RcEF3Ns: A Scalable and Reconfigurable Multi-FPGA Emulation Platform for 3D NoC

Jintao Zheng, Ning Wu, Lei Zhou, and Xiaoqiang Zhang

Abstract—**Taking advantage of Three Dimension (3D) Integrated Circuit (IC) technology, 3D Network-on-Chip (NoC) is becoming a popular development of high-performance System-on-Chip (SoC). To model 3D NoC and evaluate the performance fast and accurately, the emulation method is therefore critical issue. Compared to software emulation, Field Programmable Gate Array (FPGA) based emulation can offer a high speed validation process with a higher accuracy. But there still exist some difficulties such as partition, scalability and so on. This paper proposed a novel multi-FPGA emulation platform, RcEF3Ns (Reconfigurable Emulation on multi-FPGA for 3D NoCs), based on Xilinx Virtex-6 FPGAs. Our layered approach employs a single FPGA to manage vertical transaction independently, covering bus and network communication mechanism. All the parameters can be reconfigured on-chip to model 3D NoC architecture without re-synthesizing. The experiments show RcEF3Ns' speedups of 1500-18200 without sacrificing accuracy when compared to software HDL simulators.**

Index Terms—**3D NoC Emulation, Multi-FPGA, Scalable, Reconfigurable**

I. Introduction

THE combination of NoC and Through Silicon Via(TSV), 3D Network-on-Chip (NoC), which achieves lower transmission latency and power consumption, higher transistor density and platform bandwidth, has becoming a popular and important development of high-performance System-on-Chip(SoC)[1-2]. Designing 3D NoC implies making several design choices, such as topology structure, switching mechanism, the flow control policies, routing algorithm and traffic pattern. Fast and accurate performances evaluation for 3D NoC is therefore becoming a critical issue.

As SoC complexity increases with 3D Integrated Circuit(IC) technology scaling, existing general emulation platforms, software-based or hardware-based, are far from being suited for large systems. Several SystemC based environments have been proposed to perform 3D NoC modeling and simulation at the system level [3-4]. However, these software-based solutions compromise between cycle accuracy and simulation speed. On the other side, researchers also proposed several mature FPGA-based platforms for NoC infrastructure exploration such as DART [5], NoCem [6] and AcENoCs [7], which are for 2D NoC only.

In fact, simulating 2D NoC with few cores is feasible in a reasonable time, but for a large-scale 3D NoC, emulation speed is prohibitive. Emulating 3D NoC on FPGA therefore becomes a very popular way for interconnect validation, as it can offer a high speed validation process without compromising accuracy [8]. Case study for the deployment of 2D NoC structure to 3D for the generation of large emulation platform on FPAG has been presented in [9]. However, due to the resource limitation of reconfigurable devices, emulation 3D NoC on multi-FPGA also suffers from partition problem and scalability issue and so on.

Although the emulation speed of FPGA can be attractive, a long time to synthesize and generate the programming file is needed with specialized tools. Besides, in such schemes, off-chip delays cause performance bottleneck and size of 3D NoC is limited by available FPGA I/O pins and additional overhead for inter-FPGA communication adapters [10]. Typical multi-FPGA prototyping platforms employ a technique known as wire-multiplexing for inter-FPGA interconnections, in order to overcome the resource and pin limitation [11].

In this paper, we proposed a scalable and reconfigurable multi-FPGA platform, RcEF3Ns (Reconfigurable Emulation on multi-FPGA for 3D NoCs), which is a target for large scale 3D NoC. Our approach has four main advantages:

- Great degree of scalability.
- Cycle-accuracy and high speed.
- Cover most key design aspects of 3D NoC.
- Reconfiguration on-chip without re-synthesizing.

Configuring 3D NoC parameters on-chip and employing a single FPGA to manage vertical transaction enhance RcEF3Ns' flexibility, making it an effective platform for validating 3D NoC designs early in design process. Experiments have been conducted to analyze the behavior of various scale 3D NoC under uniform traffic conditions. RcEF3Ns can be widely applied to providing quick performance measurement and design tradeoffs estimation of 3D NoC.

