A REVIEW OF DESIGN APPROACHES FOR ENHANCING THE PERFORMANCE OF NOCS AT COMMUNICATION CENTRIC LEVEL

MISBAH MANZOOR,* ROOHIE NAAZ MIR,† AND NAJEEB-UD-DIN HAKIM‡

Abstract. As the trend of technology shrinking continues a vast amount of processors are being incorporated in a limited space. Due to this almost half of the chip area in Multi-Processor Systems-on-Chips (MPSoCs) is under interconnections, which pose a big problem for communication. Network-on-Chips (NoCs) evolved as a significant scalable solution for removing wiring congestion and communication problem in MPSoCs. NoCs provide the advantage of customized architecture, increased scalability and bandwidth. NoC is a structured framework where communication is the prime concern. In this review paper we present an overview of research and design approaches in the communication centric areas of NoCs. Here we have tried to discuss and iterate most of the available work done for communication in 2D NoCs. This paper gives the insight of different attributes and performance parameters of NoCs. Further it gives a detailed description of how topology, flow control and routing mechanisms can affect the qualitative aspects (performance) of NoCs. It then explains how various attributes of routing can help in increasing the efficacy of NoCs. Subsequently a brief review of different simulators used for NoCs is given. All of this is provided based on the survey of academic, theoretical and experimental approaches presented in the past. Finally some suggestions for future work are also given.

Key words: Network-on Chips (NoCs), 2D NoCs, Performance, Throughput, Latency

AMS subject classifications. 68M10

1. Introduction. As the trend of Moore's law continues, hundreds of billions of transistors are getting incorporated on a single chip. As a result of this the count of transistors and processors on chips is soaring tremendously. This decreased the scalability of buses and posed various problems. Buses caused wiring congestion and delays, due to which communication suffered [1]. As such shared buses were replaced by multiple bridge buses and then they further by crossbars, but still communication demands of growing multi-chip designs could not be met. With the exponential rise of components and IPs, the communication infrastructure began to crumble. Communication in Multi-Processor System on Chips (MPSoCs) became a much costly asset than computation. Hence new communication architecture was needed and thus Network-on-chip (NoC) was welcomed. NoCs enhanced communication and provided higher bandwidth. NoC eases the communication design for System-on-chips (SoCs) as it gets rid of earlier congested and complex communication structures [2]. Table 1.1 shows the qualitative advantages of NoCs over earlier conventionally used buses. As can be seen in Fig 1.1, NoC is made up of building blocks namely routers (which implement routing algorithm), links (which make physical connection between routers), network interface (that connects core and network), processing element (functional blocks which run some application). Each router is connected to a processing element and its neighboring routers through links. The data flows over the links through routers to reach the particular destination. NoC's performance is characterized by different parameters. These parameters are briefly discussed below:

1. *Latency:* It is the time in which a packet goes from source to destination in a network. In other words it is the time frame from the packet generation and packet arrival at destination. Latency depends on topology as well as routing. A network should have lower latency for better performance.

2. *Throughput:* It is the rate of successful arrival of packets to a destination for a certain traffic pattern. It means how many packets are received in a particular simulation time. It is used to measure the

---

*Department of Electronics and Communication, National Institute of Technology Srinagar, India (`misbah@nitsri.net`)

†[2]Department of Computer Science and Engineering, National Institute of Technology Srinagar, India (`naaz310@nitsri.net`)

‡[3]Department of Electronics and Communication, National Institute of Technology Srinagar, India (`najeeb@nitsri.net`)

Table 1.1: Qualitative Advantages of NoCs over Buses

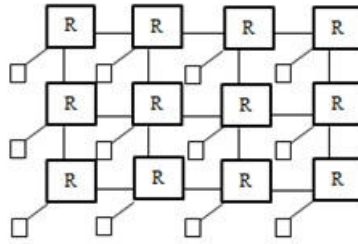| S.No	Parameter	Bus	NoC
1	Scalability	Less	More
2	Throughput	Less	More
3	Frequency	Less	More
4	Gate Count	More	Less
5	Power	More	Less
6	Area	More	Less
7	Energy Efficiency	Less	More
8	Testability	Less	More

Fig. 1.1: Basic 2D Mesh Network-on-chip

performance of entire NoC network. Maximum throughput is calculated when maximum channels get saturated in a network. A network should have high throughput for fruitful performance.

3. *Adaptivity:* It is the property through which packets can take different routes to reach the same destination. It is related to path diversity. More the number of paths to a particular destination more adaptive the routing can become. Adaptivity helps in balancing load in a network. It also provides fault tolerance. This is because there can be multiple paths to reach destination by avoiding the faults. A network should have more adaptivity for good performance.

4. *Power Efficiency:* It is the property of reducing power consumption in the network. It can be done at different levels in NoCs but usually it is done at router architectural level as it consumes more power. A network should consume low power and should be energy efficient.

5. *Fault Tolerance:* It is the network's ability to perform its intended function in presence of faults. Faults can occur either at links or nodes or cores. Fault tolerance deals with successful operations even in presence of breakdowns in the system. It is closely related with reliability. Greater the fault tolerance more reliable is the network. A system should have high fault tolerance.

1.1. Contribution and Structure of Paper. In the previous year's many survey papers have been presented. They are very useful in understanding the NoC's communication architecture and working. But there is still a lack of broad review in the communication centric areas of 2D NoCs. The main aim of the communication centric design approach is to be able to achieve greater design productivity, performance and reliability. So in order to enhance this vision, we are providing an extensive overview of research and design approaches at the communication centric level in Network on chips. This research will help the researchers to get a clear perspective of how to enhance the performance and productivity in NoCs.

We have organized our survey as follows. Section 2 presents the related work. Section 3 introduces the communication centric areas of NoCs (Topology, Flow Control and routing). It discusses different attributes of topology and different types of topologies. Further different flow control schemes presented in the past are discussed. Then routing along with its attributes is introduced and discussed. Different types of routing and their impact on certain important aspects of NoCs such as adaptivity, power, area, fault tolerance and performance of NoCs is presented. Section 4 presents and compares different simulators used for NoCs. Lastly

section 5 concludes the paper and presents some suggestions for future work.

2. Related Work. The authors in [3] presented a paper in which they discussed the performance of wires in scaled technologies. According to them under technology scaling total number of wires were growing exponentially, global wires posed serious problems to designers, as they do not scale in length. The delay across these wires remained constant. The effect of wire capacitance, resistance and inductance was described in detail. Finally they suggested that new structured communication architecture was needed. Then a new concept was introduced, which is considered as the base for understanding network on chips. It laid emphasis on routing packets rather than wires [4]. According to authors routing packets will simplify the layout and structure. It will also increase the performance and modularity. They said that topologies must balance power efficiency with wire utilization. They also explained that a well-developed network interface is essential for high performance. The survey presented in [5] stated the problems of earlier technologies followed by the evolution of network on chips. The authors briefly explained router micro architecture and routing protocols. They also explained different switching techniques followed by different architectural issues and NoC implementation. In another review [6] the basic background of NoC architecture and its design characteristics were presented. Then brief discussion of three types of router designs i.e. circuit switched router design, virtual channel router design and wormhole router design was given. The paper was concluded by considering that virtual channel designs are most efficient. The authors in [7] presented a review and discussed network on chips with its different functional layers followed by different design methodologies. They gave a brief idea of routing and arbitration techniques. Then they shared some quality of service metrics. Finally they introduced a new NoC architecture i.e. Bidirectional network on chip (BiNoC). Another survey of NoC architectures was presented which described the characteristic behavior of different NoC architectures [8]. A brief explanation of switching in different architectures followed by different types of routing was given. It was followed by some insight on different connection types in different NoC architectures and different buffer management techniques.

3. Communication Centric Areas of NoC. As NoC is all about communication between different resources, hence here communication is the supreme concern. There are broadly three communication centric areas of NoCs namely: topology, flow control and routing as shown in Fig 3.1. To target the specifications and performance of an application in NoCs, the designer has to deal with different constraints in these three communication centric areas. Each of the three above mentioned areas play an important role in the communication, as well as maintaining the performance of network since various performance parameters depend on them. These areas have a broader inner classification and relation which is shown and discussed as under. We will continue with the flow given in the Fig 3.1.

3.1. Topology. Topology defines how the network nodes are physically connected. Selecting a good network topology is very important as it has a great impact on cost, performance and fault tolerance. An efficient topology should have less latency and larger bandwidth. Topology selection depends on communication requirements of an application. Applications can use basic NoC topologies or more robust recent ones depending on their specific needs. In this section we will elaborate different types of NoC topologies, their attributes and architectural designs, which make them more robust and performance efficient [9].

Based on the networks regularity and layout, 2D NoC topologies are divided into two broad types namely Regular and Irregular topologies. In regular ones each router is connected to a processing element or core and to a fixed number of other routers e.g. Mesh, Torus, Octagon etc. In irregular ones each router is not connected to its processing element, here some routers are used for transmitting packets only e.g. fat-tree, 3 stage-butterfly etc. as shown in Fig 3.2. Regular topologies have low cost and less design time. However they suffer from scalability as certain real time applications are complicated designs. Reports show that about 60% NoCs employ mesh or torus topology [10]. To meet the complex SoC applications nowadays customized topologies with heterogeneous cores are used. In customized NoCs each router has different number of ports and other hardware internal components for implementing different topologies in different regions. It has proved to increase the performance than regular topologies [11]. Thus NoC topology modeling can be seen as a potential candidate for customized heterogeneous NoCs. Customized NoCs increase the performance by reducing delay and power overheads otherwise caused by regular NoC topologies [12].

