

# Low-Swing Signaling for Energy Efficient On-Chip Networks

by

Sunghyun Park

B.S., Electrical Engineering and Computer Science  
Korea Advanced Institute of Science and Technology (2008)

Submitted to the Department of Electrical Engineering and Computer  
Science

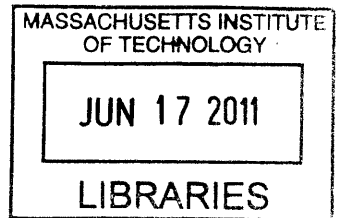
in partial fulfillment of the requirements for the degree of  
Master of Science in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2011

© Massachusetts Institute of Technology 2011. All rights reserved.



**ARCHIVES**

Author .....  
Department of Electrical Engineering and Computer Science  
May 16, 2011

Certified by .....  
Li-Shiuan Peh  
Associate Professor  
Thesis Supervisor

Certified by .....  
Anantha P. Chandrakasan  
Professor  
Thesis Supervisor

Accepted by .....  
Leslie A. Kolodziejski  
Chairman, Department Committee on Graduate Students



# Low-Swing Signaling for Energy Efficient On-Chip Networks

by

Sunghyun Park

Submitted to the Department of Electrical Engineering and Computer Science  
on May 16, 2011, in partial fulfillment of the  
requirements for the degree of  
Master of Science in Electrical Engineering and Computer Science

## Abstract

On-chip networks have emerged as a scalable and high-bandwidth communication fabric in many-core processor chips. However, the energy consumption of these networks is becoming comparable to that of computation cores, making further scaling of core counts difficult.

This thesis makes several contributions to low-swing signaling circuit design for the energy efficient on-chip networks in two separate projects: on-chip networks optimized for one-to-many multicasts and broadcasts, and link designs that allow on-chip networks to approach an ideal interconnection fabric. A low-swing crossbar switch, which is based on tri-state Reduced-Swing Drivers (RSDs), is presented for the first project. Measurement results of its test chip fabricated in 45nm SOI CMOS show that the tri-state RSD-based crossbar enables 55% power savings as compared to an equivalent full-swing crossbar and link. Also, the measurement results show that the proposed crossbar allows the broadcast-optimized on-chip networks using a single pipeline stage for physical data transmission to operate at 21% higher data rate, when compared with the full-swing networks. For the second project, two clockless low-swing repeaters, a Self-Resetting Logic Repeater (SRLR) and a Voltage-Locked Repeater (VLR), have been proposed and analyzed in simulation only. They both require no reference clock, differential signaling, and bias current. Such digital-intensive properties enable them to approach energy and delay performance of a point-to-point interconnect of variable lengths. Simulated in 45nm SOI CMOS, the 10mm SRLR featured with high energy efficiency consumes 338fJ/b at 5.4Gb/s/ch while the 10mm VLR raises its data rate up to 16.0Gb/s/ch with 427fJ/b.

Thesis Supervisor: Li-Shiuan Peh  
Title: Associate Professor

Thesis Supervisor: Anantha P. Chandrakasan  
Title: Professor





## Acknowledgments

I would first like to thank my advisors, Professor Li-Shiuan Peh and Professor Anantha Chandrakasan. It has been a great honor and pleasure to work with such big names. Their guidance and patience have allowed me to complete two projects successfully. I also thank their excellent lectures, 6.374 and 6.883, which formed the foundation of my research. Not only in technical point of view but also in a humanity aspect, they are truly good mentors. I would never forget their congratulations and nice words when my little girl Rachael was born.

I want to extend my gratitude to my group colleagues who are always willing to answer my questions: Tushar, Owen, Masood, Fred, Mahmut, Niket, Kostas, Bhavya, and Woocheol. I am grateful to Sunghyuk and Arun for their help in the chip measurement. I'd also like to thank Byungsub (Intel), Sungwon (MIT), and Jaewon (KAIST) for their technical advice in the chip tapeouts. I thank Professor SeongHwan Cho for his kind teaching in my undergraduate days.

I should not forget to thank my parents, my sister, my parents-in-law, and my family. I thank my two lovely penguins, Seonghee and Seohee, for always being with me. I really appreciate Seonghee's selfless support and patience. Seohee's presence itself makes me do my best in everything. Also, I am sincerely grateful to my older sister, Haejin, for her heartening encouragement. I am pretty much sure that she will become one of the greatest doctors in Korean history. With the best truth of my heart, I would like to thank my parents. No words can express my gratitude for their unconditional trust and love. I would not be who I am today without their endless love and sacrifice.

I would like to express my appreciation to my fallen comrade, Jangho Yoon, killed in Afghanistan back in 2007. His body died, but his good-natured soul is still alive in my heart. I will walk with God so as not to make his sacrifice in vain.

Lastly again, I truly thank Mr. Doosoo Park and Ms. Soonsil Shin.



# Contents

<b>1</b>	<b>Introduction</b>	<b>13</b>
1.1	Motivation . . . . .	13
1.2	Packet-Switched On-Chip Networks . . . . .	14
1.3	Previous Work on Low-Swing Signaling . . . . .	16
1.4	Thesis Contributions . . . . .	17
<b>2</b>	<b>Low-Swing Broadcast-Optimized On-Chip Networks</b>	<b>19</b>
2.1	Broadcast-based Cache Coherence Protocols . . . . .	19
2.2	Broadcast-optimized Router Microarchitecture . . . . .	20
2.2.1	Multicast Buffer Bypassing Flow Control . . . . .	20
2.2.2	Broadcast-optimized Crossbar Switch . . . . .	21
2.3	Low-Swing Signaling in Data Path . . . . .	23
2.3.1	Tri-state RSD-based Crossbar Switch . . . . .	24
2.3.2	Design Considerations . . . . .	27
2.4	Test Chip fabrication . . . . .	29
2.5	Measurement Results . . . . .	31
<b>3</b>	<b>Clockless Low-Swing Repeaters</b>	<b>37</b>
3.1	Motivation for Clockless Low-Swing Signaling . . . . .	37
3.2	Self-Resetting Logic Repeater (SRLR) . . . . .	40
3.2.1	Introduction to Self-Resetting Logic . . . . .	40
3.2.2	Process Variation Robust SRLR Design . . . . .	41
3.3	Voltage-Locked Repeater (VLR) . . . . .	48

3.4	Test Chip Fabrication . . . . .	54
3.5	Extracted Simulation Results . . . . .	56
<b>4</b>	<b>Conclusion</b>	<b>61</b>
4.1	Thesis Summary . . . . .	61
4.2	Future Work . . . . .	62

# List of Figures

1-1	Intel 80-core TeraFLOP, 65mm CMOS, 5GHz . . . . .	14
1-2	Packet-switched on-chip networks. . . . .	15
2-1	Proposed router microarchitecture. . . . .	21
2-2	Proposed router pipeline. . . . .	23
2-3	64 bits 5x5 tri-state RSD-based crossbar switch. . . . .	24
2-4	Tri-state Reduced-Swing Driver (RSD). . . . .	25
2-5	Tri-state RSD-based crossbar switch waveforms. . . . .	26
2-6	Eye diagram with (a) 1mm-wire nominal resistance (b) 1mm-wire 3-sigma high resistance. . . . .	28
2-7	Balance between charging and discharging time caused by a delay cell	28
2-8	Test chip die photo overlaid with a layout. . . . .	30
2-9	Differential signaling wire model with full shielding. . . . .	30
2-10	Power consumption measurement circuit diagram. . . . .	32
2-11	1mm link energy comparison between full-swing and low-swing signaling.	34
2-12	2mm link energy comparison between full-swing and low-swing signaling.	34
2-13	Tri-state RSD energy measurement with various voltage swings. . . .	35
2-14	3.8GHz 1bit 5x5 crossbar switch energy measurement. . . . .	36
3-1	(a) 9x9 regular mesh on-chip network (b) 11-node irregular mesh on-chip network. . . . .	38
3-2	Inter-Symbol Interference (ISI) caused by channel loss. . . . .	39
3-3	Generic view of Self-Resetting Logic (SRL). . . . .	40
3-4	Proposed SRLR design (not yet completed). . . . .	41

3-5	Example of SRLR waveforms. . . . .	42
3-6	Oscillation-free SRLR design. . . . .	43
3-7	Example of shrinking pulses caused by process variation. . . . .	43
3-8	Alternatively repeated SRLR with two different pulse widths. . . . .	44
3-9	Error-free SRLR design in 3-sigma on-die variation. . . . .	45
3-10	Self-calibrating SRLR with a threshold voltage monitor circuit. . . . .	47
3-11	Proposed VLR without enable switches. . . . .	49
3-12	VLR operation: (a) logic High low-swing generation (b) logic Low low-swing generation. . . . .	50
3-13	Example of VLR waveforms. . . . .	51
3-14	Zero static current, process variation robust VLR design. . . . .	52
3-15	Modified VLR operation: (a) logic High (b) logic Low. . . . .	53
3-16	Test chip layout: various asynchronous on-chip interconnects. . . . .	55
3-17	Single-ended signaling wire model with full shielding. . . . .	56
3-18	SRLR extracted simulation waveforms: (a) 5Gb/s/ch (b) 1Gb/s/ch. . . . .	57
3-19	10mm SRLR Energy simulation results at low data rate. . . . .	58
3-20	10mm VLR Energy simulation results at high data rate. . . . .	58

## List of Tables





# Chapter 1

## Introduction

### 1.1 Motivation

The ever-increasing energy consumption and shrinking returns of complex uniprocessor architectures have led to the advent of many-core processor chips. With a saving in design and verification time enabled by the modular design of many-core chips, this many-core architecture trend is becoming more popular [1]. Moreover, driven by a growing number of transistors available at each new technology, current hardware roadmaps call for doubling the number of on-chip computation cores approximately every two years [2]. If this trend materializes, in at most a decade and a half, we may reach one thousand on-chip cores.

As on-chip core counts increase, designing scalable on-chip interconnection fabrics has been an essential research field. Recently, packet-switched on-chip networks are in the spotlight as the solution of the scalable and high-bandwidth communication fabric, replacing nonscalable buses and crossbar switches in many-core processors. MIT's RAW [3], UT Austin's TRIP [4], Intel's TeraFLOPS [5], Tiler's TILE64 [6], and Intel's 48-core IA-32 [7] have adopted these packet-switched on-chip networks. The Intel's TeraFLOPS is shown in Figure 1-1 as an example of the many-core processor that employs such on-chip networks.

Unfortunately, the power consumption of these on-chip networks becomes a major concern in many-core processor design. For instance, 36% and 39% of total chip

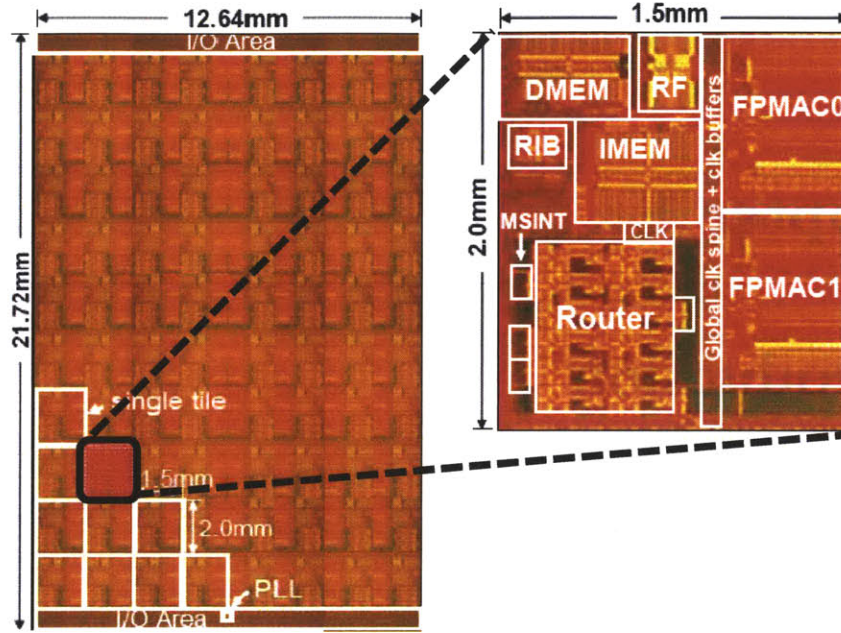


Figure 1-1: Intel 80-core TeraFLOP, 65nm CMOS, 5GHz

power are consumed by the network in MIT's Raw [3] and Intel's TeraFLOPS [5], respectively. Therefore, for further scaling of on-chip core counts along with the current processor design trend, more energy efficient on-chip networks will be required.

## 1.2 Packet-Switched On-Chip Networks

On-chip networks, as a subset of a broader class of interconnection networks, facilitate data communication between processor components such as a computation core, cache, and memory controller. The best way to design such networks is the use of dedicated wires. However, since it is impossible to have the vast amount of wiring required to directly connect all components as the number of processor components increase, many-core processors have to share and multiplex communications on wires.

Bus-based networks, which can be viewed as the simplest variation of on-chip networks, scale only to a small number of processor components [1]. The limited scalability is because bus traffic quickly reaches saturation as components counts

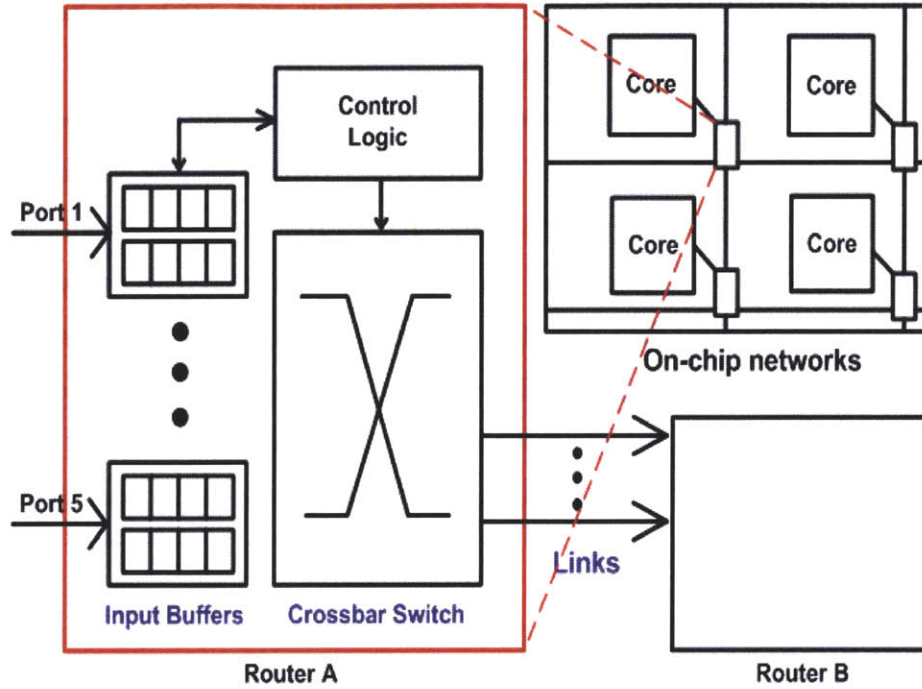


Figure 1-2: Packet-switched on-chip networks.

increase. Plus, the arbitration delay of the centralized arbiter increases as more components are added to the bus. Even though crossbars alleviate the bandwidth limit of the buses, they also suffer from poor scalability in terms of area and energy consumption.

On the other hand, packet-switched on-chip networks provide scalable bandwidth at relatively low latency overhead that correlates sub-linearly with the number of processor components in the networks [1]. As a result, packet-switched on-chip networks are fast replacing buses and crossbars in many-core processor chips. Such on-chip networks shown in Figure 1-2 have routers at every node, connected to neighbor nodes with local on-chip wiring. The router is composed of four main blocks: buffers, allocation logics, crossbar switches, and links. The buffers are responsible for storing flits throughout their duration in the router. This is in contrast to a processor pipeline that latches instructions in buffers between each pipeline stage. The allocation logic determines which flits are selected to proceed to the next stage. The crossbar switches

physically move flits from the input port to the output port, while the links play a role in the flits' physical transfer to the next router.

The primary contributors to power consumption of on-chip networks are links (39% in RAW, 31% in TRIPS, 17% in TeraFLOPS), crossbar switches (30% in RAW, 33% in TRIPS, 15% in TeraFLOPS), and buffers (31% in RAW, 35% in TRIPS, 22% in TeraFLOPS) [3] [4] [5]. Since links and crossbar switches are responsible for actual data transmission, they form the unavoidable component of network power. Furthermore, this datapath power will increase in percentage relative to control and storage circuitry power as process technology scales down [8]. Thus, it is necessary to reduce power consumption of the links and crossbar switches for attaining energy efficient on-chip networks.

### 1.3 Previous Work on Low-Swing Signaling

Low-swing signaling is now one well-known low-power design technique in both on-chip and off-chip interface circuits. In on-chip domain, the low-swing signaling will be considered as a natural choice for future on-chip communication fabrics since the wire performance benefit from CMOS process scaling does not keep up with the gate performance benefit.

The low-swing technique is based on the dependence of dynamic energy on swing voltage. Reducing the voltage swing across data path leads to reduced charging and discharging of the wire capacitance in comparison with the full-swing signaling, thereby making on-chip communication fabrics more energy-efficient and faster. Particularly, under the circumstance where it is hardly possible to reduce length of wires and their fanout by using advanced processes or architectures, the low-swing signaling is the best design technique toward getting energy efficient on-chip networks. More details of on-chip low-swing signaling are discussed in [8] [9].

Apart from conventional low-swing circuits which use a second lower supply voltage and inherent threshold voltage drop introduced in [8] [9], there have been a number of more sophisticated circuits proposed, based on linear-mode drive transistors [10]

[11], charge sharing [12] [13] [14], cut-off drivers [8] [15] [16], and channel attenuation [17] [18] [19].