The rest of the paper is organized as follows. Section II describes the scalable and configurable emulation framework

Manuscript received July 2, 2014; revised August 5, 2014. This work was supported by the National Natural Science Foundation of China under Grant (No.61376025, No.61106018), the Industry-academic Joint Technological Innovations Fund Project of Jiangsu under Grant (No.BY2013003-11), the Funding of Jiangsu Innovation Program for Graduate Education under Grant (No.KYLX_0273), and the Fundamental Research Funds for the Central Universities.

Jintao Zheng, the College of Electrical and Information Engineering, Nanjing University of Aeronautics and Astronautics (NUAA), Nanjing 210016, China(e-mail: zheng.jintao@qq.com).

Ning Wu, the College of Electrical and Information Engineering, NUAA, Nanjing 210016, China(e-mail: wunee@nuaa.edu.cn).

Lei Zhou, the College of Information Engineering, Yangzhou University, Yangzhou 225009, China(e-mail:tomcat800607@126.com).

Xiaoqiang Zhang, the College of Electrical and Information Engineering, NUAA, Nanjing 210016, China(e-mail: zxq198111@qq.com).

of RcEF3Ns. Section III presents the emulation flow of 3D NoC on RcEF3Ns. Experiments and discussion are detailed in Section V. The analysis of device resources and the evaluation speed are also detailed. Section VI contains the conclusion and perspectives.

II. Emulation Framework of RcEF3Ns

RcEF3Ns is a cycle-accuracy emulation platform which is instantiated on Xilinx Virtex-6(LX240T) FPGAs. The layered approach divides 3D NoC into two parts: the horizontal network layers (2D NoC) and the vertical crossbar. Every horizontal network layer is implemented in a FPGA while the vertical crossbar in another single one. Our layered approach employs an independent FPGA to management vertical transaction, which enhances the platform's flexibility. The multi-FPGA emulation framework of RcEF3Ns is shows in Fig.1.

A. Horizontal Network Layers

The number of horizontal network layers, as well as the number of FPGAs, is expected to be not restricted. FPGA for

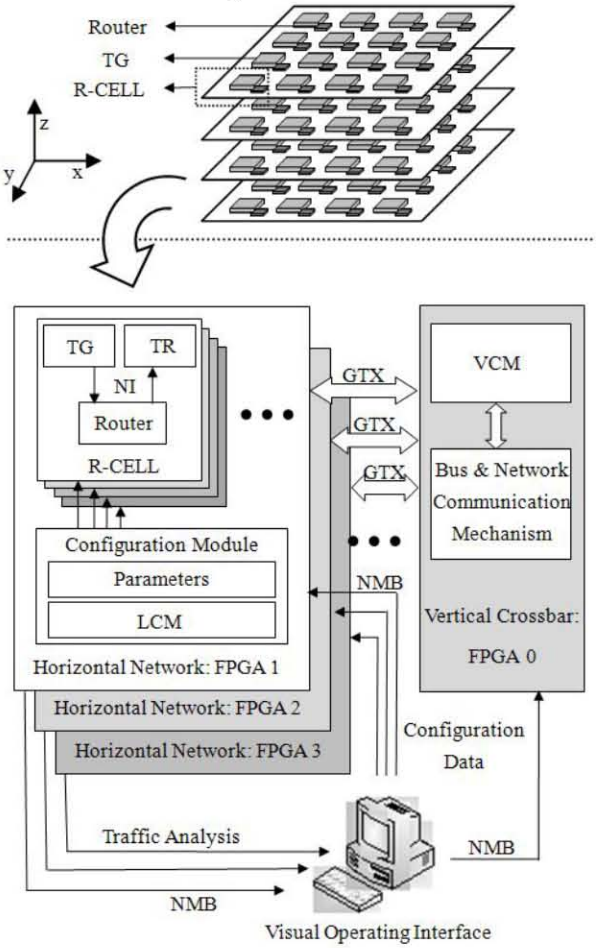


Fig. 1. The multi-FPGA emulation framework of RcEF3Ns

horizontal network is made up of a number of Routing-CELLs (R-CELLs) and configuration module. R-CELL is the elementary cell of horizontal network, which consists of Traffic Generator (TG), Traffic Receive (TR), Network Interface (NI) and a scalable and reconfigurable Router. Then we assume that the multi-FPGA architecture containing M FPGAs. Each R-CELL on FPGA has its own address (X, Y, Z). The (X, Y) coordinate is respectively the position in a 2D NoC, and (Z) is the number of FPGA.