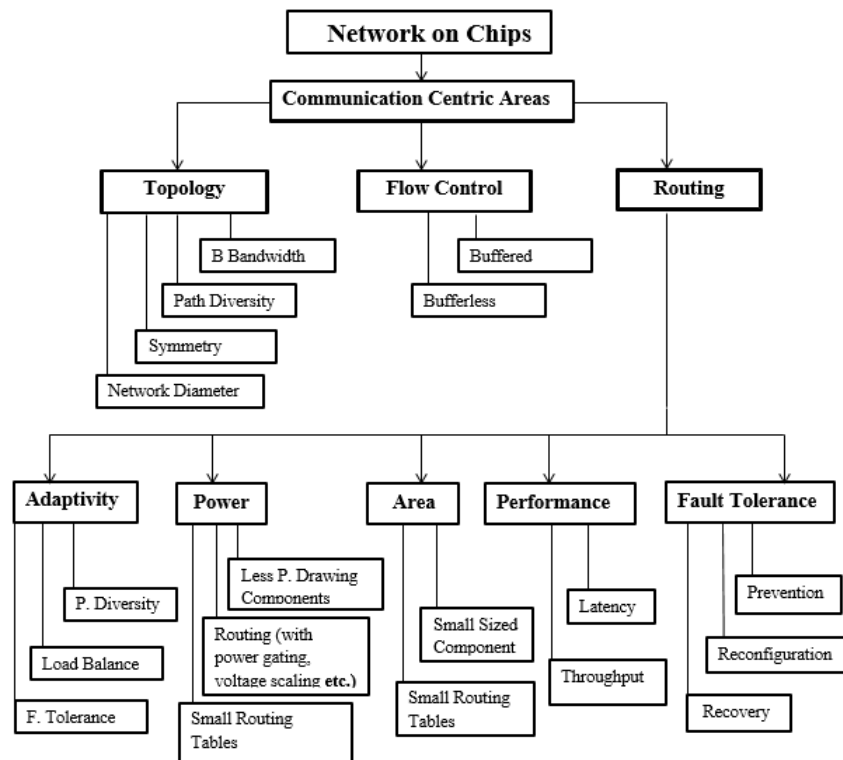A topology has various attributes few of which are discussed below.

Fig. 3.1: Communication Centric areas of Network-on-chip and their attributes

Bisection Bandwidth: Bisection of a network is a cut that divides the network and nodes into two equal halves. It is the minimum bandwidth for all the bisections in a network. Larger bisection bandwidth implies more paths between two half sub networks and thus increases throughput, performance and fault tolerance.

Pathdiversity: Path in a network is the way formed by the channels from one node to another. Path can be minimal (minimum number of hops) or maximal. The more paths or routes in a network, the more robust it is to faults. More number of minimal paths in a network, lesser the latency and better the performance. It also helps in load balancing.

Symmetry: It represents the uniformity in the structure of a network (nodes and links). It plays very important role in load balancing which in turn affects performance. A symmetric network balances load uniformly and makes routing easy as every node can share same routing space in the network.

Network Diameter: It represents the maximum number of shortest paths between all pairs of nodes in a network. It is calculated by number of hops (no. of links it has to pass to reach destination). Smaller network diameter implies smaller hop count and less latency. It helps in managing traffic flow.

Mesh is commonly adopted topology as it is simple and regular but it suffers from scalability at times. Different topologies are being introduced day by day to enhance the performance of NoCs. We are discussing some of them here and then compare their architectural advantages and performance. The work in [13] introduced Cross by Pass Mesh. It combines the regular mesh with some by pass links. The additional links are efficient for reaching longer distances easily with less hop count. These additional links improved symmetry and scalability than regular mesh and torus. Another topology called as centrally connected Mesh (C^2 Mesh) was introduced in[14]. It had four additional links connected to the center. For designing $n \times n$ (C^2) mesh first centre is found, for odd n, one center node is connected to 4 corner nodes while for even n, four center nodes are connected to four corner nodes. It increased scalability and performance as compared to simple mesh. In
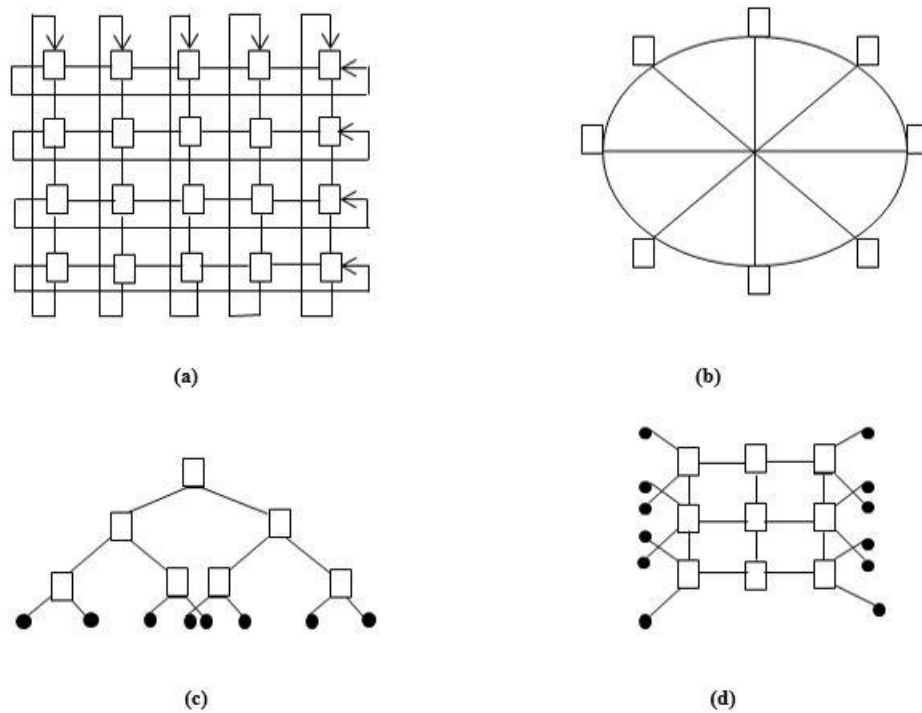
Fig. 3.2: Regular Topologies: (a) Torus (b) Octagon: Irregular Topologies (c) Fat-Tree (d) 3-Stage Butterfly

[15] the authors introduced an area and power efficient NoC named as Diagonally Linked Mesh (D-Mesh). It added additional diagonal ports to router for Quasi-minimal routing. It evenly distributed saturation load. Simulations proved diagonal links are more area and power efficient than normal links. Another novel hybrid topology was proposed in [16]. It was a combination of mesh, torus and folded torus topology. It used three types of links namely torus links, folded tori links and mesh links. These reduced hop-count and increased performance in comparison to regular mesh, regular torus and regular folded torus. Then Quad spare mesh topology was introduced in [17]. It was a simple mesh which was divided into smaller 2×2 meshes and each 2×2 mesh was provided with a spare router for fault tolerance. It was specially used for router and link failures. It re-configures itself in case of failures. It increased fault tolerance with little increase in latency. Another new topology called as Multi Level Mesh was put forward in [18]. This topology differs from simple 2D Mesh by the fact that it has several meshes that share common routers. It resulted in less latency, less power consumption than regular 2D mesh. The authors in [19] proposed a new NoC topology named as PentaNoC. It actually involves cascading any number of pentagon shape blocks using various wrap around links. This reduced hop count and increased path diversity which in turn resulted in good performance. A new NoC topology called as SlimNoC was presented in [20]. It was a scalable, energy and area efficient topology. It had a diameter of 2 which reduces buffer area which again in turn reduces power consumption and also gives better performance. A new star type NoC topology to overcome latency issues of long distance data transmission was introduced in [21]. Here a $n \times n$ mesh is taken and n is assumed as a multiple of 3. Then it is divided into many 3×3 sub meshes and central node of each is connected to the diagonal nodes. These central nodes constitute a second level mesh. Choosing central nodes for long distance communication reduces the latency. It reduced hop count and gave much better performance for large sized NoCs. Table 3.1 highlights the architectural advantages and performance improvements using different topologies.

