

An Integrated Architecture for Global Interconnects in a Gigascale System-on-a-Chip (GSoC)

Payman Zarkesh-Ha and James D. Meindl

Microelectronics Research Center, Georgia Institute of Technology, Atlanta, GA 30332-0269, USA

Abstract

An integrated architecture for global interconnects in a gigascale system-on-a-chip (GSoC) is presented using the models for global signal, clock, and power supply wiring networks. Based on the models for wiring resource demand, noise limit, and bandwidth requirement, an interconnect design plane is proposed. The new design plane demonstrates the limits imposed on global on-chip interconnect physical dimensions for future technology generations.

Introduction

Global interconnects are commonly considered as a key potential bottleneck to the advancing performance of future integrated systems [1]. The complete global interconnect architecture of a digital system implemented on a single chip consists of *signal*, *clock*, and *power supply* distribution networks (Fig. 1). Since the signal, clock and power supply networks utilize the same interconnect stacks, it is imperative to integrate the complete interconnect architecture in order to design and optimize the global interconnects for a gigascale system-on-a-chip (GSoC). In this paper an integrated architecture for global interconnects in a GSoC is presented using the models for *wiring resource demand*, *noise limit*, and *bandwidth requirement*.

a) **Wiring Resource Requirement** – The total required area for global wiring networks is often limited by the GSoC area. Using the models for wiring requirements of signal, clock, and power supply distribution networks, a limit on the global interconnect physical dimensions is defined.

b) **Wiring Bandwidth Requirement** – Since clock is the fastest signal in a GSoC, it requires a high bandwidth interconnect so that the rise time is preserved when the clock signal passes through the interconnect. Therefore, the bandwidth of global clock distribution often defines the speed limit of the interconnect architecture.

c) **Wiring Noise Limit** – Interconnect coupling noise, or cross-talk, is often a major concern for global wires. Unlike the interconnect bandwidth, the peak cross-talk noise prediction usually needs an accurate distributed RLC approach.

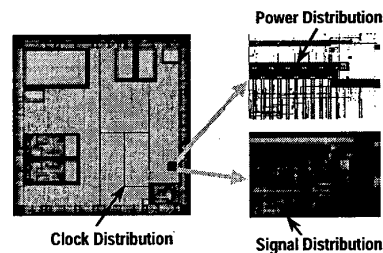


Figure 1 – Signal, clock and power supply wires in a real chip utilizing the same interconnect architecture

Based on the models for wiring resource demand, noise limit, and bandwidth requirement, an interconnect design plane is proposed. The new design plane demonstrates the limits imposed on global on-chip interconnect physical dimensions for future technology generations. It is shown that the size of the design region is vanishingly small, as a result of scaling to deep sub-micron technology assuming two global wiring levels.

Wiring Resource Requirement

Figure 2 illustrates the wiring breakdown for two different cases, 1) a low power design and 2) a high performance chip. As shown, the wiring requirement for signal, power and clock is highly design-dependent. For instance, the area dedicated to power supply distribution network in a high performance chip is larger than that in a low power design.

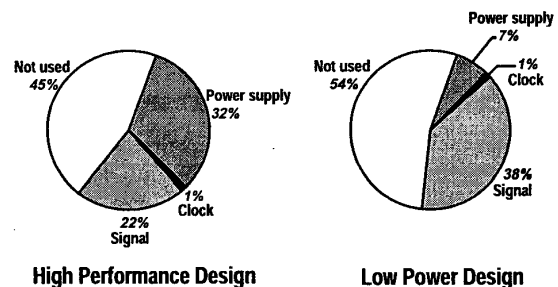


Figure 2 – Wiring area breakdown in a low power design and a high performance chip

Therefore, accurate models for wiring requirements of signal, clock, and power supply networks are required in order to define a limit on the global interconnect physical dimensions.

The total required area for the wiring networks is often limited by the SoC area, A_{SoC} ,

$$A_{Signal} + A_{Clock} + A_{Power} \leq A_{SoC}, \quad (1)$$

where A_{Signal} , A_{Clock} , and A_{Power} are the area of signal, clock and power supply networks, respectively.