TG and Router are reconfigurable. TG can generate various traffic model inject to the network through NI, such as uniform traffic pattern, hotspot traffic pattern and fully customized traffic pattern. Router is scalable for which we can define the number of ports (maximum 7 including North, South, East, West, Up, Down and Local ports), the size of buffer, the number of virtual channel and so on.

The configuration module generates R-CELLs with router structure parameters and links the horizontal network using Level Crossbar Matrix (LCM) on-chip. The LCM connection is shown in Fig. 2.

The LCM fully presents the interconnection of 2D NoC. The ports connection should be pairwise constraint. East port can only links to other's West port, while North port links to South port. As is shown in Fig. 2, East port of R-CELL (2, 1,

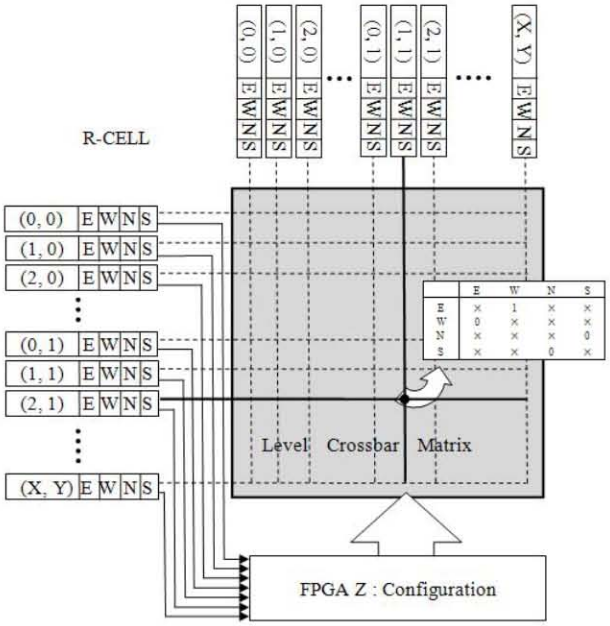


Fig. 2. The LCM connection in horizontal network layers

Z) is connected to West port of R-CELL (1, 1, Z).

B. Vertical Crossbar Matrix

For the case of vertical transaction design, bus and network communication mechanism are optional for developing a 3D NoC architecture.

In order to make the emulation platform scalability, FPGA 0 is employed to manage vertical transaction independently. All the Up and Down ports of R-CELLs are links to the big Vertical Crossbar Matrix (VCM) in FPGA 0. Besides, a number of N buses are connected to VCM for the selection of bus communication mechanism, and only Up ports are allowed to link.

Since the number of TSV is not discretionary, researchers can explore 3D NoC with different vertical channel densities by configuring VCM and disabling some linking relationship. Furthermore, additions or deletions of nodes and links can be very convenient. We can define the vertical linking relation of 3D NoC by configuring the VCM on-chip like Fig. 3.

ISBN: 978-988-19252-0-6
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

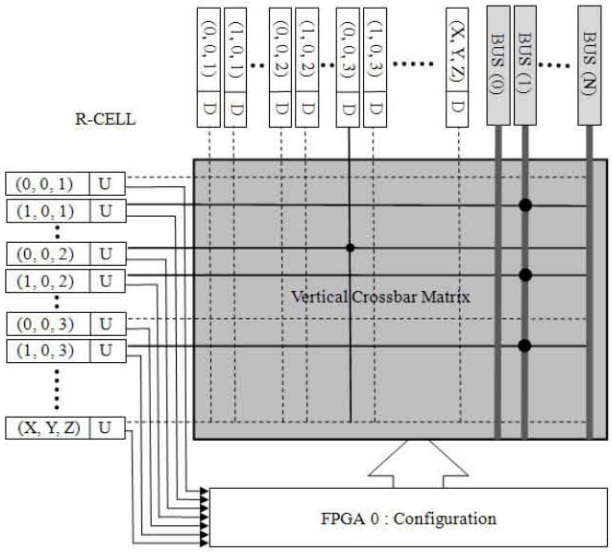


Fig. 3. The VCM connection in FPGA 0

As is shown in Fig. 3, Up port of R-CELL (0, 0, 2) is connected to Down port of R-CELL (0, 0, 3), while all the Up ports of R-CELL (1, 0, Z) are connected to BUS (1).