Low-swing drivers exploiting the linear-mode drive transistors [10] [11] are composed of PMOS pullups and pulldowns only (or NMOS pullups and pulldowns only) to obtain lower linear drive resistance even at small  $V_{ds}$ . Such designs provide much shorter propagation delay as compared to the low-swing signaling which is generated by simply lowering supply voltage, but they require an additional dedicated power supply. Even though the charge sharing-based low-swing drivers [12] [13] [14] limit voltage swing without the second power supply, they need some particular data patterns for reliable operation. The voltage swing of low-swing signaling circuits based on the cut-off drivers [8] [15] [16] is directly affected by threshold voltage variation of drivers or sensing circuits, thereby making the receiver design harder. Along with the inherent channel loss caused by RC-dominant on-chip wires, pre-emphasis techniques such as equalization [17] [18] [19] can also be employed to generate low-swing signaling. These circuits enable both higher bandwidth and lower energy consumption especially in long wires, but it is practically impossible to utilize them as links in the packet-switched on-chip networks due to their huge area overhead. For example, the 10mm 1-bit driver of [19] occupies 1760um. Moreover, typical 2D-mesh on-chip networks are likely to need parallel links covering just 1-2mm instead of direct 10mm wiring links [5], thereby limiting their feasible applications in such on-chip networks.

## 1.4 Thesis Contributions

As discussed in the previous section, existing low-swing techniques are some distance from the packet-switched on-chip networks which are increasingly prevalent in many-core processor chips. In addition, in order to efficiently support state-of-the-art cache coherence protocols and enable on-chip networks to approach ideal communication fabrics, the network requirements are now changing. This thesis focuses on filling such a gap between the architectural requirements of on-chip networks and low-swing signaling circuit design. More specifically, this work explores low-swing signaling

circuits for two separate projects: on-chip networks optimized for one-to-many multicasts and broadcasts, and link designs that enable energy and delay performance of a point-to-point interconnect of variable lengths.

Chapter 2 presents the broadcast-optimized, low-swing signaling-based on-chip networks in response to the state-of-the-art cache coherence protocols which highly depend on broadcasts and multicasts. A multicast buffer bypassing flow control reduces buffer power along the control path while a broadcast-optimized low-swing crossbar switch reduces interconnect energy in the data path. The multicast buffer bypassing flow control design is done by Tushar Krishna. The energy efficiency of the proposed low-swing data path is proven with measurement results of the test chip fabricated in 45nm SOI CMOS.

In Chapter 3, two clockless low-swing repeaters, which exploit different mechanisms to avoid Inter-Symbol Interference (ISI), are presented for the energy-efficient, asynchronous link designs. With the aid of a multi-hop buffer bypassing flow control, the proposed low-swing repeaters will enable reconfigurable on-chip networks to approach an ideal communication fabric. The concept and energy efficiency of the clockless low-swing repeaters are analyzed and compared with other kinds of asynchronous on-chip interconnects in 45nm SOI CMOS Process Design Kit (PDK).

## Chapter 2

# Low-Swing Broadcast-Optimized On-Chip Networks

### 2.1 Broadcast-based Cache Coherence Protocols

From the perspective of computer architecture, designing on-chip networks optimized for a cache coherence protocol is critical for many-core processors to achieve peak efficiency in energy, latency, and throughput performance. The cache coherence protocol is a protocol which maintains the consistency between all the caches in a system of distributed shared memory. An extreme case of such protocol designs is a full-bit directory-based protocol where the data being shared is placed in a common directory that maintains the coherence between caches. Such a protocol minimizes bandwidth demand of the on-chip networks at the cost of substantial storage overhead per block to manage many individual cores and caches. The other end of the spectrum belongs to snooping protocols. These designs do not require any directory storage, but instead broadcast all requests and invalidates, thus increasing substantial bandwidth demand of the networks. Most of recently proposed cache coherence protocols [20] [21] [22] [23] [24] [25] lie between such two extremes, to better balance the network bandwidth and coherence state storage. These designs exploit coarser directory state than the full-bit directory-based ones, and count on a combination of broadcasts, multicasts, and direct requests to maintain cache coherence. Accordingly, it is essential to de-

sign on-chip networks that efficiently provide broadcasts and dense multicasts for the support of such advanced cache coherence protocols.

## **2.2 Broadcast-optimized Router Microarchitecture**

In response to demands for the efficient broadcast support discussed in the previous section, this work presents low-swing signaling-based, broadcast-optimized on-chip networks. The proposed on-chip networks features (1) multicast buffer bypassing flow control, (2) a broadcast-optimized crossbar switch, and (3) low-swing signaling in the data path. This section covers first two of them, and in the following section, details of the last feature will be discussed.

### **2.2.1 Multicast Buffer Bypassing Flow Control**

This work is done by T. Krishna and will be briefly discussed.

Each router in packet-switched on-chip networks has its own buffers to avoid collision of flits wishing to use the same output links. The presence of such buffers, however, makes the networks far away from an ideal communication fabric which would incur only point-to-point wire delay and energy between source and destination core. Even though previous studies tried to mitigate this problem with physical express links [26] or flow control schemes [27] [28], they all work only for unicast flits. To the best of my knowledge, the bypassing scheme presented here is the first flow control that enables multicast flits to speculatively bypass the buffering pipeline stage at routers.

The proposed flow control sends lookahead signals a cycle before actual data, to pre-allocate a crossbar at intermediate routers like the unicast bypassing schemes [27] [28]. However, the lookahead signals carry more information such as multiple output port requests and multiple destinations than unicast flits do. Also, the multicast bypassing flow control generate multiple lookahead signals, one corresponding to each output port out of which the flit forks. To maximize bypassing efficiency, output port requests of the incoming lookahead signals are prioritized over the requests of



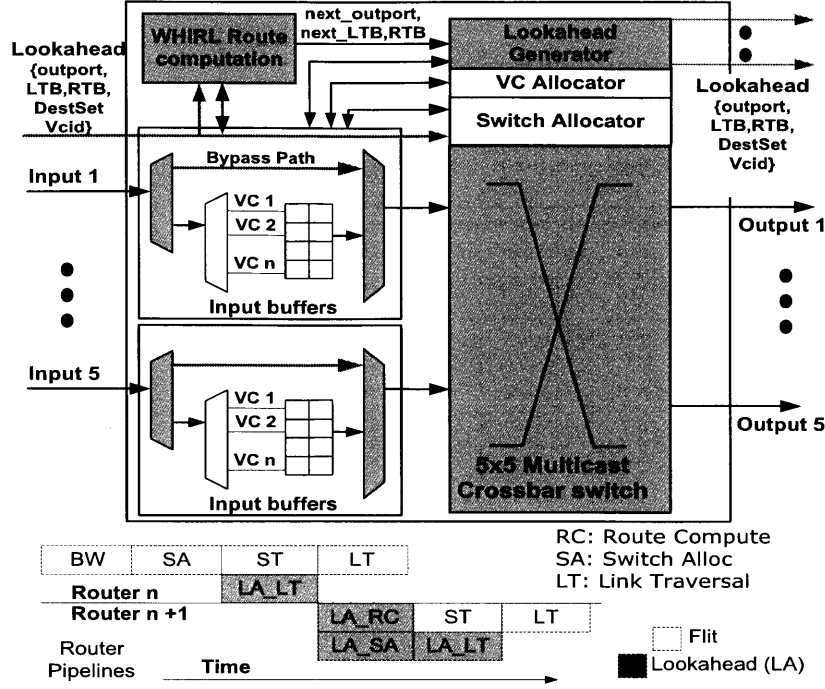


Figure 2-1: Proposed router microarchitecture.

other flits buffered at the same input port. When the lookahead wins all of its output ports, the intermediate router sets up the bypass control signals that allow the flit to connect directly to the crossbar and link, instead of getting buffered. Then, the incoming flit forks out in a single cycle, using the broadcast-optimized crossbar switch which will be described in the next subsection. The multicast bypassing flow control also supports partially successful allocation by the lookahead, in which case the incoming flit uses the bypass path and gets buffered simultaneously. The overall router microarchitecture is shown in Figure 2-1.

### 2.2.2 Broadcast-optimized Crossbar Switch

Broadcast on-chip networks add the ability for routers to fork the same flit out of multiple ports. Previous designs do this by reading the same flit out of the input buffers one-by-one and sending it out of different output ports [29], or by circling flits within a specific router and sending them one-by-one out of all requisite output ports

[30]. These approaches force broadcast/multicast flits to be queued up more in the buffers since they go serially out of each of the ports. It increases occupancy time of each buffer, which in turn increases the number of buffers required in the network to achieve a target throughput. In addition, they suffer from multiple arbitration cycles in sending out one particular flit.

These problems could be addressed by forking flits within a crossbar switch. Recent works [31] [32] presumed forking within a crossbar, but but did not propose how to realize such single-cycle multicast crossbars. Mux-based crossbars, which use multiple stages of muxes throughout the area of the crossbar to realize all possible input-to-output connections, has very high loading due to fan-out of each input to muxes corresponding to each output, thus resulting in high power consumption. Traditional pass gate-based matrix crossbars also cannot support broadcasts unless huge energy slack is available. This is because, for the broadcast flits, their input drivers should be over-sized to drive one full horizontal and all vertical wires. A possible solution to avoid such over-design is the use of adaptive input drivers. It could be implemented by using a parallel set of minimum-sized drivers, each of which connects to the input horizontal wire. M of the drivers are turned on when M-multicasts are requested, thereby providing appropriate current. However, it has propagation delay that increases with M. This latency degradation becomes worse as the crossbar size gets bigger.

This thesis proposes a broadcast-optimized crossbar switch based on tri-state drivers. In this design, input drivers only need to drive the horizontal wires. Each vertical wire has its own tri-state switch and another driver, and thus it can support unicasts and all kinds of multicasts without any over kills. The latency of the tri-state switch-based matrix crossbar is independent of M, and thus faster and more robust design than the adaptive input driver-based crossbar. Even in terms of energy, the proposed crossbar shows better performance than the adaptive input driver-based crossbar, since the capacitances of wires (about 300fJ/mm with minimum width/pitch in 45nm SOI CMOS) are an order of magnitude higher than those of the transistors (about 10fJ in 45nm SOI CMOS).

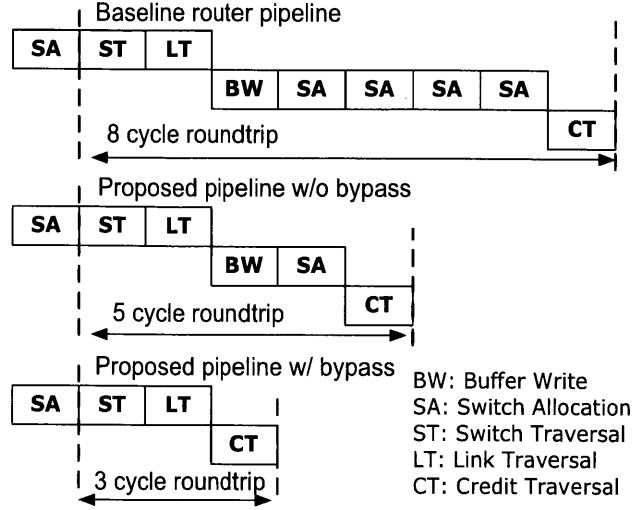


Figure 2-2: Proposed router pipeline.

Figure 2-2 describes router pipeline reduction effect achieved by the proposed router microarchitecture. The broadcast-optimized crossbar switch enables 3-cycle reduction when compared with the baselines [29] [30]. If bypass speculation also succeeds, the 5-cycle roundtrip can be once more reduced to the 3-cycle roundtrip. It is noted that this reduced router pipeline allows not only latency improvement but also energy savings by skipping up to three switch allocation processes and one buffering at each router.

## 2.3 Low-Swing Signaling in Data Path

To more aggressively increase energy efficiency of the broadcast-optimized on-chip networks described as background in the previous section, a low-swing technique has been applied to the networks. The proposed design employs reduced-swing voltage in most of its data path, from cross points of a crossbar switch to link receivers (RXs).

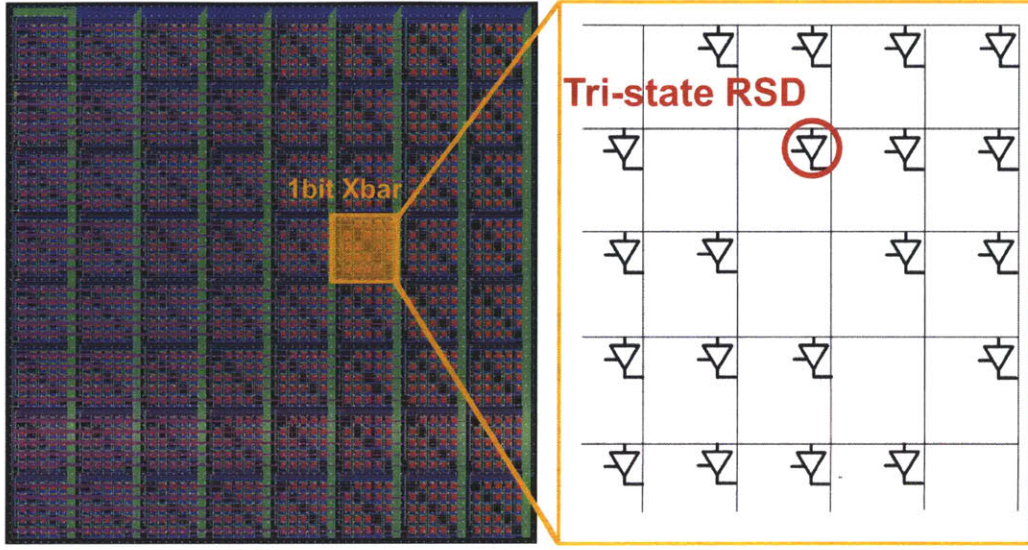


Figure 2-3: 64 bits 5x5 tri-state RSD-based crossbar switch.

### 2.3.1 Tri-state RSD-based Crossbar Switch

The proposed low-swing crossbar switch is shown in Figure 2-3. The tri-state switch and vertical wire driver, which are discussed in the Section 2.2.2, have been combined into one tri-state Reduced-Swing Driver (RSD) by stacking pass gates between VDD and Low VDD power supply rails. At the cost of the extra supply voltage, this stacked circuit design shown in Figure 2-4 has several benefits over the existing RSDs [8] [9] [12] [13] [14] [15] [16] [17] [18] [19].

First, the stacked transistors reduce off-state leakage current that is one of the major design concerns in an advanced CMOS process technology. Furthermore, when the passgate transistors are replaced with high threshold voltage transistors, the tri-state RSD includes an inherent power gating circuit. Spectre simulations in 45nm SOI CMOS show that the tri-state RSD has 18.7x smaller off-state leakage current when compared with a separate passgate switch and RSD.

Secondly, the proposed circuit, which can be viewed as the linear-mode drive transistor-based RSD type categorized in the Section 1.3, provides much bigger charging and discharging current than other RSD types such as the charge sharing-based RSDs [12] [13] [14] and cut-off drivers [8] [15] [16]. This property allows on-chip

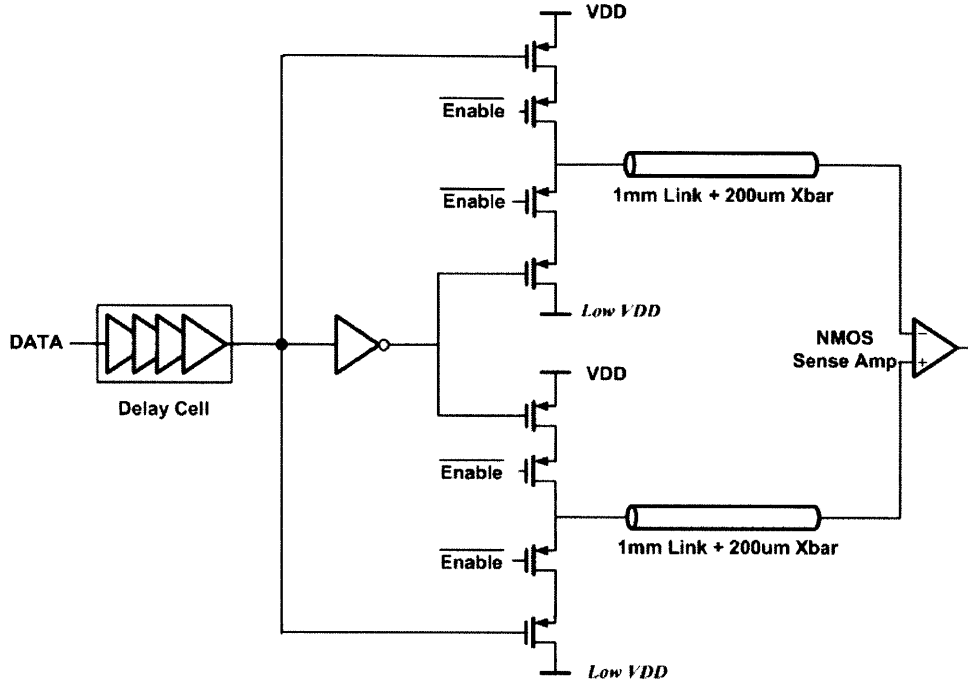


Figure 2-4: Tri-state Reduced-Swing Driver (RSD).

networks to enjoy another benefit of low-swing signaling, latency performance improvement. The tri-state RSD substantially reduces interconnect wires' delay caused by large load capacitance and resistance, enabling flits to traverse both the crossbar switch and link within a single cycle even at an aggressive clock frequency. Therefore, the 3-cycle roundtrip shown in Figure 2-2 has now been reduced to 2-cycle roundtrip. The comparison with a synthesized full-swing data path and the maximum frequency that allows the single cycle for both ST and LT will be presented later in the Section 2.5.