3.2. Flow Control. It is the mechanism that does the sequencing of data in the path from source to destination. It also allocates different resources like channels, bandwidth, buffers etc. Flow control mainly

Table 3.1: Architectural Advantages and Performance Improvements of Different Topologies

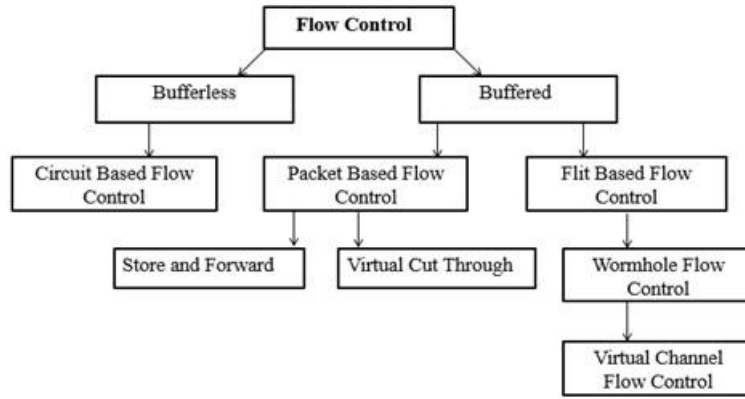| S.No	Architectural and Performance Improvements	CB Mesh	C^2 Mesh	D Mesh	Hybrid Mesh	Quad Spare Mesh	Multi Level Mesh	Penta NoC	Slim Noc	Star NoC
1	Scalability	Y	Y	Y	Y	Y	Y	Y	Y	Y
2	Network Diameter	Y	-	-	-	-	-	Y	-	Y
3	Increased Bisection Bandwidth	Y	Y	-	-	-	-	Y	-	-
4	Reduced Implementation Complexity	-	Y	-	-	-	-	-	-	Y
5	Reduced Area	-	-	Y	Y	-	-	-	Y	Y
6	Reduced Cost	-	Y	-	-	-	-	-	Y	-
7	Increased Throughput	Y	-	-	Y	Y	Y	Y	-	Y
8	Reduced Latency	Y	Y	Y	Y	Y	Y	-	Y	Y
9	Incresed Fault Tolerance	-	-	Y	-	Y	Y	Y	-	-
10	Reduced Power	-	-	Y	Y	-	-	-	Y	Y
11	Adaptivity	Y	Y	Y	-	-	-	Y	-	-

Fig. 3.3: Flow Control Mechanisms

deals with resource allocation and resolves the contention for resources. An efficient flow control mechanism can increase the propagation speed of the packets in the network. It can also help in removing deadlocks and livelocks, as it removes the long waiting periods for packets to share the network resources. Switching forms a part of flow control as it decides how the data flows within the routers and channels. Broadly switching is divided into two types namely circuit switching and packet switching. Circuit switching is a buffer less flow control mechanism whereas Packet switching is a buffered flow control mechanism as shown in Fig 3.3. Adding buffers efficiently increase the flow control. Buffers can be allocated either in terms of packets or flits. If it is allocated in terms of packets then it either forms Store and Forward or Virtual Cut Through flow control mechanisms. On the other hand if it is allocated flit wise it forms wormhole flow control mechanism. Wormhole flow control in turn includes Virtual channel flow control as it is associated with virtual channels. The most commonly used among all these is wormhole flow control as it gives better buffer utilization and less latency. It allows using idle bandwidth, in terms of virtual channels over the same physical channel.

The authors in [22] compared the buffered and bufferless flow control mechanisms and suggested various ways of how these can be optimized. Buffered flow results in high power while bufferless causes deflection, which at times can degrade network performance. Bufferless flow control mostly gives advantage at lower loads while buffered gives better performance at medium and higher loads. Buffered flow can use buffers which serve to

Table 3.2: Advantages of Different Flow Control Techniques

S.No	Architectural and Performance Improvements	Clumsy FC	Prediction Based FC	Distributed Flit Buffer FC	Improved FC	Injection Level FC	Flit Level FC	QLT FC	Fault Tolerant FC
1	Reduced Area	-	-	Y	-	-	-	Y	Y
2	Reduced Cost	-	-	-	-	Y	-	Y	-
3	Reduced Design Complexity	-	-	Y	-	-	-	-	-
4	Increased Packet Injection Rate	Y	Y	-	Y	-	-	-	-
5	Reduced Latency	-	Y	Y	-	Y	Y	-	-
6	Increased Throughput	Y	-	-	-	-	Y	-	-
7	reduced Power	-	-	-	-	Y	-	Y	Y
8	Congestion Control	-	Y	-	-	Y	-	-	-
9	Fault Tolerance	-	-	Y	Y	-	-	Y	Y
10	Better Buffer Utilization	-	Y	-	-	-	Y	Y	-

maximize energy efficiently while bufferless can use a better routing algorithm for reducing latency. A predictive closed loop flow control method was introduced in [23]. The authors here introduced a router model which tells the state of neighbor router based on amount of flits present in their input buffers. Based on this they predict availability of routers locally and as a whole globally. Using this information the packet injection rate in network can be controlled and congestion can be prevented. Thus better performance can be achieved. Further a distributed flit buffer flow control technique was introduced in [24]. The authors merged ack/nack protocol with relay stations distributed on channels. It provided advantage of using smaller router and long wires with ease in physical design. By combining these two approaches better performance was achieved. The authors in [25] introduced a flow control technique named as clumsy flow control. Their aim was to reduce the impact of deflection routing in bufferless NoC routers. So they employed pipelining mechanism which used two stages. One stage was used for calculating the output link and another for assigning that link. It resulted in less deflection with increase in performance. The authors in [26] presented an improved flow control for implementing Minimal Fully Adaptive routing in NoCs. According to duato's theory there are two virtual channels i.e. escape virtual channels (EVCs) and adaptive virtual channels (AVCs). AVCs can be used at any time but EVCs can only be used for packets which follow deadlock free algorithm. There is a constrain of atomic reallocation on AVCs. This paper enhances flow control by demolishing AVC reallocation constrain from various ports on the router. It proposed router architecture where packet exchange occurs from escape virtual channels to adaptive virtual channels non-atomically and hence performance is improved. An injection level flow control was introduced in [27] which was based on calculating the status of paths between source and destination. Destination node only sends the path information to source node and accordingly source adjusts the packet injection rate in the network. Hence injection rate is controlled at different levels in accordance with network payload and thus performance is improved. Another flit reservation flow control mechanism was presented in [28], where the impact of using control flits ahead of data flits was enhanced. Due to this buffers were reserved in advance for upcoming data flits. The advance reservation results in immediate buffer reuse unlike existing techniques which hold the buffer until credit is received. It thus reduces delay and increases throughput. Another flow control method namely Quarter load threshold for wormhole switching was introduced in [29]. Here the authors say when the network saturates, congestion occurs and the buffer state becomes full. According to them congestion can be controlled if only some buffer slots of the total buffer space are used. So they put a limit on buffer state of network and this limit was its quarter load value i.e. quarter buffers of a node should be used. As such balance was achieved between latency and throughput and performance was enhanced. The authors in [30] proposed a fault tolerant flow control for NoCs. Here they handle soft errors and recover these errors at link level. Dynamic packet fragmentation is used here. Upon error detection faulty flits are fragmented and then re-transmitted through new virtual channels. They introduced a router for this which gives 97% error coverage with little area and enhanced reliability. Table 3.2 summarizes and compares the advantages of above mentioned different flow control techniques at architectural and performance level.

3.3. Routing. Routing determines the number of possible paths a message can take to traverse to its destination. It is like a road map. A good routing algorithm should choose smaller (minimal) paths, balance the load and increase the throughput of the network. Also a good routing algorithm should have greater path diversity so that it can provide adaptivity and fault tolerance. Taxonomically routing is classified as deterministic, oblivious and adaptive. In deterministic routing always the same path is used to send the data between source and destination. In oblivious routing data is send through some random path without considering the network state. Deterministic routing is a part of oblivious routing. Adaptive routing uses the information of the network to send the data to destination. It considers factors like congestion, faults etc. and then routes the data accordingly. Routing is the backbone of communication in NoCs as it affects various qualitative metrics or attributes which are discussed below:

3.3.1. Adaptivity. Adaptive routing can be seen as the future of NoCs. Adaptive routing takes into account the network conditions by using the local information from neighbors. This information could be link or node failures, packets waiting for resources (congestion), load information etc. It increases path diversity, balances load and provides greater flexibility and fault tolerance. Adaptive routing gives better performance especially for non-uniform traffic patterns. Adaptive routing mostly results in lower latency and higher through-put along with congestion control. It greatly impacts power consumption in network.

An adaptive routing algorithm for NoCs was proposed in [31]. In it the importance of proper buffer management for increasing throughput was explained. The authors combined two attributes i.e. adaptivity and buffer management and introduced a modified XY routing algorithm. In this algorithm the paths are chosen locally by calculating the available bandwidth in each direction. The direction with highest weight or bandwidth is chosen and accordingly blocks of buffers are arranged in that route and the packet is transferred. It was compared to normal XY, OE and DYAD and gave better performance results than XY and OE. A destination adaptive routing (DAR) based on delay estimates was proposed in [32]. This routing algorithm determines global congestion than local congestion. It is done in two stages; firstly queuing delay from every node to every other node is calculated in a distributed manner. Secondly this delay contributes to the determination of ratios so that traffic can be distributed to the destination between ports of a router. Router architecture for this algorithm was also introduced. As a result of congestion awareness and adaptivity, performance was improved. Another author in [33] proposed a fully adaptive routing algorithm and region based approaches for 2D and 3D NoCs. This algorithm relies on congestion determination and adaptivity. It first detects congestion and then routes the packets. Here the network is divided into clusters and congestion is detected by a group of clusters. Each cluster consists of four routers and a fifth router named as cluster agent. Each cluster firstly gathers congestion information from local routers, and then it distributes the same to neighboring clusters. Hence each router is aware of congestion about every other router and takes the routing decision accordingly. As such performance was improved. The authors in [34] proposed an algorithm called as Adaptive look ahead algorithm. This algorithm is a combination of full adaptive and partially adaptive algorithms. It determines next two hops on a single node (next hop and look ahead hop) and routes the packet region wise. According to it a router is surrounded by four regions. If destination is in region 1, 2 or 3 then look ahead algorithm is followed else a fully adaptive algorithm is chosen. It does not take congestion into account but gives less computational complexity. It gave better results compared to XY and OE. Another adaptive routing namely Dynamic and Mixed routing (MIXROUT) was introduced in [35], which works according to the load status of the network. It actually combines an adaptive routing Multiple and Load-Balance Path Routing (MULTI) and deterministic routing (XY). When the load is high MULTI is used otherwise XY is followed. MULTI works well and mitigates congestion in high loads while XY operates well in low loads as it does not suffer from power and thermal optimization issues unlike MULTI. As such balance is maintained and performance is enhanced. An adaptive table routing, based on hierarchical clusters called as C-routing was presented in [36]. It combines cluster approach and turn model approach. Here a node does not have the cost of all the nodes in the network rather it has its own information and cluster information. First intercluster routing is followed than intracluster. It combines XY and partially adaptive routing. The former is followed in north direction and latter in east, west and south directions. This routing reduces table size and improves performance. The authors in [37] proposed an adaptive routing used specially in MPSoCs. It finds maximum shortest paths between

