Simple Design Criterion for Maximizing Data Rate in NoC Links

Anastasia Barger and David Goren
IBM Haifa Lab
kapchits@il.ibm.com
davidg@il.ibm.com
Haifa University, Mount Carmel

Avinoam Kolodny
kolodny@ee.technion.ac.il
Technion - Israel Inst. of Tech.

Abstract

A simple design criterion is presented for obtaining maximal data rate in network on chip (NoC) links. It is shown that the maximal data rate is achieved near the boundary between RC and RLC model validity domains. The criterion is applicable to various on-chip transmission line structures, including crossing lines at adjacent metal layers. Inductive effects are represented by time-of-flight through the link structure. The resulting design criterion can be expressed using only static RC wire parameters.

Introduction

On-chip packet-switched network can potentially become the preferred interconnection approach for future Systems-on-Chip (SoC). The structured layout of Network on Chip (NoC) links significantly simplifies the design and modeling problems encountered in design of global interconnect in digital integrated systems, such as delay uncertainty, crosstalk noise, interconnect power consumption and inductive effects. Examples of NoC are presented in [1, 2, 3].

A NoC link typically consists of a number of parallel signal wires of fixed width and spacing. These wires provide point-to-point connections between network routers. There is no fan-out in the wires as in general interconnect trees. In order to reduce crosstalk, some of these wires may serve as shielding wires.

The layout view of a NoC link is presented in Figure 1:

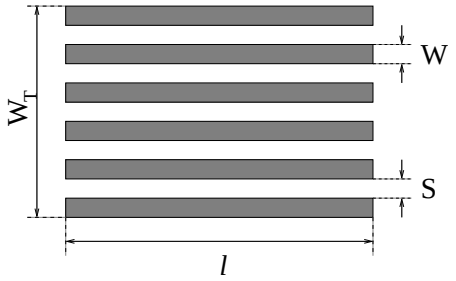

Figure 1: Link layout in an area of total width W_T and length *l*. W and S are wire width and spacing.

Link performance can be expressed by a number of metrics such as data rate [4], data flux density and bisectional bandwidth [5]. The Data Rate is one of the most appropriate metrics for NoC link:

$$\text{Data Rate} = \frac{N}{Delay} \tag{1}$$

where N is the total number of signal wires in the link and delay is the delay of a single wire, which depends on wire unit length resistance (R), capacitance (C), inductance (L) and optionally the conductance (G) for wires right above the silicon substrate [6, 9]. All these parameters in turn depend on link geometry - wire width, thickness and spacing. Consequently, these geometry parameters affect both the delay and the number of wires N within the given link area. In previous works [4], a numerical optimization procedure was presented for the link, neglecting inductance and using only RC model. For high performance links, however, the RC model is not accurate enough since the inductance can affect link delay [6]. Digital designers need simple models and methods for dealing with inductance in optimizing NoC links. Using transmission line structures (T-lines) for NoC links was proposed in [12]. By adding an underlying ground plane, dedicated current return paths are defined in these T-lines, enabling inductance calculation within a stand alone model. Potential frequency dependent phenomena in T-lines [8, 9, 10, 11] due to skin effect can greatly complicate the calculations. A simple model for the T-lines with reduced frequency dependent phenomena is also presented in [12]. The model enables simple inductance calculations avoiding use of field solvers.

A simple criterion for maximizing data rate is presented in this paper, based on analysis of propeties of RC and RLC models. The criterion, which fully considers the inductive effects in the interconnect lines and uses the simple T-line model as presented in [12], can be expressed by RC wire parameters only. An optimization of different metrics has been presented by [5] for an RLC link with repeaters.

This paper is organized as follows: Link modeling is presented in section 2. Data rate optimization using an RC model is presented in section 3. Optimization with an RLC model is presented in section 4, and simple design criterion is introduced in section 5 for a basic layout configuration. A generilized criterion for different layout configurations is given in section 6.