On the other hand, off-chip delays cause performance bottleneck, a high-speed type of inter-FPGA communication to cancel out the influence on the network latency must be carried out. Impact of multi-FPGA partitioning on the latency was discussed in [9]. Therefore the 3D NoC can give better timing performances with the appropriate partitioning. The use of a high-speed inter-FPGA communication enhances the latency on 3D NoC.

Xilinx Virtex-6 FPGA provides GTX transceiver solution for inter-FPGA communication. The GTX transceiver is highly configurable and tightly integrated with the programmable logic resources of the FPGA. The highest transfer rate of GTX can reach 6.5Gb/s per channel, while the frequency of network is 100Mhz, so the latency of our cycle-accuracy emulation can be guaranteed.

C. *Configuration Parameters*

Developing 3D NoC implies making several choices, so RcEF3Ns is designed to cover most key design aspects of 3D NoC. Parameters describing the presented 3D NoC will be written into a binary file (Configruation.bin), and transferred through NMB bus into FPGAs. NMB bus a kind of Xilinx FPGA technology that provides a simple and efficient data transfer mechanism between hosts and user designs. Fig. 4 shows the data flow of configuration parameters.

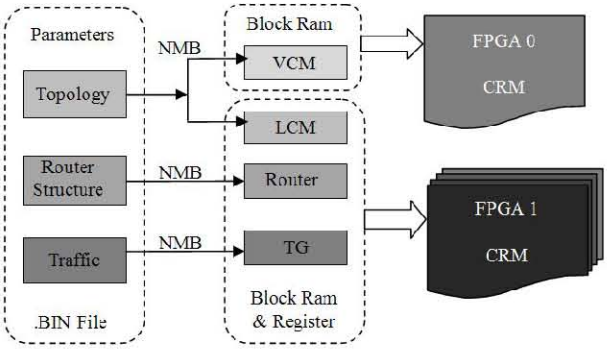


Fig. 4. Data flow of configuration parameters

Block ram and registers are common used to store configuration data and a Configuration Resolve Module (CRM) is needed to disassemble these binary data and rebuild the Resistor Transistor Logic (RTL) model of the represented 3D NoC.

We classify three kinds of parameters to describe a whole 3D NoC system: topology parameters, router structure parameters and traffic parameters.

Topology parameters are used to describe the form of interconnection of 3D NoC. In particular, the number of nodes and layers, the interconnection of each R-CELLs' port are all needed to establish 3D topology. The interconnection of 3D NoC is instantiated as M LCM and one VCM. Due to the limitation of reconfigurable devices, we deploy 16 R-CELLs and a LCM to determine the interconnection of topology in each horizontal FPGA. The FPGA resource utilization will be evaluated in Section V.

As introduced before, router in R-CELL is a scalable and reconfigurable component. Router structure parameters are detailed in Table I.

There are seven ports on the router of R-CELL, however, it is scalable to activate or inactivate some of them according to the design intention but except for the Local port. The number of virtual channel and the size of buffer are optional too, which should not be over 4 and 8. Besides, there are typical choices for reset type, routing type, arbiter type and switch type.