Thirdly, the combined RSD allows voltage swing to be reduced from cross points of the crossbar switch to link RXs, not from link TXs to link RXs, maximizing benefits of low-swing signaling. When the area of a crossbar switch increases, or the length of links decreases, this benefit gets larger. Consequently, the proposed design that employs low-swing signaling from the cross points will reap more benefits in a future CMOS process since the core-to-core distance in many-core processors becomes

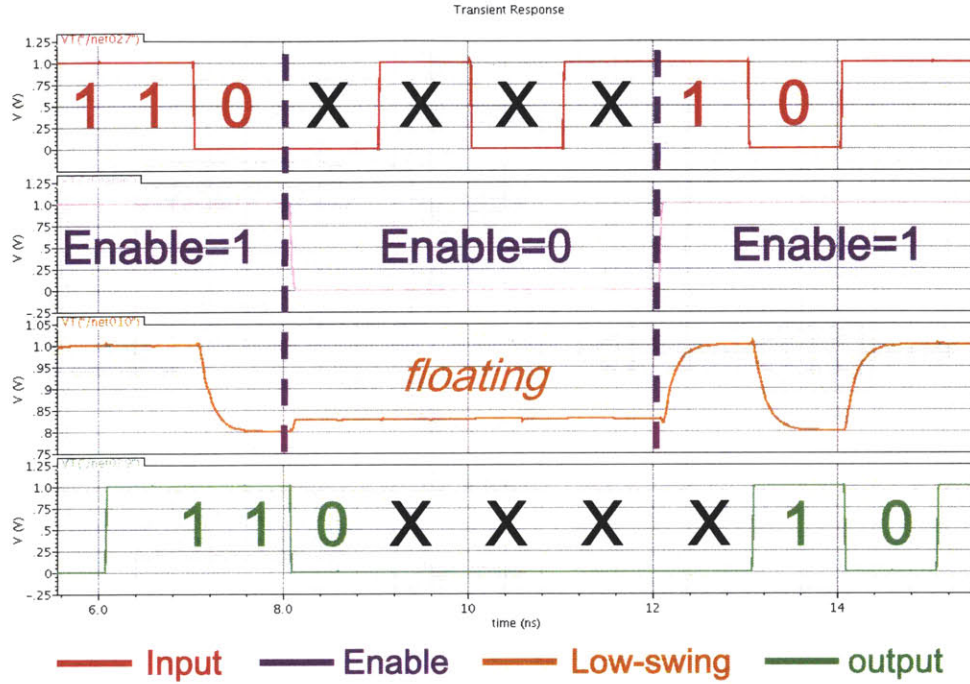


Figure 2-5: Tri-state RSD-based crossbar switch waveforms.

shorter as the process scales down.

Fourth, this design allows single transistor passgates to operate normally in SOI CMOS process where such passgates do not work due to the parasitic bipolar effect [33]. This is because one end of the passgate is directly connected to the 1mm-long wire that has large load capacitance. In such a design, the bipolar leakage exerts little effect on link functionality. Figure 2-5, which is an example of the tri-state RSD-based crossbar switch waveforms simulated with an extracted netlist in 45nm SOI CMOS PDK, shows proper functionality of the single transistor passgates.

The last benefit I would like to emphasize is that its simpler design and more reliable character than pre-emphasis drivers [17] [18] [19], which generate low-swing signaling by channel attenuation, may allow it to be integrated into the CAD synthesis flow, along with elaborate floor plan and routing. To explore such potential, a low-swing crossbar synthesis project based on the tri-state RSDs is now being led by another student, Chia-Hsin Owen Chen, as an extension of this work.

### 2.3.2 Design Considerations

The main concern of designing low-swing signaling circuits is reduced noise margin. According to [34], most of on-chip interconnect noise is generated by three sources: (1) crosstalk coupling noise, (2) channel attenuation, and (3) RX offset.

The crosstalk coupling noise comes from neighbor full-swing aggressors such as computation cores and router logics. The simplest way to reduce this coupling is the use of shielding wires, and consequently, this work adopts such shielding means. The physical layout of data and shielding wires will be described in the Section 2.4.

The channel attenuation is also becoming one of the major noise sources in on-chip wires as CMOS process scales down. It is especially critical when driving long and narrow wires due to their large parasitic resistance. Pre-emphasis techniques such as equalization [17] [18] [19] have been proposed to cancel out the channel loss, but they all suffer from huge area overhead and wire variation vulnerability. Fortunately, since the length of links is just 1mm in this work, such channel attenuation is easily compensated at the cost of a little over-sizing in RSD design. It is another merit of 2D mesh topologies. Two eye diagrams shown in Figure 2-6 reveals the channel attenuation noise caused by 1mm-wire resistance variation.

The last major noise source, RX offset, is the most crucial factor in the proposed design. Monte-Carlo simulations in 45nm SOI CMOS and 90nm bulk CMOS show that 3-sigma offset of traditional sense amplifiers [9] to be about 120mV and 90mV, respectively. To lower these RX offset voltages, compensation techniques [35] [36] [37] can be used at the expense of design complexity. However, this work employs conventional sense amplifiers without such offset compensation techniques for simplicity. Details of voltage swing decision for balance between energy efficiency and reliability will be discussed with measurement results in Section 2.5.

Another design issue of the tri-state RSD is synchronization between an enable signal and data signal. Even though these two signals are simultaneously generated by router pipeline, the enable signal needs to pass through several big drivers while the data signal arrives earlier. This is because, in matrix crossbar switches, the



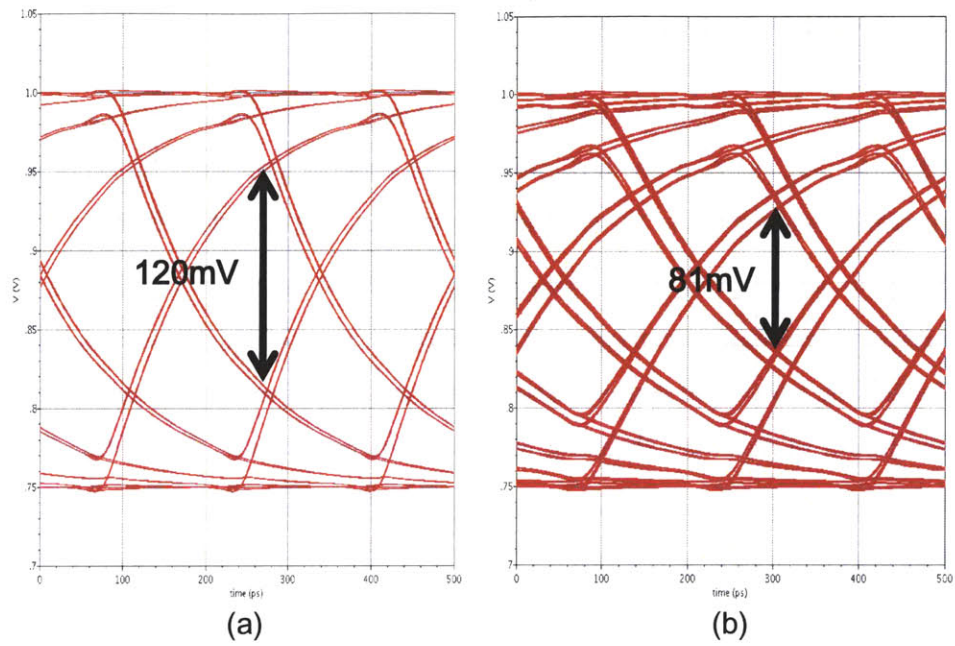


Figure 2-6: Eye diagram with (a) 1mm-wire nominal resistance (b) 1mm-wire 3-sigma high resistance.

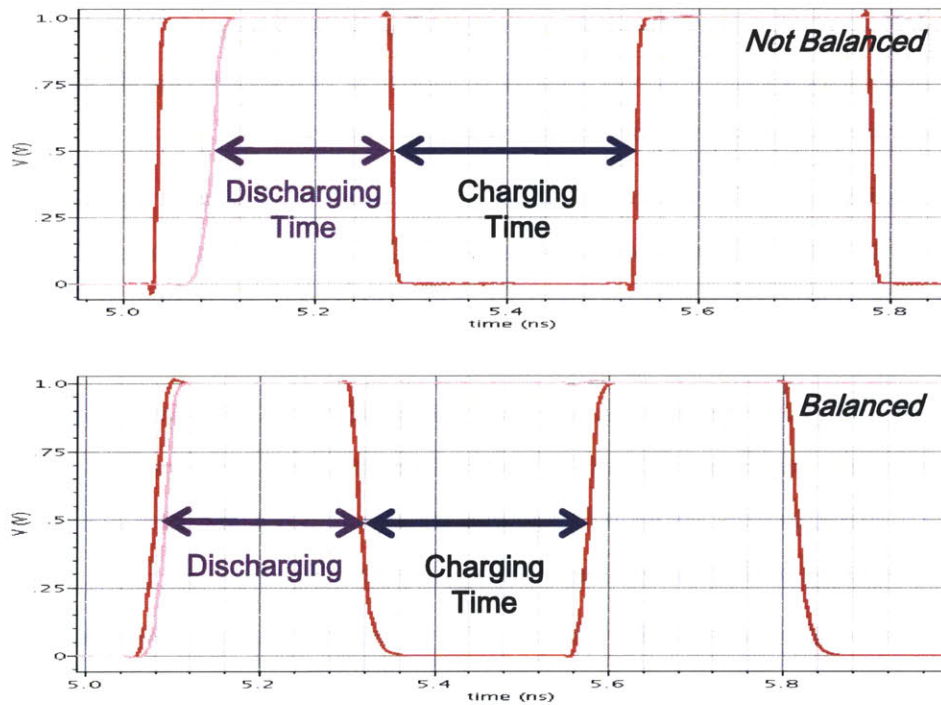


Figure 2-7: Balance between charging and discharging time caused by a delay cell



data signal drives only its corresponding 1-bit crossbar while the enable signal has to drive all of 64 1-bit crossbars. This mismatch causes the imbalance between charging and discharging time on link wires with some sequences such as data=010... and enable=011..., thus increasing inter-symbol interference (ISI). To avoid such ISI, a delay cell consisting of four minimum inverters has been added in the input signal path. The effect of the delay cell is shown in Figure 2-7.

Apart from an extra supply voltage demand, the tri-state RSD-based crossbar switch requires a relatively large area budget. As shown in Figure 2-8, the proposed crossbar switch occupies 26% of overall router area while the SWIFT [38] crossbar takes 20% of its router. This is because the tri-state RSD-based crossbar houses its RSD at every cross point, whereas traditional passgate crossbars described in the Section 2.2.2 have their drivers only at the link TXs. In the case of a 5x5 crossbar without u-turn, the tri-state RSD-based crossbar has 20 RSDs, and on the other hand, the passgate crossbars have 5 drivers. However, the passgate crossbars cannot be directly applied to some advanced fabrication processes like SOI CMOS where the single transistor passgate no longer works [33]. Therefore, design decision of low-swing crossbars should be made out of consideration for various design metrics such as crossbar size (64bits or 128bits), input port counts (5x5 or 6x6), area budget, and fabrication process.

## 2.4 Test Chip fabrication

To explore the energy efficiency and performance of the proposed on-chip networks, a proof-of-concept chip has been fabricated in 45nm SOI CMOS. Figure 2-8 shows its die photograph overlaid with a layout to outline the regions of each design blocks.

The test chip includes 16 routers on 4x4 mesh network topology. Each router has its own Network Interface Circuit (NIC) that houses a traffic injector and a traffic receiver. The traffic injector generates uniform random traffic by using a Pseudo Random Bit Sequence (PRBS) and receives arbitrary injection rate information via a scan chain from an off-chip interface. Data flits carry their generation time in the

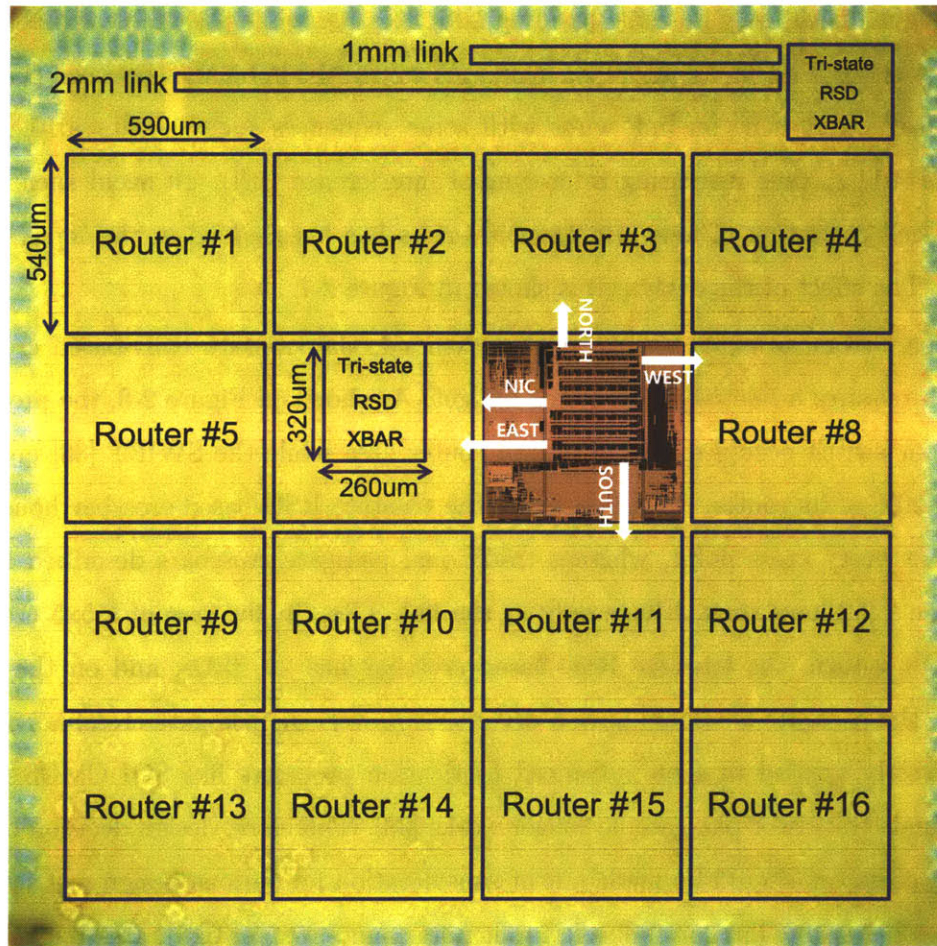


Figure 2-8: Test chip die photo overlaid with a layout.

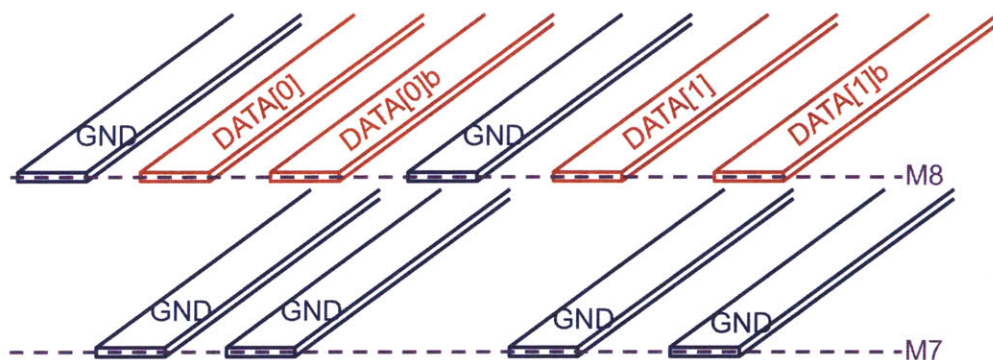


Figure 2-9: Differential signaling wire model with full shielding.

data field for the latency calculation. The traffic receiver accepts incoming flits and computes both the number of accepted packets and total packet latency.

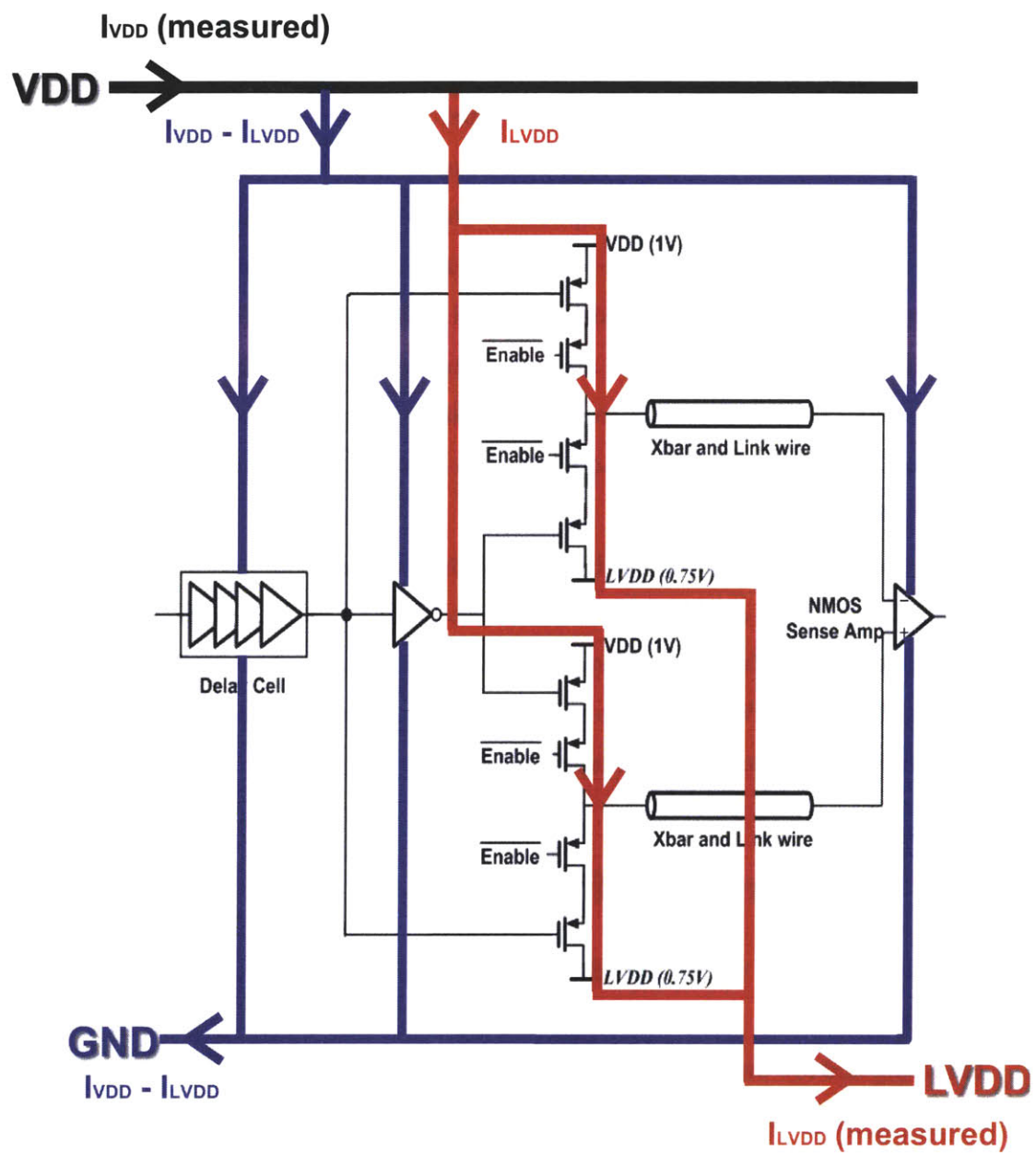
In addition to the network, the test chip contains a stand-alone 64-bit crossbar switch connected to 1mm-wire links and 2mm-wire links. Since the length of links in the network is much shorter than 1mm due to area limitation, it is necessary to test the stand-alone crossbar to explore the impact of low-swing signaling on actual many-core processors whose core-to-core distance is 1mm or 2mm.

