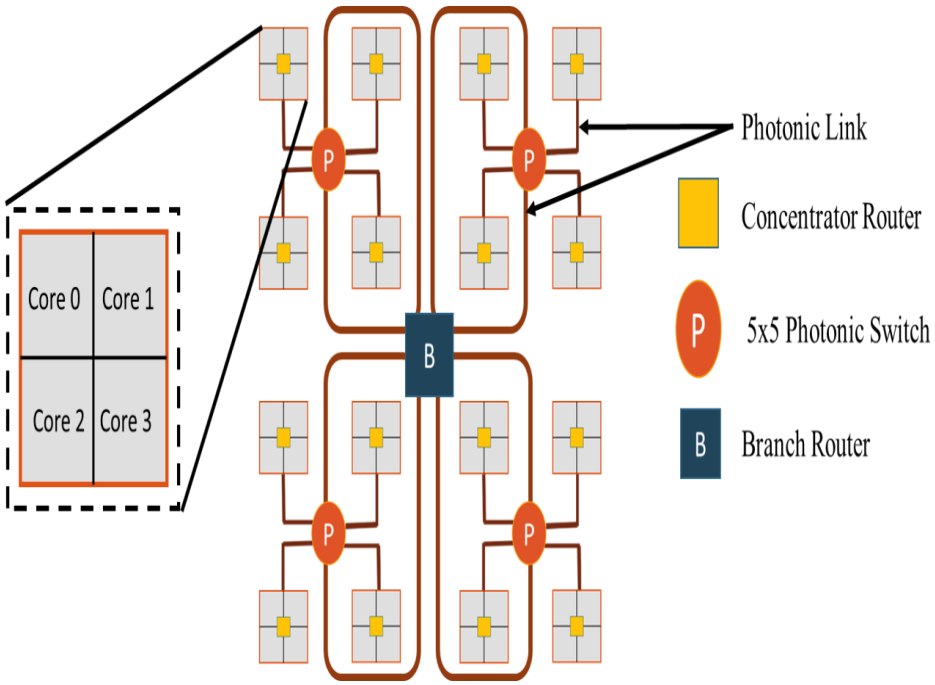
hierarchical networks. Another advantage of hierarchical architecture is it leads to a smaller photonic diameter compared to non-hierarchical architectures.

To overcome the problem of high contention among shared resources, we implement a wavelength routing scheme using micro-ring resonator (MRR) based photonic switches. Every source-destination pair is assigned with a unique set of wavelengths. To support high bandwidth requirements, we use broadband MRRs for wavelength routing. Using broadband based MRR helps in switching multiple wavelengths per MRR. To reduce the static power consumption of the proposed architecture, we minimize the photonic hardware usage, compared to other state-of-the-art photonic architectures [21] [22] [23]. Also to minimize the thermal heating power, we use passive broadband MRRs in all the photonic switches [37].

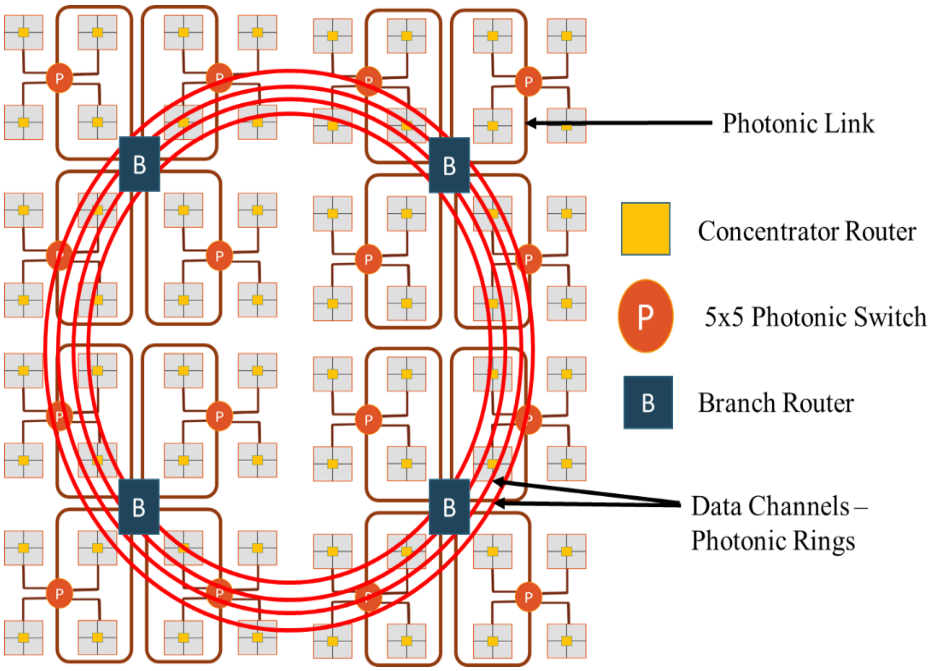
The three levels of communication hierarchy in NOVA are:

- A broadband MRR based photonic switch at the first level.
- Silicon-nitride arrayed waveguide grating (AWG) based hybrid router at the second level.
- Configurable photonic rings at the last level of hierarchy.

In the following section, we describe the system level architecture of NOVA, explaining in detail about the 1) arrangement of processor cores, 2) micro-architecture of the concentration containing four processor cores and a hybrid router. We then provide a detailed discussion on the different modules involved in three levels of hierarchy and illustrate their interactions. We finally present the architectural variations that we carried out by considering different photonic switches and multiple configuration schemes of the photonic rings.



(a)



(b)

Figure 5.1: Architectural view of (a) 64 and (b) 256 core NOVA system

5.1 System level architecture

NOVA is a scalable and configurable architecture supporting multiple network sizes. We have targeted 64-core and 256-core CMP in this work. Figure 5.1 shows a detailed architectural view of both 64 and 256-core NOVA system. In both configurations, a group of 16 cores have a dedicated on-chip memory controller. NOVA is a hybrid architecture that uses electrical links for short distance transfers and photonic interconnects for long distance transfers.

Figure 5.1 (a) shows a 64-core system with 16 concentrators connected using photonic switches at the first level of hierarchy and AWG based router (*branch router*) at the second level. Every concentrator has a group of four processing cores connected using a hybrid router. Figure 5.1 (b) is a 256-core system with 64 concentrators. Similar to the 64-core system, even this has photonic switches at the first level, *branch routers* at the second level. Extending on this, a 256-core system has photonic rings at the last level of hierarchy. While the 64-core system uses a single off-chip mode-locked laser as the light source, the 256-core system uses four off-chip mode-locked lasers [21].

NOVA uses four way concentration, where a single hybrid router is attached to four processing cores. Each concentrator is given a unique ID and every processor core in the network is also associated with a unique ID. A communication packet is thus identified by two unique IDs, one defining the processor it originated from and the second stating the concentrator to which the processor is attached. NOVA makes use of traditional electrical links for transferring data between cores, which belong to the same concentrator. Every concentrator has a 5×5 hybrid router, which uses four of its input/output pairs to connect four processing cores and uses its multi-buffered fifth pair of ports to connect to the photonic switches that assist in inter-concentrator communication [22]. Every concentrator module is also equipped with modulator

drivers, modulators, photo-detectors and trans-impedance amplifiers for photonic-electrical and electrical-photonic conversion. As explained in Chapter 2, all these photonic elements assist in modulation and demodulation of photonic data. A more detailed description about the hybrid router is given in the next section.

Photonic communication is used if the destination core is not in the same concentration as that of source core. Four concentrators are connected using a photonic switch to form a *minor-quadrant* in NOVA. Essentially a *minor-quadrant* has 16 processing cores. This forms the first level of hierarchy in the network and can function independent of other *minor-quadrants* for *local* data transfers. Four *minor-quadrants* are connected using a *branch router* forming a *major-quadrant* in NOVA. This forms the second level of communication in the hierarchy. Finally, the photonic rings form the last level of communication, assisting packet movements between different *major-quadrants.* These three levels form the backbone of our communication network. Interference among local and global traffic is avoided in NOVA owing to its inherent hierarchical nature.

NOVA is a highly structured network that allows all the photonic switches in the network to function and facilitate the transfer of *local* data independently. Photonic switches in NOVA use wavelength routing for transfer of data. As described in Chapter 2, wavelength routing uses static mapping of wavelengths for a source-destination pair. This avoids any contention among communication packets going to different destinations and therefore requires no arbitration schemes. Since NOVA is not a circuit switched network, no time is wasted in setting up of end-end path before the data transfer. Wavelength routing simplifies the network routing scheme and unlike a circuit switched network, saves the additional hardware required for control logic

circuitry used in path setting [23]. The simplicity of the routing scheme helps in achieving extremely low latencies compared to other photonic architectures [38] [40].

Any photonic NoC topology requires careful consideration of waveguide layout, in absence of which there might be increased crosstalk in the network. Waveguide crossings in NOVA are restricted to only within photonic switch structures and are avoided in photonic links that connect them. We have considered losses to due waveguide crossing in NOVA and have faithfully modeled these losses in our simulation.

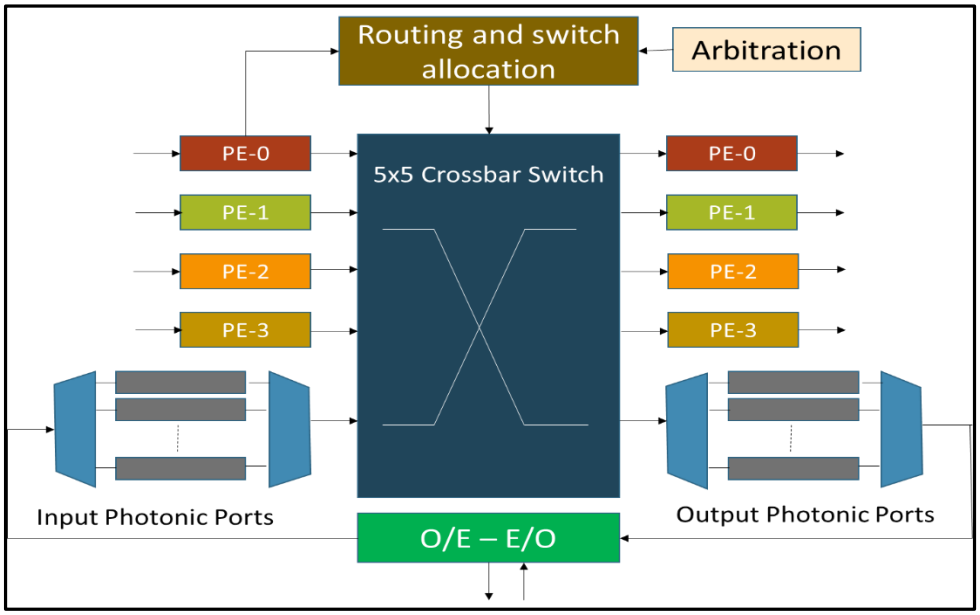


Figure 5.2: Micro-architecture of the 5×5 hybrid router showing four ports connected to local processing elements (PE). Fifth port is a multi-buffered port connected to a photonic switch.

5.2 Hierarchical communication levels - NOVA

In this section, we first describe the hybrid router inside a concentrator module of NOVA. We then dive into the details of other modules, which are part of different hierarchical levels. While describing the modules' micro-architecture, we also explain the routing and arbitration schemes employed at every hierarchical level and illustrate the interaction between levels.

5.2.1 Hybrid router

Figure 5.2 shows a detailed view of the 5×5 hybrid router. It is a four stage pipelined router with the following pipeline stages: 1) buffer write/route computation, 2) region validation/switch allocation, 3) switch traversal, and 4) link traversal. The first four ports of the router are connected to four processing elements (cores) and the fifth port is used for communication with cores that belong to a different concentrator. As shown in the figure, the photonic ports are connected to O-E and E-O conversion module that does the wavelength assignment and converts packets between electrical/photonic domain. A basic round robin arbitration scheme is used for allocating the output ports for intra-concentrator communication. For inter-concentrator communication, because wavelength routing is used, all four processing cores can send the data simultaneously, provided no two cores are sending a data packet to the same destination concentrator.

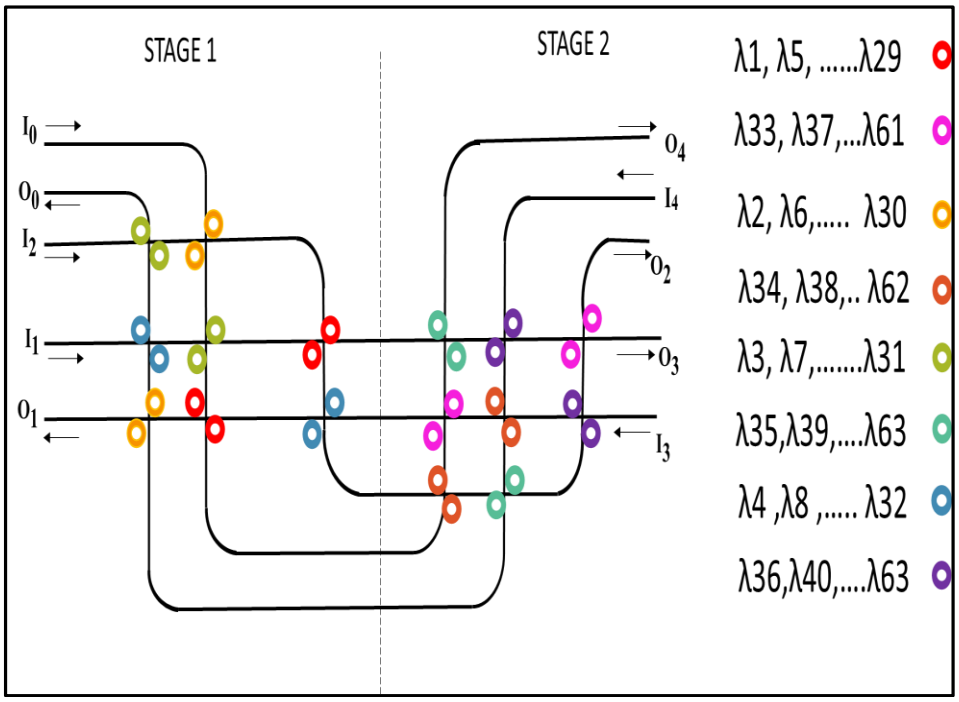


Figure 5.3: A detailed view of photonic switch in NOVA. Every input/output pair in the switch supports 16 wavelengths

5.2.2 Photonic switch – 1st level hierarchy

Figure 5.3 shows a detailed view of the photonic switch used in NOVA. It is a 5×5 switch with four pairs of input-output ports connected to four concentrators and the fifth pair connected to *branch router*. The photonic switch is a modified version of the wavelength router [37]. Unidirectional waveguides form the input/output ports of the switch. U-turns are not allowed in this photonic switch. The photonic switch allows simultaneous data transfers in all five directions and requires no arbitration during the communication process. This is a non-blocking switch and uses passive broadband MRRs that require less amount of tuning power [42]. Broadband MRRs are used to reduce the total number of micro-ring resonators in the network. Each broadband MRR is capable of switching eight wavelengths simultaneously [56]. Usage of broadband MRRs in the photonic switch has two advantages 1) high bandwidth support and 2) reduction in number of stages of the photonic switch which in turn reduces the waveguide crossings.