Using the stochastic net-length distribution for global interconnects in a SoC [2], the required area for signal wiring network is estimated based on the Rent's rule parameters of the system, K_{eq} and P_{eq} , as shown below

$$L_{tot} = \sum_{m=2}^{N_n} N_{Net}(m) \cdot L_{av}(m), \quad (2)$$

where

$$N_{Net}(m) \approx \frac{K_{eq} N_m (m^{P_{eq}-1} - (m+1)^{P_{eq}-1})}{m+1},$$

$$L_{av}(m) \approx (0.5\sqrt{m+1}) \frac{m-1}{m+1} \sqrt{A_{SoC} \left(\eta_p + \frac{N_m}{m} (1 - \eta_p) \right)},$$

$N_{Net}(m)$ is the number of m -terminal nets, $L_{av}(m)$ is the average length of m -terminal nets, and η_p is the placement efficiency.

The area for clock distribution is often negligible compared to the chip area as shown in Fig. 1.

Based on a mesh-structure power distribution and assuming a uniform load distribution, A_{Power} can be approximated by the expression shown below [3]

$$A_{Power} = \frac{2m-1}{m^2} \cdot A_{SoC}, \quad (3)$$

where

$$m \approx \frac{16\delta V_{DD}^2 H}{P_{tot} \rho_w} \text{ for peripheral wire-bond system,}$$

$$m \approx \frac{16\delta V_{DD}^2 H n_{pad}}{P_{tot} \rho_w} \text{ for area-array bonding pads,}$$

P_{tot} is the total chip power dissipation, n_{pad} is the number of power supply pads, δ is the maximum allowable IR-drop, H is metal thickness, ρ_w is metal resistivity, and V_{DD} is the power supply voltage.

To verify the simplified models, IR-drop profile is simulated using finite element method and the results are shown in Fig. 3. It is seen that the IR-drop reduces significantly by using area-array bonding pads instead of peripheral wire-bonds.

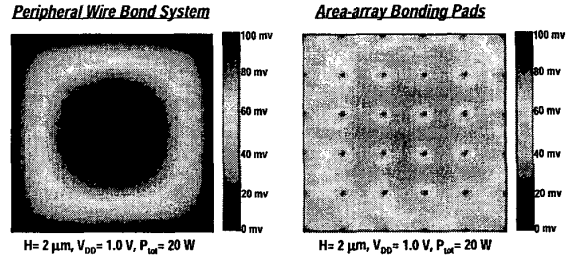


Figure 3 – IR-drop profile in peripheral wire-bond and area-array bonding pad systems using finite element method

Using the simplified models for A_{Power} , it is also shown that the power supply wiring area reduces significantly by using area array bonding pads instead of peripheral wire-bonds as illustrated in Fig. 4.

Wiring Bandwidth Requirement

Since the clock signal is the fastest signal in a SoC, it requires a high bandwidth interconnect so that the rise time is preserved when the clock signal passes through the interconnect. Therefore, the bandwidth of global clock distribution often defines the speed limit of the interconnect architecture.

Figure 5 illustrates the SPICE simulations for the frequency response of three interconnects with the length of 1, 2, and 3 cm. As shown, there is very little difference in bandwidth when a distributed RC-circuit representation is used instead of the accurate distributed RLC approach. This phenomenon has been also observed in [4]. Therefore, the wiring bandwidth requirement can be expressed simply by

$$f_{Clock} \leq f_{-3dB} \approx \frac{1}{2\pi r_c l^2}, \quad (4)$$

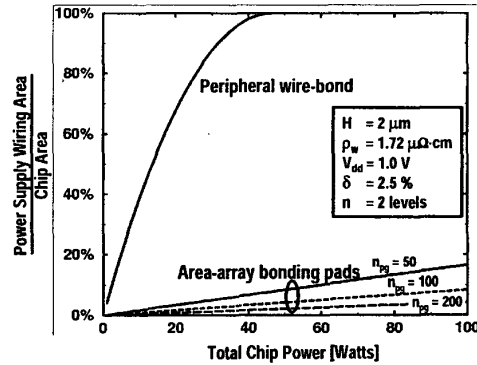


Figure 4 – Power distribution area versus total chip power

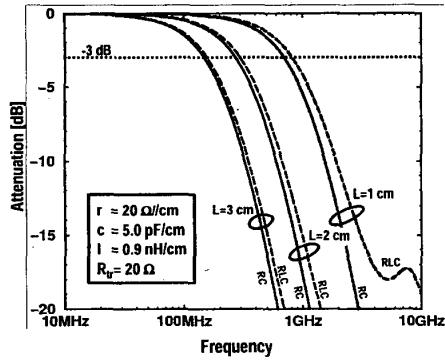


Figure 5 – Frequency response of an interconnect using HSPICE

where r and c are the interconnect resistance and capacitance per unit length and l is the length of the interconnect. The more detailed expression for bandwidth requirement is shown in Table I.