Figure 2-9 shows wire model diagram used as links in both the 4x4 networks and stand-alone crossbar switch. As mentioned in the Section 2.3.2, this work employs full shielding wires to minimize crosstalk coupling noise. To create RC-dominant wire channel, the values of 0.15um and 0.30um are selected as the wire width and space, respectively.

To reduce cost of high-speed I/Os, all configurations are setup by slow I/Os only except the reference clock. Data read and write operations of the scan chain are done through the slow I/Os. Also, an on-chip clock generator is implemented with a Voltage Control Oscillator (VCO) in case the high-frequency external clock is contaminated by parasitic inductance of bonding wires.

## 2.5 Measurement Results

This section proves the energy efficiency of the proposed low-swing data path with experiment results. Figure 2-11 and Figure 2-12 show energy comparison between full-swing and low-swing signaling on 1mm link and 2mm link, respectively. The low-swing signaling was generated by the tri-state RSD described in Section 2.3.1 and the equivalent full-swing driver was designed with the same maximum data rate as the low-swing data path. Equation 2.3 shows the energy calculation equation used in Figure 2-11 and Figure 2-12. In this equation, only two terms,  $I_{VDD}$  and  $I_{LVDD}$ , are measured from the test chip.  $I_{VDD}$  is the total current drawn from nominal power supply (1.0V), and  $I_{LVDD}$  is the current sinking to the second supply voltage (0.75V). Figure 2-10 shows details of the power measurement circuit diagram.



**Full-swing Path Power =  $VDD \times (I_{VDD} - I_{LVDD})$**

**Low-Swing Path Power =  $(VDD - LVDD) \times I_{LVDD}$**

Figure 2-10: Power consumption measurement circuit diagram.

$$E_{total} = E_{full-swing\ path} + E_{low-swing\ path} \quad (2.1)$$

$$= I_{full-swing} \times VDD \times \frac{1}{f} + I_{low-swing} \times (VDD - LVDD) \times \frac{1}{f} \quad (2.2)$$

$$= \{(I_{VDD} - I_{LVDD}) \times VDD + I_{LVDD} \times (VDD - LVDD)\} \times \frac{1}{f} \quad (2.3)$$

The 1mm 250mV tri-state RSD-based link enables 58% - 64% energy reduction and the 2mm link shows 66% - 71% energy savings when compared with their equivalent 1V full-swing links. The measured TX energy efficiency is almost the same (within 5%) as the simulated efficiency, but the measured RX energy efficiency is 7% - 15% higher than the simulation results due to the nonideal clock duty ratio. Since energy benefits of low-swing signaling come from reduction in dynamic power of link wires, the 2mm low-swing link show higher energy efficiency than the 1mm low-swing link. On the other hand, the 1mm link shows 1.9x higher bandwidth than the 2mm link with the identical tri-state RSD design. This is because both wire parasitic capacitance and resistance of 2mm wires are two times bigger than them of 1mm wires, resulting in 4 times larger RC time constant. To increase the bandwidth of the 2mm link, a bigger tri-state RSD will be required at the cost of energy and area.

Measurement results show that the tri-state RSD-based crossbar switch enables 51% - 58% energy savings as compared to an equivalent full-swing data path. Since additional power is dissipated to drive gate capacitors of RSDs even in non-activated cross points, the energy efficiency of the low-swing crossbar is about 10% lower than the RSD-based links. It can be seen that such 10% energy overhead is the expense of reduced latency in router data traversal pipeline stages. As discussed in Section 2.3, the tri-state RSDs located at cross points allow low-swing data path to be longer than the design where RSDs are housed at link TXs, thereby resulting in the combined ST and LT at aggressive data rate. The test chip experiment shows the proposed low-swing crossbar allows the combined ST and 1mm-LT to operate at 3.8GHz clock frequency. This data rate is 21% higher than the 1V full-swing driver optimized for energy-delay product.



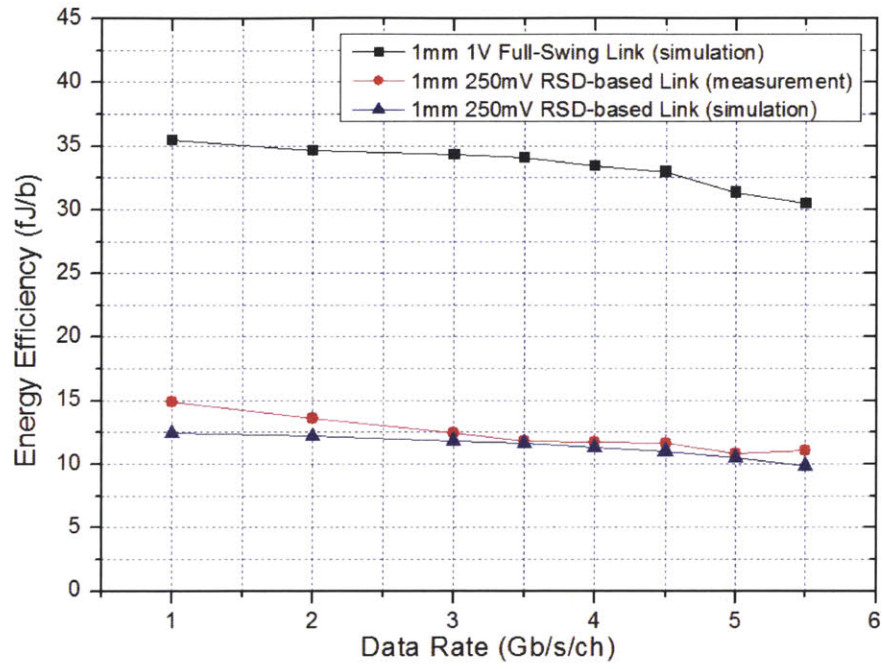


Figure 2-11: 1mm link energy comparison between full-swing and low-swing signaling.

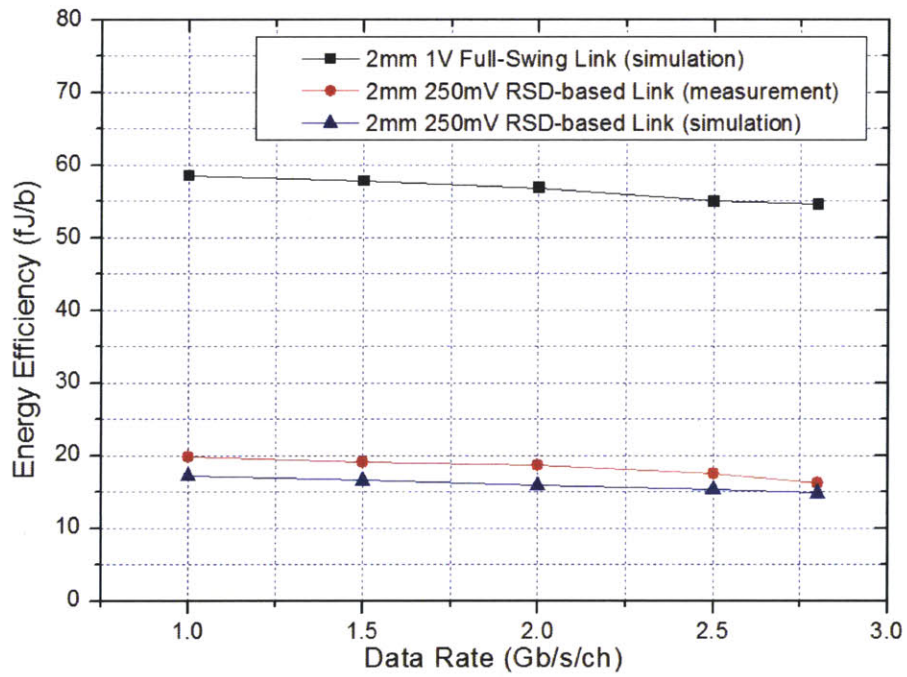


Figure 2-12: 2mm link energy comparison between full-swing and low-swing signaling.

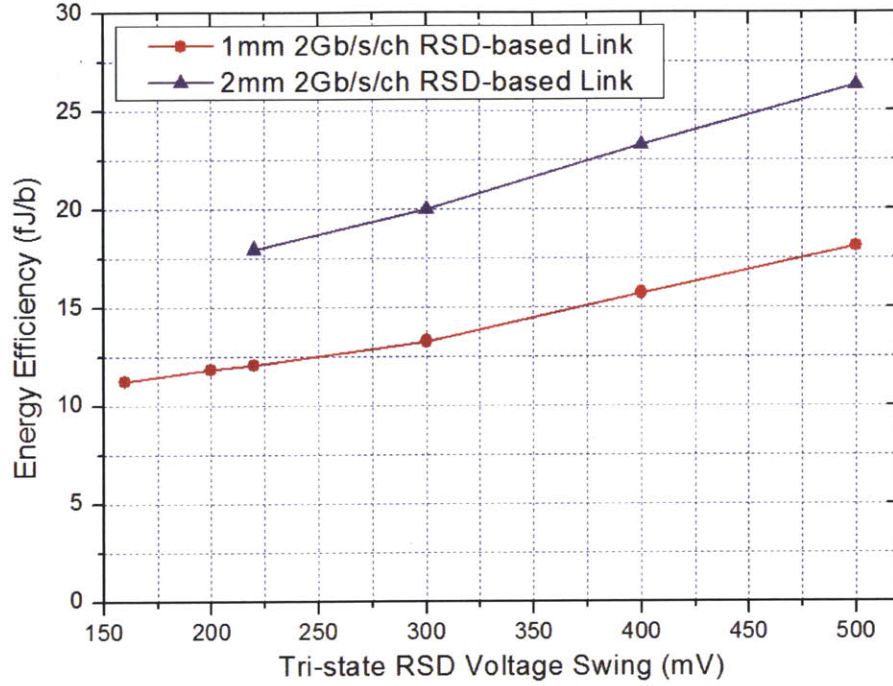


Figure 2-13: Tri-state RSD energy measurement with various voltage swings.

Figure 2-13 shows energy efficiency characteristics with variations of voltage swing. It is noted that low-swing signaling energy benefits are not too high when the voltage swing goes down to 160-240mV level. This is because the fixed link RX energy becomes dominant as the link TX energy gets smaller. Considering the voltage swing which endures  $3\text{-}\sigma$  sense amplifier offset is about 210mV, it is concluded that the 250mV voltage swing is the best design choice for balance between energy efficiency and reliability.

Figure 2-14 shows power measurement results of the tri-state RSD-based 5x5 crossbar switch with changes in multicast counts: a unicast, 2 multicast, 3 multicast, and broadcast. At 3.8GHz clock frequency, the crossbar sense amplifiers (RXs) consume 48uW regardless of multicast counts. On the other hand, the tri-state RSDs (TXs) power dissipation linearly increases as the number of multicasts increases. Actually, the slope of the TX power linear curve is much bigger than the measured slope. Due to a design mistake on the test support circuit where other 4 input ports are in a

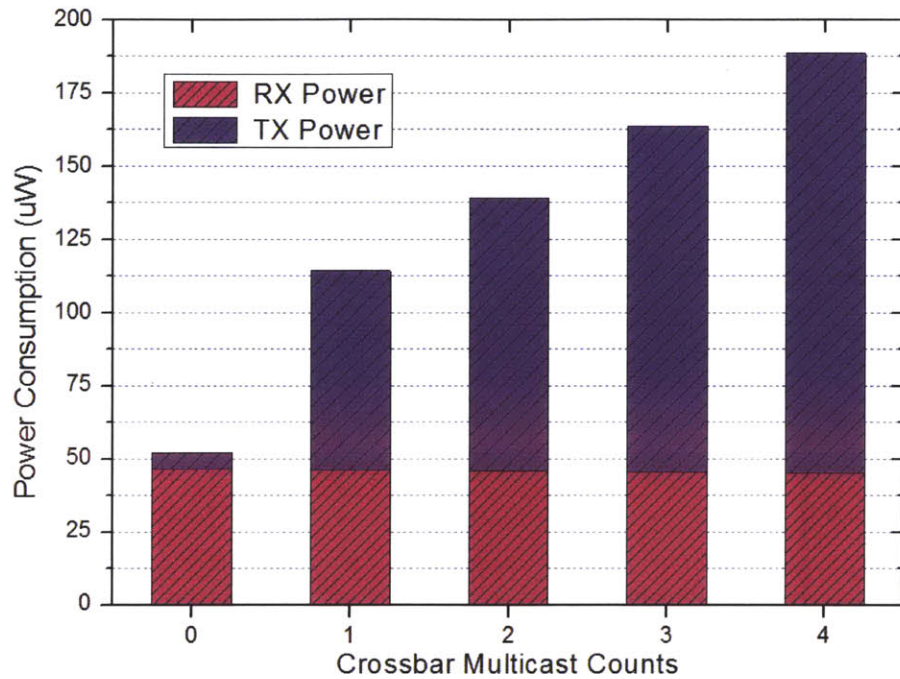


Figure 2-14: 3.8GHz 1bit 5x5 crossbar switch energy measurement.

floating state, such high impedance inputs oscillate under the influence of coupling with the activated input port. As a result, the measured TX power consumption shows fixed offset regardless of multicast counts. Fortunately, this problem will not appear at the 4x4 network side since all the input ports in the network are connected to static flip-flops allowing the input ports to always stay in a low impedance state.



# Chapter 3

## Clockless Low-Swing Repeaters

This chapter explores low-swing link designs that enable energy and delay performance of a point-to-point interconnect of variable lengths. With the support of a multiple-hop buffer bypassing flow control, such links will make on-chip networks approach an ideal communication fabric.

### 3.1 Motivation for Clockless Low-Swing Signaling

An ideal communication fabric would incur only dynamic energy and delay of wires between the source and destination core. But dedicated global point-to-point wires between all nodes do not scale [39], and hence, the networks that multiplex and share wires are widely accepted to be the way forward.

If existing low-swing signaling circuits [8] [9] [12] [13] [14] [15] [16], which employ a reference clock to convert low-swing signal to full-swing signal at their receivers (RXs), are applied to such networks, they cannot make the best use of low-swing benefits on both energy and latency. This is because the intermediate routers, which just play a role as a multiplexer, waste many clock cycles and energy to convert low-swing signal to full-swing signal at all the data traversal pipeline stages. To easily understand this disadvantage, consider an example of the 9x9 regular mesh on-chip networks in Figure 3-1 (a) with the assumption that some flits travel only between selected 11 nodes shown in Figure 3-1 (b). When the clocked low-swing circuits

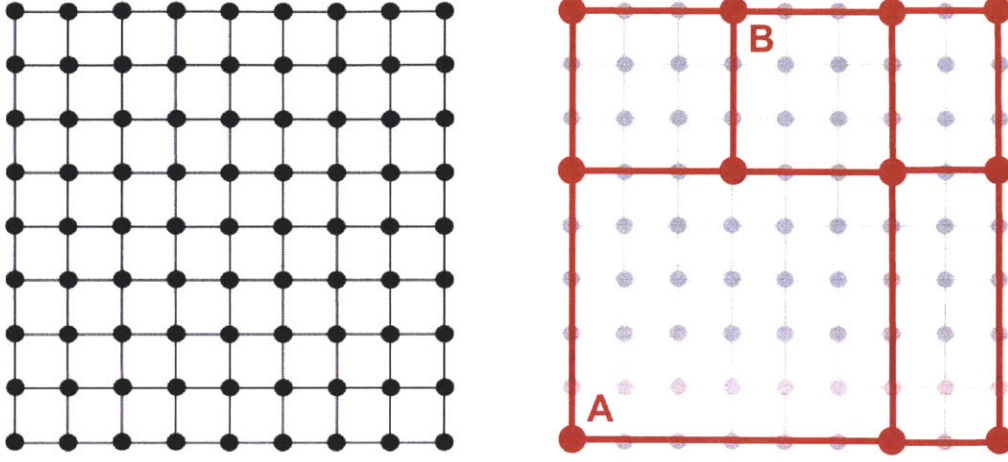


Figure 3-1: (a) 9x9 regular mesh on-chip network (b) 11-node irregular mesh on-chip network.

are used for the flits that travel from node A to B, it takes 11 cycles, instead of 3, consuming unnecessary energy at every intermediate router. On the other hand, clockless repeaters may take 3 cycles if the data transmission on the repeater scheme is fast enough.