For a 256-core NOVA there are 16 photonic switches and in a 64-core NOVA, the network consists of four photonic switches. Figure 5.3 shows the basic version of photonic switch where every input/output pair supports switching of 16 wavelengths. The switch consists of two stages with each stage supporting eight wavelengths. Table 5.1 shows the wavelength assignment between every pair of input/output ports in the switch. Since U-turns are not allowed in the switch, respective columns in the table are marked as not applicable (N/A). The input/output ports of the photonic switch that are connected to the *branch-router* are configured to support 16/32/64 wavelengths. More details on this architectural variation is presented in the next section.

To facilitate photonic transfers, every concentrator in the network supports 64 wavelength dense wavelength division multiplexing (DWDM). If a packet currently residing in the hybrid

router is destined to a processor core belonging to a different concentrator, then the hybrid router allocates the photonic output port to the corresponding packet. It then converts the packet to photonic domain and sends it over the photonic switch to the destination processor core. *Branch router* is used only if the destination concentrator of the packet falls in a different *minor-quadrant* compared to packet's source concentrator.

Table 5.1: Wavelength assignment in the photonic switch

Input / Output	**O_0 (λ)**	**O_1 (λ)**	**O_2 (λ)**	**O_3 (λ)**	**O_4 (λ)**
I_0	N/A	N = 1, X = 0; λ = (N + (X * 4)) $\forall$ X $\leq$ 15	N = 2, X = 0; λ = (N + (X * 4)) $\forall$ X $\leq$ 15	N = 3, X = 0; λ = (N + (X * 4)) $\forall$ X $\leq$ 15	N = 4, X = 0; λ = (N + (X * 4)) $\forall$ X $\leq$ 15
I_1	N = 4, X = 0; λ = (N + (X * 4)) $\forall$ X $\leq$ 15	N/A	N = 1, X = 0; λ = (N + (X * 4)) $\forall$ X $\leq$ 15	N = 2, X = 0; λ = (N + (X * 4)) $\forall$ X $\leq$ 15	N = 3, X = 0; λ = (N + (X * 4)) $\forall$ X $\leq$ 15
I_2	N = 3, X = 0; λ = (N + (X * 4)) $\forall$ X $\leq$ 15	N = 4, X = 0; λ = (N + (X * 4)) $\forall$ X $\leq$ 15	N/A	N = 1, X = 0; λ = (N + (X * 4)) $\forall$ X $\leq$ 15	N = 2, X = 0; λ = (N + (X * 4)) $\forall$ X $\leq$ 15
I_3	N = 2, X = 0; λ = (N + (X * 4)) $\forall$ X $\leq$ 15	N = 3, X = 0; λ = (N + (X * 4)) $\forall$ X $\leq$ 15	N = 4, X = 0; λ = (N + (X * 4)) $\forall$ X $\leq$ 15	N/A	N = 1, X = 0; λ = (N + (X * 4)) $\forall$ X $\leq$ 15
I_4	N = 1, X = 0; λ = (N + (X * 4)) $\forall$ X $\leq$ 15	N = 2, X = 0; λ = (N + (X * 4)) $\forall$ X $\leq$ 15	N = 3, X = 0; λ = (N + (X * 4)) $\forall$ X $\leq$ 15	N = 4, X = 0; λ = (N + (X * 4)) $\forall$ X $\leq$ 15	N/A

5.2.3 Branch router – 2nd level hierarchy

A detailed view of the AWG based *branch router* is as shown in Figure 5.4. *Branch router* in NOVA is used for inter *minor-quadrant* transfers. It is a 4×4 router connecting four photonic switches. AWGs are used at every input port of the *branch router* and help in demultiplexing of incoming DWDM photonic signals. The branch router has 16 input buffers, four for each *minor-*

quadrant. As shown in Figure 5.4, there are four sets of multiple buffers in the router and each set has four buffers one for each concentrator inside a *minor-quadrant*. Following sections describe in more detail about the different modules of *branch router* and explain the routing process involved:

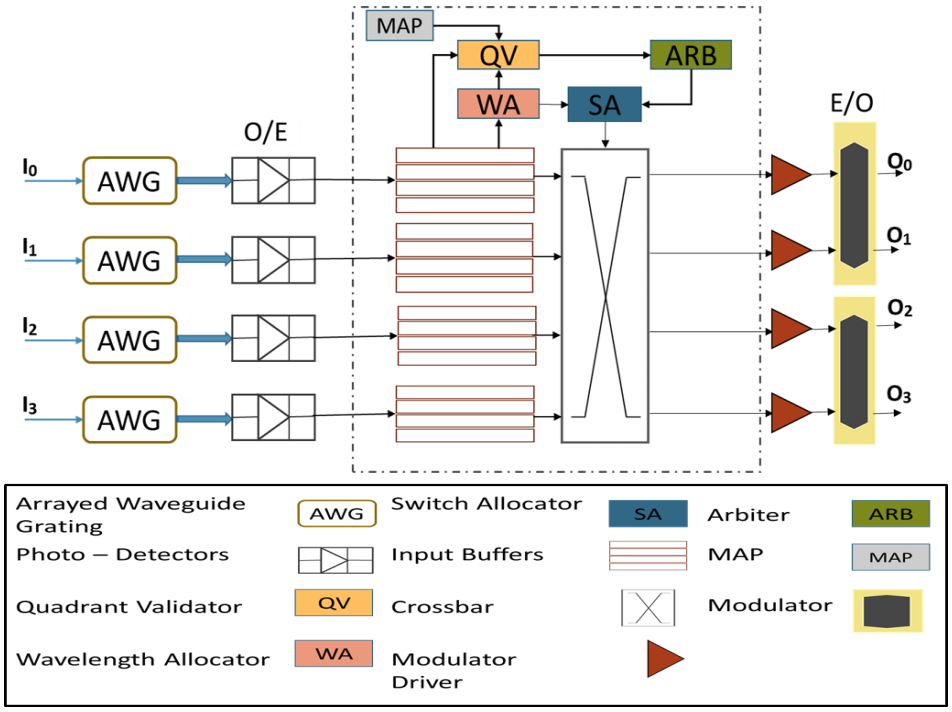


Figure 5.4: Detailed architecture of branch router in NOVA

AWG: AWG is a multi-wavelength receiver fabricated on silicon. A logical view of an AWG is as shown in Figure 5.5. AWG is a device which can separate (demultiplex) or combine (multiplex) photonic signals, encoded using different wavelengths [57]. It is usually built as a planar light-wave circuit. The light coming from an input waveguide first enters a multimode waveguide section, then propagates through several single-mode waveguides to enter a second multimode section, and finally onto the output waveguide. Wavelength filtering in AWG is based on an interference effect and the different photonic

path lengths in the single-mode waveguides. Silicon-nitride based AWGs have high fabrication tolerance and provide high bandwidth support [58]. Owing to its advantages, we have used a low loss silicon-nitride AWG, integrated with hybrid silicon photo-detectors in NOVA.

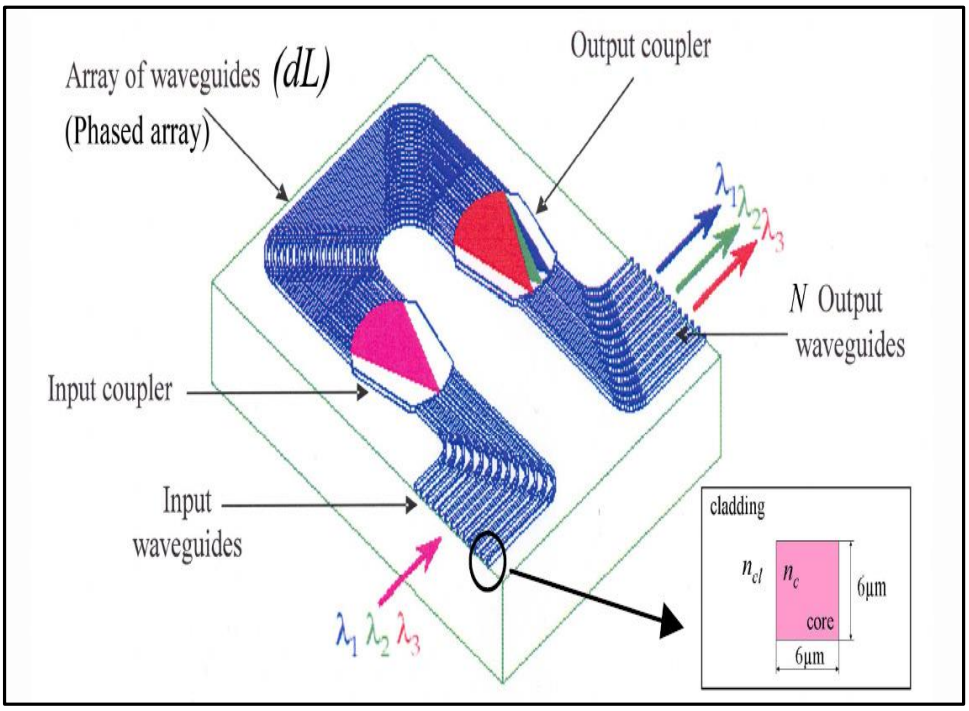


Figure 5.5: Structure of an arrayed waveguide grating. Adopted from [58]

AWG has been used in our architecture to simplify the process of demultiplexing and allotting an input buffer for incoming flits. A photonic switch that is connected to the *branch router* can route data from its four different concentrators. Because of wavelength routing, different data streams occupy the DWDM bandwidth and are being transferred to the *branch router* simultaneously. In the absence of the AWG device, there has to be an electrical circuit which decodes the incoming flits and stores them into respective input buffers without misaligning them. AWG integrated with series of silicon photo-detectors

helps in avoiding the additional electrical circuit for directing the incoming flits to respective input buffers. If the incoming flits get misaligned, then routing them to destined concentrators becomes highly difficult. Therefore it is imperative to maintain the flit arrival order in the router and AWG helps in achieving the same.

Routing: Below are the modules of branch router that assist in the routing process:

- Arbiter
- Switch allocator
- Wavelength allocator
- Quadrant validator
- Map table

As with the first level hierarchy, we adapt wavelength routing mechanism at the *branch router* also. *Branch router* converts the incoming photonic flits to electrical domain and stores them in the local buffers. The flits present inside different buffers request access for output ports of the *branch router* to move towards their respective destinations. The router has four output ports connected to four photonic switches and each output port supports 64 DWDM.

As soon as a flit arrives at the input buffer, route computation process begins and is done using two modules 1) quadrant validator 2) wavelength allocator. The map table inside the router is six bits wide and it maps 16 different concentrators to four *minor-quadrants* (in a 64-core CMP). First the source and destination information present in the header flit is read and the quadrant validator module then allocates a respective output port to the flit using the information in map table. Meanwhile, the wavelength allocator module allocates a particular set of wavelengths based on the destination concentrator ID.

Once the routing computation is done, the arbiter module scans all the requests in the current cycle and detects any contention that may arise. The only case a contention can occur is when two different flits are trying to reach the same destination concentrator. Contention does not arise if the flits are going to different concentrators, even though these concentrators are a part of same *minor-quadrant*. In case of contention between two flits, a round robin arbitration scheme is employed to resolve the contention.

Once a particular output port and respective wavelengths are allocated to a flit, and there are no other flits contending for the same destination concentrator, then the switch allocator module allocates the switch to the requesting input port. The electrical flits are then converted to photonic domain using modulator drivers and modulators, present inside the router. The converted flits are sent to the photonic switch on the destination side, which then routes the flits to respective concentrators. The whole process of routing the incoming flits to different destinations takes 4-6 cycles inside the *branch router* that includes photonic to electrical conversion, wavelength and quadrant validation, arbitration and switch traversal. For a 64-core NOVA, only one *branch router* exists in the network. In a 256-core NOVA, there are four *branch routers*. The traffic that is intra *major-quadrant* does not interfere with the global traffic in a 256-core NOVA system. The hierarchical nature of NOVA results in high throughput due to reduced congestion among network requests and reduced interference of local and global traffic compared to non-hierarchical architectures.

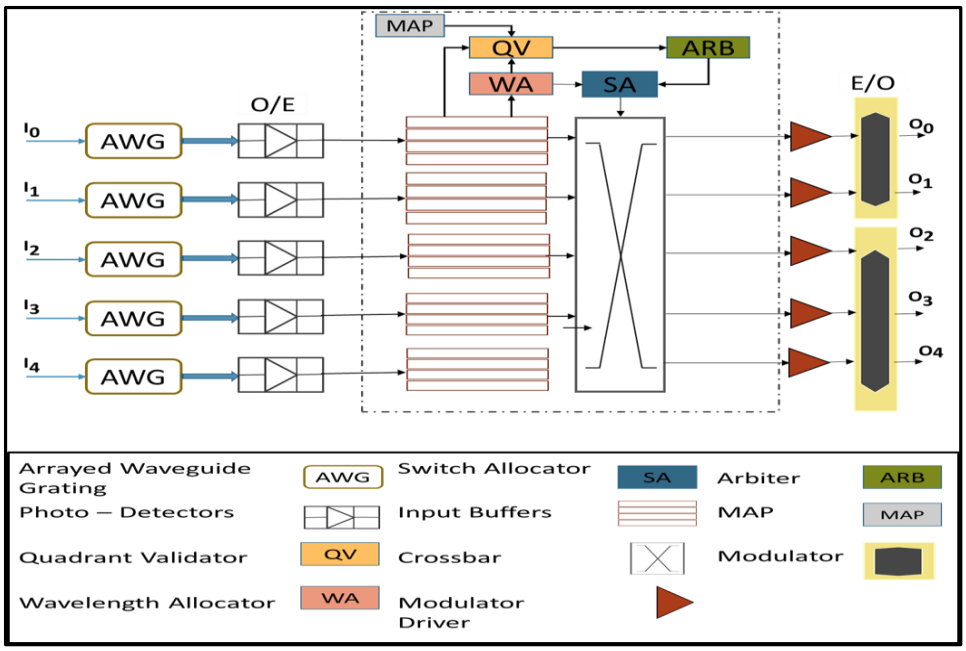


Figure 5.6: Modified branch router with additional ports and buffers

5.2.4 Photonic rings – 3rd level hierarchy

For the 256-core system, we extend the 64-core NOVA architecture and add another level of hierarchy – photonic rings. These photonic rings help in transfers between two *major-quadrants.* In a 256-core CMP, the *branch router* is modified to support the traffic moving across *major-quadrants*. The modified *branch router* is as shown in Figure 5.6. As shown in the figure, additional pair of ports are added to the router to handle the global traffic. The number of buffers in the router is also increased to accommodate the additional traffic. Map table inside the *branch router* is slightly modified for a 256-core NOVA system. Additional bit is added to the map table to indicate inter *major-quadrant* transfers. The information in the map table is used by the quadrant validator module to allocate the fifth output port that connects to the photonic rings.

