DesignCon 2015

High-speed Low-power On-chip Global Signaling Design Overview

Xi Chen, NVIDIA Corporation
xich@nvidia.com

John Wilson, NVIDIA Corporation
johnwilson@nvidia.com

John Poulton, NVIDIA Corporation
jpoulton@nvidia.com

Rizwan Bashirullah, University of Florida
rizwan@tec.ufl.edu

Tom Gray, NVIDIA Corporation
tgray@nvidia.com

This research was, in part, funded by the U.S. Government. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Government.

Abstract

On-chip global signaling in modern SoCs faces significant challenges due to wire pitch scaling and increasing die size. Conventional on-chip synchronous CMOS links have already hit a performance wall in power and latency. Although approaches based on custom low-swing equalized serial-link techniques can yield improvements, strict power/silicon budgets and non-ideal in-situ conditions of large SoCs make their design much more challenging than simply transitioning off-chip signaling technologies to on-chip. Therefore, a holistic approach to the on-chip global signaling problem is required. We present analyses and solutions that take into account channel design, low power transceiver circuits, clocking architectures, and power supply considerations.

Authors Biography

Xi Chen is currently a Research Scientist at NVIDIA, Santa Clara, CA. His research interests include high-speed signaling and clocking, 3D IC design methods, and analog mixed-signal ICs. He received the B.S. and M.S. degrees in electrical engineering from Zhejiang University, Hangzhou, China, in 2003 and 2006, respectively, and the Ph.D. degree in electrical engineering from North Carolina State University, in 2011.

John Wilson is a member of the Circuits Research Group at NVIDIA in Durham, NC. His research interests include: low-power circuit design, signal integrity, chip package co-design, and 3D-ICs. In 2003 he received the Ph.D. in Electrical Engineering from North Carolina State University. From 2003 to 2006 he was a Research Professor leading projects that in the areas of advanced packaging, low-power capacitive & inductive coupled transceivers for 3D-ICs, and low-power on-chip global signaling. From 2006 to 2012, while with Rambus in Chapel Hill, NC, he worked on low-power, high-speed I/O circuit & signaling techniques, and methods to mitigate signal integrity problems in memory interfaces that use single-ended signaling. He has authored and co-authored approximately 50 publications, and has over 10 awarded patents and filed patent applications.

John W. Poulton is currently a Senior Distinguished Scientist at NVIDIA, Durham, NC, where he is working on low-energy on- and off-chip signaling. He received the B.S. degree from Virginia Polytechnic Institute and State University, Blacksburg, VA, USA, in 1967, the M.S. degree from the State University of New York, Stony Brook, NY, USA, in 1969, and the Ph.D. degree from the University of North Carolina, Chapel Hill (UNCCH), NC, USA, in 1980, all in physics. From 1981 to 1999, he was a researcher with the Department of Computer Science, UNCCH, where from 1995 he held the rank of Research Professor. From 1999 to 2003 he was Chief Engineer with Velio Communications, where he developed multi-gigabit chip-to-chip signaling systems. From 2003 to 2009 he was a Technical Director with Rambus, Inc., Chapel Hill, where he led an effort to build power-efficient multi-gigabit I/O systems, demonstrating a system in 2006 with the lowest energy per bit published up to that time. He is a Fellow of the IEEE.

Dr. Rizwan Bashirullah received the B.S. degree in Electrical Engineering from the University of Central Florida and the M.S. and Ph.D. degrees in Electrical Engineering from North Carolina State University (NCSU), in 1999 and 2004, respectively. He joined the Department of Electrical and Computer Engineering at the University of Florida in 2004, where he is currently a tenured Associate Professor. His research interests are in mixed-signal circuits for biomedical applications, power delivery systems and on-chip/off-chip signaling subsystems. He has authored or coauthored about 100 referred technical abstracts and papers in conferences and journals. Dr. Bashirullah received the 2005 National Science Foundation (NSF) Early Career Development Award, the 2010 University of Florida Inventor Recognition Award and the 2011 Microwave Magazine Best Paper Award of the IEEE Microwave Theory and Techniques Society (MTT-S).