Wiring Noise Limit

Interconnect coupling noise, or cross-talk, is often a major concern for global wires. Unlike the interconnect bandwidth, the peak cross-talk noise prediction usually needs an accurate distributed RLC approach. Figure 6 shows the SPICE simulations for the peak cross-talk noise using distributed RC and RLC approaches. As shown, the peak cross-talk noise in an RLC line is often higher than that of RC line. It has been shown in [5] that the worst case peak cross-talk noise using a distributed RLC model can be simplified to

$$\frac{V_n}{V_{dd}} \approx \frac{\pi}{4} \frac{c_m}{c_{gnd} + c_m} \leq \text{Noise limit}, \quad (6)$$

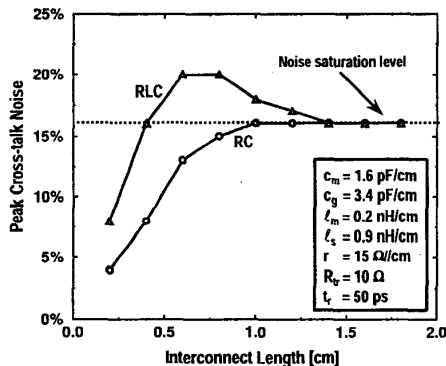


Figure 6 – Peak cross-talk noise versus wire length using HSPICE

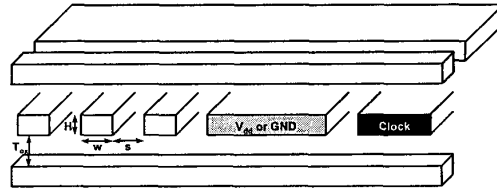


Figure 7 – A simplified orthogonal interconnect system showing signal, clock and power supply wires

Table I – Simplified expressions for global interconnects

Equation Description	Simplified Expressions for Global Wiring Requirements in Terms of w , s , H , and T_{ox}
Wiring resource requirement	$(w+s) \sum_{m=2}^{N_m} N_{net}(m) \cdot L_w(m) \leq 0.5 \left(1 - \frac{P_{mD_w}}{165V_{dd}^2 H n_{ps}} \right)^2 A_{soc}$ <p>where: $N_w(m) = \frac{K_w N_m (m^{n-1} - (m+1)^{n-1})}{m+1}$ $L_w(m) = (0.5\sqrt{m+1} - \frac{m-1}{m+1}) A_{soc} \left(\frac{N_m}{m} (1 - \eta_m) \right)$</p>
Wiring bandwidth requirement	$f_c \leq \frac{1}{4\pi\rho_w \epsilon_o \epsilon_r (\sqrt{HT_{ox}} + \sqrt{ws}) A_{soc}}$
Wiring noise limit	$\frac{\pi}{4} \frac{\sqrt{ws}}{(\sqrt{HT_{ox}} + \sqrt{ws})} \leq \% \text{Noise}$

where c_m and c_{gnd} are the mutual and ground capacitance per unit length, respectively.

Also, as the length of the interconnect increases, the cross-talk noise using both RC and RLC models, saturates at a level given by [6]

$$\frac{V_n}{V_{dd}} \approx \frac{1}{2} \frac{c_m}{c_{gnd} + c_m}, \quad (7)$$

where c_m and c_{gnd} are the mutual and ground capacitance per unit length, respectively. From the comparison between (6) and (7), it is shown that the worst case peak cross-talk noise of distributed RLC model is about 57% more than that of RC model.

Complete Global Interconnect Design Plane

Assuming the interconnect architecture of Fig. 7 for signal, clock, and power distribution network and based on the models for wiring resource requirement, wiring bandwidth demand, and wiring noise limit given in Table I, a design plane for global interconnects is defined.

As shown in Fig. 8, for a design in 180 nm technology generation with two global wiring levels, the forbidden zone of operation lies external to the bounded region defined by the three wiring limits. Moreover, the design points for interconnect physical dimensions for minimum pitch, minimum aspect ratio, and maximum speed can be obtained from the design plane.