A wavelength routing based single write multiple read (SWMR) waveguide configuration is implemented on the photonic rings of NOVA. Usage of SWMR avoids global arbitration on the rings. NOVA has four photonic rings configured using wavelength based SWMR (W-

SWMR) scheme. In this configuration, on a particular photonic ring, one of the *branch routers* writes its data and the remaining *branch routers* read from this ring. Figure 5.7 shows the working of W-SWMR arbitration scheme in NOVA. In the figure, *branch router of* quadrant 1 (Q1) is the writer and all other routers are readers.

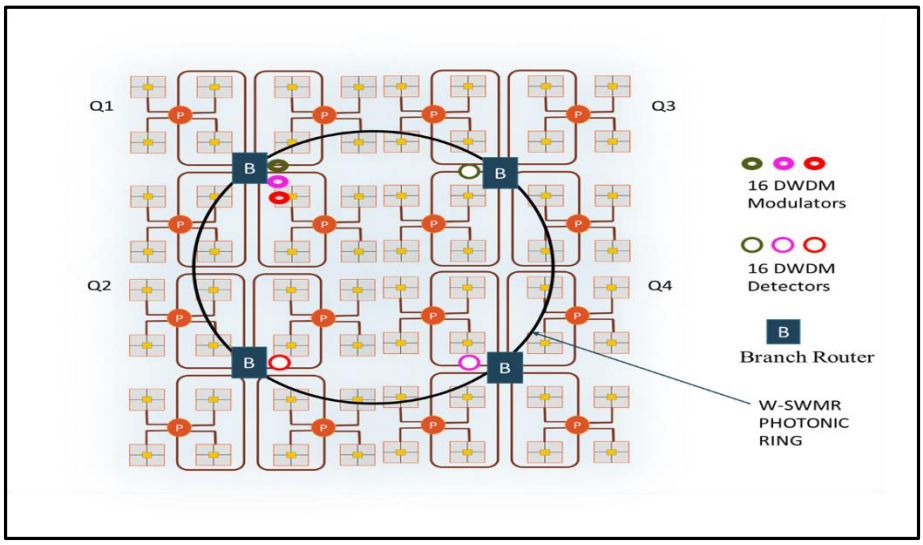


Figure 5.7: Working of W-SWMR. Only one ring is shown for brevity

Since this is a wavelength based SWMR, pre-determined wavelengths are assigned to each reader on the ring. A flit currently residing in the buffer of a *branch router*, whose destination concentrator belongs to a different *major-quadrant*, is assigned the fifth output port of the router. The fifth port connects to the photonic rings at the last level in NOVA. The arbiter in the *branch router*, checks for any contention between flits inside different buffers. Since wavelength routing is used on the rings, a contention can occur only if two flits are destined to same *major-quadrant*. In such a scenario, a round robin arbitration scheme is employed to resolve the contention. A maximum of three input buffers can win the arbitration for transfer via photonic rings, provided the flits inside those three input buffers are destined to different *major-quadrants*. The detectors at the receiving side of the ring are in "off" state by default. To switch on the detectors at the

receiving side, a wake-up signal is sent by the writer, following which the data is sent to the reader using the assigned wavelengths. Once the receiver detects a tail flit of the incoming packet, it turns off its detectors in the following cycle to save power consumption.

In this section, we explained the micro-architecture of all the modules involved in NOVA. We presented the routing and arbitration schemes implemented at different hierarchical levels. In the following section, we will discuss the architectural variations of NOVA.

5.3 Architectural variations of NOVA

In this section, we describe the architectural variations of NOVA that we carried out to arrive at the optimal configuration of the architecture for a given set of input applications and CMP platform. These configurations enabled interesting tradeoffs between performance and power consumption of the proposed architecture.

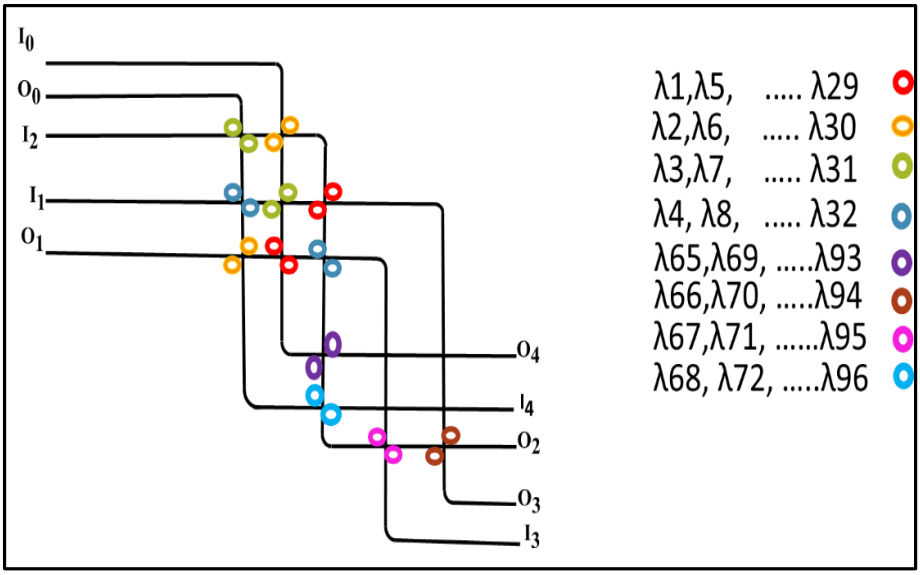


Figure 5.8: Modified photonic switch. Port 2 that is connected to the branch router can send and receive 32 wavelengths simultaneously.

5.3.1 Photonic switch variation

In the previous section, we presented the micro-architectural details of the photonic switch at the first level of hierarchy in NOVA. In this section, we will discuss the modified version of the baseline photonic switch. As we ran various workloads from PARSEC benchmark on our architecture, we observed that inter *minor-quadrant* traffic was higher than the intra *minor-quadrant* traffic [59]. To meet the high inter *minor-quadrant* traffic requirements, we re-designed our baseline photonic switch and added more broadband MRRs on the pair of ports that were connected to the *branch router*. Figure 5.8 shows a detailed architectural view of the modified switch. As shown in the figure, modified switch supports 32 wavelengths in the pair of ports (port - 2) connected to the *branch router*. With this modification, all the concentrators connected to the *branch router*, send and receive data from it, using 32 wavelengths. Table 5.2 shows the wavelength allocation with additional wavelength support in port - 2. Figure 5.9 shows another version of the baseline switch with 64 wavelength support for the pair of ports connected to branch router. The additional wavelengths helped in reducing the latency of the data packets to/from the branch router, which in turn increases the overall bandwidth of the system.

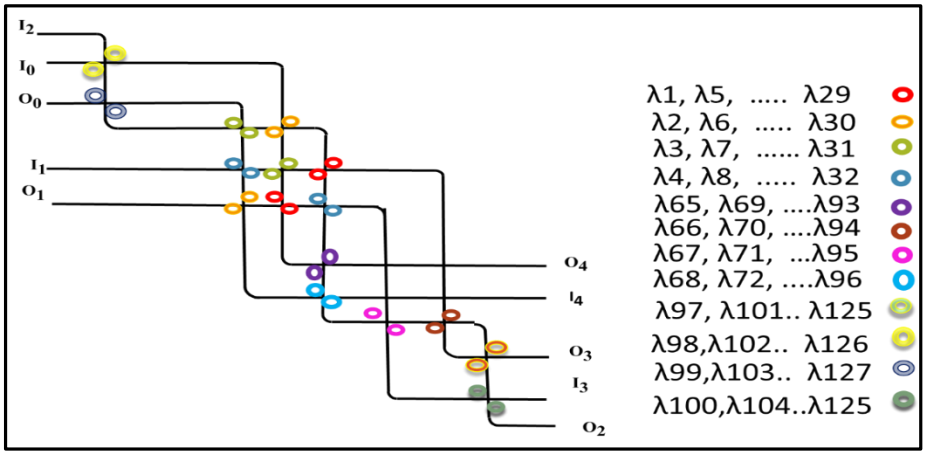


Figure 5.9: Modified photonic switch. Port 2 that is connected to the branch router can send and receive 64 wavelengths simultaneously.

Table 5.2: Wavelength assignment in the modified photonic switch supporting 32 wavelengths in port 2

Input / Output	**O_0 (λ)**	**O_1 (λ)**	**O_2 (λ)**	**O_3 (λ)**	**O_4 (λ)**
I_0	N/A	N = 1, X = 0; λ = N + (X * 4) $\forall$ X $\leq$ 15	**N = 2, X = 0; λ = N + (X * 4) $\forall$ X $\leq$ 31**	N = 3, X = 0; λ = N + (X * 4) $\forall$ X $\leq$ 15	N = 4, X = 0; λ = N + (X
I_1	N = 4, X = 0; λ = N + (X * 4) $\forall$ X $\leq$ 15	N/A	**N = 1, X = 0; λ = N + (X * 4) $\forall$ X $\leq$ 31**	N = 2, X = 0; λ = N + (X * 4) $\forall$ X $\leq$ 15	N = 3, X = 0; λ = N + (X
I_2	**N = 3, X = 0; λ = N + (X * 4) $\forall$ X $\leq$ 31**	**N = 4, X = 0; λ = N + (X * 4) $\forall$ X $\leq$ 31**	**N/A**	**N = 1, X = 0; λ = N + (X * 4) $\forall$ X $\leq$ 31**	**N = 2, X = 0; λ = N + (X**
I_3	N = 2, X = 0; λ = N + (X * 4) $\forall$ X $\leq$ 15	N = 3, X = 0; λ = N + (X * 4) $\forall$ X $\leq$ 15	**N = 4, X = 0; λ = N + (X * 4) $\forall$ X $\leq$ 31**	N/A	N = 1, X = 0; λ = N + (X
I_4	N = 1, X = 0; λ = N + (X * 4) $\forall$ X $\leq$ 15	N = 2, X = 0; λ = N + (X * 4) $\forall$ X $\leq$ 15	**N = 3, X = 0; λ = N + (X * 4) $\forall$ X $\leq$ 31**	N = 4, X = 0; λ = N + (X * 4) $\forall$ X $\leq$ 15	N/A

5.3.2 Various configurations schemes on the photonic rings

In this section, we will discuss about the various configurations schemes implemented on the photonic rings at the last level of hierarchy in NOVA. The previous section explained about the baseline configuration on the rings that employed W-SWMR technique. We will discuss three other configurations of the photonic rings in this section:

- Reservation assisted – SWMR
- Token ring – MWSR
- Token ring – MWMR

Reservation assisted – SWMR: In this section, we will describe the reservation assisted SWMR configuration that we implemented on the photonic rings. Figure 5.10 (a) shows the working of reservation based SWMR in NOVA. In the figure, the *branch router* associated with quadrant 1 (Q1) has write access over the ring and the remaining *branch routers* read from the ring. Only one ring is shown in the figure for brevity.

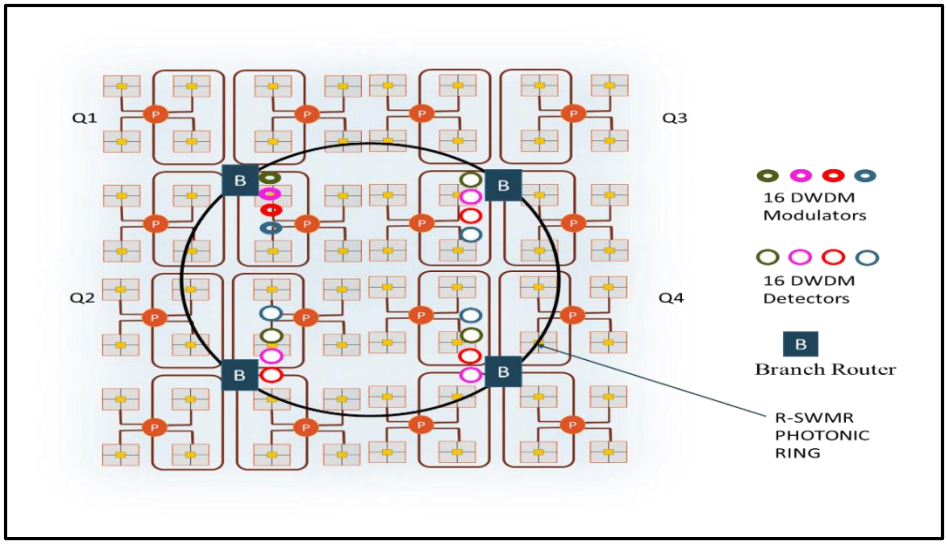
In this technique the photonic rings are divided into two logical channels – a reservation channel and a data channel. Naturally, communication over the rings is divided into reservation phase and data phase. In the reservation phase, the source *branch router* sends a request over its SWMR ring to the destination *branch router* using a narrow two bit wide reservation channel. This reservation request is read by all the *branch routers* that are configured as readers on that ring. Based on the request, only one of the *branch router* switches on its detectors. The remaining *branch routers* ignore the request. Data phase follows the reservation phase and in this phase source *branch router* sends data to the destined router. Similarly, all the other *branch routers* are sending data simultaneously across different *major-quadrants* using their respective SWMR rings. When the transfer of an entire packet is complete, the receiving side turns off its detectors. This is done once the receiving side detects an incoming tail flit.

Token ring – MWSR: In this section, we describe the MWSR configuration implemented on the photonic rings. In this technique, on a particular ring, data can be written by any of the three *branch routers* in NOVA, but is read only by the remaining *branch router*. Because multiple contenders are trying to access the data channel for writing, this configuration requires global arbitration for contention resolution. A token ring based arbitration is implemented for this configuration of rings [21]. A wavelength that is not a part of data stream, is designated to be the arbitration wavelength. A *branch router* that currently has the arbitration token, has the rights over the data channel. All the other *branch routers* wait for the token to be released by the current router. Once the entire packet is sent by the current router, it releases back the token into the ring and other routers can try to gain access over the ring. This arbitration technique has four photonic rings, with

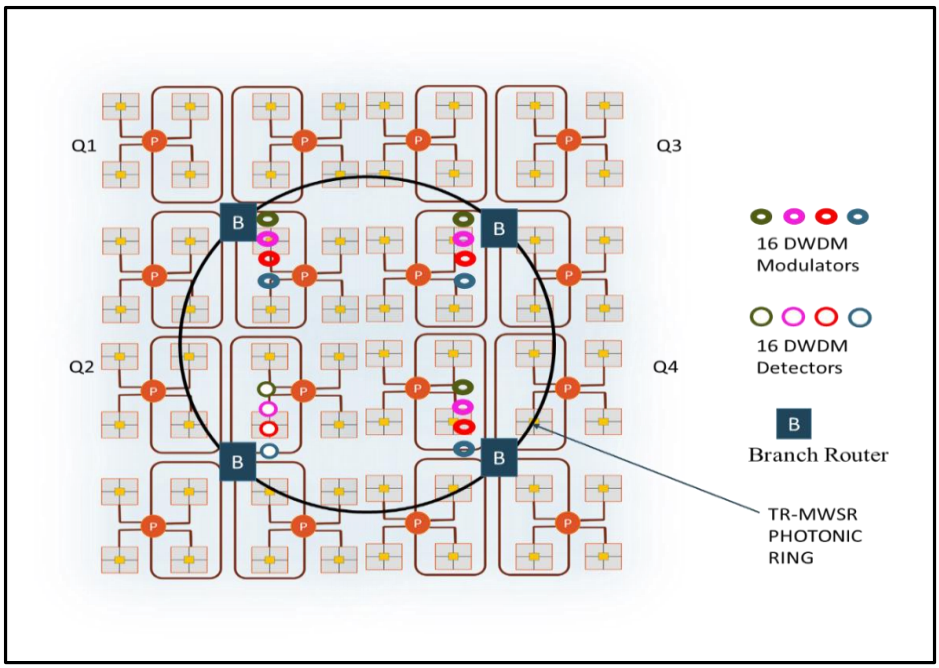
each ring having a designated reader. Figure 5.10 (b) shows the working of token ring based MWSR in NOVA. In the figure, *branch router* associated with quadrant 2 (Q2) is the only reader and the remaining routers have write access over the ring.