To eliminate such overhead and move toward to the ideal communication fabric, we propose applying low-swing signaling to repeater insertion. Considering the built-in shared wires and multiplexers for the packet-switched on-chip networks, the low-swing application to the repeater system could be viewed as a natural choice. This is because the pre-emphasis technique [17] [18] [19], an alternative to cancel out Low Pass Filter (LPF) characteristic of the channel, requires different driver sizes, receiver designs, and even bias current as wire length changes, thereby resulting in poor scalability.

However, it is unclear how to design such low-swing repeaters that can achieve high energy efficiency and fast transmission without the reference clock, at a small footprint. Channel attenuation caused by RC-dominant on-chip wires cannot be directly used for generating low-swing signaling due to inter-symbol interference (ISI) described in Figure 3-2. In order to avoid the ISI, the low-swing repeaters should

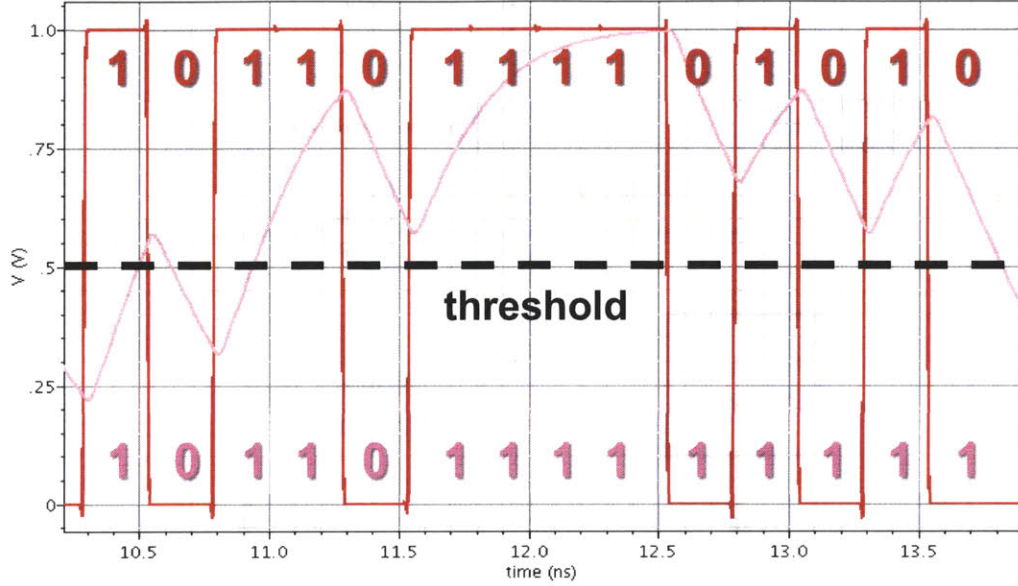


Figure 3-2: Inter-Symbol Interference (ISI) caused by channel loss.

provide some mechanism to maintain the input voltage of the repeater cell within a certain level. At the same time, such input voltage has to be energy-efficiently repeated without a clock.

This thesis proposes two separate low-swing repeater designs that satisfy such requirements of the clockless low-swing signaling. At the first design named Self-Resetting Logic Repeater (SRLR), the input voltage of repeater cells is reset to zero as soon as the repeaters cell recognize the input logic value. In other words, data are transmitted by voltage-limited pulses. On the other hand, at the second design named Voltage-Locked Repeater (VLR), the input voltage is maintained near the RX threshold voltage by feedback circuits and keeper transistors. As shown in the following sections, both of the proposed low-swing repeaters provide high speed on-chip communication without a reference clock, achieving energy and delay of a single point-to-point link of variable hop counts. Moreover, they feature digital-intensive properties, requiring no differential signaling and bias current. This characteristic provides higher wire density and potential for being integrated into the CAD synthesis flow.

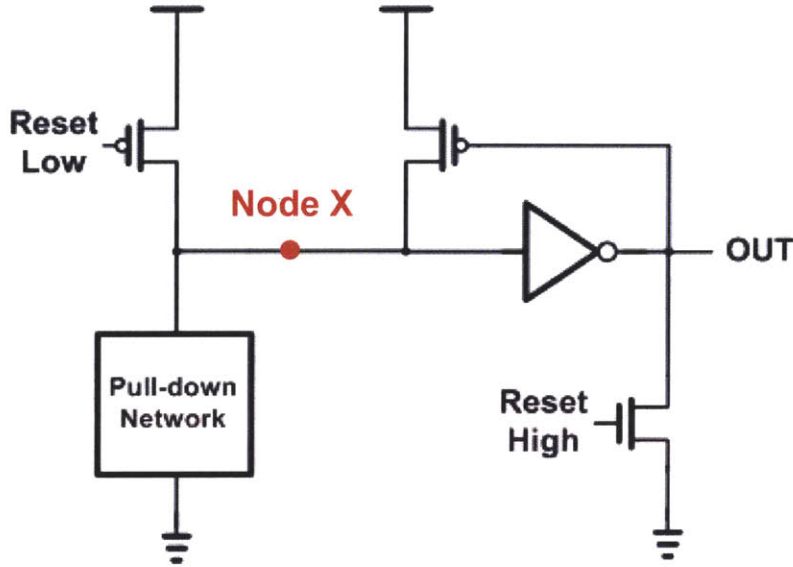


Figure 3-3: Generic view of Self-Resetting Logic (SRL).

## 3.2 Self-Resetting Logic Repeater (SRLR)

### 3.2.1 Introduction to Self-Resetting Logic

Self-Resetting Logic (SRL) provides a design solution where clocking overhead is minimized. Figure 3-3 shows a generic view of such a SRL gate. A pull-down network receives  $n$  input data pulses, and an output provides a pulse if the pull-down logic becomes TRUE. The reset signal is implemented as two separate pulses, Reset Low (RL) and Reset High (RH). RL signal is used to reset the input stage, while RH is used to reset the output stage after the output data has been propagated.

The basic operation of the SRL gate is as follows. The gate is initially in its standby state where power consumption is zero. Upon receiving input data, some switching occurs and an output pulse is generated. After the output pulse has reached a defined width, and provided that the inputs become inactive, the gate will be reset, going to its standby state again. Based on how the reset signals, Reset Low (RL) and Reset High (RH), are generated and used, SRL circuits have been actively studied in the context of SRAM designs with very short cycle times [40] [41] [42] [43]. However,

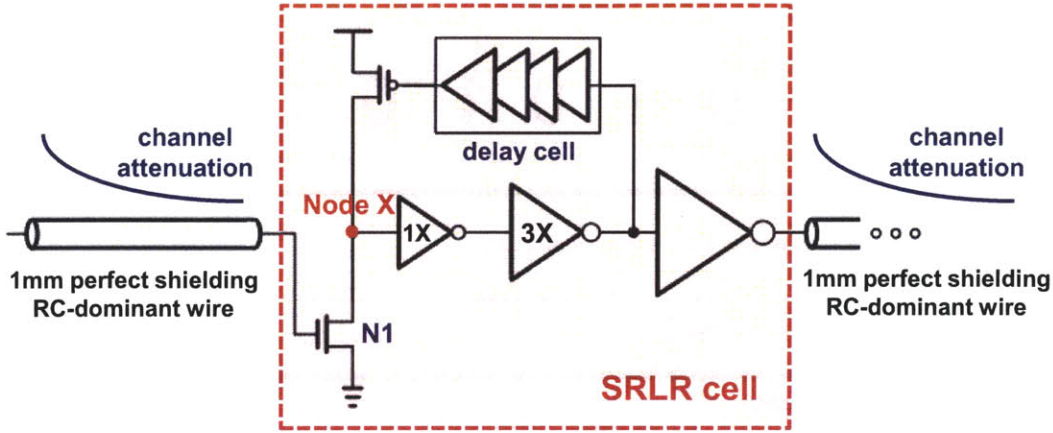


Figure 3-4: Proposed SRLR design (not yet completed).

such SRL concepts have never been applied to the domain of on-chip interconnects yet.

The best benefit of SRL circuits is zero static current. The SRL gates consume power only when their output changes. Another benefit of SRL is that node X shown in Figure 3-3 is a high impedance (floating) state in the standby mode. Indeed, this property plays a critical role in generating low-swing signaling in repeater systems. Details will be discussed in the following Section.

### 3.2.2 Process Variation Robust SRLR Design

Figure 3-4 shows the first phase of process variation robust SRLR design. The input NMOS (N1) corresponds to the pull-down network in the general SRL gate block diagram described in Figure 3-3, and Reset Low (RS) signal is locally generated, requiring no additional control logics. When an pulse arrives at N1, the node X is discharged and output voltage of the SRLR cell becomes high. The node X is again charged when a reset signal comes back though a delay cell. Finally, the node X is in a floating state, generating another pulse at the output. In order to provide an adequate pulse width without an increase in latency, the delay cell is located in a resetting feedback path. An example of SRLR waveforms is shown in Figure 3-5 to



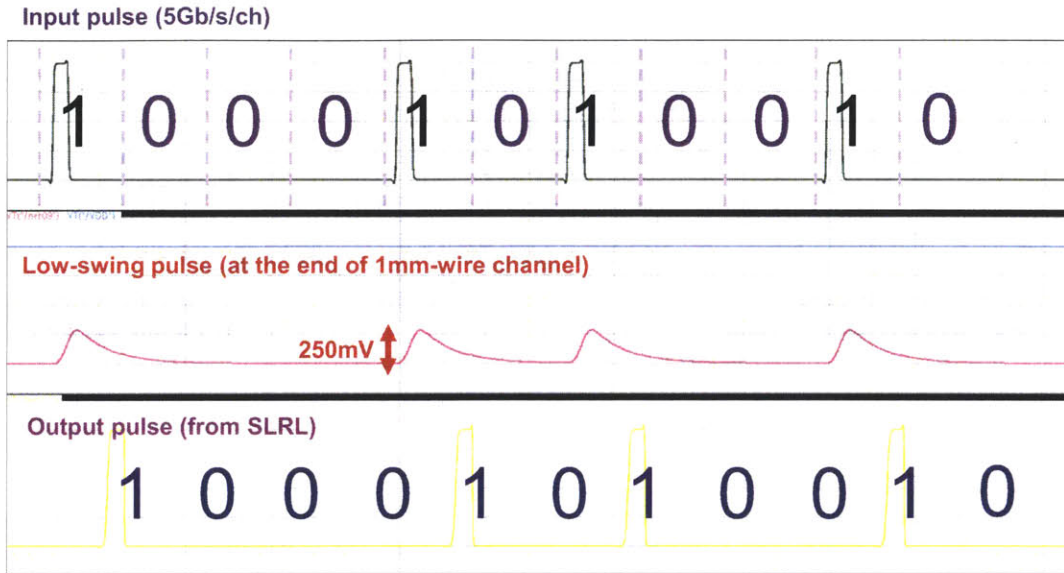


Figure 3-5: Example of SRLR waveforms.

help such an operation. Since the proposed circuit works with pulses, extra effort should be devoted to generating such pulses from synchronized data.

The high impedance node X in Figure 3-4 SRLR design is slowly discharged by the leakage current of the receiver NMOS (N1). For receiver sensitivity, N1 size is relatively large so that the leakage current of N1 is bigger than that of the smallest keeper PMOS. In other words, an off-state resistance of N1 is smaller than that of the keeper PMOS. When the node X voltage goes down to the threshold voltage of the 1X inverter by N1 leakage current, the SRLR generates a pulse even though there are no input signals. As a result, SRLR will oscillate with unnecessary energy consumption. To maintain the node X (or SRLR output) at logic high (or logic low) in the standby mode, another NMOS (N2) has been added to the reset path. This modification is shown in Figure 3-6. N2 keeps the node X voltage at  $V_{DD} - V_{th}$  after output pulses are generated, thereby preventing the node X voltage from being discharged by the leakage current below the threshold voltage of 1X inverter. It is noted that as long as N1 is turned off, the added NMOS (N2) does not cause any static currents. N2 just provides the amount of charge loss caused by N1 leakage current.

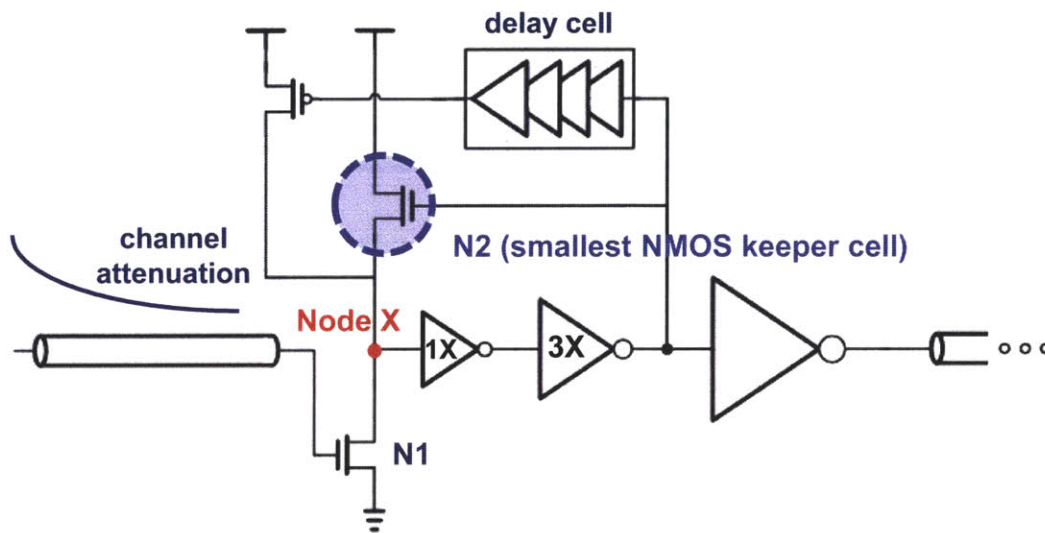


Figure 3-6: Oscillation-free SRLR design.

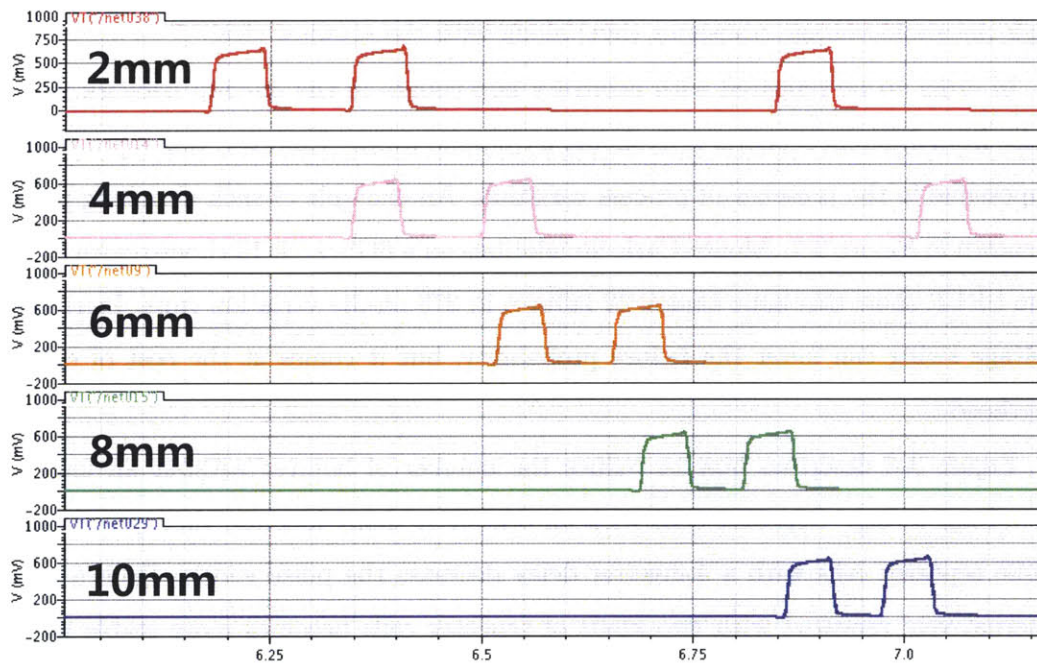


Figure 3-7: Example of shrinking pulses caused by process variation.

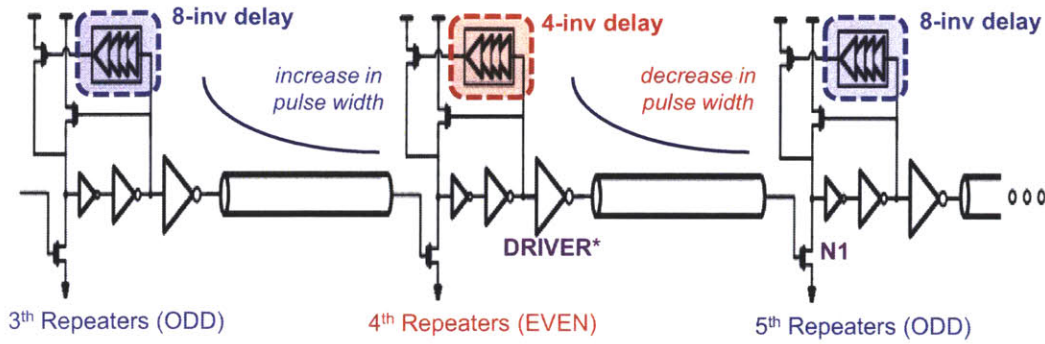


Figure 3-8: Alternately repeated SRLR with two different pulse widths.