C. Thomas Gray joined the Circuits Research group of NVIDIA, Durham, NC, in 2011, where is currently Director of Circuit Research and leads activities related to high speed signaling, low energy memories, and variation tolerant clocking and power delivery. Prior to NVIDIA, he worked on various transceiver design projects, high speed memory links, and high speed serial links for applications such as Ethernet, Fibre Channel, Infiniband, OIF, and PCI Express as a system architect at Nethra Imaging, ARM, Cadence, and IBM. He received the B.S. degree in computer science and mathematics from Mississippi College, Clinton, MS, USA, in 1988, and the M.S. and Ph.D. degrees in computer engineering from North Carolina State University, Raleigh, NC, USA, in 1990 and 1993, respectively.

1. Introduction

In modern SoC design, challenges for on-chip global signaling are pressured by two factors. First, increasing die size produces longer communication distances. Second, routing wire pitch, which reduces with smaller process nodes, makes the resistance of a given wire length scale up rapidly. To deal with these challenges, automatic P&R methods have to insert numerous repeaters and re-timers for global signaling. While compute cost decreases with smaller process nodes, power and silicon area used for on-chip global communication tends to increase with each product generation. Also, highly resistive global wires with a large number of buffer stages have significant insertion delay, and create difficulties for architecture and design.

Bundled-data wiring channels, sometimes called "fabrics", are often used where high-bandwidth, long-distance data movement is required. Custom designed repeater placement and fully reserved metal routing resource can minimize the frequency of buffering and re-clocking, and push the performance of CMOS signaling up to its limit. However, the fundamental physical problem that prevents further energy reduction in fabrics is the CV^2F energy needed to charge and discharge the wire capacitance. Moreover, fully utilized routing channels for global signaling are consuming resource from other functional circuits, and this cost is also scaling up rapidly with every new process node. Conventional synchronous CMOS signaling, even with careful custom design, faces significant challenges in supporting the performance requirements of products in the future.

Using high-speed serial link techniques for on-chip global signaling shows promise in solving some of the problems mentioned above. One big difference between on-chip and off-chip signaling design is the channel. On-chip metallization generally has much higher sheet resistance compared to package or PCB traces because of metal thickness. Therefore, on-chip high-speed link design has to explore the best bandwidth potential from the available on-chip metallization, model and design the high-speed channels with existing power delivery networks to optimize the performance. One approach to solving the signaling energy problem is simply to drive long signal wires with very low voltage swings. Since the energy required to drive a line scales with V^2, this approach can, in principle, produce large energy savings. The challenge to implementing this low-swing signaling idea on-chip is how to generate the accurate and stable voltage values in the environment which has very limited voltage headroom and noisy power rails (no I/O VDD or analog VDD is available).

To increase the signaling rate, many have adopted equalization to overcome the basic RC delay of a wire. Using equalization not only increases the effective bandwidth of the wire but also greatly reduces delay. However, with very limited voltage headroom and energy budgets, equalization techniques for on-chip signaling have to be as simple as possible, and robust enough to operate reliably in the very non-ideal on-chip environment. Overall, compared to traditional off-chip serial links, there are many new challenges to be considered for on-chip global signaling application.

2. On-chip Signaling Channel

Channel design and optimization based on available metal resources is essential for on-chip signaling. Typically, signaling over the thicker metal of the upper layers enables longer distances between repeaters (i.e. fewer repeaters for a fixed distance), which reduces the energy and delay per-unit length. To provide high cross sectional bandwidth, we prefer to build wiring channels in the thickest metal layers on-chip which are available for orthogonal routing, for instance, the two metal layers just below the RDL (redistribution layer).

One consideration for practical channel design is that the signal channels have to be routed within the gaps of existing power delivery network. Power integrity is already one of the most challenging tasks in modern SoC design. So, normally a large portion of the high-level metal layers are occupied by regular and dense power/ground (P/G) conductors. Therefore, the high-speed channels must be interleaved between these conductors. Fortunately, this is not totally bad for signaling channel design, because we can use these P/G conductors to act as shields to reduce cross-talk. The P/G shielding also provides current return paths for signals and enables low-swing single-ended signaling to be used for long distance on-chip communication. Although more vulnerable to cross-talk, single-ended signaling has the potential of working at lower supply voltage and consuming less power compared to differential signaling. Figure 1 shows the illustration of the high-speed signaling channels built between P/G grids.