Token ring – MWMR: In this section, we describe a MWMR configuration employed on the photonic rings. In this configuration, on the sender's side, the ring behaves as a MWSR and on the receiver side it behaves as SWMR ring. Global arbitration is required on the senders' side and we use similar arbitration scheme as in the previous implementation to resolve any contentions. Figure 5.10 (c) shows the working of token ring based MWMR in NOVA. As can be seen from the figure, all the four *branch routers* have both read and write access over the ring.

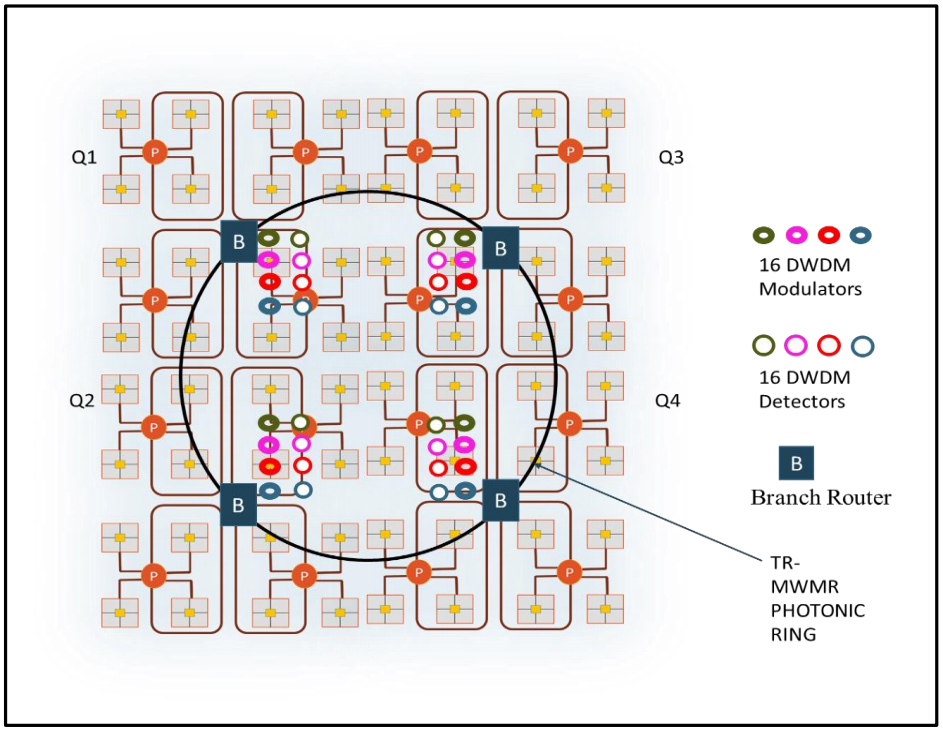
The results and the detailed performance impact of different configuration/arbitration schemes is discussed in the next chapter.



(a)



(b)



(c)

Figure 5.10: Different configurations of photonic rings in NOVA (a) R-SWMR (b) TR-MWSR (c) TR-MWMR

In this chapter, we explained the detailed architectural implementation of NOVA. We also explained about the variations that we designed and implemented to arrive at an optimal configuration of the network for given set of input applications. In the following chapter, we will discuss the experimental setup and results of NOVA.

6 EXPERIMENTS AND RESULTS

In this chapter, we present experimental results and evaluate the efficiency of our proposed photonic network-on-chip (NoC) architecture – NOVA. The first two sections present the experimental setup and details of the power estimation models used. The subsequent sections present our experimental results, including comparisons with electrical-mesh (EMesh) based NoC and three other state-of-the-art photonic NoCs – Corona, Firefly, and Flexishare [21] - [23].

6.1 Experimental setup

To validate the proposed architecture and to conduct experimental analysis, we extensively modified and developed a SystemC based NoC simulator at the cycle-accurate granularity [61]. We considered a die area of 400 mm^2 and targeted a 32 nm process technology. The photonic rings in NOVA connect four *branch-routers* present in four different *major-quadrants* of the network, which results in a ring diameter (~10 mm) smaller than the chip edge width. Based on geometric calculations of the interconnects (photonic rings) in the network, time needed for light to travel one complete circle on the photonic ring in 256-core NOVA was estimated to be 3 clock cycles at 5 GHz clock frequency. Similarly, we estimated 8 clock cycles as the time taken by the light to make one complete circle in Corona, Flexishare, and Firefly, based on the geometric calculations of respective interconnect lengths [21] – [23]. A packet size of 512 bits was considered across all architectures in all our simulations. In the next subsection, we describe in detail about the workloads we considered for our experiments.

Table 6.1: Classification of PARSEC benchmarks based on workload [62]

Application	Representation	Workload Type
Blackscholes	BS	Compute intensive
Bodytrack	BT	Compute intensive
Vips	VI	Compute intensive
Dedup	DU	Compute intensive
Freqmine	FQ	Memory/Compute intensive
Ferret	FR	Memory/Compute intensive
Fluidanimate	FA	Memory/Compute intensive
X264	X264	Memory/Compute intensive
Streamclusters	SC	Memory intensive
Canneal	CA	Memory intensive
Facesim	FS	Memory intensive

Table 6.2: Multi-program workloads

Application combination
SC-BT-BS-VI
BT-DU-FA-FQ
CA-X264-FR-FQ
SC-BT-BS-X264
CA-FS-FR-FA
BT-BS-VI-DU
X264-FR-FQ-FA
SC-FS-CA-BS
BT-X264-FQ-FA

Workloads: We used both - synthetic traffic patterns and real workloads to explore the architectural performance under diverse traffic conditions. We profile 11 different applications from the PARSEC benchmark suite [62]. We use different combinations of these applications on our simulator, to form a multi-programmed workload environment. We use such multi-programmed workloads to compare our proposed architecture with other state-of-the-art architectures. We profiled all of the applications and divided them into three categories based on their memory intensities, as shown in Table 6.1. Compute intensive benchmarks spend most of their time computing and less time communicating with memory; whereas memory intensive applications spend most of their time in communicating with memory and are latency tolerant. The final category of the benchmarks are those that are hybrid in nature. As shown in Table 6.2, we have considered nine different multi-application workloads, each combining four different application. These combinations of applications are run on different *major-quadrants* in NOVA. The application combinations are chosen in order to depict an emerging massively parallel CMP with several co-running applications. To test the scalability of NOVA, we modeled and simulated different network sizes, targeting 64-core and 256-core CMPs. A group of four concentrators having four cores each, run a specific application in a 64-core CMP and a group of 16 concentrators run a specific application in a 256-core CMP. As an example, the SC-BT-BS-X264 workload combines the parallelized implementations of Streamclusters (SC), Bodytrack (BT), Blackscholes (BS), and X264; each of them executing on a different *major-quadrant* in NOVA. These combinations provide a good mixture of applications that are diverse in their memory requirement. Application traces were extracted from gem5 full-system simulator [63].

6.2 Power estimation model

In this section, we present a detailed power estimation model of NOVA. The power consumed in NOVA can be divided into two parts: the power consumed by the electrical network and the power consumed by the photonic components. The static and dynamic power consumption of electrical routers is based on results obtained from DSENT simulator [64]. DSENT is a NoC modeling tool for rapid design space exploration of electrical and electro-photonic networks. To model the power consumption parameters for photonic components in NOVA, we adopt the results from RMLP [65] [74]. Based on the network bandwidth utilization, a CMP will produce temperature variations across its surface area. Such an unbalanced thermal profiles need to be considered in the scope of micro-ring resonator structures that can go off-resonance with thermal variations and require thermal tuning to maintain proper functioning of the device. Energy values for thermal heaters integrated to micro-ring resonators and energy dissipation parameters of various other components in NOVA are obtained from several published works [53] [64] [76]. Table 6.3 shows the energy consumption and delay parameters of photonic components in NOVA.

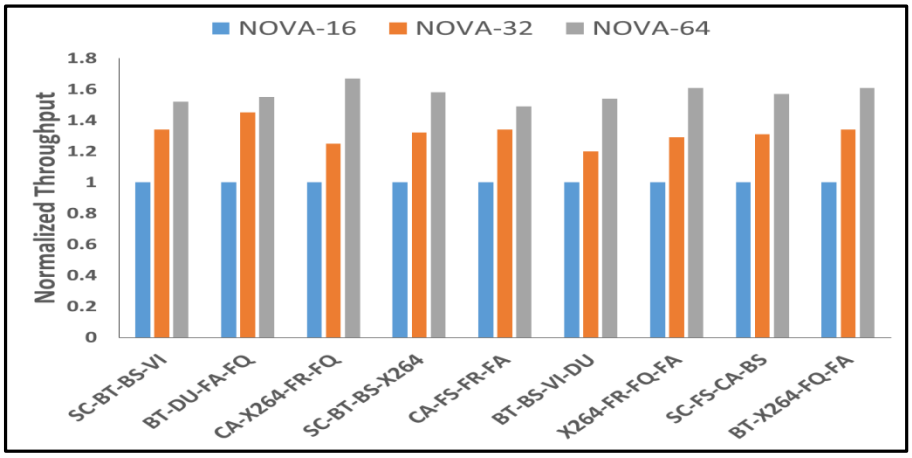
Table 6.3: Energy consumption parameters for photonic components [53] [64]

Component	Delay (ps)	Dynamic energy in (pJ/bit)	Thermal tuning energy (fJ/bit/heater)
Modulator driver	9.5	90	16
Modulator	3.1		
Waveguide	15.6 per/mm	-	-
Photo Detector	0.22	60	16
Receiver	4.0		

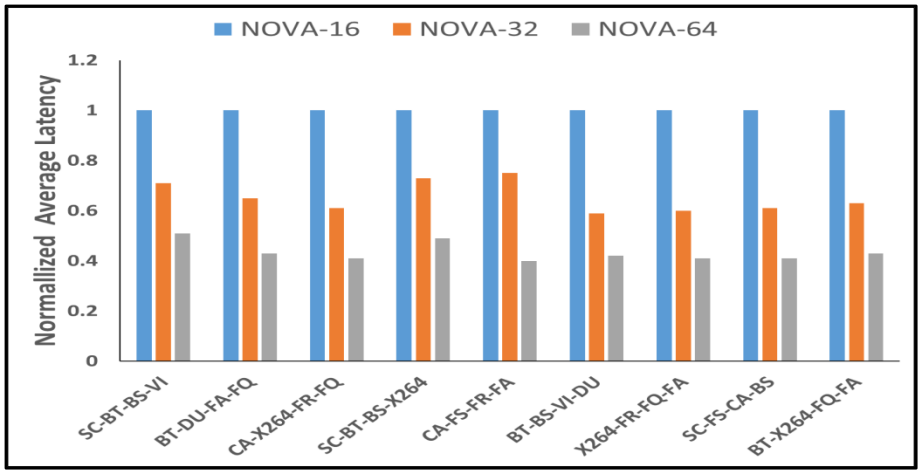
Table 6.4: Losses of photonic components [66] [67]

Photonic loss type	Loss (in dB)
Chip coupling	1
Waveguide propagation per cm	1
Waveguide coupler/splitter	0.5
Bending Loss per bend	0.005
Micro-ring through	0.02
Laser power	2.1W

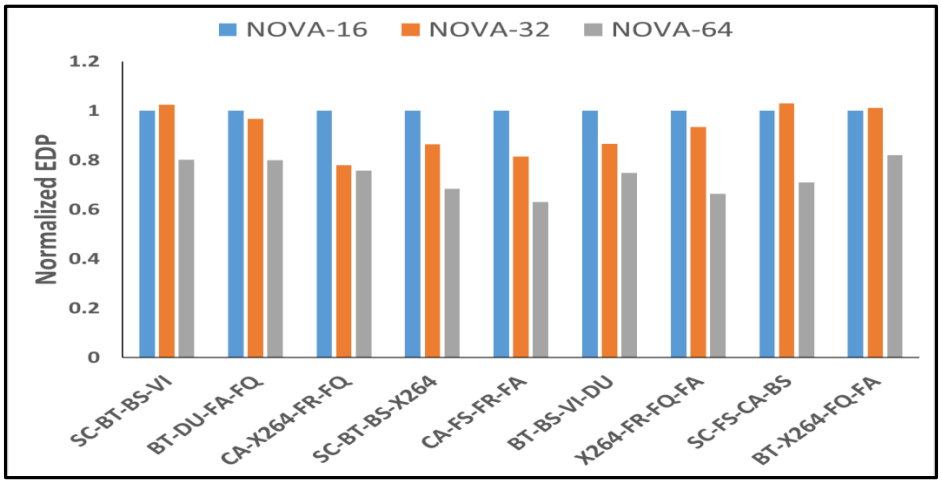
To compute the laser power consumption, we calculated photonic losses in various components in our architecture. The cumulative losses of all the photonic components set the required optical laser power budget and the corresponding electrical laser power. We assumed the laser efficiency as 30%. Per component loss values in NOVA are based on several published works [66] [67]. To reliably detect the photonic data, photo-detectors need to have a minimum responsivity. The photo-detector responsivity and TIA gain values are considered as 0.7 A/W and 4 KΩ respectively as per the work done on integrated silicon photonic transmitters and receivers [77]. Table 6.4 shows the loss values of various photonic components in NOVA.



(a)



(b)



(c)

Figure 6.1: Improvement in terms of (a) throughput, (b) latency, and (c) EDP of NOVA-64 compared to NOVA-32 and NOVA-16.