The NMOS keeper transistor (N2) holds the node X voltage at  $V_{DD} - V_{th}$ , instead of  $V_{DD}$ , enabling standby voltage to be lower than the case of a PMOS keeper design. Monte-Carlo simulations show that the process variation effect of N2 is negligible as compared to that of N1 whose size is much bigger than N2. Consequently, the NMOS keeper design allows smaller discharging time than the PMOS keeper case with little increase in process variation effect. To minimize contention with N1, N2 employs a high threshold voltage transistor (hvt) along with the smallest size.

In order to be repeated with arbitrary hop counts in the on-chip network, every pulse from different repeater cells should have same width. However, this is practically impossible in the presence of process variation. An example of such shrinking pulses is shown in Figure 3-7. Monte-Carlo simulations on a 6Gb/s/ch 10 times-repeated 10-mm SRLR show this issue causes 22 failures in 273 on-die variation runs. Increasing voltage swing can lower the failure probability, but it comes at the cost of energy overhead.

Figure 3-8 describes how to reduce the number of failures without an increase in voltage swing. Repeater cells with an 8-inverter delay increase their pulse width while repeater cells with a 4-inverter delay decrease the pulse width. Such a pulse alternatively repeated with two kinds of repeater cells becomes more stable in the process variation environment. Monte-Carlo analysis with the same condition as the circuit in Figure refsrlr2 shows that this simple modification significantly improves process variation performance, from 22/273 to 5/273 failure probability.





when input data have consecutive ones (data = 111111) and all the 5 errors come from these consecutive 1's cases. To sum up, when drivers are designed to endure 3-sigma high threshold voltage of receiver NMOS (N1), the NMOS driver-based SRLR shows more stable performance than the inverter driver-based SRLR, resulting in no errors in 273 on-die variation runs on the 6Gb/s/ch 10-mm wire channel.

However, the proposed SRLR design in Figure 3-9 still suffers from the die-to-die variation. In general, two techniques have until now been employed to reduce such die-to-die variation failures: an Adaptive Body Bias (ABB) and Adaptive Supply Voltage (ASV).

ABB could be the best solution for the process variation robust SRLR design since the process variation performance of SRLR is mostly limited by one transistor, receiver NMOS (N1). ABB can simply lower (or increase) threshold voltage of N1, to restore functionality (or reduce voltage swing). However, the tuning range of ABB is limited because of junction leakage current [44]. Above all things, ABB can be applied only to some limited fabrication processes that provide triple n-wells. Unfortunately, since this work is based on SOI CMOS process, ABB technique is not able to help SRLR endure die-to-die variations.

On the other hand, ASV can be applied to almost any kind of fabrication processes. The implementation of ASV, however, is more difficult than ABB. In ABB, once a body voltage reaches the desire level, it experiences only small perturbation from leakage current. Thus, the power supply for the body bias does not need to be strong. In contrast, ASV needs to accommodate a large and sudden current withdrawal from transistors. Hence, ASV usually requires on-chip voltage regulators. Switched-capacitor DC/DC converters can be used for the energy-efficient on-chip voltage regulators.

This thesis proposes another design technique to mitigate the die-to-die variation impact, leaving ABB and ASV applications to SRLR as future work. Figure 3-10 shows the proposed self-calibrating SRLR with a threshold voltage monitor circuit. In the threshold voltage monitor circuit which consists of current source [45] and the same transistor as N1, its output voltage ( $V_{ref}$ ) decreases when threshold voltage

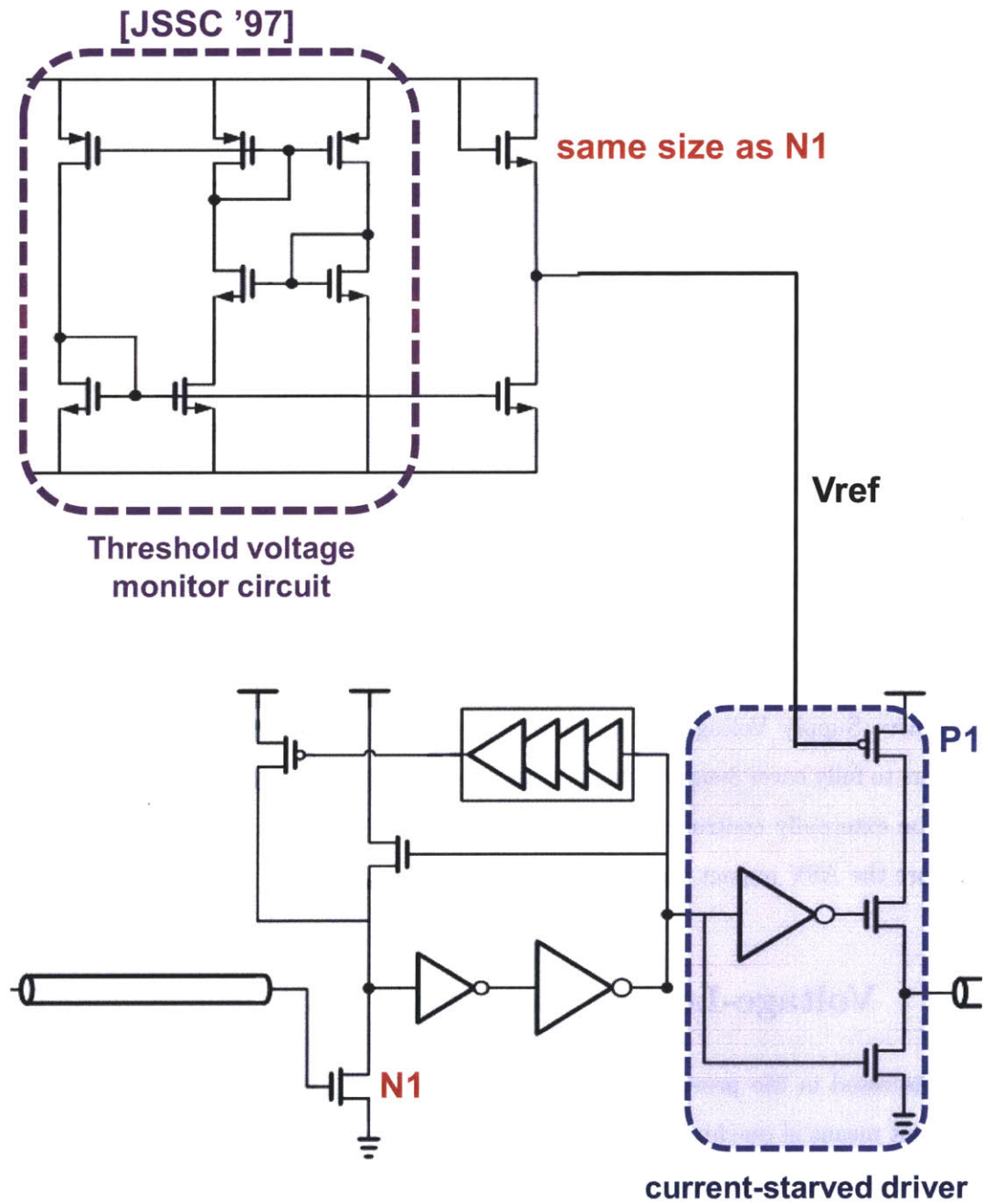


Figure 3-10: Self-calibrating SRLR with a threshold voltage monitor circuit.

of  $N_1$  increases. This output voltage is directly applied to a stacked PMOS ( $P_1$ ) in the current-starved driver to adjust charging current. Even though the equation of  $V_{ref}$  described in [45] is not perfectly applied to SOI CMOS process, the inverse tendency between threshold voltage of  $N_1$  and  $V_{ref}$  is still valid. Plus, since tail current generated by the monitoring circuit is tolerant of its PVT variation even in SOI CMOS, the  $V_{ref}$  is stable with respect to PVT variation of the monitoring circuit itself.

The self-calibrating SRLR can be considered as a background calibration circuit that provides real-time monitoring ability. Accordingly, the proposed design automatically adjusts voltage swing not only to initial process variation, but also to other variations such as temperature or aging. More importantly, this self-calibrating SRLR does not require any post-silicon test cost. On the other hand, this analog-controlled scheme does not fully cover 3-sigma die-to-die variation due to lack of sensitivity. Both die-to-die and on-die Monte-Carlo analysis reveals that there are still 4 failures in 500 runs.

Instead of over-designing SRLR drivers, this work assumes that SRLR employs Adaptive Supply Voltage (ASV) with a separate power supply rail for its NMOS drivers to fully cover 3-sigma die-to-die process variation. The separate supply voltage will be externally controlled by off-chip regulators in the test chip, to conveniently explore the ASV impact on process variation.

### 3.3 Voltage-Locked Repeater (VLR)

As discussed in the previous section, a self-resetting technique provides an energy-efficient means of on-chip communication, requiring no reference clock. Such a pulse-based data transmission, however, limits its data rate since both charging and discharging should be completed within a single cycle. Even though Spectre simulations on 45nm SOI CMOS show the SRLR supports up to 6.4Gb/s/ch data rate, it will decrease as the core-to-core distance gets larger than 1mm or interconnect wires become more narrower in a future fabrication process. Or on-chip networks may need

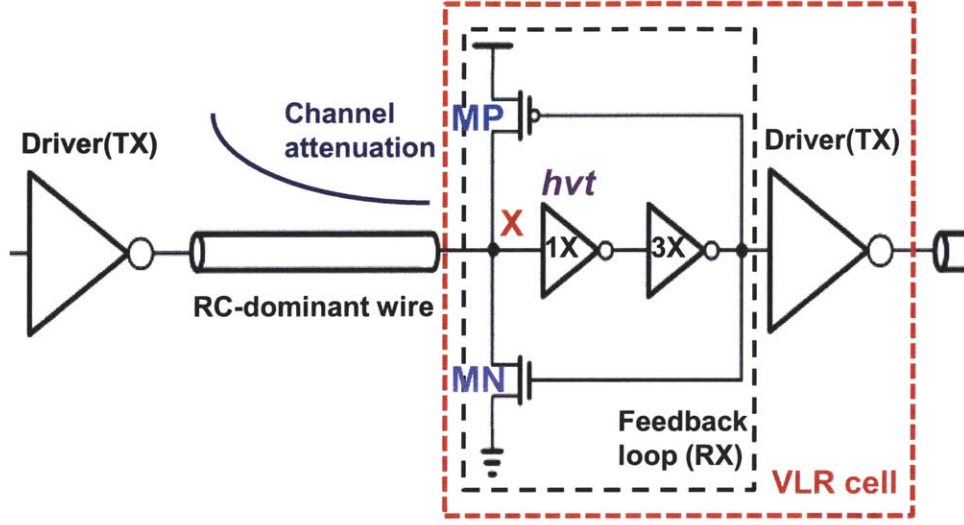


Figure 3-11: Proposed VLR without enable switches.

higher bandwidth and much lower latency links at fixed wire density for the Double Data Rate (DDR) and reconfigurability. To achieve such link requirements on the built-in shared wires and multiplexers in the packet-switched on-chip networks, this section introduces another clockless low-swing repeater system named Voltage-Locked Repeater (VLR). Figure 3-11 shows the proposed VLR circuit design.

Figure 3-12 describes how VLR generates low-swing signaling. To the best of my knowledge, all the existing low-swing signaling circuits [8] [9] [12] [13] [14] [15] [16] [17] [18] [19] limit the voltage swing at their TX. VLR, on the other hand, generates low-swing signaling at RX through a highly-resistive channel. When the node X voltage exceeds threshold voltage of the first inverter (1X), the logic High signal starts to travel through the lower feedback path (3X and MN). After the returning signal turns MN on, the node X is discharged and it is finally locked at some voltage level. The specific voltage level is determined by wire resistance and on-state resistance of MN. The logic Low (0V) operation is similar to the logic High (1V) operation. In this case, the node X voltage is locked by the upper feedback path (3X and MP) and its value is determined by wire resistance and on-state resistance of MP.

Figure 3-13 shows an example of VLR waveforms. Two keeper transistors, MP and

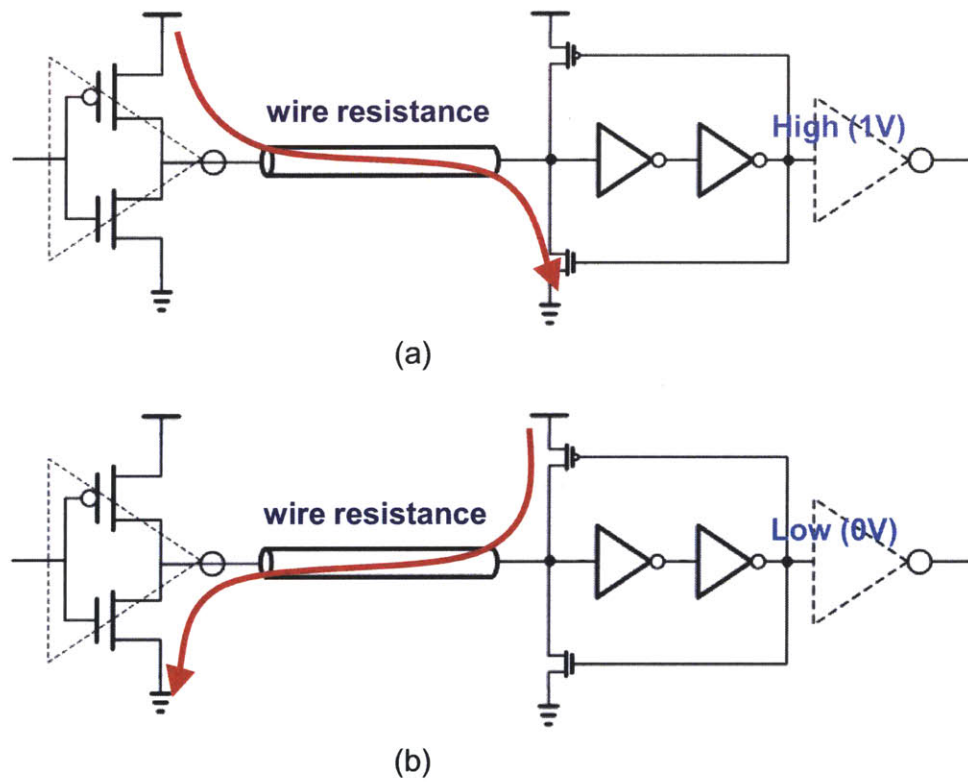


Figure 3-12: VLR operation: (a) logic High low-swing generation (b) logic Low low-swing generation.



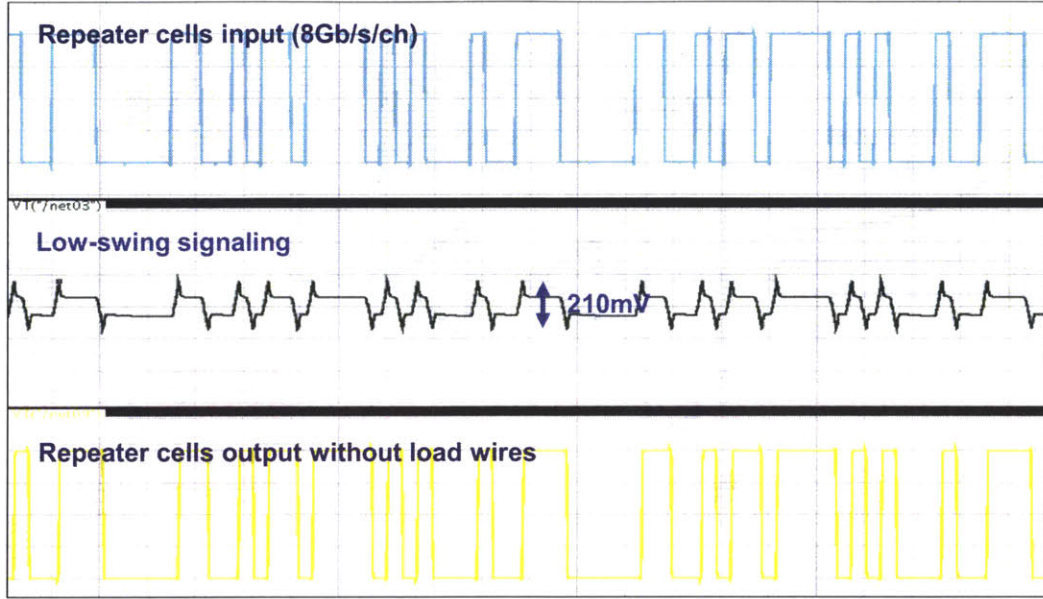


Figure 3-13: Example of VLR waveforms.

MN, allow voltage swing to remain near threshold voltage of the first inverter (1X), minimizing Inter-Symbol Interference (ISI) caused by channel attenuation. Since both charging and discharging currents are provided by linear mode transistors whose  $V_{ds}$  stays between 400mV and 600mV, VRL features higher data rate than cut-off drivers [8] [15] [16] and other linear-mode drivers [10] [11]. Spectre simulations in 45nm SOI CMOS show that VLR supports up to 15Gb/s/ch data rate on a 10-mm wire channel.

More importantly, VLR features better process variation robustness than SRLR. Voltage overshoots observed at the low-swing signaling waveform (the second window) in Figure 3-13 increase a noise margin of VLR. Since such overshoots are generated by inherent delay to lock the node X voltage, VLR does not require any other additional circuits to compensate process variation. Monte-Carlo analysis on a 10mm-wire channel proves that this simple VLR design shown in Figure 3-11 by itself causes no errors in 273 runs (3-sigma) on-die variation.