Link modeling

The cross section view of the suggested NoC link is shown in Figure 2:

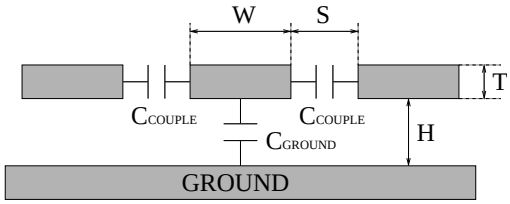

Figure 2: Basic link cross section. A solid ground plane resides under the signal lines, to ensure the current return path.

Each one of the R, L, C link parameters is a function of the link cross section geometry - width W, spacing S, wire thickness T, and insulator thickness H. Increasing the width, for example, decreases wire resistance but increases the ground capacitance C_{GROUND}. Increasing the spacing between the wires decreases the coupling capacitance C_{COUPLE}. The resistance of the link wires can be approximated by DC resistance and wire capacitances for the link can be calculated from [14].

Link optimization using RC delay model

The delay of distributed RC line driven by an ideal driver (zero output impedance) at the near end, and open termination at the far end can be presented by [7]:

$$delay = 0.37 \cdot RC \tag{2}$$

where R and C are the total resistance and capacitance of the wire. While T and H parameters are fixed for each metal layer in a given process technology, the parameters W and S can be chosen by the link designer. The data rate can be expressed by W and S, and the numerical methods can be applied to find the W_{OPT} and S_{OPT} parameters which maximize the data rate function of the link [4].

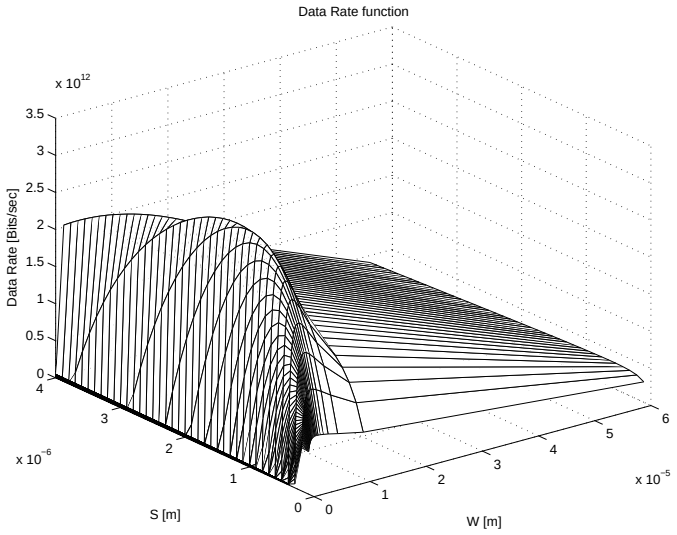

Figure 3: RC data rate vs. S and W.

Figure 3 shows the RC data rate as a function of W and S. The data rate demonstrates the trade off between the number of wires in the link and the delay in each wire. For a link with large wire width and spacing, the delay of each wire is small, due to the small resistance and coupling capacitance, but the total number of the wires is also small and hence the data rate is small. For a link with very thin wires, the total number of the wires is increased, but the delay of each wire increases due to both larger resistance and larger coupling capacitances, and hence the data rate is small. It can be seen in Figure 3 that the data rate can be maximized between these two extreme cases.

Link optimization using RLC delay model

Inductive effects on the link output are exemplified on Figure 4, using RC and RLC models. The differences in delay, slew rate and signal integrity can be easily seen. Under the assumption of ideal driver in the input and open termination output, the inductance of the wire increases the propagation delay and decreases the rise time of the signal [6].