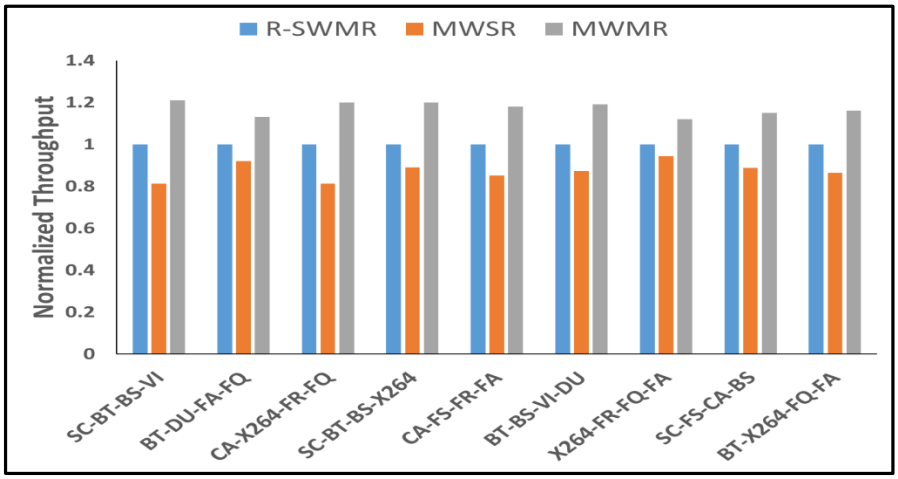
6.3 Impact of architectural variations

In this section we discuss the performance impact of architectural variations in NOVA. We first discuss on the improvements achieved by using different photonic switches at the first level of hierarchy in the network. We then discuss on the impact of using different configurations of the photonic rings at the last level of hierarchy in NOVA.

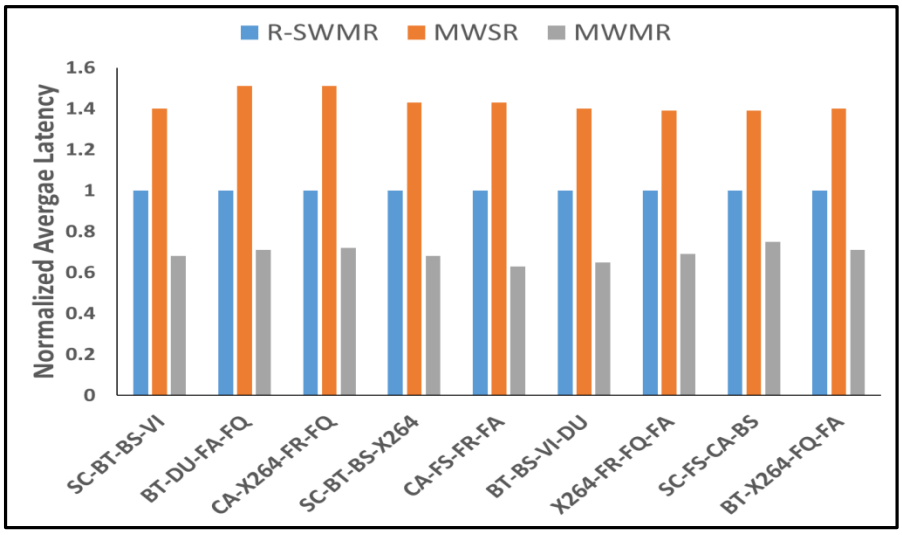
After running various benchmark applications from PARSEC on our network, we observed that the network traffic between *minor-quadrants* dominated the *intra-minor-quadrant* traffic. To satisfy the bandwidth requirements of real multi-program applications, we redesigned the photonic switches to support additional bandwidth. In particular, the ports of the photonic switch that were connected to the *branch router* were modified to support additional wavelengths. The baseline photonic switch supports switching of 16 wavelengths, and the two modified switches support 32 and 64 wavelengths. Additional bandwidth support is achieved by using 8 and 16 extra modulators for 32 and 64 wavelengths respectively. Figure 6.1 (a) shows the improvement in throughput when using three different photonic switches that support 16, 32, and 64 wavelengths. There is an improvement of 50% going from NOVA-16 to NOVA-64 in terms of throughput. This increase in throughput can be attributed to the increase in the number of modulators in the communication path. Figure 6.1 (b) shows that there is an improvement of 55% in terms of latency going from NOVA-16 to NOVA-64. This is because in NOVA-16, to transfer a flit of 64 bits it takes 4 cycles whereas in NOVA-64 it takes just one cycle. Finally, as shown in Figure 6.1 (c) NOVA-64 does better in terms of energy-delay product (EDP) compared to NOVA-16 with as much as 30% reduction on an average.

In the next set of experiments, we evaluated the performance impact by implementing different configurations of photonic rings present at the last level of hierarchy. As explained in Section 5.3.2, three different configurations were implemented on the photonic rings in NOVA. First technique uses a reservation based SWMR configuration where a reservation phase precedes the actual transfer of data. In the second technique, a token ring based MWSR is employed on the photonic rings. In the last technique a token ring based MWMR configuration is used on the rings. Token ring based MWMR scheme performs better compared to other two

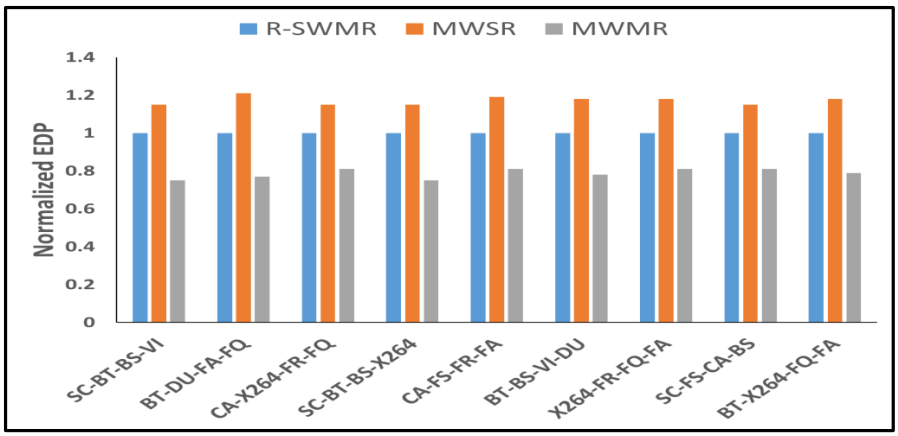
techniques. Two sets of experiments were performed to observe the impact of varying configurations on the ring. In the first experiment, 25% of the total memory requests were sent over the ring, to the concentrators in a different *major-quadrant*. Figure 6.2 (a) – (c) show the impact of different configurations on the rings with 25% of the total memory requests going over the ring. Token ring based MWMR has an average improvement of 20% in terms of throughout over reservation based SWMR, 40% improvement in the latency and 25% improvement in EDP.



(a)

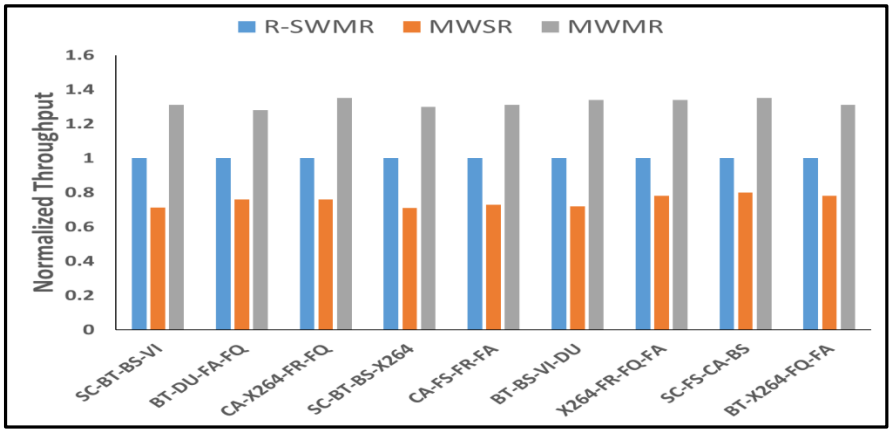


(b)



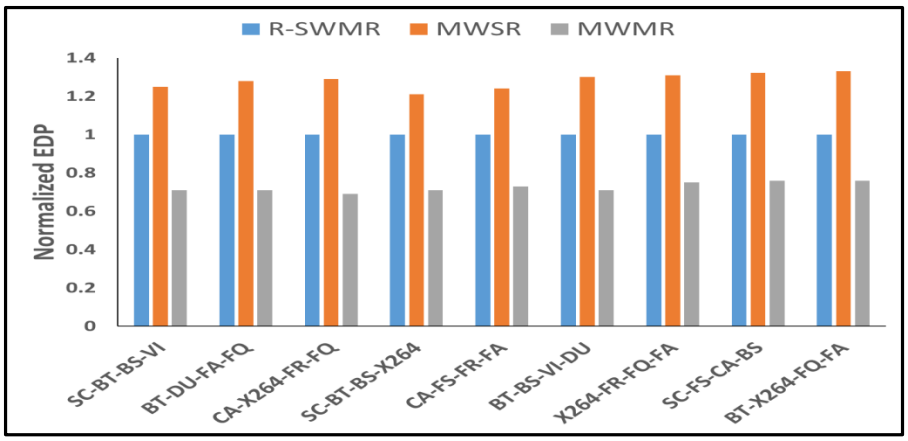
(c)

Figure 6.2: Improvement in terms of (a) throughput, (b) latency, and (c) EDP with different photonic ring configurations in NOVA – R-SWMR, MWSR, and MWMR (25% global traffic)



(a)

(b)



(c)

Figure 6.3: Improvement in terms of (a) throughput, (b) latency, and (c) EDP with different photonic ring configurations in NOVA – R-SWMR, MWSR, and MWMR (50% global traffic)

In the second set of experiments, 50% of the total memory requests were sent over the ring, to the concentrators in a different *major-quadrant*. Figure 6.3 (a) – (c) show the impact of different configurations on the rings. Token ring based MWMR has an average improvement of 35% in terms of throughout over reservation based SWMR, 40% improvement in terms of latency and 21% improvement in terms of EDP. As multiple writers are contending for a single ring in MWSR, there is a wait time associated with every transfer in this technique and this leads to reduced performance compared to other two schemes. The total ring modulators on the photonic rings is kept constant across different configurations in NOVA. This helps the token ring based MWMR perform better compared to other two schemes.

6.4 Comparison with other state-of-the-art NoCs

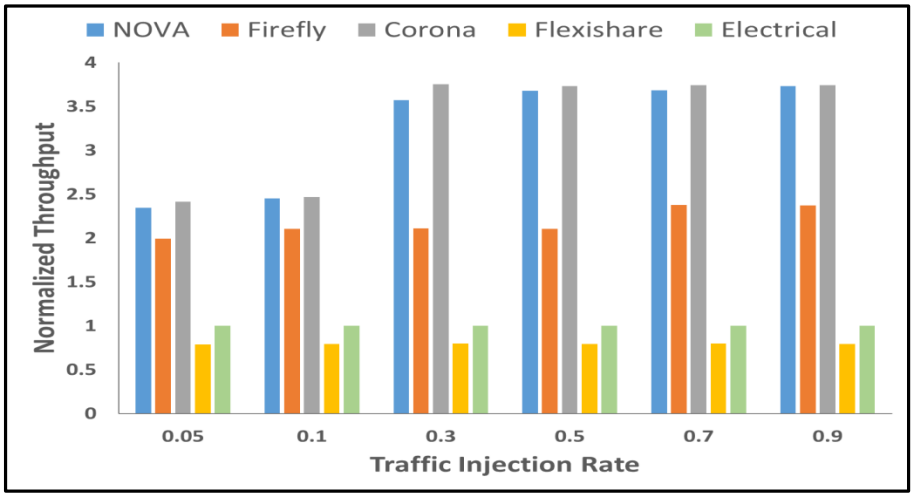
In this section, we will compare NOVA with EMesh, Corona, Firefly, and Flexishare architectures [21] – [23]. We have made our best effort to carefully implement every feature of these architectures, based on the description in their respective papers for a meaningful

comparison. In the first set of experiments we compare 64-core CMPs of all architectures and in the second set of experiments we compare 256-core CMPs of all architectures.

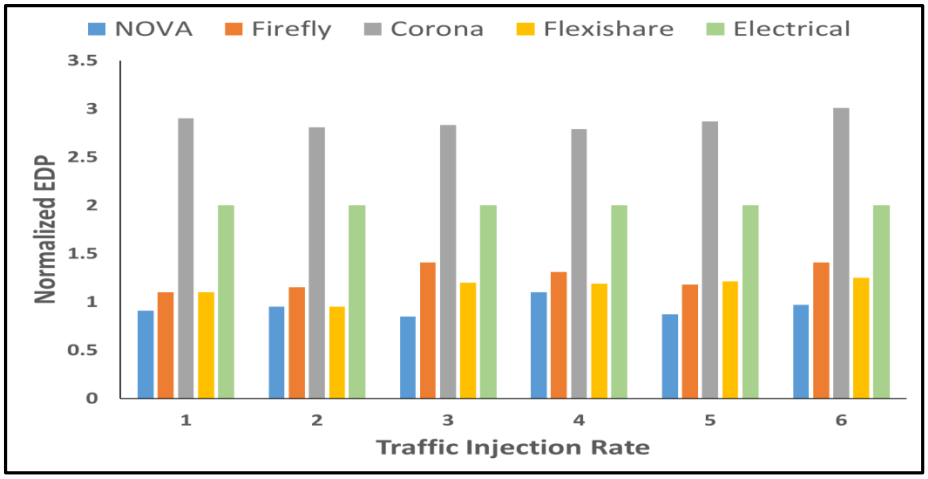
Figures 6.4 (a)-(d) show the throughput, EDP, power, and latency results for 64-core CMP architectures, running synthetic uniform random traffic pattern on their networks. From Figure 6.4 (a), it can be observed that NOVA has nearly 4× throughput compared to Flexishare. This is because of high network congestion in Flexishare, which handles the global traffic by serially injecting arbitration wavelengths to different MWMR data waveguides. NOVA has nearly 3.5× throughput compared to EMesh NoC, owning to its high bandwidth and low latency photonic links. NOVA achieves comparable throughput to Corona which has significantly high photonic hardware. Table 6.3 shows the photonic hardware comparison of all architectures with NOVA. Firefly achieves comparable throughput to NOVA at low packet injection rates and quickly saturates at higher packet injection rates. From Figure 6.4 (b), it can be seen that NOVA has 50% improvement in terms of EDP, compared to EMesh and an average 2.2× improvement over Corona. Figure 6.4 (c) shows the power consumption of NOVA, Corona, Firefly, and Flexishare normalized to EMesh network. From the figure it can be seen that NOVA has about 90% improvement over EMesh, 3.6× improvement over Corona. This can attributed to the fact that NOVA has much less power consumption compared to Corona owing to reduced photonic hardware. Figure 6.4 (d) shows the average packet latency of all the architectures and it can be clearly seen that NOVA has better performance than Firefly, Flexishare, and EMesh, which saturate much earlier owing to several factors such as network congestion and inefficient arbitration schemes leading to starvation of nodes.