TABLE I
ROUTER STRUCTURE PARAMETERS

Parameter	Value	Comments
Num_port	0-7	Number of ports of router
Num_vcs	1-4	Number of virtual channel
Depth_buffer	1-8	Size of buffer
Router_addr	(X, Y, Z)	Coordinate of R-CELL
Port_id	0-7	Code of each port(0-Local, 1-East, 2-West, 3-North, 4-South, 5-Up, 6-Down)
Reset_type	0	Synchronous reset
	1	Asynchronous reset
Routing_type	0	XYZ routing algorithm
	1	Adaptive Dimension Order Routing algorithm
	2	Routing table algorithm
	3	Default routing algorithm
Arbiter_type	0	Round-robin arbiter
	1	First-Come-First-Sever arbiter
Switch_type	0	Wormhole switch
	1	Virtual cut-through switch

Traffic modeling defines the structure of data transmission from initiators to targets. The proposed TG is aimed at faithfully replicating traffic patterns. According to [12], traffic modeling is defined by 3 parameters: spatial distribution of targets, injection rate, and packet length. RcEF3Ns apply for the same approach and Traffic parameters are detailed in Table II.

Pro_dest[0:num_nodes] is an array about the probabilities

ISBN: 978-988-19252-0-6
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

TABLE II
TRAFFIC PARAMETERS

Parameter	Value	Comments
Pro_dest [0:num_nodes]	0-1	Packet spatial distribution
Inj_rate	0-1	Injection rate(flits/cycle /node)
Pkt_length	2-32	Packet size(flits)
Pkt_length_type	0	Fixed packet size
	1	Random packet size
Warm_time	10^3	Warming up time(cycles)
Emu_time	10^6	Emulation time(cycles)

that source nodes communicate with every other nodes. For uniform traffic pattern, the probabilities are all the same. In addition, the fully customized traffic pattern sufficiently reflects on the scalability of TG.

D. Traffic Analysis

Traffic flows generated by TG are injected into 3D NoC and then received by TR. TR collects every received packet and extracts latencies and hops. The packet contains a five-flit header. The first and the second flit represent the coordinate of the source and destination R-CELL respectively. The number of flits in the packet payload is stored in the third flit. The forth flit is the packet routing hops which is set 0 at first and plus one when pass a router. To evaluate the latency performance, the fifth flit counts the time T_S that packet start to be sent through NI into the network. The format of packets in RcEF3Ns is shown in Fig. 5.

After all the packets are received, TR can figure out the number of flits, the latency and the hops of packets

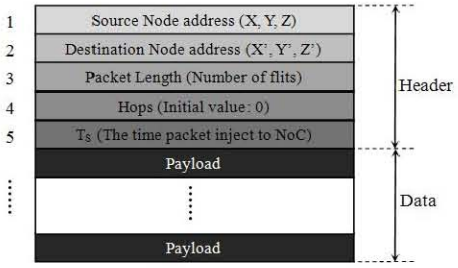


Fig. 5. Format of packets in RcEF3Ns

transferring in the whole 3D NoC by extracting the header part. TG and TR are under the same synchronous clock to guarantee the latency is correct.

When the emulation time ends up, all the TRs will send the statistical figures to Traffic Analysis (TA), which is a global statistics module on the computer. According to the configuration data and emulation result, TA computes injection rate, throughput rate, network latency and utilization of links for the proposed experimental 3D NoC.

III. EMULATION FLOW OF RCEF3NS

RcEF3Ns provides automatic emulation flow for 3D NoC design. Researchers can easily realize their exploration and verify on the proposed platform. Benefit from the scalable and reconfigurable design intent, slight changes can be quickly accepted, and network can be rebuilt on-chip without re-synthesizing, which can greatly reduce the developing period for 3D NoC.

A. Visual Operating Interface

Offering greater convenience for researchers to exploit the emulation platform, a visual operation interface is carried out on computer. RcEF3Ns abstracts and digitalizes all the parameters, and graphs the emulation results on the interface.

All the parameters configured will be sent to multi-FPGA platform through NMB bus. Each FPGA stores the configuration data in the block ram by classified address.

After the FPGA emulation process, all the statistics data will be sent back to the computer through NMB bus. Statistics reports and evaluation graph will be created for the performance analysis.

B. Emulation Flow

First step to start RcEF3Ns is configuring every FPGA using the common programming files of R-CELLS, LCM and VCM, which we have synthesized and generated before. The emulation flow of RcEF3Ns is depicted in Fig. 6.