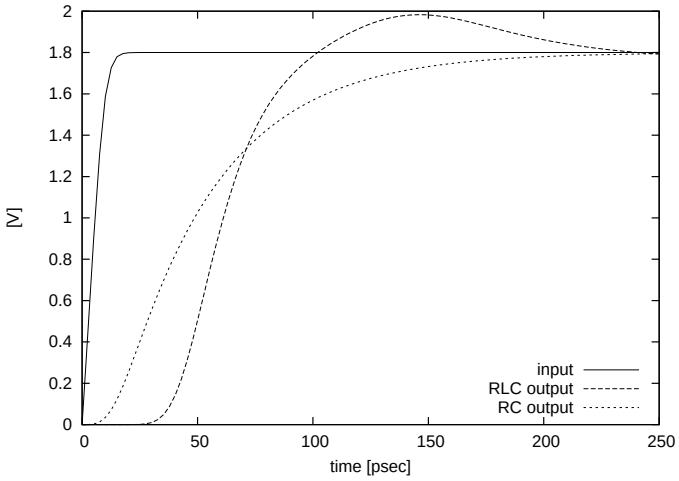

Figure 4: Simulated output waveforms using RC and RLC link models. The 50% delay is longer in the RLC model, while the signal rise time is shorter.

Under these conditions, the delay of RLC wire is given by [6]:

$$delay = \sqrt{LC} \cdot (e^{-2.9(\alpha_{asym})^{1.35}} + 0.74\alpha_{asym}) \tag{3}$$

where L and C are the total link inductance and capacitance correspondingly and α_{asym} is defined as:

$$\alpha_{asym} = \frac{1}{2} R \sqrt{\frac{C}{L}} \tag{4}$$

where R is the total resistance of the wire. For highly resistive wires, where the inductance is not important, (3) approaches the RC delay as in (2). For highly inductive wires, the RLC delay approaches the time of flight delay which is limited by the speed of light in the propagation media. The inductance in (3) is assumed in this paper to be the high frequency limit inductance (L_∞) following the approach suggested in [12]:

$$L_\infty = \frac{\tau_{of}^2}{C} \tag{5}$$

The link inductance can hence be inferred from the link capacitance in this case, by substituting the propagation time:

$$\tau_{of} = \frac{l\sqrt{\varepsilon_r}}{c_0} \tag{6}$$

where ε_r is relative dielectric constant of insulator material and c_0 is the speed of light in vacuum. Equations (5) and (6) are valid for the case of Figure 2 which does not include crossing lines. The resistance was assumed to be the DC resistance (R_{DC}) following [12].

Figure 5 presents RC delay, RLC delay and τ_{of} delay of the link as a function of wire width for the case of S=W. It can be

seen, that for the small wires width the RC delay model is the same as the RLC model. In this region, the inductance can be ignored because of high wire resistance. As the width increases, the difference between the two models becomes more and more significant. The RLC model approaches the τ_{of} delay limit, while the RC model becomes unrealistic.

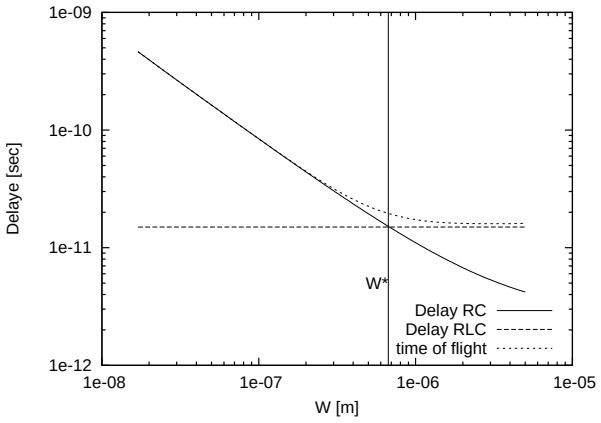

Figure 5: RC, RLC delay and time of flight vs. wires width for S=W. T=H=1μ, *l*=2mm.

The RC and RLC delay begin to split from each other at the point where the RC model still provides a good approximation for the wire delay, keeping the advantages of a simple model and avoiding the inductance calculation. This occurs near wire width W^*, where the RC delay curve intersects the τ_{of} asymptote: $0.37R_{DC}C = \tau_{of}$. Hence,

$$R_{DC}C = 2.7\tau_{of} \tag{7}$$

This is a sensible design point for the wire width, since any further increase in wire width does not give meaningful benefit in delay because of the time of flight limit, and any decrease in wire width increases the delay and reduces the data rate in this region as can be seen in Figure 6 below. The RC model is still valid at this point, as it is follows from the figure of merit for onset of inductance effects [13]:

$$R > 2\sqrt{\frac{L}{C}} \tag{8}$$

However, if we replace L by L_∞ and R by R_{DC} following [12] we get that:

$$R_{DC} > 2\sqrt{\frac{L_\infty}{C}} = 2\frac{\tau_{of}}{C} \tag{9}$$

Since L is larger than L_∞ by a factor which in practise is rarely higher than 2, it follows from (9) that the RC model is valid in (7).