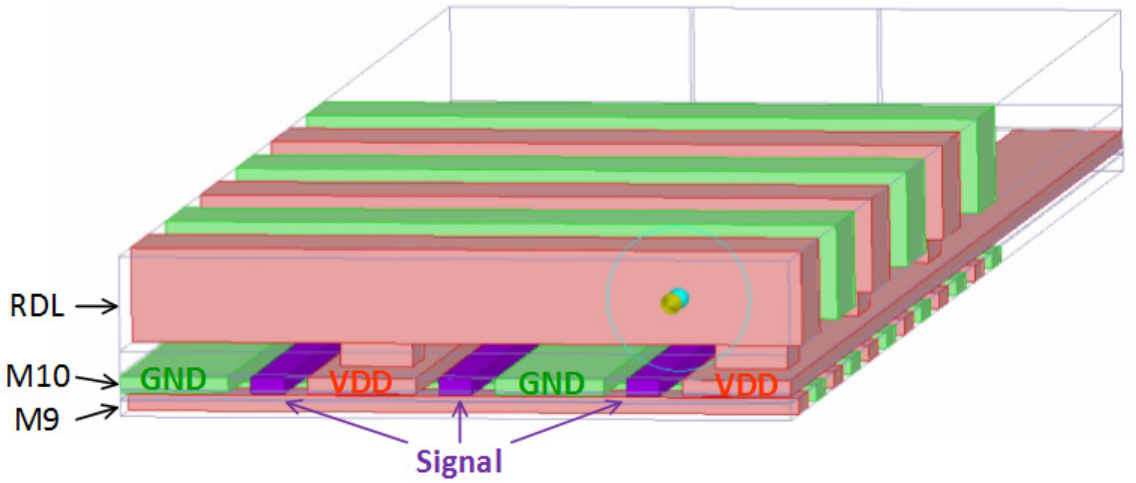


Figure 1: High-speed channels interleaved between power/ground conductors

For high-speed on-chip signaling, we have designed the wiring channel to operate at the highest possible data rate to achieve the maximum cross sectional bandwidth ultimately limited by transistor circuit performance (e.g. ~16Gbps @ 28nm node over corners). For best energy efficiency, we design the channel according to this speed requirement and try to reach the maximum length before amplification is needed. Because the on-chip

channel bandwidth is almost always RC dominated, a thicker metal layer is normally preferred due to its lower resistance and foundries usually offer different high-level metal options. To make good judgment, it is worthwhile to study the channel performance for all available metal options.

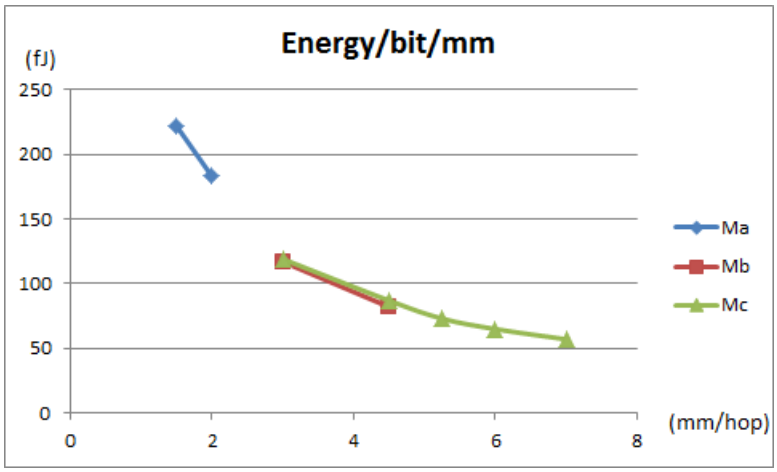
Comparisons of channel designs for three metal options are summarized in Table 1. Thicknesses are normalized to the metal option "Ma". For each metal option, the energy optimal cross sectional geometry for a single-ended channel was chosen while simultaneously considering power grid requirements. In each case, the power and ground grid conductors, which also act as shields, were made wide enough to accept a via from the RDL. As can be seen from the last column, the thicker metal will provide longer one-hop distance at full circuit speed. Here we assume that the signal voltage swing on the wire is about 100mV, and the power supply is 900mV.