For 3D NoC exploration, define application task graph and network topology using parameters and so on. While the 3D NoC partition has been done by RcEF3Ns, the researchers

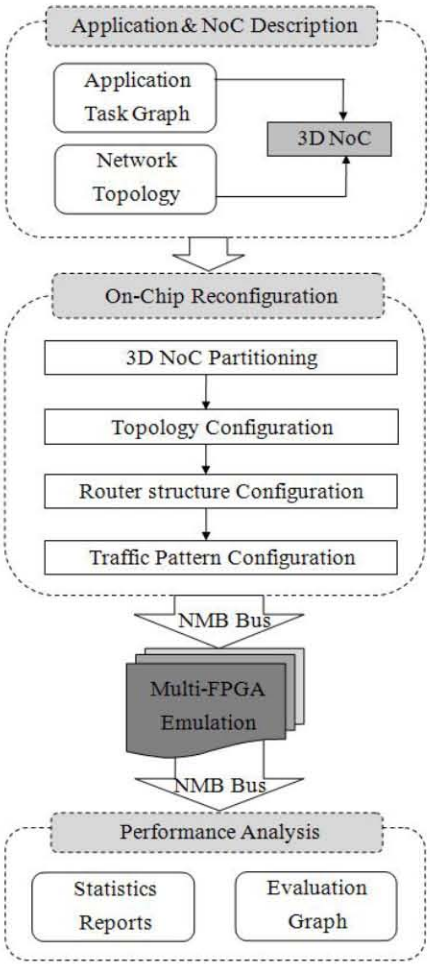


Fig. 6. The emulation flow of RcEF3Ns

only need to fill in the relevant parameters according to visual operating interface guide. All the decision will be transferred into FPGAs through NMB bus using the preset communication protocol.

After configuring all the parameters correctly, the

ISBN: 978-988-19252-0-6
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

proposed 3D NoC will be established on RcEF3Ns. Then TGs start to generated traffic flow into the network. Emulation process will be terminated according to the preset emulation time or if an error is encountered.

TRs abstract header part of received packets, and send the relative statistic results to TA through NMB bus. Then, the automated analysis module will display the emulation results in two ways: statistics reports and evaluation graph. The reports contain the number of packets sent and received, latency, hops and some other figures of each R-CELL. The evaluation graph reflects the variation tendency of throughput rate, network latency, links utilization along with the increase of inject rate.

The advantage of RcEF3Ns lies in the ease with which emulation parameters can be reconfigured for each emulation run on-chip. To modify design scheme, only by changing some parameters, the new 3D NoC can be quickly reconfigured on-chip without re-synthesizing.

IV. Experiments and Discussion

A. Device Utilization

The resource limitation of reconfigurable device affects the scale of network available. Table III shows the FPGA resource device utilization of each component in Xilinx Virtex-6 LX240T FPGA.

LUT and register are the limiting factors for implementing larger scale of 3D NoC. Each R-CELL, including TG, TR, NI and Router, needs 6081 LUTs and 2663 Registers, 4.03% and 0.88% utilization of device source respectively. So it is

TABLE III
FPGA Utilization

Components	LUT		Register	
LX240T	150720	100%	301440	100%
TG	936	0.62%	269	0.09%
TR	274	0.18%	252	0.08%
NI	170	0.11%	173	0.06%
Router	4629	3.07%	1947	0.65%
R-CELL	6081	4.03%	2663	0.88%
16 R-CELLs	97296	64.6%	42608	14.1%
LCM	14770	9.80%	9646	3.20%
VCM(for 4 layers)	39640	26.3%	33762	11.2%
VCM(for 5 layers)	61192	40.6%	81992	27.2%
VCM(for 6 layers)	89979	59.7%	119973	39.8%

available to accommodate 16 R-CELLs on each Virtex-6(LX240T) FPGA with the total LUT utilization of 64.6%.

LCM is a crossbar matrix for 16 R-CELLs' interconnection, which is under pairwise constraint. In fact, only 9.8% LUTs and 3.2% registers are needed to construct such a reconfigurable matrix.

VCM is a big crossbar matrix which is mixed with Up, Down ports and Bus. Thus, when each layer possesses 16 nodes as well as the number of Buses, VCM takes up 26.3%, 40.6% and 59.7% LUTs for 4 to 6 layers, which is exponential growth.