However, the VLR-based links suffer from poor energy-efficiency when on-chip networks' traffic is low. Even though its static current flows along with a very high

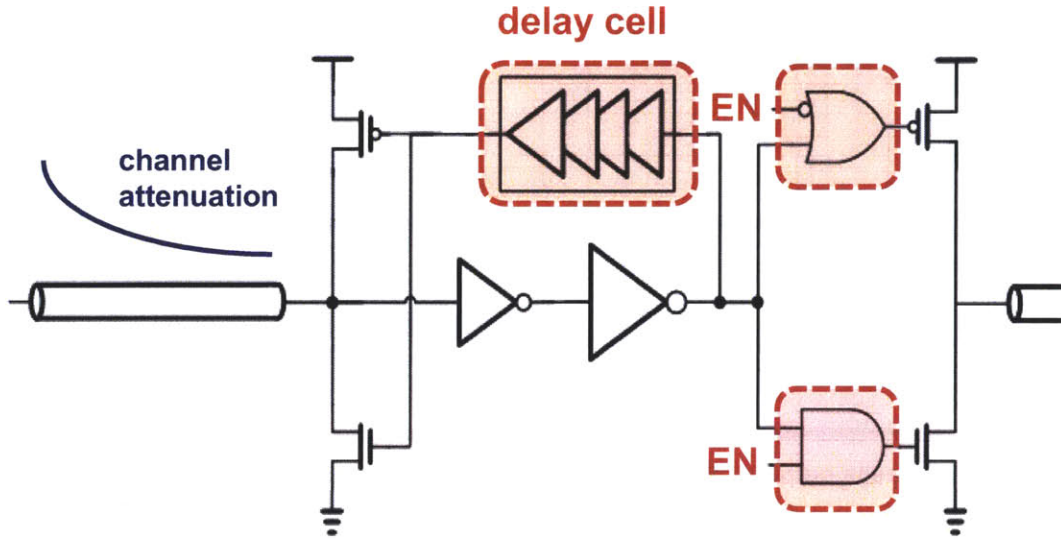


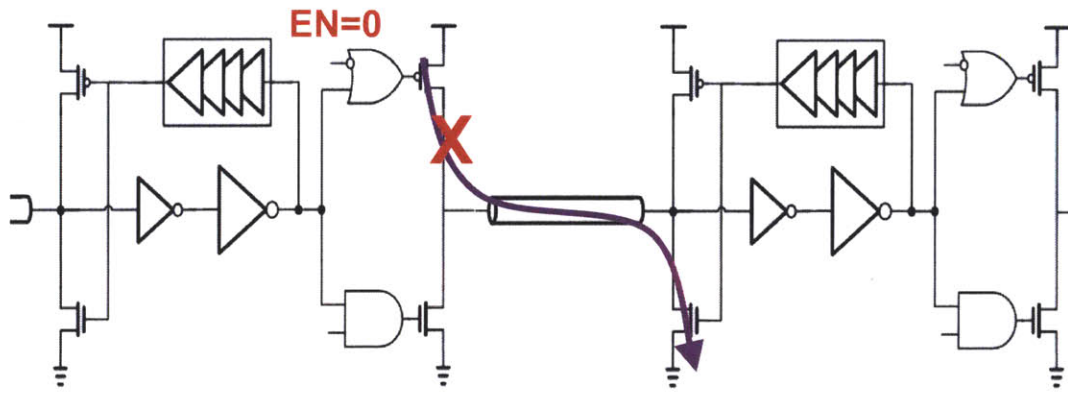
Figure 3-14: Zero static current, process variation robust VLR design.

impedance DC path, which consists of highly-resistive channel and the smallest keeper transistors shown in Figure 3-12, the static current is still 10 times bigger than its leakage current. To disconnect the DC path when links are idle, activation switches have been added to the original design. Figure 3-14 shows this modification. The activation switches add an one-gate delay to each repeater cell, but such a delay will be canceled out when the modified VLR is embedded into crossbar switches. In other words, the VLR design shown in Figure 3-14 can be directly applied to low-swing crossbar switches the way the tri-state RSD discussed in Chapter 2 does.

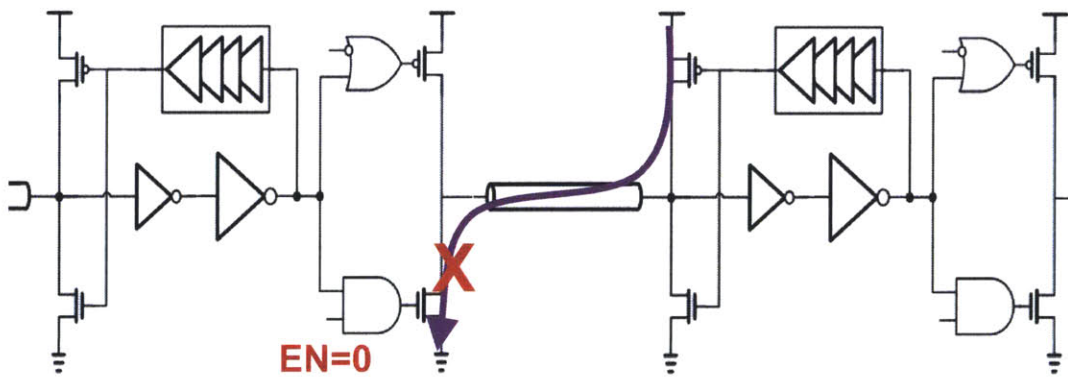
Finally, to cover on-die process variation, a delay cell has been added to the feedback path. This increased delay causes bigger overshoots, thereby resulting in a larger noise margin at the first inverter. It is noted that the increased noise margin comes at the cost of energy overhead but it is smaller than the energy overhead of over-sized drivers. This is because the delay cell does not affect the locked voltage level, leaving static voltage swing unchanged.

The operation of the completed VLR design is shown in Figure 3-15. It is essentially the same as that of the previous circuit. The only difference is the static DC path which can be disconnected by two activation gates according to switch allocation





(a)



(b)

Figure 3-15: Modified VLR operation: (a) logic High (b) logic Low.

results. For an application to the packet-switched on-chip networks, it is assumed that the EN signal comes from a switch allocator and such allocations will be done before flits start to travel by a multiple-hop buffer bypassing flow control.

### 3.4 Test Chip Fabrication

To prove the concept of SRLR and VLR in an actual silicon environment, a test chip has been fabricated in 45nm SOI CMOS process. Other kinds of asynchronous on-chip interconnects such as full-swing repeaters (FSRs), a comparator-based low-swing repeater (LSR) [11], and an equalized interconnect [17] are also implemented under the identical conditions for a fair comparison. Figure 3-16 shows the test chip layout.

Many repeater optimizations have been investigated [46] [47] [48] [49], but they are hardly practical in packet-switched on-chip networks. This is because the router-to-router distance is determined not by such data path optimizations but by the size of computation cores and caches in many-core processors. This work assumes that the router-to-router distance is 1mm and all repeater systems only employ built-in shared wires and multiplexers provided by the packet-switched on-chip networks. In other words, resource-based optimization has been applied to all the repeaters. A repeated-by-500um full-swing repeater (FSR) is included in the same die for a higher data rate comparison between full-swing and low-swing signaling.

Figure 3-17 shows a single-ended signaling wire channel used in SRLR, VLR, and two FSRs. The comparator-based LSR and the equalized interconnect employ differential signaling as shown in Figure 2-9. To explore the performance of the on-chip interconnects on a highly-resistive wire channel, the values of 0.14um and 0.16um are selected as wire width and space, respectively. To the best of my knowledge, this channel modeling has the highest coupling capacitance and wire resistance among silicon-proven on-chip interconnect studies [17] [18] [19].

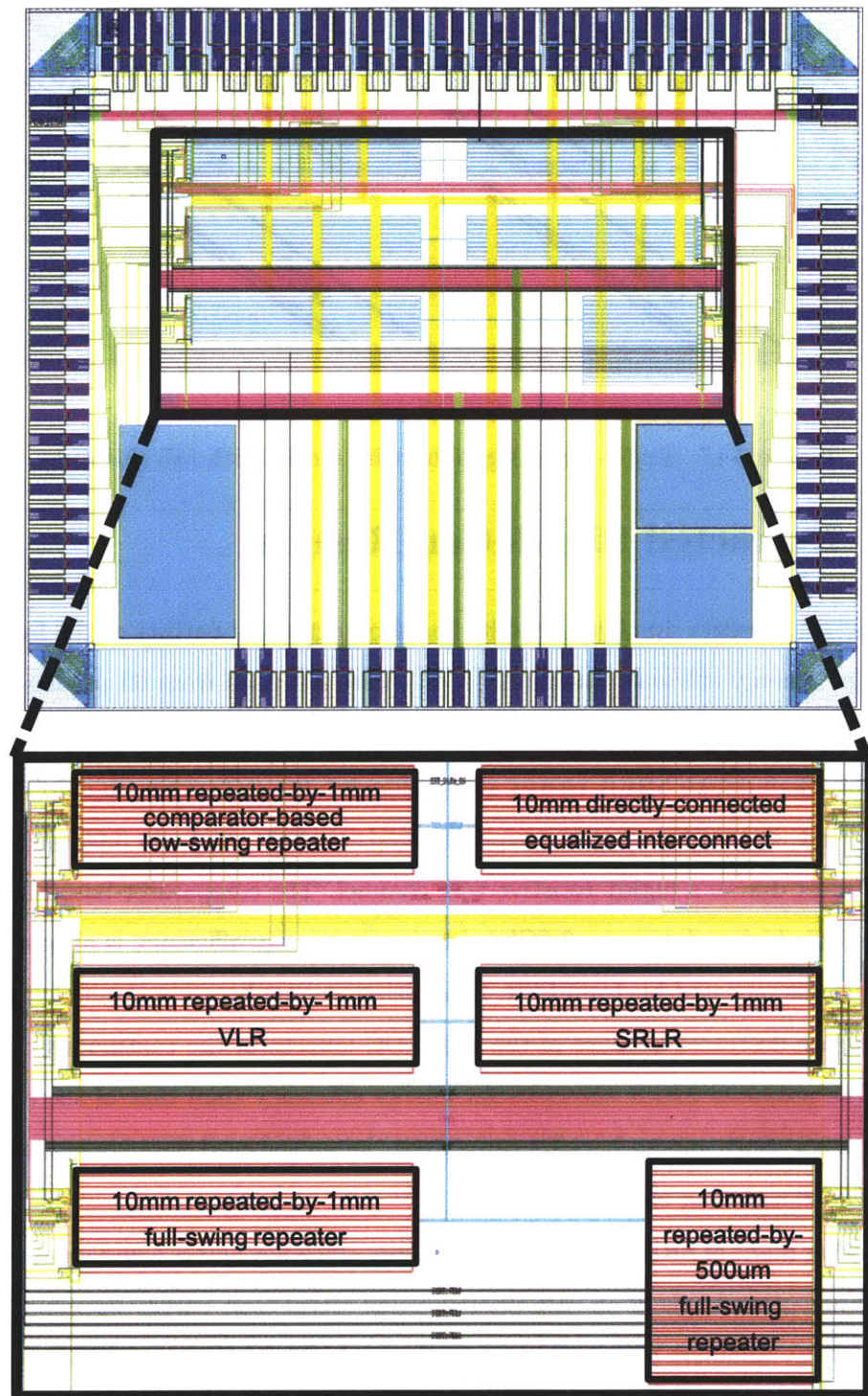


Figure 3-16: Test chip layout: various asynchronous on-chip interconnects.

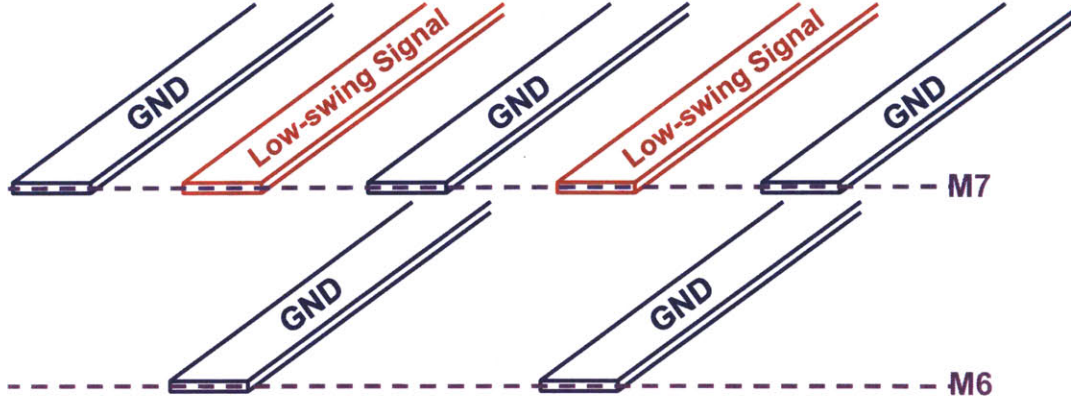


Figure 3-17: Single-ended signaling wire model with full shielding.

### 3.5 Extracted Simulation Results

This section presents simulation results based on extracted netlists of the fabricated test chip. Energy efficiency numbers (J/bit) were obtained from 60 pseudo-random inputs and normalized by data rate.

Figure 3-19 shows SRLR energy simulation results with other on-chip interconnects. First, it is noted that the repeated-by-1mm FSR supports 3.6Gb/s/ch data rate at most due to a highly-resistive wire channel. The comparator-based LSR provides a little higher data rate, 3.8Gb/s/ch, but its energy efficiency is almost similar to that of the repeated-by-1mm FSR. In addition, energy becomes significantly larger as data rate decreases because of the static DC current of the continuous comparators. The proposed SRLR shows better energy efficiency than that of the equalized interconnect. Such energy benefits come from the pulse-based data transmission and single-ended signaling of SRLR.

Another merit of SRLR is its constant energy efficiency across a wide range of data rate. While other kinds of on-chip interconnects reveal their energy optimum points, SRLR energy numbers essentially stay unchanged. This is because SRLR consumes energy only during its pulse transmission and the pulse width remains constant even when data rate varies. As shown in Figure 3-18, both SRLR pulses generated at 5GHz and 1GHz clock frequency have the same width of 75ps, which is determined



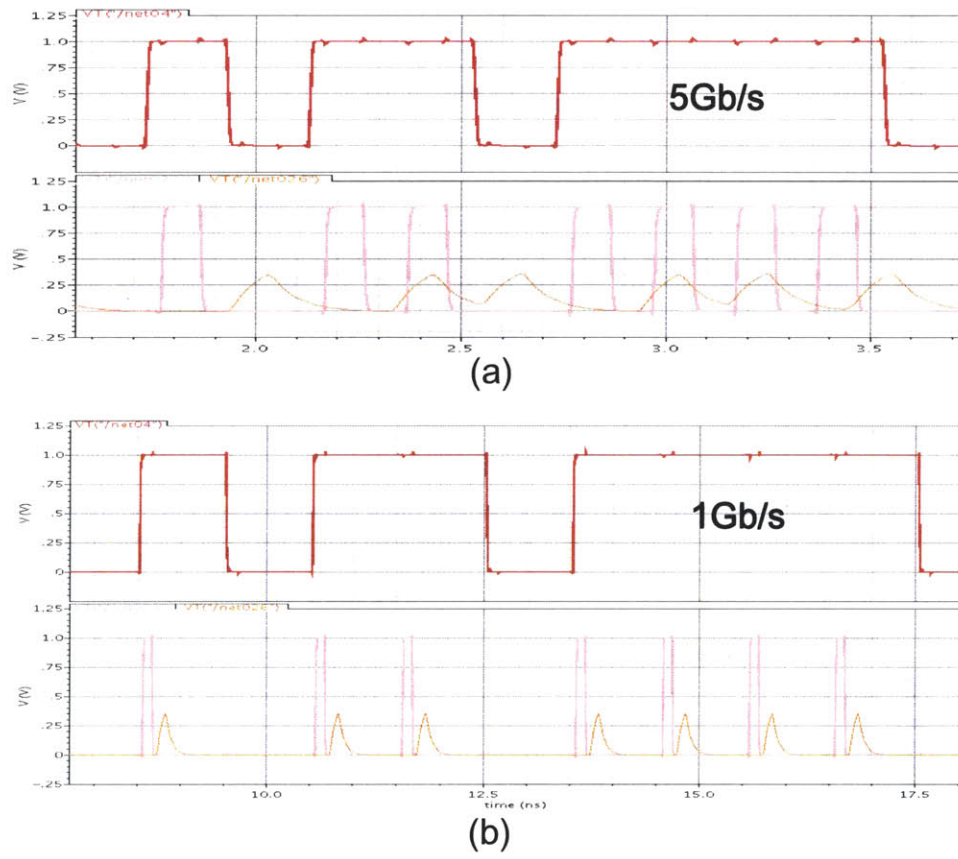


Figure 3-18: SRLR extracted simulation waveforms: (a) 5Gb/s/ch (b) 1Gb/s/ch.

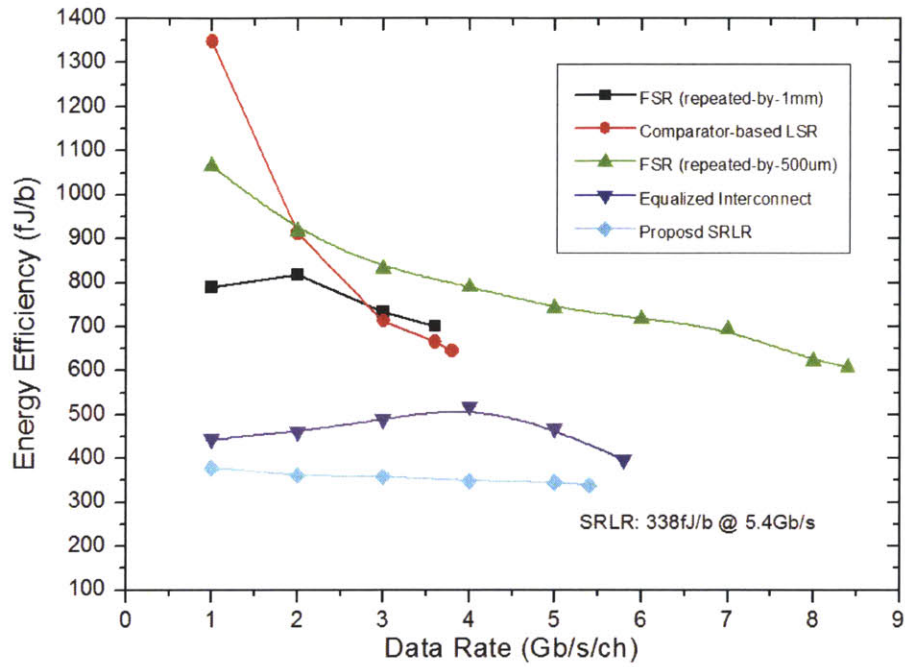


Figure 3-19: 10mm SRLR Energy simulation results at low data rate.