Table 1: Comparison of foundry metal options and channel designs

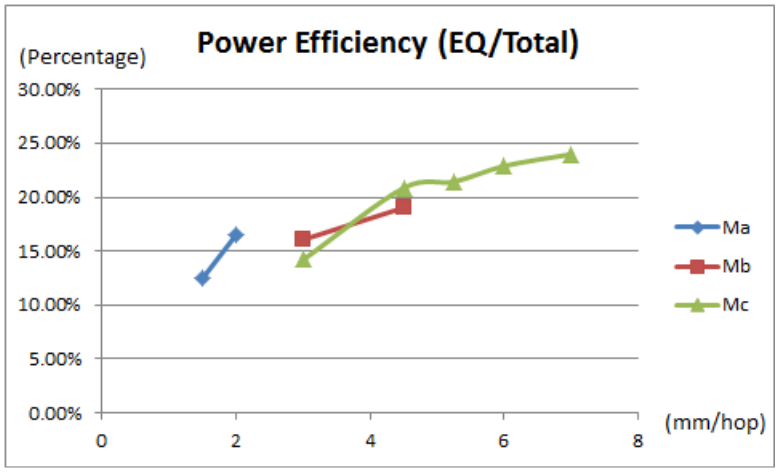
Metal Layer Options	Thickness (normalized)	Signal Width/Space	P/G Shield Width	Signal Pitch	Max Length @ 16Gbps
Ma	1x	0.5μm	3.0μm	4.5μm	2mm
Mb	1.7x	0.8μm	3.6μm	6.0μm	5mm
Mc	2.5x	1.2μm	3.6μm	7.2μm	6.5mm

To better understand the trade-offs of different metal options with respect to energy efficiency, an experimental circuit system was designed. The tested high-speed link has sixteen parallel data lanes and two source synchronous clock lanes, with six stages of repeaters and seven wire segments between the transmitter (with serializer) and the receiver (with deserializer). The whole circuit system was implemented in a 28nm process and operates at 16Gbps with 900mV typical supply. Figure 2 shows the equivalent energy per-bit per-mm length and power efficiency with different metal options. Here efficiency is defined as the ratio of the equalization power used to change the voltage polarity on the wires (theoretical minimum power) versus the total power.

In Figure 2(a), it is obvious that thicker metal can send signals over longer distances before hitting the bandwidth limit. Hence, the thicker metal option has more potential to achieve lower energy/bit/mm and also lower delay/mm. When equalization is used, actual energy/mm spent on the wire tends to increase as the wire length increases, but as long as the circuit power is still significant, the total energy/mm will benefit from longer wire length per hop. The efficiency numbers shown in Figure 2(b) are still have room for optimization, and can also be improved with the supply voltage scaling and transistor size shrinks in newer process node (ex. 16nm FinFET).



(a)



(b)

Figure 2: (a) Equivalent energy/bit/mm and (b) percentage of equalization power of an experimental circuit (16 data lanes + 2 clock lanes, seven hops) built on 28nm process with different metal options

3. Repeater and Clocking Architecture

For really long distance (>10mm) signaling at high-speed (>16Gbps), repeaters become necessary not only to re-drive the low-swing signals on the wires, but also to re-time the data and mitigate the effects of jitter accumulation. With intrinsic delay matching between clock and data lanes, source-synchronous clocking can provide much higher data rates compared to fully synchronous clocking. Because the latency and cost of deserialization and re-synchronization are too high to be used in every repeater, the clocking architecture needs to be carefully designed and evaluated to make sure all cascaded repeaters in a link can operate reliably at full rate.

Figure 3 shows the basic structure of a repeater for high-speed signaling. The amplifier is used to bring the low swing signal on the wire back to full swing and drive the following logic circuits in the repeater. Alternatively, the amplification could be achieved by using

the regeneration loop gain within a sampler, which reduces the voltage gain requirement of the linear amplifier. But this would decrease the benefits of clock and data path delay matching when there are environmental (e.g. voltage) variations and reduce the timing margin. For the best delay matching between data and clock lanes, linear amplifier is preferred to provide the full amplification for both clock and data lanes, at the cost of some constant DC current. The simplest sampler that can be built for full-swing high-speed signals consists of four latches driven by double-data-rate clocks, and performs as a 1-to-2 deserializer and a 2-to-1 serializer. Ideally, all jitter and offset accumulated in data lanes will be removed by the sampling clocks, if the clock signal quality is good enough. After the sampler, the full speed digital signal will once again drive both the pre-emphasis driver and DC driver which connect to the next channel segment, exactly the same as what happens in the very first transmitter stage.