Thus, RcEF3Ns based on Xilinx Virtex-6 LX240T FPGAs, can supports various 3D NoC with the maximum scale of 6 layers and 16 routers for each horizontal network layer. Further, to expand the capacity of RcEF3Ns with the same level of scalability and reconfigurability, larger FPGA should be considerable.

B. Emulation Speed Comparison

The experiments compare the performance of 3D Mesh with different layers of the XYZ routing algorithm under the uniform traffic pattern. The horizontal network size is 3×3 Mesh, and the number of layers varies from 2 to 6. The number of virtual channel is 2 with the buffer size of 8. The packet type is random and the packet size is 2-16 flits. The warming time set to 10^3 cycles, while the emulation time 10^6 cycles. The whole network is working at the frequency of 100 MHz.

Then, based the same RTL model of the experimental design, we compared the emulation time of HDL simulators Modelsim and RcEF3Ns. The obtained results are shown in Table IV.

As is shown in Table IV, for the software-based platform, emulation time is sharply increasing along with the layers of 3D NoC. However, without re-synthesizing, FPGA-based RcEF3Ns is much faster, whose emulation time is only concerned to the emulation cycle, and remained unaffected to any other factors.

TABLE IV
Timing Comparison

Time(s)	Modelsim	CoEF3Ns	Speedups
3×3×2	15	0.01	1500
3×3×3	32	0.01	3200
3×3×4	59	0.01	5900
3×3×5	107	0.01	10700
3×3×6	182	0.01	18200

Consequently, RcEF3Ns can reduce the time to emulation 3D NoC when compared to Modelsim tool with the same RTL model, and the speedups can be 1500-18200 as the network getting larger in our experiment.

C. Performance Analysis

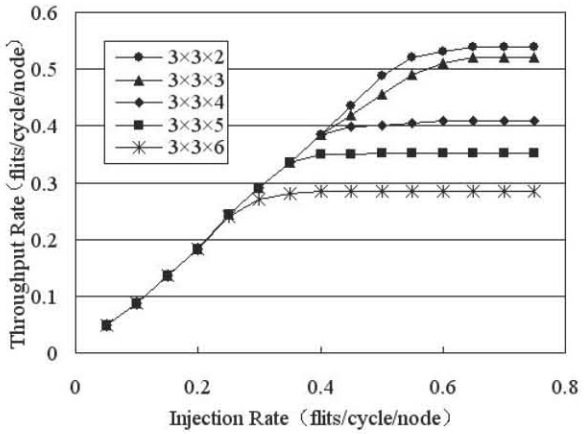
3D NoC are motivated to achieve lower transmission and higher platform bandwidth. For RcEF3Ns, layer addition or deletion can be done simply and fast for the scalability of the emulation platform.

We presented the experiment to examine the impact of the stacked layers. So we build the same horizontal network layers, and step by step to increase the stacked layers. It is quickly and simply for RcEF3Ns to rebuild the new 3D NoC. The performance comparison of throughput rate and network latency variation along with the increase of injection rate can be seen in Fig. 7.

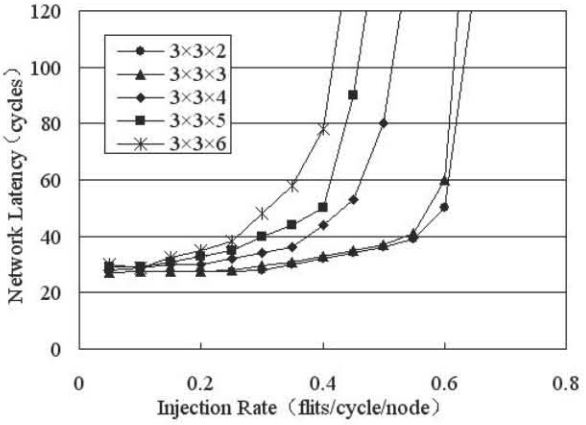
As Fig. 7 indicates, the throughput rate for each node of 3D Mesh is decreasing as the layers increases. Because the number of destination nodes to be communicated for each node increases under the uniform traffic pattern.