Table 3.3: Performance Enhancement due to adaptive routing

S.No	Architectural and Performance Improvements	AdR with Buffer Utilization	Destination Adapt. Routing	Region based AdR	Adap. Lookahead Routing	MIX Routing	C Routing	AdR For MPSoC	Centra. AdR	AdR for Reliab.
1	Less Latency	Y	Y	Y	Y	Y	-	Y	Y	Y
2	High Throughput	Y	Y	Y	Y	Y	Y	-	Y	-
3	Congestion Control	-	Y	Y	-	Y	-	-	-	-
4	Reduced Power	-	-	-	-	-	-	Y	-	Y
5	Reduced Cost	Y	-	-	-	-	-	-	Y	-
6	Reduced Area	Y	-	-	-	-	Y	-	-	-
7	Reliability and Fault Tolerance	-	-	-	-	-	Y	-	-	Y
8	Thermal Temperature Stability	-	-	-	-	Y	-	-	-	-

each source destination pair. It is an odd even based deadlock free unicast/multicast routing that works on Hamiltonian method. There are specific routing rules for even and odd rows. Since it is a hamaltonian odd even routing, it is highly adaptive and balances load evenly in the network. An interesting centralized adaptive table based routing for NoCs was introduced in [38]. It monitors traffic effectively in the network and balances load evenly. It is done by two modules namely feedback module (which monitors traffic) and control module (which decides routing path). Based on traffic congestion XY or YX routing is followed by toggling according to congestion. This method balances load very well as compared to distributive adaptive routing. The authors in [39] presented an adaptive routing algorithm which improves the reliability of NoCs. This technique reduced the effect of electron migration, hotspot carrier injection and negative bias temperature instability (NBTI) on lifetime of NoCs. It introduced a concept of packet per port i.e. a metric which balances the stress in the network. It affected the ageing of NoC components and tried to age the components evenly which in turn increase the reliability. Table 3.3 given below summarizes the performance enhancement due to different above mentioned adaptive routing techniques.

3.3.2. Power. Future SoCs require power efficient NoC architectures. Power from NoCs point of view is related to different factors like hop count, complex routing functions and power drawing components of NoCs architecture. If the number of hops taken to reach destination is large, more power will be consumed by the system. So it is preferred to choose minimal routing for minimizing power consumption. Also routing algorithms constituting complex routing tables should be avoided as they consume more power and energy. Moreover as technology is shrinking, leakage power continues to grow and leads to higher power consumption in NoCs. Continuous switching and transactions increase power consumption. Also various NoC components such as routers, buffers, crossbars etc. are all power drawing components and contribute to enhanced power consumption. So the design of power and energy efficient NoCs is the need of the hour.

A number of techniques have been presented for power efficient NoCs. Power consumption can be decreased by making more robust router and link architectures. The main components which contribute to NoC power are the routers and links. So power saving can be done in these two areas from architectural point of view. For routers techniques like buffered and buffer less architecture are applied and for links power gating and voltage scaling can be applied [40]. It was observed that routers consume a significant amount of power, particularly, when they are idle. In routers buffers are the main source of power dissipation. So an alternative was put forward in the form of bufferless routers [41]. In a router without buffers, flow control deflection algorithms are used, where in packets have to be transmitted on their arrival only. Thus, the only required buffers are few pipeline registers. They do result in low power dissipation but they decrease network performance as the load increases. As such they brought a trade off between performance and low power. The authors in [42] proposed a scalable power gating method Turn on-on Turn (TooT) for routers. In NoCs some routers remain idle for a long time, so when they wake up they consume power. (TooT)reduces wake ups and thus reduces power. It avoids powering on a router when a straight packet is forwarded or ejected and powers on only when packets come from other directions. As such it improves static power and energy. The authors in [43] reported that link

power dissipation contributes a greater portion of overall power consumption. Reducing link voltage reduces square times power consumption, so they changed the link voltage based on the communication urgency. NoC with two different kinds of links δ links and λ links were setup. The δ links used the normal voltage and the λ links used a lower voltage. The data packets with the robustness flag high are transmitted on δ links, whereas those with the robustness flag low are transmitted on λ links. This reduced power consumption. Another power saving method was proposed in [44]. It allowed transferring flits rapidly between adjacent routers in half clock cycles by utilizing both edges of the clock. This method leads to link and buffer power reduction along with latency improvement. The authors in [45] have shown the impact of routing algorithm on power and performance. Routing algorithm has impact on buffer size, hop count and router logic. Their power modeling and comparison of different routing algorithms showed more buffer depth more power drawn, more virtual channels more rapidly energy increases, more adaptive algorithm more power it consumes than deterministic ones. A new and hybrid first in first out (FIFO) architecture for reducing power in NoCs was introduced in [46]. Here the authors used a Complementary metal oxide semiconductor (CMOS) Memristor in FIFO which is a non-volatile, scalable and area efficient device. According to them for increasing performance in NoCs, FIFO depth has to be increased which in turn increases power and area. So to overcome this they used CMOS Memristor in it. They used it in RAM block. On implementation it served better results in terms of power and area than conventional FIFOS. A new slow silent virtual channel method for low power NoCs was put forward in [47]. Adding virtual channels increase throughput but as long as bandwidth is not saturated, after that leakage power starts. So they incorporated low power techniques with virtual channels. They used run time power gating for each virtual channel, sleep control methods for different wake up periods, routing techniques to overcome standby power in addition to voltage and frequency scaling. The authors in [48] introduced a new low power router architecture namely centralized buffer router. They used centralized buffers plus elastic buffered links in its router architecture. The centralized buffer in router is used only when a packet from input buffer cannot go to its output buffer. It has many pipelining stages, and the control and data information splits in the links. They also provided a deadlock avoidance method. The technique gave much improvement in power and latency.

3.3.3. Area. Area is also an important factor as far as high performance NoCs are considered. Area increases power also increases. Area can be minimized by many methods like by using smaller routing tables for routing, by minimizing the size of VLSI components and resources, by using less complicated reconfigurable designs etc. A new router architecture for area reduction and power efficiency was introduced in [49]. It presents a new virtual channel sharing technique named as partial virtual channel sharing. According to it sharing of resources is essential for minimizing chip area. Here virtual channel buffer is shared by other ports in a router based on communication needs. As such buffer utilization increases without causing significant area overhead. The authors in [50] presented an area efficient partially reconfigurable crossbar switch for NoCs. Reconfigurable NoCs usually have larger area because they have more complex crossbar switch design. Here a partially reconfigurable crossbar switch design is presented which has smaller area and less delay. It is made up of look up tables (LUTs) and reconfiguration takes place by changing these LUTs which in turn reduce area. They also proposed an algorithm for making connections in crossbar switches. A new router architecture which reduces area and increased speed in NoC architectures was introduced in [51]. The authors modified three components inside a router namely crossbar switch (replaced with less number of LUTs), buffers(instead of 16 bit used 8 bit buffers) and a decoder(instead of decoder used two input one output OR gate). The area is measured in terms of number of LUTs used. This resulted in large area improvement as compared to conventional NoC router. Ihe authors in [52] also presented an area and power efficient router for NoCs. This router used wormhole routing, with a simple deterministic algorithm, followed by flow control and decoding techniques. They used two types of crossbar switches namely multiplexer and matrix for area efficiency. They showed that multiplexer type crossbar gives better efficiency than matrix type in terms of area and power. A low area router architecture for NoCs based on HDL was proposed in [53]. Here it uses two crossbar switches instead of one. The concept is a router which has many small crossbar switches performing same function as that with a large crossbar. It consumes less area with small overhead in latency. The latency here is minimized by providing a low latency algorithm named as predominant routing algorithm. This combination reduced area to a greater extent. Again a new router architecture named as Inter-Router Dual-function Energy and Area

efficient Links (iDEAL) for area efficient NoCs was proposed in [54]. For optimizing NoCs in terms of area buffer plays an important role. Reducing number of buffers in a router reduces area and some performance. So here in this technique number of buffers is reduced to reduce area at the same time some adaptive dual function links are introduced in router architecture. They store as well as transmit data as per requirement. Hence reduction in area and compensation for performance was obtained. The authors in [55] proposed a new area and energy efficient architecture for NoCs. They took advantage of circuit switching and multistage circuit-switching network (CLOS) network and combined it. Furthermore they proposed a heterogeneous router architecture which used a combination of buffered and bufferless routers. They also incorporated lane division multiplexing in router. On implementation, combination of CLOS and circuit switched switch performed better than normal crossbar with a large reduction in area. A new area efficient reconfigurable router architecture for NoCs was introduced in [56]. Design was carried using hardware description language. It had a dedicated channel for each direction i.e. (E,W,N,S) along with it buffers and multiplexers. Each channel has 5 multiplexers. Fixed priority arbiter is used for reducing area of this router. It showed improvement in area and latency. An area efficient table based routing for irregular NoCs was presented in [57]. The authors lay stress on finding paths, which could find higher similarity by routing methods. This surpassed the problem of region based routing along with area reduction.