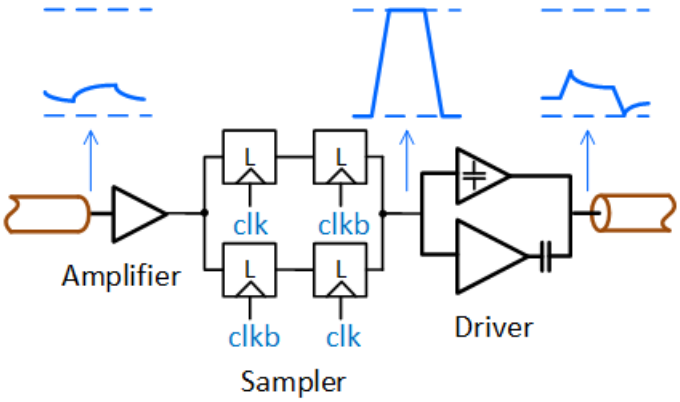


Figure 3: Circuit for a repeater in high-speed on-chip signaling

In every repeater stage, the transitions of data signals will be re-aligned to the sampling clock edges. Traditionally in a one-hop data link, the sampling clock at the receiver end is the quadrature clock (Q-clock) sent from the transmitter end. In a multi-hops source-synchronous data link, it is almost impossible to re-generate the quadrature clock in every repeater without significant power and delay overhead. The more practical method is shown in Figure 4. In this topology, both I- and Q-clocks are sent from the transmitter stage to the receivers. In each repeater, the Q-clock from previous stage samples the data and drives the local I-clock driver and then becomes the I-clock signal when arrives the next stage, while the I-clock from previous stage is simply reversed and drives the local Q-clock driver which provides the sampling clock for the next stage. This I/Q alternating clocking structure provide the timing margin for all repeaters cascaded through the data link, as long as the quadrature clocking quality is still reliable.

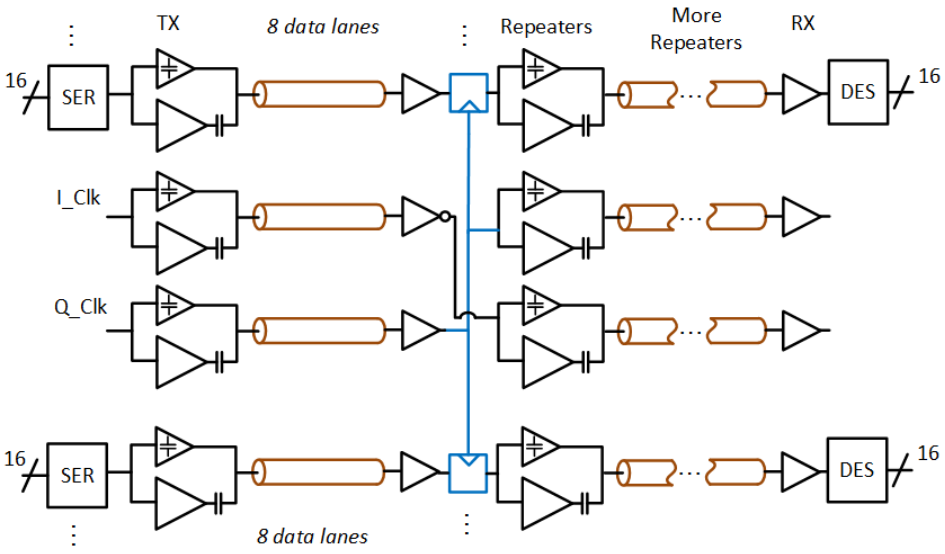


Figure 4: Architecture of a multi-hop serial link with I/Q clocks

It is worth noting that jitter in data lanes will be eliminated by the clock in every repeater, but variations in clock signal will accumulate through the link. The major cause of clock lane jitter is cross-talk from parallel data-lanes, and thicker metal layer tends to cause more significant cross-talk among wires. Figure 5 shows the eye diagrams for both the data signals and local Q-clock signal measured at the inputs of different hops. The channel used in this experiment is based on "Mb" listed in Table 1, and wire length is 4mm for each hop. These simulation results show that the jitter accumulation effects on clock signals are quite significant through multi-hops. And because the quadrature relationships of the I- and Q-clocks are un-bounded after the transmitter stage, the jitter on the I- and Q-clocks could be easily uncorrelated. Hence, the timing margin has to account for the variations in both clock lanes. For very thick metal layer like "Mc" in Table 1, the maximum wire length per hop may actually be limited by cross-talk instead of the single lane bandwidth.