Table 6.3: Photonic hardware comparison

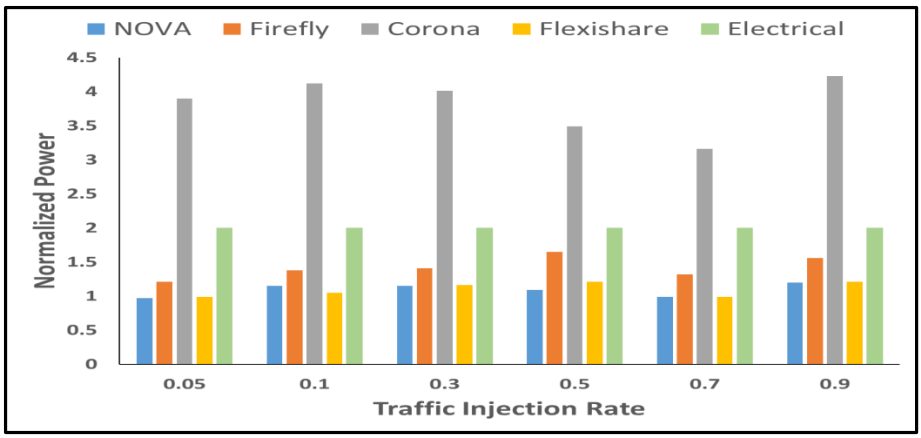
Architecture	Ring resonators
NOVA	5376
Firefly	34832
Flexishare	528960
Corona	2105344



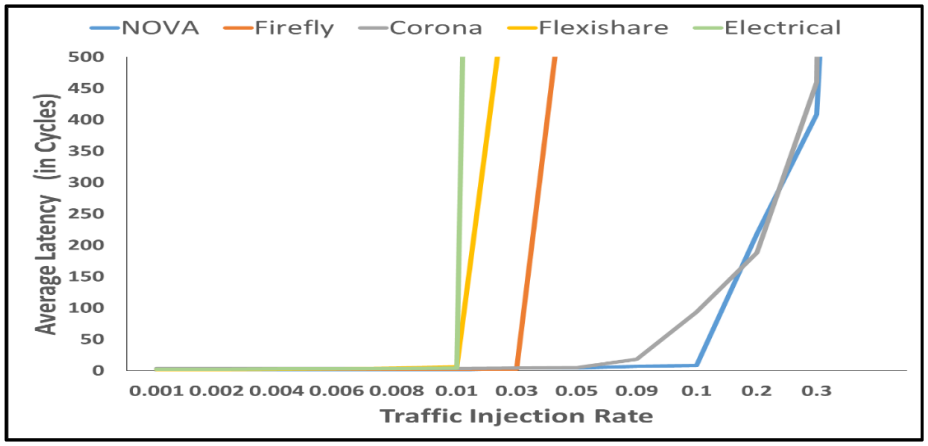
(a)



(b)

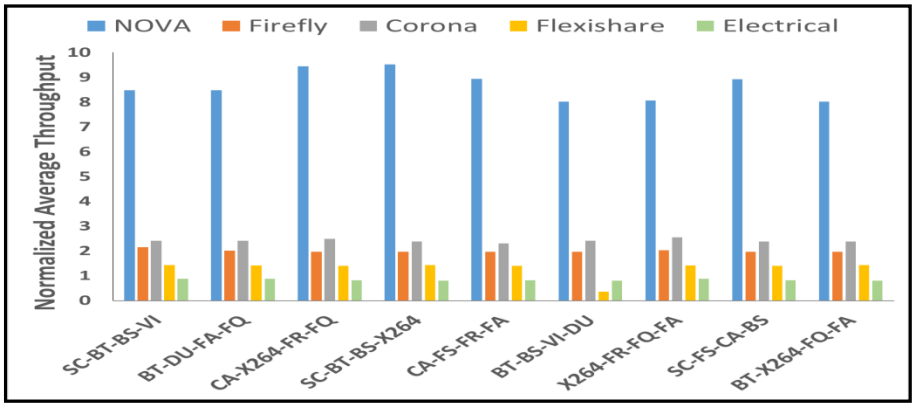


(c)

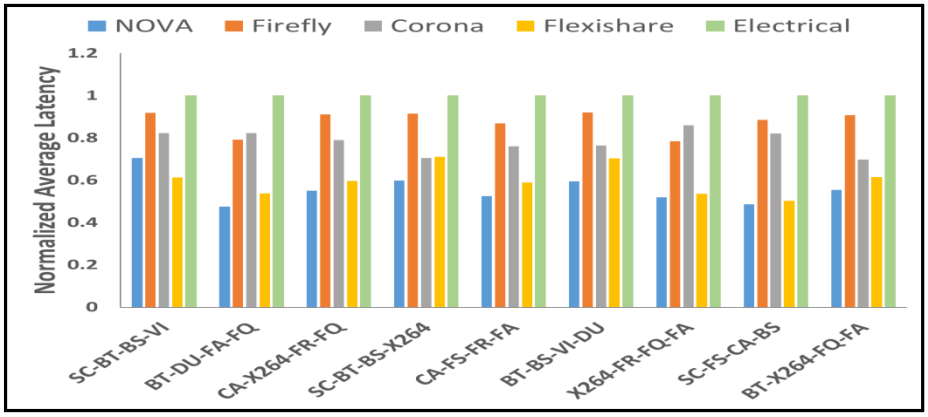


(d)

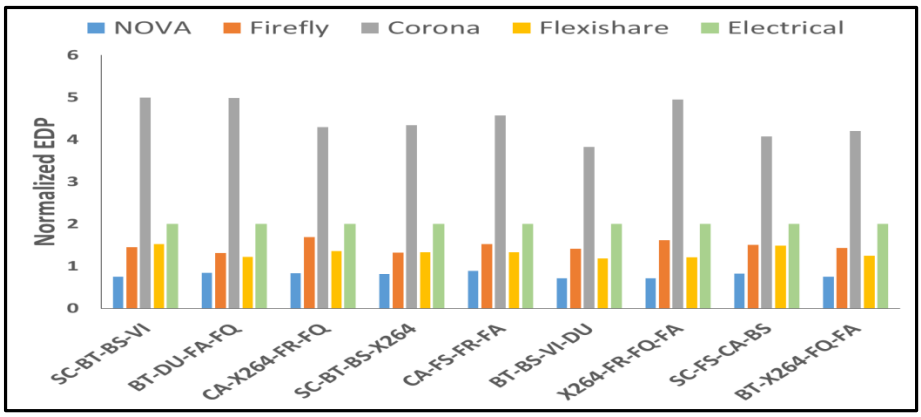
Figure 6.4: Normalized (a) throughput, (b) EDP, (c) power, and (d) average latency results comparing NOVA with other architectures for a 64-core CMP, running synthetic uniform random traffic.



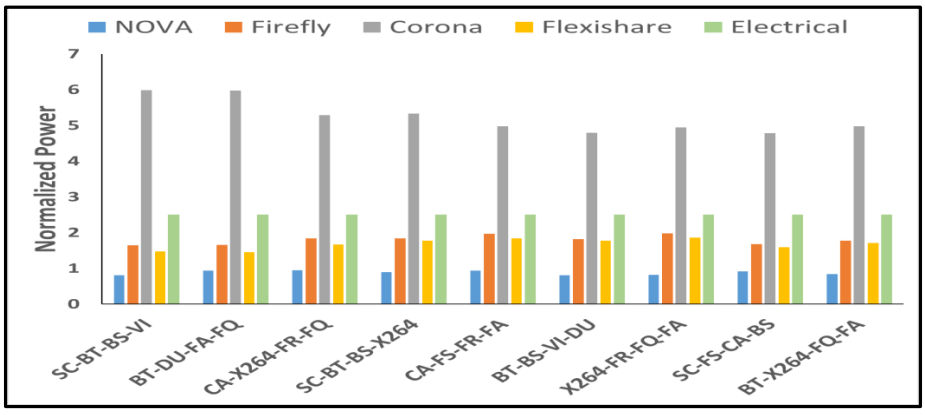
(a)



(b)



(c)



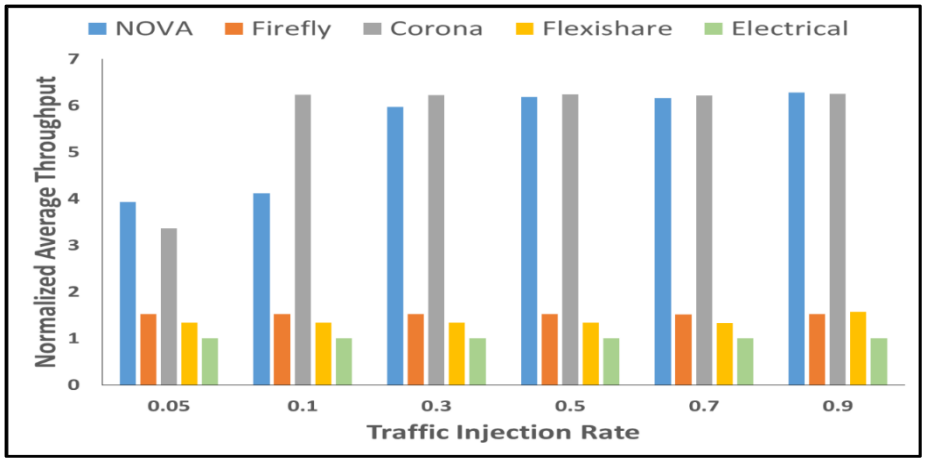
(d)

Figure 6.5: Normalized (a) throughput, (b) latency, (c) EDP, and (d) power results comparing NOVA with other architectures for a 64-core CMP. Results are shown for multi-application workloads based on PARSEC benchmarks.

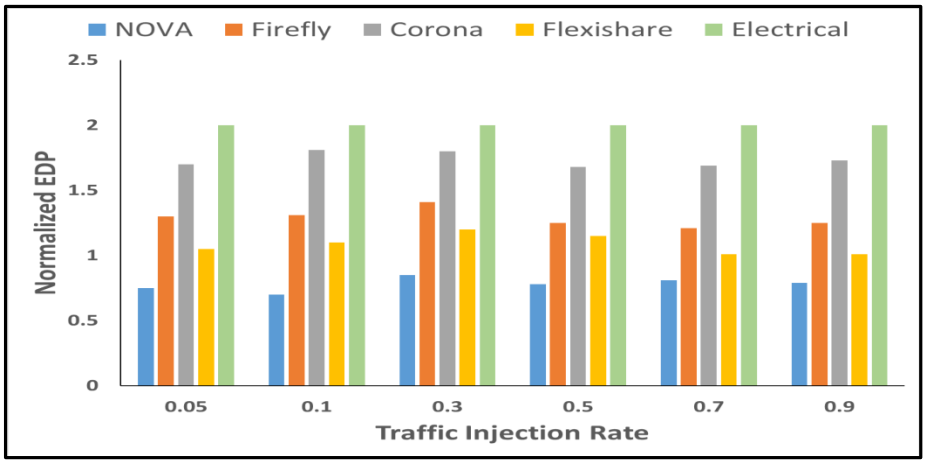
In the next set of experiments, we ran PARSEC multi-application workloads across architectures to compare the results of throughput, latency, EDP, and power for 64-core CMP. Figures 6.5 (a)-(d) show the results of this comparison study, with all results normalized with respect to the EMesh results. Owing to slow electrical links and heavy congestion on the network due to real workloads, EMesh network performance in terms of throughput is least compared to other architectures. NOVA has around 5.5× throughput compared to Flexishare as can be seen from Figure 6.5 (a). This is because Flexishare lacks the hierarchical structure and also the arbitration scheme in Flexishare allows nodes to win the arbitration only in serial manner. NOVA also provides 3.9× higher throughput than Corona. In Corona, when multiple nodes try to communicate with same destination, packets get queued up waiting for respective data waveguide that is connected to the destination. In Firefly, intra-cluster communication is achieved using electrical links. Although inter-cluster communication is done using photonic links, electrical transfers are still used to reach the respective node in the destination cluster. NOVA achieves higher bandwidth owing primarily to its hierarchical nature and also due to its use of contention free wavelength routing. In NOVA, the global traffic creates very less congestion resulting in lower average packet latency. On average, NOVA has 15%, 35%, 38%, and 49% lower average packet delay over Flexishare, Corona, Firefly, and EMesh, respectively, for different multi-application workloads. In Figure 6.5 (c), we can see the EDP comparison of all the architectures. It can be seen from the figure that NOVA has on average 1.5×, 5.1×, 1.5×, and 1.6× lower EDP compared to Firefly, Corona, Flexishare, and EMesh respectively. From Figure 6.5 (d), it can be seen that NOVA has 1.8×, 5.6×, 1.4×, and 2.1× improvements in terms of power over Firefly, Corona, Flexishare and EMesh respectively. This is largely due to the reduction in the photonic hardware in NOVA compared to other architectures.

In the next set of experiments, we explored the scalability of our architecture by considering a 256-core CMP. We compared the throughput, EDP, power and, average packet latency results of NOVA with other architectures such as Corona, Firefly, Flexishare, and EMesh. Figures 6.6 (a)-(d) show the results obtained for 256-core CMPs while running a synthetic uniform random traffic. From Figure 6.6 (a), it can be seen that NOVA has 4.13×, 4.07×, and 5.5× improvement in throughput over Flexishare, Firefly, and EMesh respectively. Figure 6.6 (b) shows the EDP results of all the architectures and it can be seen from the figure that NOVA has 3.1×, 1.8×, 1.3×, and 3.8× improvements over Corona, Firefly, Flexishare, and EMesh respectively. From Figure 6.6 (c), it can be seen that NOVA has 3.25×, 1.5×, 1.2×, and 4.5× improvements in terms of power over Corona, Firefly, Flexishare, and EMesh respectively. From Figure 6.6 (d) it can be seen that Firefly, Flexishare, and EMesh saturate much earlier than NOVA owing to added complexities of arbitration in their architectures with the increase of network size.

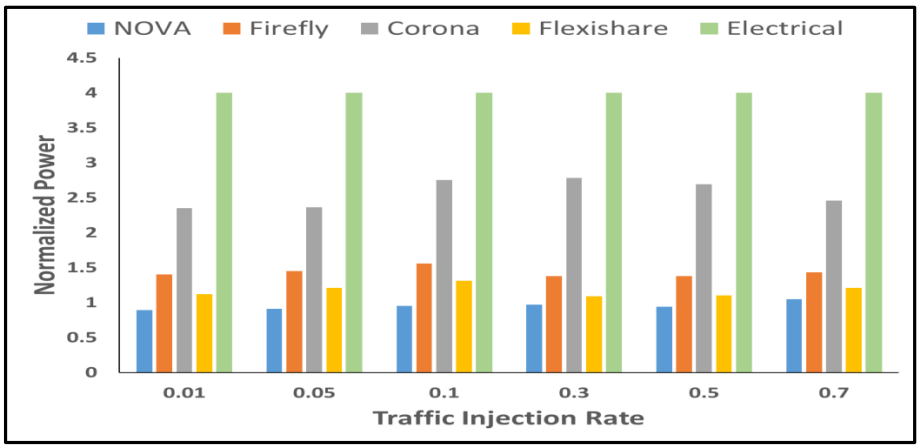
Figure 6.7 (a)-(d) show the results of normalized throughput, latency, EDP, and power of all the architectures, running PARSEC benchmark applications. It can be seen from the Figure 6.7 (a) that NOVA has on average 9.8×, 6.22×, 10.15×, and 10.95× improvement in throughput over Firefly, Corona, Flexishare and EMesh. Figure 6.7 (b) shows the normalized latency of all the architectures and it can observed that NOVA has 55%, 40.04%, 42.41% and 90% reduction over Firefly, Corona, Flexishare, and EMesh respectively. Figure 6.7 (c) shows that on average NOVA has 5.1×, 10.05×, 2.62×, and 11.02× reduction in EDP product compared to Firefly, Corona, Flexishare and EMesh networks. From Figure 6.7 (d), it can be seen that NOVA has 4.7×, 11.01×, 3.2×, and 11.3× improvements in terms of power over Firefly, Corona, Flexishare and EMesh respectively.