3.3.4. Performance. Performance in NoCs can be determined by latency and throughput. By achieving low latency or high throughput or both we can enhance the performance. Lower latency is much needed as we are in fast communication era. By designing low latency networks we can enhance the speed and also get high throughput. We can enhance performance by working on different NoC architectures, using adaptive routing algorithms, using smaller routing tables, routing architectures (buffers, allocators) etc.

An adaptive routing called zigzag routing was introduced in [58]. In it packets move in alternate x and y directions. Initially distance to destination is calculated and compared. The data is first send in the direction which has greater distance until the distance in both the dimensions becomes the same, then data alternates between x and y dimensions. It decreases latency and power and increases the performance. The authors in [59] presented an FPGA based NoC architecture. They also made it easy to interface it with a bus protocol like wishbone. They made it flexible for implementing into any topology and used wormhole routing to reduce latency. This NoC architecture was optimized for Vertex 5 FPGA and achieved low latency, high throughput and low area. A new NoC architecture for high throughput and high performance called as High Throughput Butterfly Fat Tree (HTBFT) was proposed in [60]. It is modification of butterfly fat tree architecture. It gives efficient performance in parallel machines. It has a four array tree with switches connected to four down links and two uplinks. Its switch architecture is less complex than normal butterfly tree structure (BFT). It was synthesized in Xilinx ISE and was seen that throughput increased with marginal latency changes as compared to BFT architecture. Another NoC router architecture was proposed in [61], which helped in achieving low latency and high throughput. The authors modified normal router by injection and ejection ports, spilited a packet into two halves, implemented two different DOR routing methods along some changes in hardware design. All the changes combined to increase the path diversity and increased performance as well. Likewise another high throughput router for NoCs was introduced in [62]. Stress was laid on bond between neighboring routers. The authors designed a router in which buffer and allocators were modified. They also put forward neighbor flow regulation algorithm which worked in coordination with modified structures. Simulation results show improvement in throughput and latency due to coordination between neighboring routers. X-Network which is a high performance wormhole switching network was proposed in [63]. In this network a router is shared by four processing elements and is also in contact with other routers in different directions (E,W,N,S). As a result of this hop count decreases. It was compared to conventional NoC architecture on different routing algorithms and traffic patterns. It performed better and hence increased performance. In [64] proposed high throughput router architecture for NoCs. There are different components in router architecture which affect the performance, switch allocator being one among them. They modified the arbiters of first stage switch allocators. They replaced simple round robin arbiter with a predefined priority based round robin arbiter in first stage which in turn reduces the contention in the second stage. As a result throughput increases linearly with number of virtual channels. It also reduces design complexity and latency. A new heterogeneous topology for achieving low latency and high throughput was proposed in [65]. They combined tree and mesh topology
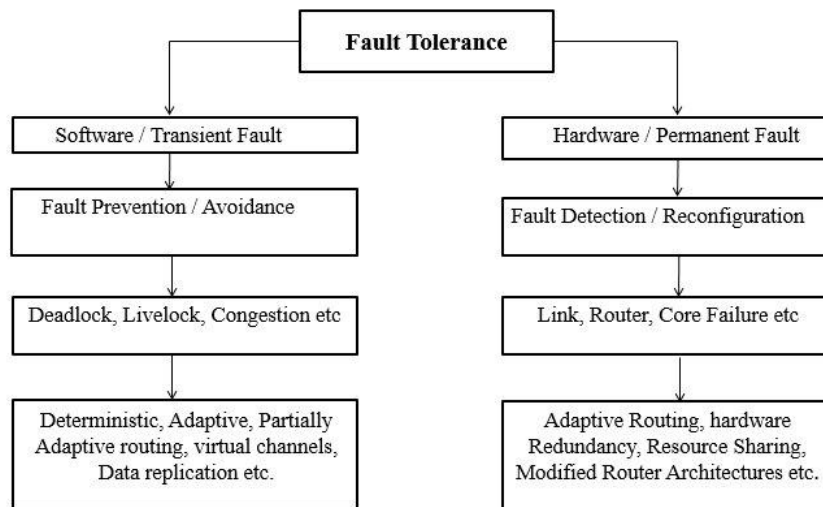
Fig. 3.4: Fault Tolerant Methods

to design a hybrid topology. Tree network provided low latency and mesh network a high throughput. As a result they obtained the advantages of both. They also provided an algorithm based on hop count to manage contention and latency. The algorithm was tested for different traffic patterns and gave very good performance and energy consumption.

3.3.5. Fault Tolerance and Reliability. It has been observed as the number of processors and transactions increase in MPSoCs, a large number of components fail during manufacturing and integration and another huge lot during the operations. It happens because faults may occur in different forms at different levels. This leads to a decrease in performance and reliability. Thus fault tolerance is required at different levels of abstraction. Adopting fault tolerant approaches is imperious for maintaining the required and appropriate performance of NoCs. Fault tolerance increases the reliability of the system. Reliability is a measure to determine how many times the network works correctly performing its task of delivering messages. To design a reliable network its error handling capability should be increased.

Fault tolerant routing is classified on the basis of faults it tolerates. First category includes faults that can be prevented or avoided. These are software faults like deadlock, livelock, congestion etc. The second category includes faults which need detection and reconfiguration i.e., hardware faults like link, node or core failures. For the prior one, routing techniques which avoid these faults are designed and for the latter one different routing techniques plus reconfiguration is needed (fault tolerance and reconfiguration) as shown in Fig 3.4. Reconfiguration and recovery are also as important as prevention. It includes methods and means to bring back a system from faulty to functioning state. It involves the circuit components like routers, re-configuring routing paths, re-configuring using various redundancies like component or information redundancies etc. The most basic fault tolerant routing used for deadlock prevention is deterministic XY routing, where a packet is first routed horizontally until it reaches the column of destination and then routed vertically. YX follows the opposite of XY, it routes packet first vertically until it reaches destination row and then horizontally. The authors in [66] introduced turn models which were partially adaptive in nature. They restricted certain turns in certain directions to break the cycles which otherwise created deadlocks. They presented negative first turn model (restricted positive to negative turns), west first turn model (restricted west last turns) and north last turn model (restricted north first turns). Their simulations showed that partially adaptive algorithms perform better for non-uniform traffic than deterministic ones. In [67] the author introduced odd even turn model which is fully adaptive and deadlock free. It restricted some of the locations where some of the turns were forbidden. There are separate routing rules for even and odd columns and it resulted in uniform adaptivity. Results showed

Table 3.4: Type of Fault Tolerance provided by different techniques

S.No	Architectural and Performance Improvements	XFA	Basic TM	Odd Even TM	Non-Mini. Fully AdR Using VCT	Selec. Exten-sion Rout. Alg. Based TM	Fault. Tol-erant Dead-lock free Rout.	Fault. Tolera. Reconf. NoC	Traffic Aware Re-conf. Archi-tec-ture	Reliable NoC Archi-tecture	Fault. Toler. NoC Relia. Improve-ment
1	Deadlock Freedom	Y	Y	Y	Y	Y	Y	-	Y	-	-
2	Congestion Control	Y	-	-	-	Y	Y	-	-	-	Y
3	Livelock Freedom	-	Y	Y	-	-	Y	-	Y	-	-
4	Router Failure	-	-	-	-	-	-	Y	Y	-	Y
5	Link Failure	-	-	-	-	-	Y	-	-	Y	-