Another possible clocking architecture only sends the I-clock through the link, as shown in Figure 6. In this architecture, the local clock buffer needs to generate the 0.5UI delay for correct data sampling. Since there is no longer a Q-clock, the sampler's timing margin only needs to be budgeted for jitter accumulation within one clock lane. The local clock buffer delay will vary over process-voltage-temperature (PVT), so it should be carefully designed to make sure the inserted delay always meets setup and hold time requirements at all conditions. Because the data are re-timed in every repeater, this local delay variation will not impact the timing of following stages. In real implementations, it may be worthwhile to distribute differential clocks (I and Ib) to minimize the offset accumulation, even the data lanes are single-ended.

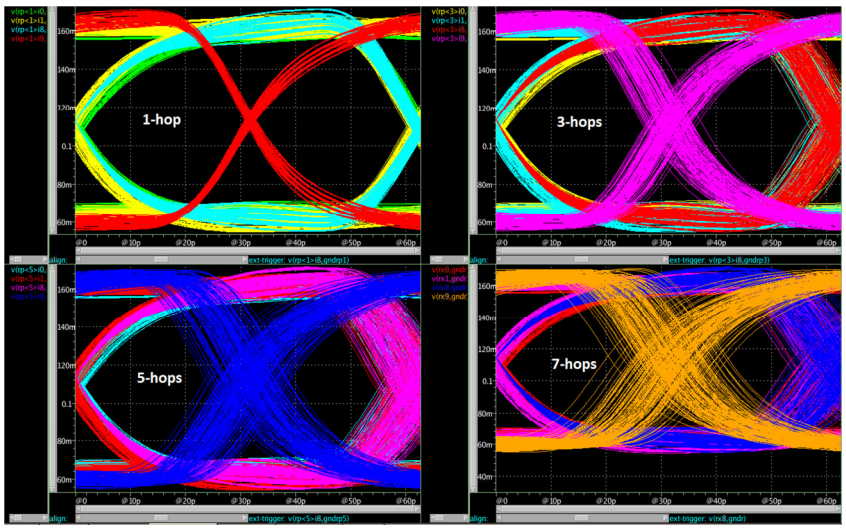


Figure 5: Eye diagrams of data and Q-clock signals at repeaters & receiver inputs (“Mb”, 4mm/hop)

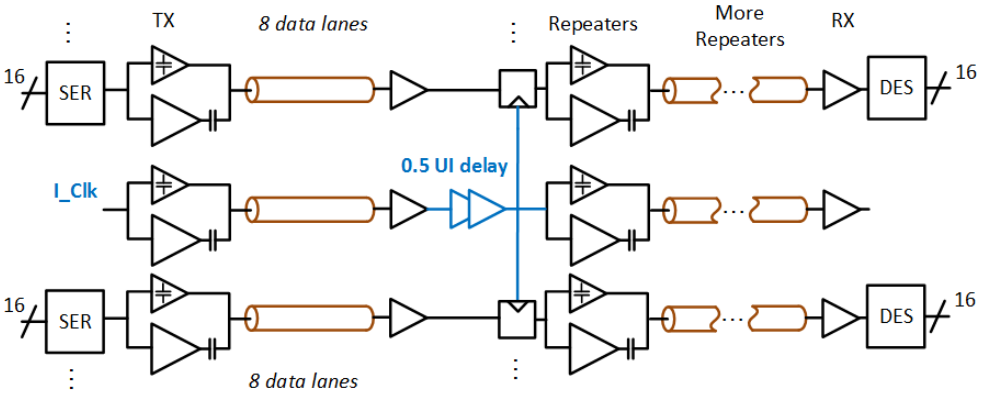


Figure 6: Architecture of a multi-hop serial link with single-ended I-clock only

4. Circuit Design Considerations

The biggest challenge for high-speed on-chip signaling circuit design is power efficiency. Low-power is critical for potential applications requiring thousands of data lanes and many TB/s of bandwidth, and this highlights a major distinction when comparing off-chip links to on-chip links. Off-chip links are evaluated in terms of pJ/bit and the most efficient links consume from 0.5-5 pJ/bit. However, on-chip links need to be implemented such that they consume only 10s of fJ/bit/mm. The distinction is that on-chip link energy is normalized to interconnect length because on-chip interconnect bandwidth reduces quadratically with length, where-as off-chip interconnect basically behaves like an LC transmission line, where bandwidth reduces roughly in proportion to the square-root of length. This means we need to use the simplest circuit structure to

generate the low-swing signal, minimize the number of transistors that are toggling, and design the circuit to operate reliably over anticipated variations.