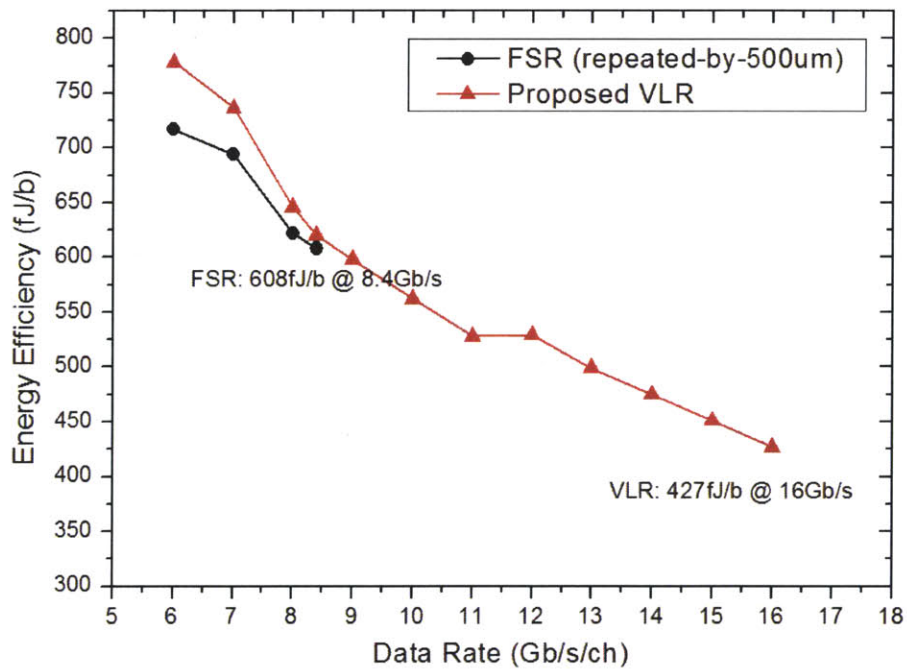


Figure 3-20: 10mm VLR Energy simulation results at high data rate.

by a pulse generator regardless of data rate. It is also noted that even equalized interconnects show a non-monotonic energy efficiency curve with data rate variation since their pre-emphasis TX and clocked RX have different energy optimum points.

Extracted simulation results on the 10mm VLRs are shown in Figure 3-20. It is noteworthy that VLR supports up to 16Gb/s/ch data rate on a highly-resistive wire channel while the repeated-by-500um FSR provides 8.4Gb/s/ch at most. Even considering nonidealities such as clock jitter, repeater jitter, and power supply fluctuation, VLR will be able to support 10Gb/s/ch Double Date Rate (DDR) with 5GHz clock by a wide margin. It is also an interesting observation for VLR to obtain the highest energy efficiency at its maximum data rate. As discussed in Section 3.3, most of VLR energy is consumed by the on-state DC path consisting of a highly-resistive wire channel, a smaller keeper transistor, and a driver transistor, resulting in relatively constant energy consumption. Consequently, the VLR energy numbers normalized by data rate become smaller as data rate increases.

Other link metrics such as channel characteristic, BER-based functionality, and eye sensitivity will be measured from the fabricated test chip due to accuracy and simulation time issues.





# Chapter 4

## Conclusion

### 4.1 Thesis Summary

This thesis concentrates on bridging the gap between architectural requirements of packet-switched on-chip networks and low-swing signaling circuit design. To be specific, this work explores low-swing signaling circuits for broadcast-optimized on-chip networks and link designs which enable energy and delay performance of a point-to-point interconnect of variable lengths.

In Chapter 1, the packet-switched on-chip networks are briefly introduced and their power consumption is analyzed to justify low-swing application to those on-chip networks. Also, existing low-swing signaling circuits are classified into four types: linear-mode RSDs, charging sharing-based RSDs, cut-off RSDs, and channel attenuation-based RSDs.

Chapter 2 presents broadcast-optimized, low-swing signaling-based on-chip networks to efficiently support advanced cache coherence protocols. First, architectural design of the proposed networks is introduced to closely explore their requirements. In response to such requirements, tri-state RSD-based crossbar switch is proposed and thoroughly investigated. The low-swing crossbar switch features (1) an inherent power gating circuit, (2) higher bandwidth driven by linear-mode RSD, (3) a longer low-swing signaling data path from crossbar's cross points to link RXs, (4) SOI-friendly circuit design, and (5) potential to be integrated into the CAD synthesis

flow. Fabricated in 45nm SOI CMOS process, the RSD-based crossbar switch enables 55% power reduction as compared to a full-swing data path. Moreover, the proposed crossbar allows the broadcast-optimized on-chip networks using a single pipeline stage for physical data transmission to operate at 21% higher data rate, when compared with equivalent full-swing networks.

In Chapter 3, an ideal communication fabric is first discussed and link design requirements of such an interconnection fabric are explored. Based on those requirements, two clockless low-swing repeater systems, SRLR and VLR, are proposed and analyzed. This thesis devotes significant emphasis on achieving process variation robust designs of SRLR and VLR. The pulse-based SRLR features very high energy efficiency while the over-driven VLR shows higher data rate than SRLR. Simulated in 45nm SOI CMOS, the 10mm SRLR consumes 338fJ/b at 5.4Gb/s/ch, whereas the 10mm VLR raises its data rate up to 16.0Gb/s/ch with 427fJ/b.

## 4.2 Future Work

The tri-state RSD-based crossbar switch relies on the second supply voltage to maintain pullup and pulldown transistors in linear-mode. Accordingly, an on-chip DC-DC converter capable of efficiently delivering power at the second supply voltage level will be required to enable the proposed circuit to be a more integrated subsystem. Also, since such on-chip power conversion provides digitally-controlled adaptive voltage swing ability, it can be applied to process variation calibration circuit design.

Process variation aware link design is also an essential research topic. Even though SRLR and VRL employ some ideas to compensate the variation effect, they are not able to fully cover 3-sigma on-die and die-to-die process variation due to their single-ended characteristic. To maintain benefits of the single-ended characteristic such as higher energy efficiency or wire density, an on-die variation calibration technique will be needed.

The best application of the clockless low-swing repeaters is reconfigurable on-chip networks. In those kinds of on-chip networks, network connectivity is changed

according to the system requirements. Accordingly, links in the reconfigurable on-chip networks should support asynchronous data transmission so that SRLR and VLR can be directly applied to such networks. In order to reduce architectural design complexity, the clockless low-swing repeaters will have fixed latency with PVT variation. It will be another interesting research topic.



# Bibliography

- [1] Natalie E. Jerger and Li-Shiuan Peh. Synthesis Lectures on Computer Architecture: On-Chip Networks. *Morgan & Claypool Publishers*, 2009.
- [2] Josep Torrellas. How to build a useful thousand-core manycore system? *IEEE International Symposium on Parallel & Distributed Processing*, pages 1–4, 2009.
- [3] M. B. Taylor et al. The raw microprocessor: A computational fabric for software circuits and general-purpose programs. *IEEE/ACM International Symposium on Microarchitecture (MICRO)*, 22(2):25–35, 2002.
- [4] P. Gratz et al. On-chip interconnection networks of the TRIPS chip. *IEEE/ACM International Symposium on Microarchitecture (MICRO)*, 27(5):41–50, 2007.
- [5] Hoskote et al. A 5-GHz mesh interconnect for a teraflops processor. *IEEE/ACM International Symposium on Microarchitecture (MICRO)*, 27(5):51–61, 2007.
- [6] Shane Bell, Bruce Edwards, et al. TILE64™ Processor: A 64-core SoC with Mesh Interconnect. *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2007 IEEE International*, 67(68):88–89, 2008.
- [7] Jason Howard, Saurabh Dighe, et al. A 48-Cire IA-32 Message-Passing Processor with DVFS in 45nm CMOS. *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2010 IEEE International*, pages 108–109, 2010.
- [8] H. Zhang, V. George, and Jan M. Rabaey. Low-Swing On-Chip Signaling Techniques: Effectiveness and Robustness. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 8:264–272, 2010.
- [9] Jan M. Rabaey, Anantha P. Chandrakasan, and Borivoje Nikolic. Digital Integrated Circuits: A design perspective. *Prentice Hall, 2nd Edition*, 1998.
- [10] H. Kojima et al. Half-swing clocking scheme for 75-percent power saving in clocking circuitry. *IEEE Journal of Solid-State Circuits (JSSC)*, pages 432–435, April 1995.
- [11] Ron Ho. On-Chip Wires: Scaling and Efficiency. *PhD thesis, Stanford University*, August 2003.

- [12] E.D. Kyriakis-Bitzaros. Design of low power CMOS drivers based on charge recycling. *IEEE International Symposium on Circuits and Systems*, pages 1924–1927, June 1997.
- [13] M. Hiraki et al. Data-Dependent Logic Swing Internal Bus Architecture for Ultralow-Power LSIs. *IEEE Journal of Solid-State Circuits (JSSC)*, pages 397–402, April 1995.
- [14] H. Yamauchi et al. An Asymptotically Zero Power Charge-Recycling Bus Architecture for Battery-Operated Ultrahigh Data Rate ULSIs. *IEEE Journal of Solid-State Circuits (JSSC)*, pages 423–431, April 1995.
- [15] R. Golshan et al. A novel reduced swing CMOS BUS interface circuit for high speed low power VLSI systems. *IEEE International Symposium on Circuits and Systems*, pages 351–354, May 1994.
- [16] B.-D. Yang et al. High-Speed and Low-Swing On-Chip Bus Interface Using Threshold Voltage Swing Driver and Dual Sense Amplifier Receiver. *European Solid-State Circuit Conference*, pages 144–147, September 2000.
- [17] R. Ho, I. Ono, F. Liu, et al. High-Speed and Low-Energy Capacitive-Driven On-Chip Wires. *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2007 IEEE International*, pages 412–413, 2007.
- [18] E. Mensink, D. Schinkel, Klumperinck, et al. A 0.28pJ/b 2Gb/s/ch Transceiver in 90nm CMOS for 10mm On-chip interconnects. *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2007 IEEE International*, pages 314–315, 2000.
- [19] B. Kim and V. Stojanovic. A 4Gb/s/ch 356fJ/b 10mm equalized on-chip interconnect with nonlinear charge-injecting transmit filter and transimpedance receiver in 90nm CMOS. *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2009 IEEE International*, pages 66–67, 2009.
- [20] N. Agarwal, L.-S. Peh, and N. K. Jha. In-network coherence filtering: snoopy coherence without broadcast. *IEEE/ACM International Symposium on Microarchitecture (MICRO)*, 42:232–243, 2009.
- [21] E. E. Bilir, R. M. Dickson, Y. Hu, M. Plakal, D. J. Sorin, M. D. Hill, and D. A. Wood. Multicast snooping: A coherence method using a multicast address network. *IEEE/ACM International Symposium on Computer Architecture (ISCA)*, pages 294–304, 1999.
- [22] P. Conway, N. Kalyanasundharam, G. Donley, K. Lepak, and B. Hughes. Cache hierarchy and memory subsystem of the AMD Opteron processor. *IEEE/ACM International Symposium on Microarchitecture (MICRO)*, 30:16–29, 2010.

- [23] M. M. K. Martin, P. J. Harper, D. J. Sorin, M. D. Hill, and D. A. Wood. Using destination-set prediction to improve the latency/bandwidth tradeoff in shared memory multiprocessors. *IEEE/ACM International Symposium on Computer Architecture (ISCA)*, pages 206–217, 2003.
- [24] M. M. K. Martin, D. J. Sorin, M. D. Hill, and D. A. Wood. Bandwidth adaptive snooping. *International Symposium on High-Performance Computer Architecture (HPCA)*, pages 251–262, 2002.
- [25] A. Raghavan, C. Blundell, and M. M. K. Martin. Token tenure: Patching token counting using directory-based cache coherence. *IEEE/ACM International Symposium on Microarchitecture (MICRO)*, pages 47–58, 2008.
- [26] B. Grot, J. Hestness, S. W. Keckler, and O. Mutlu. Express cube topologies for on-chip interconnects. *International Symposium on High-Performance Computer Architecture (HPCA)*, 2009.
- [27] A. Kumar, L.-S. Peh, P. Kundu, and N. K. Jha. Express virtual channels: Towards the ideal interconnection fabric. *IEEE/ACM International Symposium on Computer Architecture (ISCA)*, 2007.
- [28] A. Kumar, L.-S. Peh, and N. K. Jha. Token flow control. *IEEE/ACM International Symposium on Microarchitecture (MICRO)*, 2008.
- [29] N. Enright Jerger, L.-S. Peh, and M. Lipasti. Virtual circuit tree multicasting: A case for on-chip hardware multicast support. *IEEE/ACM International Symposium on Computer Architecture (ISCA)*, 2008.
- [30] P. A. Fidalgo, V. Puente, and J.-A. Gregorio. MRR: Enabling fully adaptive multicast routing for CMP interconnection networks. *International Symposium on High-Performance Computer Architecture (HPCA)*, pages 355–366, 2009.
- [31] F. A. Samman, T. Hollstein, and M. Glesner. Multicast parallel pipeline router architecture for network-on-chip. *Design, Automation and Test in Europe Conference and Exhibition (DATE)*, pages 1396–1401, April 2008.
- [32] L. Wang, Y. Jin, H. Kim, and E. J. Kim. Recursive partitioning multicast: A bandwidth-efficient routing for networks-on-chip. *ACM/IEEE International Symposium on Networks-on-Chip (NOCS)*, pages 64–73, May 2009.
- [33] Kerry Bernstein and Norman J. Rohrer. SOI Circuit Design Concepts. *IBM Microelectronics, Kluwer Academic Publishers*, 2001.
- [34] Byungsub Kim. Equalized On-Chip Interconnect: Modeling, Analysis, and Design. *PhD thesis, Massachusetts Institute of Technology*, February 2010.
- [35] N. Verma and A. P. Chandrakasan. A High-Density 45nm SRAM Using Small-Signal Non-Strobed Regenerative Sensing. *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2008 IEEE International*, pages 380–381, 2008.

- [36] I. Arsovski and R. Wistort. Self-referenced Sense Amplifier for Across-chipvariation Immune Sensing in High-performance Content-Addressable Memories. *IEEE Custom Integrated Circuits Conference (CICC)*, pages 453–456, September 2006.
- [37] M. Qazi, K. Stawiasz, L. Chang, and A. P. Chandrakasan. A 512kb 8T SRAM Macro Operating Down to 0.57V with An AC-Coupled Sense Amplifier and Embedded Data-Retention-Voltage Sensor in 45nm SOI CMOS. *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2010 IEEE International*, pages 350–351, 2010.
- [38] T. Krishna, J. Postman, C. Edmonds, L.-S. Peh, and P. Chiang. SWIFT: A SWing-reduced Interconnect For a Token-based Network-on-Chip in 90nm CMOS. *IEEE International Conference on Computer Design (ICCD)*, pages 439–446, 2010.
- [39] S. Heo and K. Asanovic. Replacing global wires with an on-chip network: A power analysis. *IEEE International Symposium on Low Power Electronics and Design (ISLPED)*, pages 369–374, 2005.
- [40] R. Heald and J. Holst. A 6 ns cycle, 256-Kb cache memory and memory management unit. *IEEE Journal of Solid-State Circuits (JSSC)*, 28(11):1078–1083, November 1993.
- [41] T. Chapell, B. Chapell, and S. Schuster et al. A 2 ns cycle, 3.8 ns access 512-Kb CMOS ECL SRAM with a fully pipelined architecture. *IEEE Journal of Solid-State Circuits (JSSC)*, 26(11):1577–1585, November 1991.
- [42] R. Heald and K. Shin et al. 64-KByte sum-addressed-memory cache with 1.6-ns cycle and 2.6-ns latency. *IEEE Journal of Solid-State Circuits (JSSC)*, 33(11):1682–1689, November 1998.
- [43] W. Hwang, R. V. Joshi, and W. H. Henkels. A 500-MHz, 32-word 64-bit, eight-port self-resetting CMOS register file. *IEEE Journal of Solid-State Circuits (JSSC)*, 34(1):56–67, January 1999.
- [44] G. Paci, D. Bertozzi, and L. Benini. Effectiveness of adaptive supply voltage and body bias as post-silicon variability compensation techniques for full-swing and low-swing on-chip communication channels. *Design, Automation and Test in Europe Conference and Exhibition (DATE)*, pages 1404–1409, April 2009.
- [45] H.J. Oguey and D. Aebischer. CMOS current reference without resistance. *IEEE Journal of Solid-State Circuits (JSSC)*, 32(7):1132–1135, 1997.
- [46] C. Y. Wu and M. Shiau. Accurate speed improvement techniques for RC line and tree interconnections in CMOS VLSI. *IEEE International Symposium on Circuits and Systems*, pages 1648–1651, May 1990.



- [47] M. Nekili and Y. Savaria. Optimal methods of driving interconnections in VLSI circuits. *IEEE International Symposium on Circuits and Systems*, pages 21–23, May 1992.
- [48] C. Y. Wu and M. Shiau. Delay models and speed improvement techniques for RC tree interconnections among small-geometry CMOS inverters. *IEEE Journal of Solid-State Circuits (JSSC)*, 25:1247–1256, October 1990.
- [49] S. Dhar and M. A. Franklin. Optimum buffer circuits for driving long uniform lines. *IEEE Journal of Solid-State Circuits (JSSC)*, 26:32–40, January 1991.