that adaptive algorithms give better performance for non-uniform traffic patterns. Another deadlock free fully adaptive routing was presented in [68]. It uses non minimal paths. It avoids deadlocks using turn models but with the help of arbitration and output selection support of routers. It allowed packets to select an output port in minimum time after arrival, even if it is not the shortest path. Also throughput improvement was achieved. The author in [69] introduced a deadlock free XY fully adaptive (XFA) routing. It is an adaptive form of XY routing and sends data through minimal paths. It gives less latency and better throughput than regular XY routing for non-uniform traffic patterns. A novel method of increasing the adaptivity and deadlock free property of turn models was presented in [70]. The authors redesigned odd even and a Low Weight and Highly Adaptive Routing (LEAR) algorithms. They released some turn restrictions of adaptive algorithms. They also used a mixed switching technique, mixing virtual cut through and wormhole switching. This provided different paths and logical separation of data along with fault tolerance and latency improvement. The authors in [71] proposed fault tolerant scheme which is deadlock and livelock free. This scheme also combats congestion and link failures. For deadlock prevention a timeout scheme is used along with packet reinjection by a virtual source. An algorithm for livelock avoidance is also given. This technique proved more reliable and efficient for removing deadlocks and livelocks. A reconfigurable fault tolerant technique was presented in [72] which can handle single or multiple router failures. Here routers are modified which recover the cores connected to faulty routers. It uses routing tables for network updates. If a router fails, it gets detached to the core and the routers surrounding this faulty router use their unused ports and attain the abandoned core and maintain the performance. In [73] the authors proposed a reconfigurable rerouting algorithm for faulty links. Here a test is carried out for detecting faults, once a fault is detected every switch reconfigures itself for finding rerouting alternative paths to bypass faulty links. It uses transaction level model platform for simulation. The work in [74] presented a fault tolerant reconfigurable NoC architecture. It is used for saving cores attached to faulty routers. Here a core is attached to two routers, master router (original router) and slave router (neighboring router). If a master router fails, the slave router reconfigures and saves the core. This fault tolerant architecture reduced latency and improved reliability as compared to normal mesh. The authors in [75] presented a reconfigurable fault tolerant technique for bypassing faulty routers. It modified the router structure and used Dynamic XY (DyXY) routing which makes it robust for deadlock and livelock also. Its router contains additional reconfiguration control unit which helps in reconfiguration. It uses two virtual channels and an adaptive minimal routing for bypassing the faulty routers. Table 3.4 given below summarizes the type of fault tolerance provided by different techniques.

4. Network on Chip Simulators. This section presents some of the commonly used simulators for Network on chips. Fig 4.1 below presents a general block diagram of a NoC simulator. A general NoC simulator should atleast have an application package, a simulation package and a performance model. To select a particular NoC tool the user must know what operating system or installation platform he needs, what kind of topology he needs to work on, what traffic patterns he needs, what output performance parameters he needs to calculate and accordingly select a particular tool. Given below we have presented some of the commonly used NoC tools. Their comparison is presented in Table 4.1.
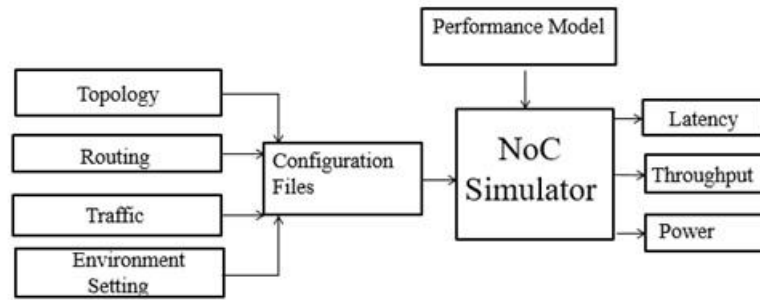
Fig. 4.1: Block Diagram of a general Network on Chip Simulator

NS 2 It is an open source event driven simulator. It is used for research in communication networks and simulating Network on chips. It uses C++ and oTcl language. The simulation process in NS 2 includes topology generation, model development, node and link configuration, execution, performance analysis and graphical visualization. It provides simulation in wired as well as wireless networks.

Nostrum It is a highly customizable Network on chip simulator. It has a command line interface. In addition to routing, finding performance parameters it also supports application mapping to networks. The user has the flexibility to configure different simulation parameters. It provides best effort and guaranteed communication results. It is a packet switched communication platform.

Nirgam It is an open source cycle accurate event driven simulator for Network on chips. It uses System C language and takes log file input for simulation. Earlier it was developed and used for 2D mesh and torus but now it is extended to 3D. It very flexible and allows the user to configure various parameters according to need like no. of virtual channels, buffer size, clock frequency etc.

Noxim It is a system C based cycle acuurate Network on chip simulator. It has a command line input interface. It is also a highly configurable tool and allows configuration of parameters by user. It uses Orion power model to calculate the power. It supports both wired and wireless NoC architectures.

DARSIM It is a parallel, highly configurable cycle accurate Network on chip simulator. It has a parallel simulation engine which enhances the speed and synchronization. It also supports parameterized table based designs. It supports static, oblivious and adaptive routing. It simulates both 2D and 3D mesh.

Gem 5 It is an open source discrete event driven simulator for computer system architecture, system level architecture and microarchitecture level in Network on chips. It takes a command line input. It provides the flexibility of rearranging, parameterizing, extending and replacing the components according to the need.

NoCTweak It is an open source System C based Network on Chip simulator. It works in command line. It is a very flexible, parameterizable tool and provides faster simulation speed. It calculates performance (latency, throughput, power) as well as energy of NoCs. It also facilitates calculation of area, timing and router components.

Booksim It is a cycle accurate simulator for network on chips. It offers the capability of modifying a large number of network parameters. It also allows network component designs especially router microarchitecture. It is capable of implementing almost all the algorithms. It is one of the most widely used simulators. Table 6 given below gives the comparison of NoC simulators presented.

5. Conclusion. This paper emphasizes the role of NoCs for communication in many core MPSoCs. Different performance parameters are discussed as they determine the efficiency of NoCs. Then stress is laid on communication centric areas, which include topology, flow control and routing. They being the backbone for effective NoC communication, their different types, their attributes, and their impact on performance are discussed. Different types of routing for increasing adaptivity, performance, fault tolerance and reducing power

Table 4.1: Comparison of NoC Simulators

S.No	Simulator	Year	Company	Installation Platform	Language	Topology	Traffic Pattern	Performance Metrics
1	NS 2	1995	DARPA	Windows with Cygwin	C++ and oTcl	Wide range	Synthetic	Latency and Throughput
2	Nostrum	2005	Royal Institute of Technology Stockholm	Linux	System C	2D mesh and Torus	Synthetic	Latency, Throughput and Link Utilization
3	Nirgam	2007	University of Southampton UK and MNIT Jaipur	Linux	System C	2D Mesh and Torus	Synthetic and Embedded	Latency, Throughput and Power
4	Noxim	2010	Catagne University	Linux	System C	2D Mesh	Synthetic	Latency, Throughput and Power
5	DARSIM	2010	Massachusetts Institute of Technolgy	Linux	C++	2D and 3D Mesh	Trace driven injector or Cycle level MIPS Simulator	Latency and Throughput
6	Gem 5	2011	Joint Collaboration of AMD, ARM, HP, MIPS, Princeton, MIT and Universities of Michigan,Texas and Michigan	Linux, Solaris, MacOS	C++	Wide range	Synthetic and Embeded	Throughput and Energy
7	NoCTweak	2012	ST Microelectronics, UC Micro, Intel and University of California, Davis	Linux	System C	Mesh, Torus and Ring	Synthetic and Embeded	Latency, Throughput, Power and Energy
8	Booksim	2013	National science foundation and Stanford Pervasive Parallelism Laboratory	Linux	C++	Wide range	Synthetic	Latency and Throughput

and area are discussed. Also some NoC simulators are briefly discussed. By analyzing the work done in the past and the survey presented above, we have come to the conclusion that performance of NoCs can be enhanced by working in communication centric areas. Analyzing in depth different attributes of each communication centric area major improvement in performance can be achieved. Also we should lay stress and effort on following research areas in future.

1. NoC topology modeling for irregular networks can be interesting area to work.
2. Customized topology modeling for heterogeneous NoCs needs research.
3. For flow control an important side to look and do research is fairness (schemes which fairly allocate resources in the network).
4. Flow control is closely related to router microarchitecture and pipelining, so these can also be important areas of research.
5. Customizing routing algorithms which would enhance quality of service (QoS) metrics (less latency, high throughput,congestion control, increased adaptivity etc.) can be an area of research.
6. Power aware routing (routing algorithms and techniques which reduce network activity for reducing power) can be taken as area of highly needed research.
7. Fault tolerance and reliability at communication centric level (fault tolerant topologies, fault tolerant flow control schemes, fault tolerant routing) can be taken as an area of research.
8. NoC traffic modeling can be future area of research so that realistic performance of NoCs can be obtained and analyzed.
9. Also development of different real traffic simulation platforms and tools for giving the on ground performance of NoCs can be taken as an important research area.

REFERENCES

[1] S. Sharma, C. Mukherjee and A. Gambhir, *A Comparison of Network-on-chip and Buses*, Proceedings of National Conference on Recent Advances in Electronics and Communication Engineering, pp. 1-7, 2014.

[2] M. Ali, M. Welzl and M. Zwicknag, *Networks on Chips: Scalable Interconnects for Future Systems on Chips*, Proceedings of IEEE Computer Society Annual Symposium on VLSI, pp. 240-245, 2002.

[3] R. Ho, K.W. Mai and M.A. Horowitz, *Future of Wires*, Proceedings of IEEE, pp. 486-492, 2000.

[4] W.J. Dally and B. Towles, *Route Packets, not Wires: On-Chip Interconnection Networks*, Proceedings of the 38th Design Automation Conference (IEEE Cat. No.01CH37232), pp. 684-689, 2001.

[5] A. Agarwal and R. Shankar, *Survey of Network on Chip (NoC) Architectures and Contributions*, Journal of Engineering, Computer and Architecture. Vol.3, No. 1, pp. 1-15, 2009.