Another difficulty for on-chip signaling circuit design is power supply noise. Unlike chip-to-chip signaling which usually has a dedicated I/O VDD, the on-chip signaling circuit will most likely share the power supply with large amount of noisy digital circuits. The noisy power supply not only degrades circuit performance, but also causes common-mode voltage mismatches among stages in the serial link and may introduce errors.

As mentioned in the previous section, the amplifier is an important circuit block in every repeater stage to recover the low-swing signals. Working at the highest possible data rate and with a very limited energy budget, amplifier performance is very sensitive to random process variation. The mismatches between amplifier and bias circuit (in single-ended signaling), or between differential pair (for differential signaling) create voltage offset, which will cause timing margin reduction in the following sampling stage and hence lower the data rate and the energy efficiency. Unfortunately, random process variation is an even more severe problem especially with shrinking feature size.

Many different methods have been explored to handle the circuit design challenges mentioned above. In this section, we discuss one signaling style which will be referred to as “pulse mode”. Basically, pulse mode signaling means only the transitions of the data will be sent along the wires and last about one UI, with a positive pulse for 0-to-1 transition and a negative pulse for 1-to-0. When a sequence of unchanging data bits is sent, the driver output will only send the first bit and then return to the common mode voltage until next transition.

The pulse mode driver schematic is shown in Figure 7(a). It is basically just two AC-coupling equalizers running in parallel with one of the inputs reversed and then followed by a 1-UI delayed. Equalization (EQ) is a very useful technique to increase the effective bandwidth of the wire and greatly reduces the RC delay. The most commonly used EQ for on-chip low-swing links is transmitter EQ, because it is simple and easy to “overdrive” the wire to the supply voltage. Difficulties of using transmitter EQ include how to match the DC and AC paths delays and also how to adaptively optimize the strength ratio of DC and AC drivers. Although the high resistance nature of the on-chip wires make the driver energy basically AC dominated, balancing DC and AC drivers still produces lots of overhead for local logic circuits and clock distribution. Pulse mode signaling totally removes the DC driver from the transmitter side, therefore avoiding these design difficulties and improving the energy efficiency (ex. compared to Figure 2(b)). Pulse mode signaling also avoids the common-mode variation problem caused by VDD mismatches between neighboring repeaters, therefore, providing better tolerance to the power supply noise.

Figure 7(b) shows the amplifier circuit for pulse mode signaling. Because of the DC-balanced nature of pulse mode signaling, it is possible to use a self-biased amplifier with appropriate compensation. The self-bias loop gain can attenuate the amplifier offset caused by random process variation to an acceptably low level, without the requirement

of extra tuning circuits and calibration time. Also, the self-biased amplifier tends to have better tolerance to VDD variation compared to replica-biased design. A latch stage is needed for the pulse mode amplifier to bring the signal back to CMOS mode before the sampler.

Figure 7(c) and (d) show the basic sampler circuit design and signal eye diagrams for the proposed single-ended pulse mode signaling.