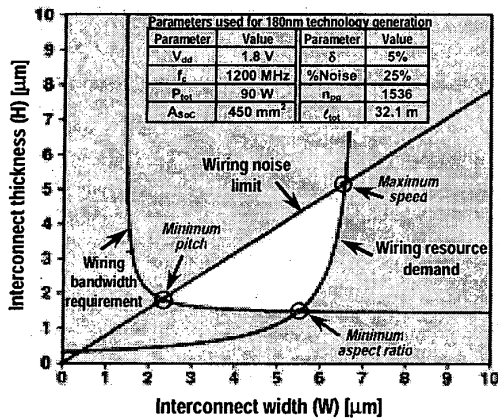


Figure 8 – A complete interconnect design plane

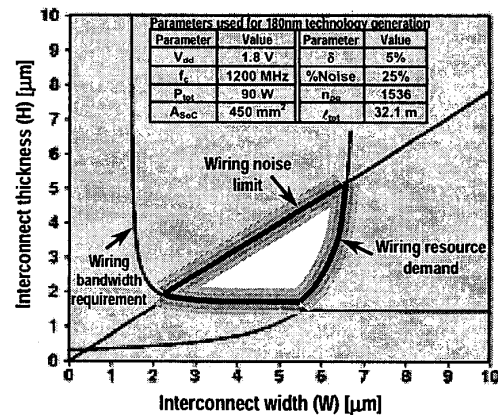


Figure 10 – The impact of interconnect variations on the design plane

The proposed interconnect design plane can be used to enhance the understanding of limits associated with the interconnect architecture in the future generations of technology.

For instance, Fig. 9 illustrates the design planes for 180, 130, 100 and 50 nm generations of technology based on the ITRS [7]. As shown, assuming two global wiring levels, the size of this design region appears to be almost vanishingly small, as a result of scaling of technology into the deep sub-micron range.

Also, the impact of interconnect process variations can be studied by the proposed design plane. For instance, Fig. 10 shows that the design region for physical dimensions of an interconnect in 180nm technology is smaller when the interconnect variations is taken into account.

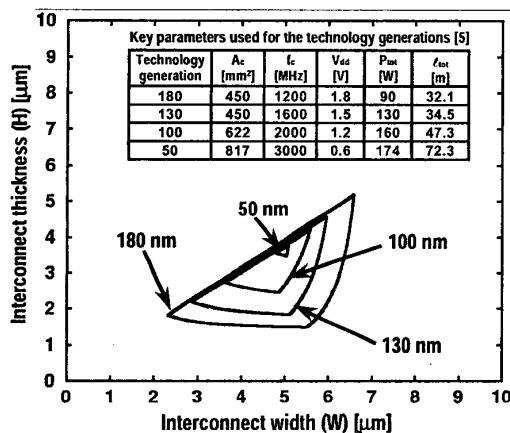


Figure 9 – Interconnect design plane for 180, 130, 100, and 50 nm generations of technology

Conclusion

An integrated architecture for global interconnects in a gigascale system-on-a-chip (GSoC) is presented in this paper using the models for global signal, clock, and power supply wiring networks. Based on the models for wiring resource demand, noise limit, and bandwidth requirement, an interconnect design plane is proposed. The new design plane demonstrates the limits imposed on global on-chip interconnect physical dimensions for the future generations of technology. It is shown that the size of the design region is vanishingly small as a result of scaling to deep sub-micron technology assuming two global wiring levels. Also, the impact of interconnect process variations is studied using the developed design plane.

References

- [1] D. Sylvester, and K. Keutzer, "Getting to the bottom of deep submicron II: a global wiring paradigm," Proc. of *Int. Symp. on Phys. Design*, pp. 193-200, April 1999.
- [2] P. Zarkesh-Ha and J.D. Meindl, "Stochastic net-length distribution for global interconnects in a heterogeneous system-on-a-chip," Proc. of *Symp. on VLSI Tech.*, pp. 44-45, June 1998.
- [3] P. Zarkesh-Ha and J.D. Meindl, "An integrated architecture for global interconnects in a gigascale system-on-a-chip (GSoC)," to be presented at *Symp. on VLSI Tech.*, June 2000.
- [4] A. Deutsch *et al.*, "High-speed signal propagation on lossy transmission lines," *IBM Journal of Res. and Dev.*, pp. 601-615, 1990.
- [5] J.A. Davis and J.D. Meindl, "Compact distributed RLC models for multilevel interconnect networks," Proc. of *Symp. on VLSI Tech.*, pp. 165-166, June 1999.
- [6] T. Sakurai, "Closed-form expressions for interconnect delay, coupling, and cross-talk in VLSI's," *IEEE Trans. On Electron Devices*, pp. 118-124, Jan. 1993.
- [7] *International Technology Roadmap for Semiconductors*, Semiconductor Industry Association, 1999.