[6] Bharti B. Sayankar and S. S. Limaye, *Overview of Network on Chip Architecture*, 2nd National Conference on Information and Communication Technology, pp. 33-36, 2011.

[7] Wen-Chung Tsai, Ying-Cherng Lan and Yu-Hen Hu, *Networks on Chips: Structure and Design Methodologies*, Journal of Electrical and Computer Engineering. pp. 1-15, 2012.

[8] Muhammad Athar Javed Sethi, Fawnizu Azmadi Hussin and Nor Hisham Hamid, *Survey of Network on Chip Architectures*, Science International Journal. Vol.27, No.5, pp. 4133-4144, 2015.

[9] S. Kumar, A. Jantsch, J.-P. Soininen, M. Forsell, M. Millberg, J. Oberg and K. Tiensyrja; A. Hemani, *A Network on Chip Architecture and Design Methodology*, Proceedings of the IEEE Computer Society Annual Symposium on VLSI, pp. 1-8, 2002

[10] Salma Hesham; Jens Rettkowski; Diana Goehringer; Mohamed A. Abd El Ghany, *Survey on Real-Time Networks-on-Chip*, IEEE Transactions on Parallel and Distributed Systems, Vol. 2, No. 3, pp. 1500 – 1517. 2016.

[11] Bjerregaard Tobias and Mahadevan Shankar, *Survey on Real-Time Networks-on-Chip*, ACM Computing Surveys, Vol. 38, No. 1, pp. 1-51, 2016.

[12] Konstantinos Tatas, Kostas Siozios, Dimitrios Soudris and Axel Jantsch, *Designing 2D and 3D Network-on-Chip Architectures* TEX, Springer Publications New York Heidelberg Dordrecht, London, 2014.

[13] Usman Ali Gulzari, Sheraz Anjum, Shahrukh Aghaa, Sarzamin Khan and Frank Sill Torres, *Efficient and Scalable Cross-by-pass-Mesh Topology for Networks-on-Chip*, IET Computer Digital Technology, Vol. 11, No. 4, pp. 140-148, 2017.

[14] Lalit Kishore Arora and Rajkumar, C^2 *Mesh : New Interconnection Network Topology Based on 2D Mesh*, 3rd IEEE International Advance Computing Conference, pp. 282-286, 2012.

[15] Chifeng Wang, Wen-Hsiang Hu, Seung Eun Lee and Nader Bagherzadeh, *Area and Power Efficient Innovative Network-on-Chip Architecture*, 18th Euromicro Conference on Parallel, Distributed and Network-based Processing, pp. 533-539, 2015.

[16] K. Swaminathan, Sandeep Gopi, Rajkumar, G. Lakshminarayanan and Seok-Bum Ko, *A Novel Hybrid Topology for Network on Chip*, IEEE 27th Canadian Conference on Electrical and Computer Engineering, pp. 1-6, 2014.

[17] Yu Ren, Leibo Liu, Shouyi Yin, Qinghua Wu, Shaojun Wei and Jie Han, *A VLSI Architecture for Enhancing the Fault Tolerance of NoC using Quad-spare Mesh Topology and Dynamic Reconfiguration*, IEEE International Symposium on Circuits and Systems, pp. 1-19, 2013.

[18] Mohsen Saneei, Ali Afzali-Kusha and Zainalabedin Navabi, *Low-latency Multi-Level Mesh Topology for NoCs*, IEEE 18th International Conference on Microelectronics, pp. 36-39, 2006.

[19] Ahlem Boudellioua andNasser Alzeidi, *PentaNoc: A New Scalable and self-similar NoC Architecture*, The 5th International Workshop on Design and Performance of Networks on Chip, pp. 358-364, 2018.

[20] Maciej Besta, Syed Minhaj Hassan, Sudhakar Yalamanchili, Rachata Ausavarungnirun, Onur Mutlu and Torsten Hoefler, *SlimNoC: A Low-Diameter On-Chip Network Topology for High Energy Efficiency and Scalability*, Proceedings of the Twenty-Third International Conference on Architectural Support for Programming Languages and Operating Systems, pp. 43-55, 2018.

[21] Kuan-Ju Chen, Chin-Hung Peng and Feipei Lai, *Star-Type Architecture with Low Transmission Latency for a 2D Mesh NOC*, IEEE Asia Pacific Conference on Circuits and Systems, pp. 919-921, 2010.

[22] George Michelogiannakis, Daniel Sanchez, William J. Dally and Christos Kozyrakis, *Evaluating Bufferless Flow Control for On-Chip Networks*, Proceedings of the 4th ACM/IEEE International Symposium on Networks-on-Chip, pp. 1-8, 2010.

[23] U.Y. Ogras and R. Marculescu, *Prediction-based Flow Control for Network-on-Chip Traffic*, 43rd ACM/IEEE Design Automation Conference, pp. 839-844, 2016.

[24] Nicola Concer, Michele Petracca and Luca P. Carloni, *Distributed Flit-Buffer Flow Control for Networks-on-Chip*, Proceedings of the 6th IEEE/ACM/IFIP International Conference on Hardware/Software Co-design and System Synthesis, pp. 215-220, 2008.

[25] S. Anu Karpaga1 and D. Muralidharan, *High Throughput Pipelining NoC using Clumsy Flow Control*, Indian Journal of Science and Technology, Vol. 9, No. 29, pp.1-5, 2016.

[26] Alireza Monemi, Chia Yee Ooi, Muhammad Nadzir Marsono and Maurizio Palesi, *Improved Flow Control for Minimal Fully Adaptive Routing in 2D Mesh NoC*, , Proceedings of 9th International Workshop on Network on Chip Architectures, pp. 1-6, 2016.

[27] Minghua Tang and Xiaola Lin, *Injection Level Flow Control for Network-on-Chip (NoC)*, Journal of Information Science and Engineering, Vol. 27, No. 3, pp. 527-544, 2011.

[28] Li-Shiuan Peh and W.J. Dally, *Flit-Reservation Flow Control*, Proceedings of the 6th International Symposium on High-

Performance Computer, pp. 73-84, 2011.

[29] Minghua Tang and Xiaola Lin, *Quarter Load Threshold (QLT) flow control for wormhole switching in mesh based Network-on-Chip*, Journal of System Architectures, Vol. 56, No. 2, pp. 452-462, 2010.

[30] Young Hoon Kang, Taek-Jun Kwon and Jeffrey Draper, *Fault-Tolerant Flow Control in On-Chip Networks*, Fourth ACM/IEEE International Symposium on Networks-on-Chip, pp. 79-86, 2010.

[31] Syeda Tanjila Atik, M.M. Imran, Julkar N Mahi, Jenia A Jeba, Z I Chowdhury and M S Kaiser, *An Adaptive Routing Algorithm for on-chip 2D Mesh Network with an Efficient Buffer Allocation Scheme*, International Conference on Computer, Communication, Chemical, Material and Electronic Engineering, pp. 1-4, 2018.

[32] Rohit Sunkam Ramanujam and Bill Lin, *Destination-Based Adaptive Routing on 2DMesh Networks*, ACM/IEEE Symposium on Architectures for Networking and Communication Systems, pp. 1-12, 2010.

[33] Masoumeh Ebrahimi , *Fully adaptive routing algorithms and region based approaches for two dimensional and three dimensional networks on chip*, IET Computer Digital Technology, Vol. 7, No. 6, pp. 264–273, 2013.

[34] Abhilash Menon, Lian Zeng, Xin Jiang and Takahiro Watanabe, *Adaptive Look Ahead Algorithm for 2-D Mesh NoC*, IEEE International Advanced Computing Conference, pp. 299-302, 2015.

[35] Gaoming Du, Dayi Liang, Yukun Song and Duoli Zhang, *A Dynamic and Mixed Routing Algorithm for 2D Mesh NoC*, International Conference on Anti-Counterfeiting Security and Identification, pp. 1-4, 2014.

[36] Manas Kumar Puthal; Virendra Singh; M.S. Gaur; Vijay Laxmi, *C-Routing: An Adaptive Hierarchical NoC Routing Methodology*, IEEE/IFIP 19th International Conference on VLSI and System-on-Chip, pp. 1-6, 2011.

[37] Poona Bahrebar and Dirk Stroobandt, *Adaptive Routing in MPSoCs using an Efficient Path-based Method*, International SoC Design Conference, pp. 31-34, 2013.

[38] Ran Manevich, Israel Cidon, Avinoam kolodny and Isask'har Walter, *Centralized Adaptive Routing for NoCs'*, IEEE Computer Architecture Letters, Vol. 9, No. 2, pp.57-60, 2010.

[39] Juman Alshraiedeh and Avinash Kodi, *An Adaptive Routing Algorithm to Improve Lifetime Reliability in NoCs Architecture*, , IEEE International Symposium on Defect and Fault Tolerance in VLSI and Nanotechnology Systems, pp. 1-4, 2016.

[40] Emmanuel Ofori-Attah and Michael Opoku Agyeman, *A Survey of Low Power NoC Design Techniques*, Proceedings of 2nd International Workshop on Advanced Interconnect Solutions and Technologies for Emerging Computer Systems, pp. 22-27, 2017.