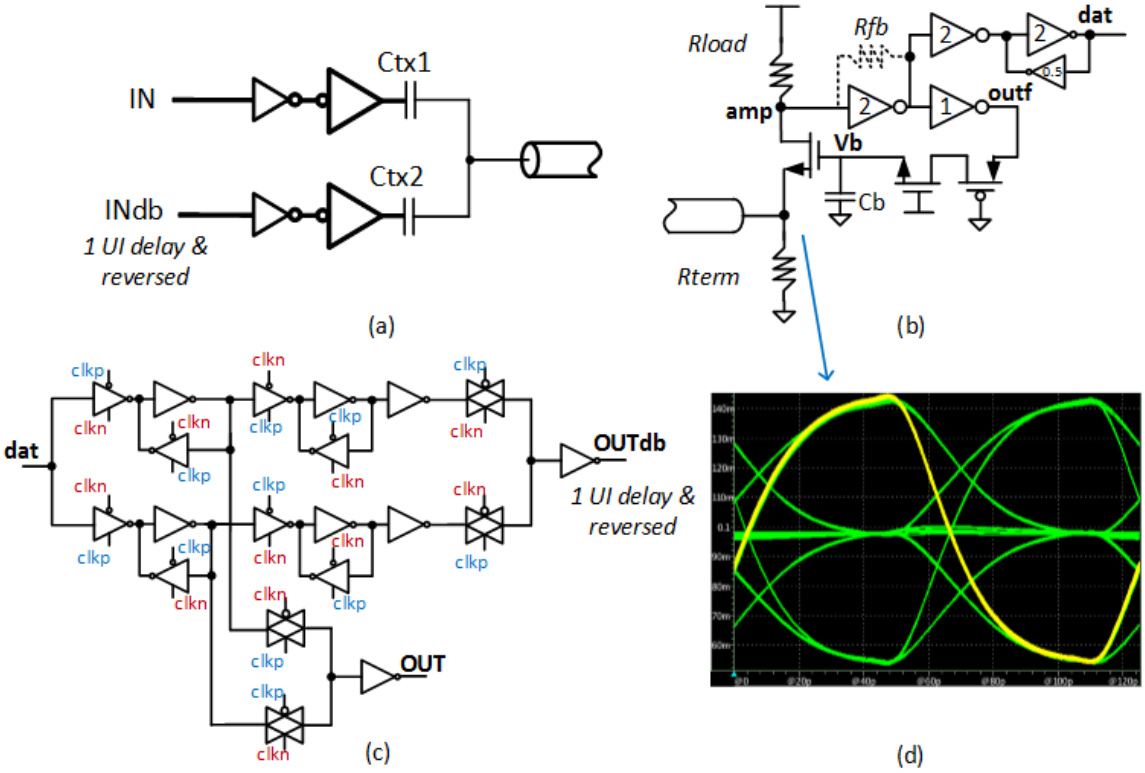


Figure 7: Schematics and waveform of pulse-mode signaling: (a) AC-coupling pulse-mode driver, (b) self-biased amplifier, (c) sampler for pulse-mode signaling, and (d) eye diagram of voltage signals

5. Power Supply Noise Effects

Unlike the off-chip I/O transceiver, the on-chip signaling circuits have to share the noisiest power supply with core logic circuits. This makes accurate data recovery challenging in a multi-hop data link. Study of the power supply noise effects on high-speed on-chip serial links is very helpful in learning about the performance potential in real SoC environments.

Table 2 shows experimental results of a 7-hop serial link. The test circuits are built in a 28nm process with 900mV typical supply. The VDD noise comes from sinusoidal voltage signals which have frequency of 510MHz. Different columns are for varying data

rate and different rows are for changing peak-to-peak VDD noise amplitudes. The number (if not "Pass") in each blank means the number of hops from the starting transmitter that data can go through before any error bit is detected. Generally, the results are quite intuitive in that either higher data rate (smaller timing margin) or larger noise amplitude could cause the data signals fail earlier in the link.

This may be a simplified case compared to the real on-chip environment, but it is typical enough to provide us some knowledge of the designs limitations. To better understand the relation of system performance to power supply noise, more cases with different noise frequency and noise patterns need to be explored. The major effects of VDD noise are: high frequency noise tends to break the delay tracking function of the source-synchronous clock and cause some bit errors, and large VDD voltage variations will cause timing-offset accumulation (i.e. duty-cycle distortion or quadrature distortion) in the clock lanes and make the clock distribution fail after several stages.

Table 2: On-chip data link performance at various data rates and VDD noise amplitudes

Vnpp/Rate	13Gb/s	14Gb/s	15Gb/s	16Gb/s	17Gb/s	18Gb/s	19Gb/s	20Gb/s
150mV	Pass	Pass	Pass	Pass	Pass	Pass	Pass	Pass
200mV	Pass	Pass	Pass	Pass	Pass	Pass	6	4
250mV	Pass	Pass	Pass	Pass	Pass	6	3	3
300mV	Pass	Pass	Pass	Pass	4	3	1	1
350mV	Pass	6	3	3	3	1	1	1
400mV	6	3	3	1	1	0	0	0

6. Conclusion

In this paper, we present the challenges and possible solutions to implementing a high-speed low-power serial link for on-chip global signaling. Methods for custom channel, circuits, and clocking architecture designs are discussed. Compared to traditional CMOS signaling and off-chip serial link designs, the multi-hops source-synchronous serial link, using high-level thick metal channels and simple equalization technique, shows promising potential to fulfill the performance requirements in future SoC products.