Network throughput reaches a certain extent after tends to saturation. When the number of layers varies from 2 to 6 with

ISBN: 978-988-19252-0-6
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)



(a) Throughput rate



(b) Network latency

Fig.7. Performance comparison of throughput and latency

the horizontal network 3×3 Mesh, the saturation throughput rates are 0.5411, 0.5202, 0.4082, 0.3523 and 0.2946 flits/cycle/node respectively, which reflects a downtrend as stacked layers increased. The network latency stays a relative stabilization, but sharply grows up if the network tends to saturation, especially for the larger scale 3D NoC.

V. Conclusions

In this paper we have presented a study of a multi-FPGA platform for emulating large 3D NoC, and RcEF3Ns is proposed to solve the scalability issue and allow high emulation speed. A single FPGA is employed to manage vertical transaction independently which enhance RcEF3Ns' flexibility. Researchers can model their 3D NoC simply and fast, by reconfiguring parameters on-chip without re-synthesizing.

RcEF3Ns shows speedups when compared to software HDL simulators especially for the larger system without sacrificing accuracy. Further, we will consider in future implementations different clocking sources and serial interfaces for RcEF3Ns as we expect to greatly improve the emulation efficiency.

References

[1] Rahmani, A-M., et al. "Congestion aware, fault tolerant, and thermally efficient inter-layer communication scheme for hybrid NoC-bus 3D architectures."*Networks on Chip (NoCS), 2011 Fifth IEEE/ACM International Symposium on*. IEEE, 2011.

[2] Chao, Chih-Hao, et al. "Traffic-and thermal-aware run-time thermal management scheme for 3D NoC systems." *Networks-on-Chip (NOCS), 2010 Fourth ACM/IEEE International Symposium on*. IEEE, 2010.

[3] Pasricha, Sudeep. "A framework for TSV serialization-aware synthesis of application specific 3D networks-on-chip." *VLSI Design (VLSID), 2012 25th International Conference on*. IEEE, 2012.

[4] Liu, Cheng, et al. "Vertical interconnects squeezing in symmetric 3D mesh network-on-chip." *Proceedings of the 16th Asia and South Pacific Design Automation Conference*. IEEE Press, 2011.

[5] Wang, Danyao, Natalie Enright Jerger, and J. Gregory Steffan. "DART: A programmable architecture for NoC simulation on FPGAs." *Networks on Chip (NoCS), 2011 Fifth IEEE/ACM International Symposium on*. IEEE, 2011.

[6] Schelle, Graham, and Dirk Grunwald. "Onchip interconnect exploration for multicore processors utilizing FPGAs." *2nd Workshop on Architecture Research using FPGA Platforms*. 2006.

[7] Lotlikar, Swapnil, Vinayak Pai, and Paul V. Gratz. "AcENoCs: A configurable HW/SW platform for FPGA accelerated NoC emulation." *VLSI Design (VLSI Design), 2011 24th International Conference on*. IEEE, 2011.

[8] Wolkotte, Pascal T., Philip KF Holzenspies, and Gerard JM Smit. "Fast, accurate and detailed NoC simulations." *Networks-on-Chip, 2007. NOCS 2007. First International Symposium on*. IEEE, 2007.

[9] Fresse, Virginie, et al. "Case study: Deployment of the 2D NoC on 3D for the generation of large emulation platforms." *Rapid System Prototyping (RSP), 2012 23rd IEEE International Symposium on*. IEEE, 2012.

[10] Kouadri-Mostefaoui, A., F. Rousseau, and Frederic Petrot. "Large scale on-chip networks: An accurate multi-FPGA emulation platform." *Proc. of 11th EUROMICRO Conference on Digital System Design Architectures Methods and Tools (DSD'08)*. 2008.

[11] Babb, Jonathan, et al. "Logic emulation with virtual wires." *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on* 16.6 (1997): 609-626.

[12] Mello, Aline, et al. "Virtual channels in networks on chip: implementation and evaluation on hermes NoC." *Proceedings of the 18th annual symposium on Integrated circuits and system design*. ACM, 2005.

ISBN: 978-988-19252-0-6
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)