NOVA
Firefly
Corona
Flexishare
Electrical
Normalized Average Throughput
0
1
2
3
4
5
6
7
0.05
0.1
0.3
0.5
0.7
0.9
Traffic Injection Rate

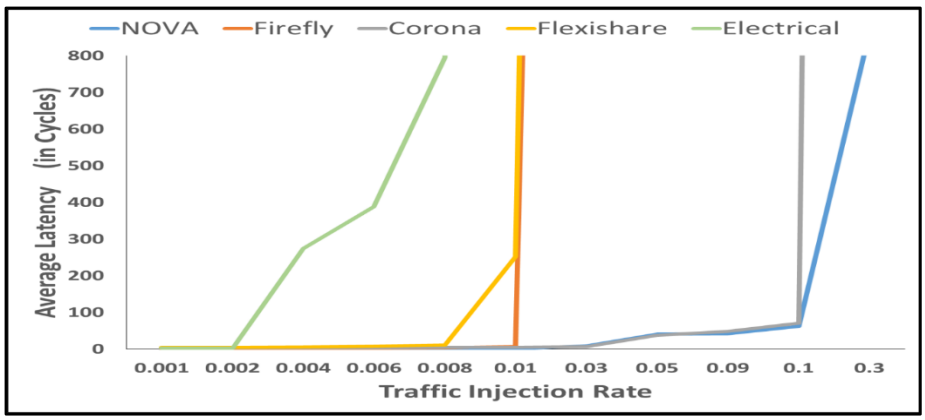
(a)


NOVA
Firefly
Corona
Flexishare
Electrical
Normalized EDP
0
0.5
1
1.5
2
2.5
0.05
0.1
0.3
0.5
0.7
0.9
Traffic Injection Rate

(b)

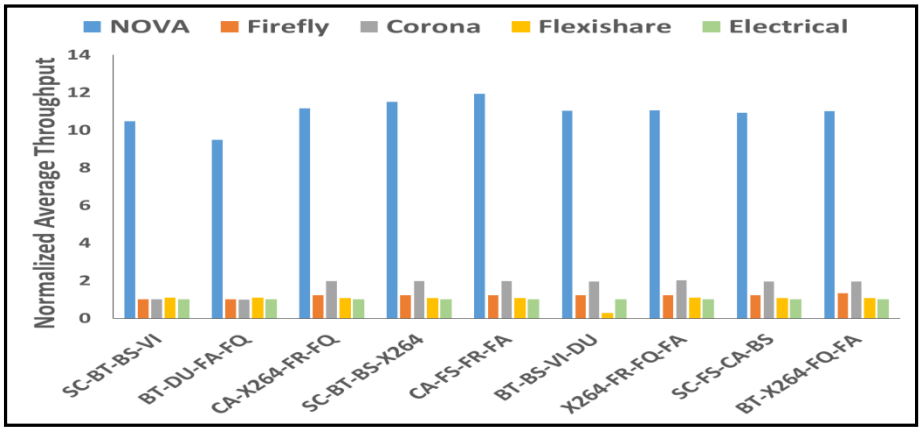

NOVA
Firefly
Corona
Flexishare
Electrical
Normalized Power
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
0.01
0.05
0.1
0.3
0.5
0.7
Traffic Injection Rate

(c)

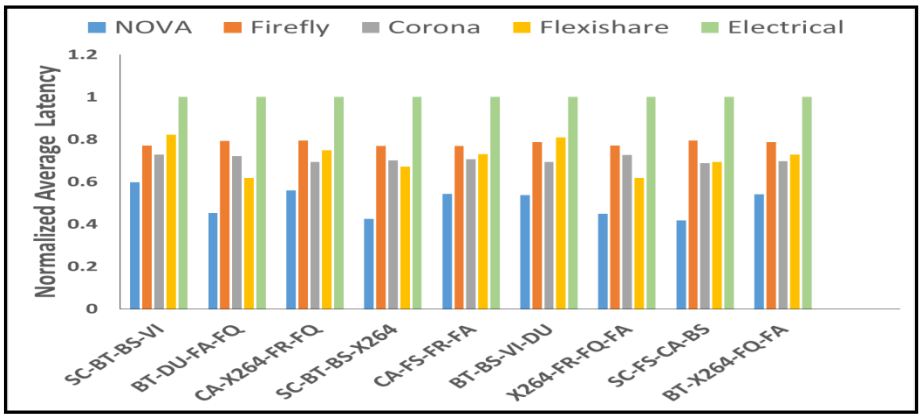


(d)

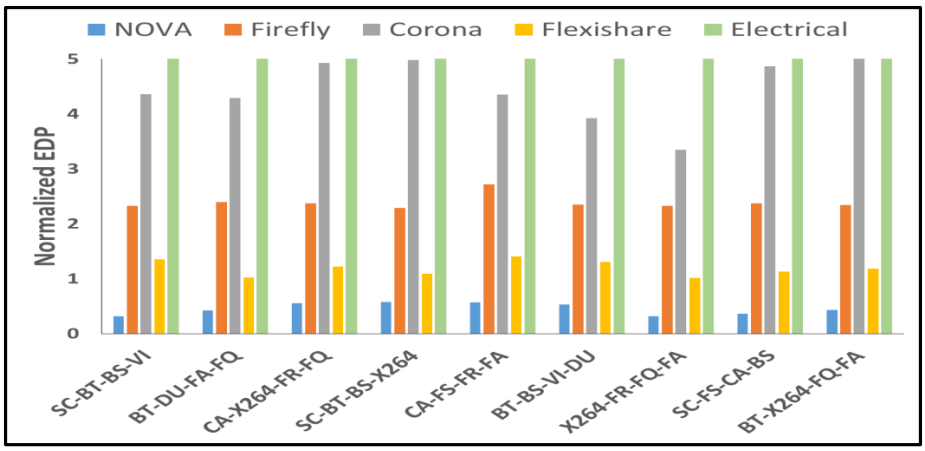
Figure 6.6: Normalized (a) throughput, (b) EDP, (c) power, and (d) average packet latency results of NOVA compared with other architectures for a 256-core CMP while running synthetic uniform random traffic



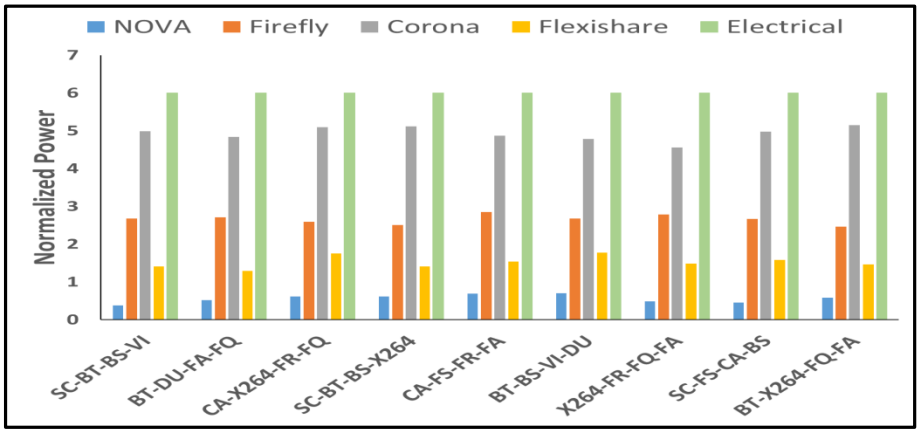
(a)



(b)



(c)



(d)

Figure 6.7: Normalized (a) throughput, (b) latency, (c) EDP, and (d) power results of NOVA compared against other architectures for a 256-core CMP. Results are shown for multi-application workloads based on PARSEC benchmarks. All results are normalized to EMesh.

Our proposed architecture was shown to achieve an average improvement of 6.1×, 55%, 5×, and 5.9× in terms of throughput, average latency, EDP, and power respectively, as compared to the other previously proposed photonic architectures and across different network sizes. Given its scalability and superior performance across various workload types, we believe that our approach is an attractive option for future multi-core NoC-based systems executing multiple and diverse applications.

7 CONCLUSION

In this chapter we summarize the key contributions of the thesis and highlight the improvements of our proposed photonic NoC architecture over existing state-of-the-art photonic NoCs. We also present some of the possible future extensions to our proposed work.

7.1 Summary

As the number of processing elements on a die are rapidly growing, interconnects are becoming increasingly crucial for overall system performance. Interconnects play a significant role in determining maximum achievable performance as inter-processor and processor-memory communication are dependent on it. The main contribution of this thesis is design of novel hierarchical electro-photonic NoC architectures - NOVA. Chapter 1 introduced this thesis, highlighted some of the challenges faced by conventional electrical NoCs, and described in detail the motivation to pursue this area of research. Chapter 2 provided the essential background relevant to this research area including some basics on photonic NoC architectures. In Chapter 3, the thesis optimization goals and parameters were specified. A representative subset of relevant research in this area was presented in Chapter 4 as related work, along with a discussion on some of the existing issues and challenges in the photonic NoC area. To address some of the issues described in Chapter 4, we designed and implemented a hierarchical electro-photonic architecture and carried out several experiments to validate our proposal. Chapter 5 presented the micro-architectural details of our proposed architecture and explained in detail about different architectural variations designed to arrive at an optimal configuration of the network for a given

set of input applications. Chapter 6 analyzed the experimental results of our simulations, presented a comparison study of NOVA with other state-of-the-art architectures.

7.2 Conclusion

In conclusion, our proposed hierarchical electro-photonic NoC addresses the challenges of high power dissipation, network congestion, and inefficient arbitration schemes prevalent in photonic NoCs [21] - [23]. NOVA uses contention free wavelength routing of data, which requires no arbitration and supports extremely low latencies. NOVA has three distinct hierarchical levels of communication, supporting high-bandwidths. Owing to its hierarchical nature NOVA avoids high interference of local and global traffic in the network. NOVA has reduced power dissipation compared to other photonic NoCs due to reduced photonic hardware usage. Experimental results show that NOVA achieves an average improvement of 6.1×, 55%, 5×, and 5.9× in terms of throughput, average latency, EDP, and power respectively. Also the margin of improvement over other architectures when the network size was scaled to 256-core was better than 64-core CMP which is strong indicator of scalability support in the proposed architecture. Given its scalability and superior performance across various workload types, we believe that our proposed architecture is an attractive option for future multi-core NoC-based systems executing multiple and diverse applications.

7.3 Future work

As discussed in Chapter 4, a significant amount of research is being done for NoC based multi-core systems. In Chapter 5, we presented a novel photonic NoC architecture to resolve some of the challenges related to on-chip communication. However, this work can certainly be

extended to improve achievable system performance. Some of the possible research directions to extend this work are presented below:

- **Off-chip communication:** Extending the work presented in this thesis to off-chip communication can be a possible future direction that can be explored further. In the off-chip electro-photonic architecture development work, we can integrate memory modules to NoC fabric using photonic interconnects to further improve the system performance.
- **Dynamic reconfiguration:** We can further increase the available bandwidth by reconfiguring the network at run time by monitoring the bandwidth availability and applying different reconfiguration algorithms.
- **Thermally resilient photonic NoC**: Another possible future direction for our work could be to develop a thermally resilient photonic NoC architecture. Variations in temperature can cause a change in the refractive index and can potentially disrupt the operation of photonic devices. With thermally resilient architecture, we can further gain in terms of power while reducing bit error rates.

Photonic NoC has huge potential to enable high performance gains with appropriate system design tradeoffs. The above mentioned directions are not exhaustive by any means and represent some of the multiple ways in which future research can alleviate bottlenecks in NoC for multi-core chip platforms.

REFERENCES

[1] A. E. Willner, “Mining the optical bandwidth for a terabit per second”, in Proceedings IEEE, vol. 34, no.6, 1997, pp. 32-41.

[2] Intel FAER series lecture, “ Introduction to Multi-Core”, http://www.faer.ac.in/treach/pdf/Introduction_to_Multi_Core.pdf. Last accessed on 11/3/14.

[3] R. Ho, K.W.Mai, and M.A.Horowitz, “The future of wires”, in Proceedings IEEE, vol. 89, no. 4, Apr 2001, pp. 490-504.

[4] “Tile64 processor”, http://www.tilera.com/sites/default/files/productbriefs/TILE-Gx8072_PB041-04_WEB.html. Last accessed on 11/3/14.

[5] “Oracle sparc servers”, http://www.oracle.com/technetwork/server-storage/sun-sparc-enterprise/documentation.html. Last accessed on 11/3/14.

[6] “Intel’s SCC”, https://www-ssl.intel.com/content/www/us/en/research/intel-research.html. Last accessed on 11/3/14.

[7] “AMD FX series processors”, http://sites.amd.com/us/promo/processors/Pages/fx-processor.aspx. Last accessed on 11/3/14.

[8] W. J. Dally and B. Towles, “Principles and Practices of Interconnection Networks”, Morgan Kaufmann, 2003.

[9] S. Koohi and S. Hessabi, “Power efficient nanophotonic on-chip network for future large scale multiprocessor architectures”, in Proceedings NANOARCH, Mar 2011.

[10] R. Morris and A. Kodi, “Exploring the design of 64- and 256-core power efficient nanophotonic interconnect”, in Quantum Electronics Journal, vol. 16, no 5, Sep 2010, pp. 1386 - 1393.

[11] S.Bahirat and S.Pasricha, “Exploring Hybrid Photonic Networks-on-Chip for Emerging Chip Multiprocessors”, in CODES+ISSS, Oct 2009, pp. 129-136.

[12] ITRS Technology Working Groups, “International Technology Roadmap for Semiconductors (ITRS)”, http://public.itrs.net. Last accessed on 11/3/14.

[13] K.Bergman, L.P.Carloni, A.Biberman, J.Chan, and G.Hendry, “Photonic Network-on-chip Design”, in Integrated Circuits and Systems series, ed., Springer May 2013.

[14] M.Petracca, B.G.Lee, K.Bergman, and L.P.Carloni, "Photonic NoCs: System-Level Design Exploration", in Proceedings MICRO Aug. 2009.

[15] M. Gnan, S.Thorns, D.Macintyre, R.De La Rue, and M.Sorel, "Fabrication of low-loss photonic wires in silicon-on-insulator using hydrogen silsesquioxane electron-beam resist", in Electronics Letters, vol. 44, no. 2, Jan. 2008, pp. 115–116.

[16] M.J.Shaw, J.Guo, G.A.Vawter, S.Habermehl, and C.T.Sullivan, "Fabrication techniques for low-loss silicon nitride waveguides", Micromachining Technology for Micro-Optics and Nano-Optics III, vol. 5720, no. 1, June 2005, pp. 109–118.

[17] A.Kodi and A.Louri, "Rapid reconfigurable and scalable all-photonic interconnect for distributed shared memory processors", in Lightwave Technology Journal, vol. 22, May 2004, pp. 2101-2110.