Figure 6 compares the data rate of RC and RLC link models for the case of S=W. Both RC and RLC models show that the data rate of the link can be optimized but they predict the optimum in different locations. The RLC optimum appears at a smaller W than the RC optimum. As the wire width increases a significant difference between the two models can be seen. It should be noted again that using the simple RC model in the region of wide wires can be extremely misleading. The RLC optimum occurs near the point W^* defined above, where increasing the wire width no longer improves the wire delay due to time of flight limit.

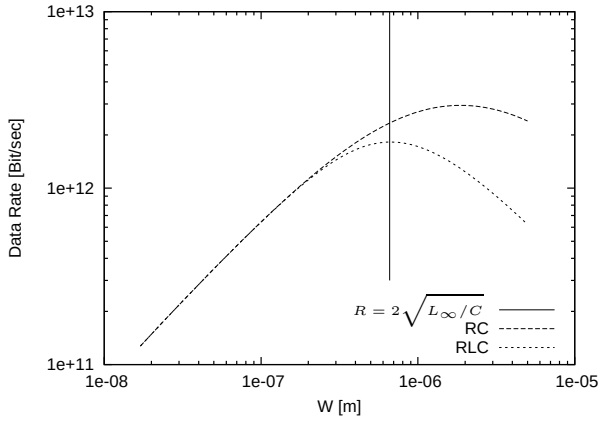

Figure 6: RC and RLC data rate vs. wires width for S=W. Solid line shows the boundary between RC and RLC models.

Optimization criterion for more general link layouts

NoC links normally have crossing lines above or below the signal lines (parallel wires going in perpendicular direction to the link). Several variants of the basic layout are therefore depicted in Figure 7:

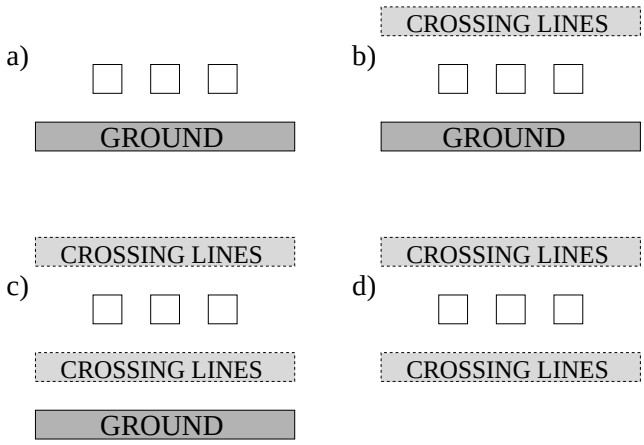

Figure 7: General link cross section: a) basic structure, b) structure with crossing lines above the link signal layer, c) structure crossing lines underneath the link signal layer, d) coplanar structure with crossing lines above and below the link

For capacitance extraction, crossing lines are often assumed to act as a ground plane , so that, crossing wires increase link capacitance [15]. However, crossing wires do not provide a current return path in the direction of the link, and hence they do not reduce the inductance (unlike a true solid plane). Hence

the effect of crossing lines is to slow down the wave propagation velocity in the link (slow wave effect).