[41] Emmanuel Ofori-Attah, Washington Bhebhe and Michael Opoku Agyeman, *Architectural Techniques for Improving the Power Consumption of NoC-Based CMP*, Journal of Low Power Electronics and Application, Vol. 7, No. 2, pp. 1-24, 2017.

[42] Hossein Farrokhbakht, Mohammadkazem Taram, Behnam Khaleghi and Shaahin Hessabi, *TooT: an efficient and scalable power-gating method for NoC routers*, Tenth IEEE/ACM International Symposium on Networks-on-Chip, pp. 1-8, 2016.

[43] Andrea Mineo; Maurizio Palesi; Giuseppe Ascia; Vincenzo Catania, *Runtime Online Links Voltage Scaling for Low Energy Networks on Chip*, 16th Euromicro Conference on Digital System Design, pp. 941-944, 2013.

[44] Anastasios Psarras, Junghee Lee, Pavlos Mattheakis, Chrysostomos Nicopoulos and Giorgos Dimitrakopoulos, *A Low-Power Network-on-Chip Architecture for Tile-based Chip Multi-Processors*, International Great Lakes Symposium on VLSI, pp. 335-340, 2016.

[45] Mahdiar Ghadiry,Mahdieh Nadi, M.T. Manzuri and Dara Rahmati, *Performance and Power Analysis of Routing Algorithms on NOC*, IEEE Transactions on Computers, Vol. 54, No. 8, pp. 321-327, 2005.

[46] Mohammed E. Elbtity and Ahmed G. Radwan, *Small Area and Low Power Hybrid CMOS-Memristor Based FIFO for NoC*, International Symposium on Circuits and Systems, pp. 1-5, 2018.

[47] Hiroki Matsutani, Michihiro Koibuchi, Daihan Wang and Hideharu Amano, *Adding Slow-Silent Virtual Channels for Low-Power On-Chip Networks*, Second ACM/IEEE International Symposium on Networks-on-Chip, pp. 23-32, 2008.

[48] Syed Minhaj Hassan and Sudhakar Yalamanchili, *Centralized buffer router: A Low Latency, Low Power Router for High Radix NoCs*, Seventh IEEE/ACM Intenational Symposium on Networks-on-Chip, pp. 1-8, 2013.

[49] K.S. Kashyap, B.B. Sayankar and P. Agrawal , *Power and Area Efficient NOC Router Through Utilization of Idle Buffers*, International Journal Computer, Engineering and Application, Vol. 6, No. 1, pp. 31-39, 2014.

[50] Chin Hau Hoo and Akash Kumar, *An Area-Efficient Partially Reconfigurable Crossbar Switch With Low Reconfiguration Delay*, International Conference on Field Programmable Logic and Applications, pp. 400-406, 2012.

[51] V. Tiwari and K. Khare, *Design Of Speed and Area Efficient NoC Architecture By Integrating Switches With Simplified Decoder And Reduced Buffers*, International Journal of Science, Technology and Research, Vol. 8, No. 10,pp. 2990-2994, 2019.

[52] Sweta Sahu and Harish M. Kittur, *Area and Power Efficient Network on Chip Router Architecture*, Proceedings of 2013 IEEE Conference on Information and Communication Technologies, pp. 855-859, 2013.

[53] M.S. Suraj, D. Muralidharan and K. Seshu Kumar, *A HDL based reduced area NOC router architecture*, IEEE International Symposium on Computer Architecture, pp. 241-250, 2008.

[54] Avinash Karanth Kodi, Ashwini Sarathy and Ahmed Louri, *iDEAL: Inter-Router Dual-function Energy and Area-efficient Links for Network-on-Chip (NoC) Architecture*, International Symposium on Computer Architecture, pp. 241-250, 2008.

[55] Anuja Naik and Tirumale K. Ramesh, *Effcient Network on Chip (NoC) using heterogeneous circuit switched routers*, International conference on VLSI Systems, Architectures, Technology and Applications, pp. 1-6, 2016.

[56] Mayank Kumar, Kishore Kumar, Sanjiv Kumar Gupta and Yogendera Kumar, *Area Efficient Architecture for Network on Chip (NoC) based Router*, , International Journal of Science, Research and Development, Vol. 4, No. 1, pp. 1464-1467,

2016.

[57] Rafael Gonçalves Mota, Jarbas Silveira, Jardel Silveira, Lucas Brahm, Avelino Zorzo, Filipo Mór and César Marcon, *Efficient Routing Table Minimization for Fault Tolerant Irregular Network-on-Chip*, IEEE International Conference on Electronics, Circuits and Systems, pp. 632-635, 2016.

[58] Pedro Valencia, Eric Muller and Nan Wang, *ZigZag: An Efficient Deterministic Network-on-chip Routing Algorithm Design*, 8th IEEE Annual Information Technology, Electronics and Mobile Communication Conference, pp. 1-5, 2017.

[59] Sudhir N. Shelke and Pramod B. Patil, *Low-Latency, Low- Area Overhead and High Throughput NoC Architecture for FPGA Based Computing System*, , International Conference on Electronic Systems, Signal Processing and Computing Technologies, pp. 53-55, 2014.

[60] Mohamed A. Abd El ghany, Magdy A. El-Moursy and Mohammed Ismail, *High Throughput Architecture for High Performance NoC*, International Symposium on Circuits and Systems, pp. 2241-2242, 2009.

[61] Yoon Seok Yang, Hrishikesh Deshpande, Gwan Choi and Paul Gratz, *Exploiting Path Diversity for Low-Latency and High-Bandwidth with the Dual-path NoC Router*, International Symposium on Circuits and Systems, pp. 2433-2436, 2012.

[62] Weiwei Fu, Jingcheng Shao, Bin Xie, Tianzhou Chen and Li Liu, *Design of a High-Throughput NoC Router with Neighbor Flow Regulation*, IEEE 14th International Conference on High Performance Computing and Communications, pp. 493-500, 2012.

[63] Xiaofang Wang and Leeladhar Bandi, *X-Network: An Area-Efficient and High-Performance On-Chip Wormhole-Switching Network*, 12th IEEE International Conference on High Performance Computing and Communications, pp. 362-368, 2018.

[64] Pengzhan Yan, Shixiong Jiang and Ramalingam Sridhar, *A High Throughput Router with a Novel Switch Allocator for Network on Chip*, 28th IEEE international System on chip Conference, pp. 160-163, 2015.

[65] Sungju Han ,Jinho Lee and Kiyoung Choi, *Tree-Mesh Heterogeneous Topology for Low-Latency NoC*, Proceedings of 2014 International Workshop on Network-on-Chip Architectures, pp. 19-24, 2014.

[66] C.J. Glass and L.M. Ni, *The Turn Model for Adaptive Routing*, Proceedings of 19th Annual International Symposium on Computer Architecture, pp. 277-287, 1992.

[67] Ge-Ming Chiu, *The Odd-Even Turn Model for Adaptive Routing*, IEEE Transactions on Parallel and Distributed. Systems, Vol. 11, No. 7, pp. 729-738, 2000.

[68] Yuri Nishikawa, Michihiro Koibuchi, Hiroki Matsutani and Hideharu Amano, *A Deadlock-free Non-minimal Fully Adaptive Routing Using Virtual Cut-through Switching*, Fifth IEEE International Conference on Networking, Architecture, and Storage, pp. 431-438, 2010.

[69] Monika Gupta and S.R.Biradar , *XFA Routing Algorithm for Network on Chip*, International Journal of Advanced Engineering Research Science, Vol. 1, No. 1, pp. 1-5, 2014.

[70] Yongqing Wang, Liquan Xiao, Sheng Ma, Zhengbin Pang and Kefei Wang , *Selective Extension of Routing Algorithms Based on Turn Model*, 22nd Euromicro International Conference on Parallel, Distributed, and Network-Based Processing, pp. 174-177, 2014.

[71] Vahid Janfaza and Elaheh Baharlouei , *A New Fault-Tolerant Deadlock-Free Fully Adaptive Routing in NOC*, IEEE International Symposium on Defect and Fault Tolerance in VLSI and Nanotechnology Systems, pp. 16-19, 2016.

[72] Navonil Chatterjee, Priyajit Mukherjee and Santanu Chattopadhyay , *A Strategy for Fault Tolerant Reconfigurable Network-on-Chip Design*, 20th International Symposium on VLSI Design and Test, pp. 1-2, 2016.

[73] Armin Alaghi, Mahshid Sedghi, Naghmeh Karimi, Mahmood Fathy and Zainalabedin Navabi , *Reliable NoC Architecture Utilizing a Robust Rerouting Algorithm*, Proceedings of IEEE East-West Design and Test Symposium, pp. 1-4, 2009.

[74] Amir Ehsani Zonouz, Mehrdad Seyrafi, Arghavan Asad, Mohsen Soryani, Mahmoud Fathy and Reza Berangi , *A Fault Tolerant NoC Architecture for Reliability Improvement and Latency Reduction*, 12th Euromicro Conference on Digital System Design, Architectures, Methods and Tools, pp. 473-480, 2009.

[75] Poona Bahrebar and Dirk Stroobandt, *Traffic-aware Reconfigurable Architecture for Fault-tolerant 2D Mesh NoCs*, Special Interest group on Embeded Systems, Vol. 15, No. 3, pp. 25-30, 2018.