[18] C.Kochar, A.Kodi, and A.Louri, "Nd-Rapid: a multidimensional scalable fault-tolerant optoelectronic interconnection for high performance computing systems", in Optical Networking Journal, vol. 6, no.5, Apr 2007, pp.231-245.

[19] M. Tan, P.Rosenberg, Y.Jong-Souk, M.McLaren, S.Mathai, T.Morris, K.Pei, J. Straznicky, N.Jouppi, and S.Wang, "A high-speed optical multi-drop bus for computer interconnections", in Proceedings HPI, 2008, pp. 3-10.

[20] Y.Hoskote, S.Vangal, A.Singh, N.Borkar, and S.Borkar, "A 5-GHz mesh interconnect for a teraflops processor", in Proceedings MICRO, vol. 27, no.5, Sept 2007, pp. 51-61.

[21] D.Vantrease, R.Schreiber, M.Monchiero, M.McLaren, N.P.Jouppi, M.Fiorentino, A.Davis, N.Binkert, R.G.Beausoleil, and J.H.Ahn, "Corona: System implications of emerging nanophotonic technology", in Proceedings ISCA, June 2008, pp. 153 – 164.

[22] Y.Pan, P.Kumar, J.Kim, G.Memik, Y.Zhang, and A.Choudhary, "Firefly:Illuminating future network-on-chip with nanophotonics", in Proceedings ISCA, June 2009, pp. 429-440.

[23] Y.Pan, J.Kim and G.Memik, "Flexishare: Channel sharing for an energy efficient nanophotonic rossbar", in Proceedings HPCA, Jan 2010, pp. 1 – 12.

[24] J.Chan, G.Hendry, A.Biberman, K.Bergman, and L.P.Carloni, "PhoenixSim: A simulator for physical-layer analysis of chip-scale photonic interconnection networks", in Proceedings DATE, Mar 2010.

[25] G.Hendry, *Architectures and Design Automation for Photonic Networks On Chip*, Ph.D. Thesis. Columbia University, 2011.

[26] F. Xia, L.Sekaric, and Y.Vlasov, “Ultracompact optical buffers on a silicon chip”, in Nature Photonics, vol. 1, 2006, pp. 65–71.

[27] M.Gnan, S.Thorns, D.Macintyre, R.De La Rue, and M.Sorel, “Fabrication of low-loss photonic wires in silicon-on-insulator using hydrogen silsesquioxane electron-beam resist”, in Electronics Letters, vol. 44, no. 2, Jan 2008, pp. 115–116.

[28] J.Cardenas, C.B.Poitras, J.T.Robinson, K.Preston, L.Chen, and M.Lipson, “Low loss etchless silicon photonic waveguides”, in OSA Optics Express, vol. 17, no. 6, Oct 2009, pp. 4752–757.

[29] Q.Xu, S.Manipatruni, B.Schmidt, J.Shakhya, and M.Lipson, “12.5 Gbit/s silicon micro-ring silicon modulators”, in Proceedings CLEO, May 2007.

[30] F.E.Doany, B.G. Lee, S.Assefa, W.M.J.Green , M.Yang, C.L.Schow, C.V.Jahnes, S. Zhang, J. Singer, V.I.Kopp, J.A.Kash, and Y.A.Vlasov, “Multichannel High-Bandwidth Coupling of Ultra-Dense Silicon Photonic Waveguide Array to Standard-Pitch Fiber array”, in Lightwave Technology Journal, vol. 29, 2011, pp. 475-482.

[31] D.Taillaert, H.Chong, P.Borel, L.Frandsen, R.M.De La Rue, and R.Baets, “A compact two-dimensional grating coupler used as a polarization splitter”, in IEEE Photonic Technology Letters, vol. 15, no. 9, Sep. 2003, pp. 1249–1251.

[32] C.Gunn, “CMOS photonics, for high-speed interconnects”, in Proceedings MICRO, vol. 26, no. 2, Mar. 2006, pp. 58–68.

[33] S. J. McNab, N.Moll, and Y.A.Vlasov, “Ultra-low loss photonic integrated circuits with membrane-type photonic crystalwaveguides”, in Optics Express Journal, vol. 11, no. 22, Nov 2003, pp. 2927–2939.

[34] T. Shoji, T.Tsuchizawa, T.Tatanabe, K.Yamada, and H.Morita, “Low loss mode size converter from 0.3 um square Si wire to single mode fibers”, in Electronic Letters, vol. 38, no. 25, Dec 2002, pp. 1669–1670.

[35] X.Zheng, S.Lin, Y.Luo, J.Yao, G.Li, S.Djordjevic, J.Lee, H.Thacker, I.Shubin, K.Raj, J.Cunningham, and A.Krishnamoorty, “Efficient WDM Laser Sources Towards Terabyte/s Silicon Photonic Interconnects”, in Lightwave Technology Journal, vol 31, no 24, Dec 2013, pp. 491-502.

[36] Intel’s Silicon Laser, http://www.intel.com/content/dam/www/public/us/en/documents/intel-research/Silicon-Laser_WhitePaper.pdf. Last accessed on 11/3/14.

[37] X.Tan, M.Yang, L.Zhang, Y.Jiang, and J.Yang, “A Generic Optical Router Design for Photonic Network-on-Chips”, in Lightwave Technology Journal, vol. 30, no 3, Mar 2012, pp. 368-376.

[38] G.Hendry, J.Chan, S.Kamil, L.Oliker, J. Shalf, L.Carloni, and K. Bergman, "Silicon nanophotonic network-on-chip using TDM arbitration", in IEEE Symposium on High Performance Interconnects (HOTI), Aug. 2010, pp. 88–95.

[39] J.Chan and K.Bergman, "Photonic interconnection network architectures using wavelength-selective spatial routing for chip-scale communications", in Optical Communications and Networking Journal, vol. 4, no 3, Mar 2012, pp. 189 – 201.

[40] Y.Yaoyao, X.Jiang, W.Xiaowen, Z.Wei, L.Weichen, and N.Mahdi, "A Torus-Based Hierarchical Optical-Electronic Network-on-Chip for Multiprocessor System-on-Chip", in ACM Journal of Emerging Technologies, vol. 8, no. 5, Feb 2012, pp. 430-456.

[41] J. Goodman, F.Leonberger, K.Sun-Yuan, and R.Athale, "Optical interconnects for VLSI systems", in Proc of IEEE, vol. 72, no. 7, July 1984, pp. 850-866.

[42] E.Carrera and R.Bianchini, "OPTNET: A cost-effective optical network for multiprocessors", in Proceedings ICS, ACM Press, June 2008.

[43] Z.Chen, G.Huaxi, Y.Yingtang, and F.Dongrui, "A Hierarchical Optical Network-On-Chip Using Central-Controlled Subnet and Wavelength Assignment", in Lightwave Technology Journal, vol. 32, no. 5, Mar 2014, pp. 930-938.

[44] H.Li, H.Gu and Y.Yang, "A Hierarchical Cluster-based Optical Network-on-Chip", in ICFCC, May 2010, pp. 823 –827.

[45] N. Kirmanet, and J.Martinez, "A power-efficient all-optical on-chip interconnect using wavelength-based oblivious routing", in Proceedings ASPLOS Mar 2010.

[46] N. Kirman, M. Kirman, R.K.Dokania, J.F.Martnez, A.B.Apsel, M.A.Watkins, and D.H. Albonesi, "Leveraging Optical Technology in Future Bus-based Chip Multiprocessors", in Proceedings MICRO Dec. 2006, pp. 492-503.

[47] R.Morris, *PROPEL: Power & Area-Efficient, Scalable Opto-Electronic Network-on-Chip*, Master's Thesis. Ohio University, 2009.

[48] C.Li, M.Browning and P.V.Gratz, and S.Palermo, "LumiNOC: A Power-Efficient, High-Performance, Photonic Network-on-Chip", in IEEE Transactions on CAD, vol.33 no.6 May 2014, pp. 512-517.

[49] R.Morris, A.K.Kodi, and A.Louri,"Dynamic Reconfiguration of 3D Photonic Networks-on-Chip for Maximizing Performance and Improving Fault Tolerance", in Proceedings MICRO Dec. 2012.

[50] R,Morris, A.Kodi, and A.Louri, "System-Level Modeling and Analysis of Thermal Effects in Optical Networks-on-Chip", in Proceedings MICRO, Dec 2012.

[51] P.Hamdeni, N.Jerger, and S,Hessabi, "QuT: A Low-Power Optical Network-on-Chip", in Proceedings NOCS, Sep 2014.

[52] A.Shacham, K.Bergman, and L.P.Carloni, "Photonic Networks-on-Chip for Future Generations of Chip Multiprocessors", in IEEE Transactions on Computers, vol.57, no 9, June 2008, pp.1246-1260.

[53] A.Shacham, K.Bergman, and L.P.Carloni, "On the Design of a Photonic Network-on-Chip", in NOCS, 2007, pp. 53-64.

[54] Y.Xu, J.Yang, and R.Melhem, "Tolerating Process Variations in Nanophotonic On-chip Networks", in Proceedings ISCA, June 2012, pp.142-152.

[55] A.Qouneh, L.Zongqi, M.Joshi, W.Zhang, X.Fu, and L.Tao, "Aurora: A Thermally Resilient Photonic Network-on-Chip Architecture", in ICCD, Sep 2012, pp.379-386.

[56] A.Biberman, P.Dong, B.G.Lee, J.D.Foster, M.Lipson, and K.Bergman, "Silicon Micro-ring Resonator-Based BroadbandComb Switch for Wavelength-Parallel Message Routing", in Proceedings LEOS Oct 2007. pp. 474 – 475.

[57] M.Piels, J.F.Bauters, M.LDavenport, M.J.R.Heck, and J.E.Bowers, "Low-Loss Silicon Nitride AWG Demultiplexer Heterogeneously Integrated With Hybrid III–V/Silicon Photodetectors", in Lightwave Technology Journal , vol. 32, no 4, Mar 2014, pp. 817 – 823.

[58] B.Larsen, L.Nielsen, K.Zenth, L.Leick, C.Lund, L.Anderson, and K.E.Mattsson., "A low-loss, silicon-oxynitride process for compact optical devices", in Proceedings ECOC, Apr 2003.

[59] C. Bienia, S.Kumar, J.Singh, and K.Li, "The PARSEC Benchmark Suit: Characterization and Architectural Implications", in PACT, Oct 2008.

[60] T. Pimpalkhute, *Heterogeneous prioritization for network-on-chip based multi-core systems*, Master's Thesis, Colorado State University, 2013.

[61] "SystemC initiative", www.systemc.org. Last accessed on 02/24/15

[62] C.Bienia, S.Kumar, J.Singh, and K.Li, "The PARSEC Benchmark Suit: Characterization and Architectural Implications", in PACT, Oct 2008.

[63] N.Binkert, B.Beckmann, G.Black, S.Reinhardt, A.Saidi, A.Basu, J.Hestness, D. Hower, T.Krishna, S.Sardashti, R.Sen, K. Sewell, M.Shoaib, N.Vaish, M.Hill, and D. Wood, "The Gem5 Simulator", ACM SIGARCH Computer Architecture News, vol. 39, no 2, May 2011, pp. 1-7.

[64] C.Sun, C.Chen, G.Kurian, L.Wei, J.Miller, A.Agarwal L.Peh, and V.Stojanovic, "DSENT - A Tool Connecting Emerging Photonics with Electronics for Opto-

Electronic Networks-on-Chip Modeling", in Proceedings NOCS, May 2012, pp 201 – 210.

[65] C.Chen and A.Joshi, "Runtime management of laser power in silicon- photonic multibus NoC architecture", in IEEE Journal of Quantum Electronics, vol. 19, no 2, Mar-Apr 2013, pp. 481-486.

[66] P.Grani and S.Bartolini, "Design Options for Optical Ring Interconnect in Future Client Devices", in ACM Journal of Emerging Technologies, vol.10, no 4, May 2014, pp. 370-395.

[67] S.Beamer, C.Sun, Y.Kwon, A.Joshi, C.Batten, V.Stojanovic, and K.Asanovic, "Re-Architecting DRAM Memory Systems with Monolithically Integrated Silicon Photonics", in Proceedings ISCA, June 2010, pp.129-140.

[68] J.F.Bauters, and J.R.Heck, D.John, D.Dai, M.Tein, J.Barton, A.Leinse, R.Heideman, D.Blumenthal, and J.Bowers, "Ultra-low-loss high-aspect-ratio Si3N4 waveguides", in Optics Express Journal, vol.19, no 4, Feb 2011, pp.3163-3174.

[69] S.N.Adya and I.L.Markov, "Fixed-outline Floorplanning: Enabling Hierarchical Design", in IEEE Transacations, TVLSI, vol 11, issue 6, Dec. 2003, pp.1120-1135.

[70] Dana Seyringer, "Improvement of the Channel Crosstalk in Narrow Channel Spacing Arrayed Waveguide Gratings Applying Specially Shaped Couplers", in Proceedings SPIE May 2010.

[71] R.W.Morris and A.K.Kodi, "Power-efficient and high-performance multilevel hybrid nanophotonic interconnect for multicores", in Proceedings NOCS May 2010, pp 207 – 214.

[72] D.A.B.Miller, "Device requirements for optical interconnects to siliconchips", in Proc. IEEE, July 2009, pp 1166 – 1185.

[73] "Nirgam simulator", http://nirgam.ecs.soton.ac.uk/. Last accessed on 11/3/14.

[74] T. Barwicz, H. Byun, F. Gan, C.W. Holzwarth, M.A. Popovic, P.T. Rakich, M.R. Watts, E.P. Ippen, F.X. Krtner, H.I. Smith, J.S. Orcutt, R.J. Ram, V. Stojanovic, O. Olubuyide, J.L. Hoyt, S. Spector, M. Geis, M. Grein, T. Lyszczarz, and J.U. Yoon, "Silicon photonics for compact, energy-efficient interconnects", in Optical Networking Journal, vol. 6, no. 5, 2007, pp.231-245.

[75] J. Cunningham, S. Ivan, Z. Xuezhe, P. Thierry, M. Attila, L. Ying, T. Hiren, L. Guoliang, Y. Jin, R. Kannan, and A. Krishnamoorthy, "Highly-efficient thermally-tuned resonant optical filters", in Optics Express Journal, vol.18, no 18, Sep 2010, pp. 19055-19063.

[76] A. Joshi, C. Batten, Y-J. Kwon, S. Beamer, I. Shamim, K. Asanovic, and V. Stojanovic., “Silicon-photonic clos networks for global on-chip communication”, in Proceedings NOCS May 2009.

[77] X.Zheng, D.Patil, J.Lexau, F.Liu, G.Li, H.Thacker, Y.Luo, I.Shubin, J.Li, J.Yao, P.Dong, D.Feng, M.Asghari, T.Pinguet, A.Mekkis, P.Amberg, M.Daryinger, J.Gainsley, H.Moghadam, E.Alon, K.Raj, R.Ho, J.Cunningham, and A.Krishnamoorthy, “Ultra-efficient 10Gb/s hybrid integrated silicon photonic transmitter and receiver”, in Optics Express Journal, vol.19, no 6, 2011, pp. 5172-5186.