In configurations with crossing lines we have to calculate L_∞ in a slightly different manner as described below. In order to calculate L_∞, only the structure that provide a current return path has to be accounted, i.e. the structure without crossing lines. The capacitance of this structure is denoted by the C_{RETURN}, which is the capacitance of the signal line as if the crossing lines did not exist at all. This applies both to crossing lines above the the signal lines and to the optional crossing lines lying between the signal line and the bottom dedicated ground plane. The inductance is therefore calculated using:

$$L_\infty = \frac{\tau_{of}^2}{C_{RETURN}} = \frac{\tau_{of}^{*^2}}{C} \tag{10}$$

Where τ_{of}^* considers the slow wave effect, and C is the total capacitance including the contribution of crossing lines acting as effective ground planes. The link design criterion can then be expressed using τ_{of}^*:

$$R_{DC}C = 2.7\tau_{of}^* \tag{11}$$

where

$$\tau_{of}^* = \tau_{of}\sqrt{\frac{C}{C_{RETURN}}} \tag{12}$$

Conclusions

A simple practical method has been presented for designing on-chip transmission line structures in digital circuits, such as NoC links. Data rate is maximized by adjusting wire widths and spaces until inductive behavior begins to set in. Link inductance is approximated by its high frequency limit, which can be expressed by the speed of light within the insulating dielectric material, and link capacitance. The resulting criterion for optimal link design which fully considers inductive effects and slow wave effect can be expressed only by static RC link parameters.

Acknowledgments

We would like to thank Prof. Eby Friedman for helpful discussions.

References

[1] William Dally, Brian Towles “Route Packets, Not wires: On Chip Interconnection Networks”, *Design Automation Conference*, 2001, page(s) 684-689.

[2] J. Liu, L-R Zheng, D. Pamunuwa, and H. Tenhunen “A global wire planning scheme for network-on-chip”, IV: 2003, page(s) 892-895.

[3] E. Bolotin, I. Cidon, R. Ginosar and A. Kolodny “QoS architacture and design process for cost effective network on chip”, *Journal of System Architecture, Special Issues on Network on Chip*, 50:105-128, Feb2004.

[4] D. Panumuwa, Li-R. Zheng, “Optimizing Bandwidth over Deep Sub-micron Interconnects”, *IEEE International Symposium on Circuits and Systems 2002*, ISCAS 2002, Vol.4, page(s): IV-193-IV-194.

[5] A. Naeemi, R.Venkatesan, and J.D.Meindl “System-on-a chip global interconnect optimization” IEEE 2002, page(s) 399-403.

[6] Y.I. Ismail and E.G. Friedman “On-chip inductance in high speed integrated circuits”, Norwell, MA: Kluwer, 2001, page(s): 247- 256.

[7] H. B. Bakoglu, Circuit Interconnects and Packaging for VLSI, *Addison-Wesley, Pub. Co. 1998.*

[8] D. Goren et al. “An interconnect-aware methodology for analog and mixed signal design, based on high bandwidth (over 40 GHz) on-chip transmission line approach”, DATE02, 2002.

[9] R. Gordin et al. “Modeling of on-chip transmission lines in high-speed A&MS design -the low frequency inductance calculation”. SPI02, May 2002.

[10] D. Goren et al. “On-chip interconnect-aware design and modeling methodology, based on high bandwidth transmission line devices”, *IEEE DAC 2003 conference*, Anaheim, 2003.

[11] D. Goren, R. Gordin, and M. Zelikson “Modeling methodology for on-chip coplanar transmission lines over the lossy silicon substrate”, *IEEE Signal Propagation on Interconnect conference*, Siena, 2003.

[12] A. Barger, D. Goren, A. Kolodny, “Design and Modeling of Network-on-Chip Interconnect using Transmission Lines”, IEEE, ICECS2005.

[13] Y. I. Ismail, E. G. Friedman, J. L. Neves, “Figures of Merit to characterize the importance of On-Chip Inductance,” *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol.7, No.4, December 1999, page(s):442-449.

[14] T. Sakurai and K. Tamaru “Simple formulas for two- and three-dimensional capacities” *IEEE Transactions on Electron Devices*. Ed-30(2), Feb 1983.

[15] Shyh-Chyi Wong, G. W. Lee, D. J. Ma “Modeling of Interconnect Capacitance, Delay and Crosstalk in VLSI”, *IEEE Transactions on Semiconductor Manufacturing*, Vol.13, No.1, Feb 2000.