Impact of On-Chip Interconnect on the Performance of 3-D Integrated Circuits With Through Silicon Vias: Part I

Vachan Kumar, *Member, IEEE*, Hanju Oh, *Student Member, IEEE*, Xuchen Zhang, *Student Member, IEEE*, Li Zheng, Muhannad S. Bakir, *Senior Member, IEEE*, and Azad Naeemi, *Senior Member, IEEE*

***Abstract*— Circuit-level models are developed to determine the upper bound on the performance of a 3-D IC link with through silicon vias (TSVs). It is shown that the performance of a 3-D link is limited not only by the on-chip interconnect RC, driver resistance, and TSV capacitance, but also by the current carrying capacity of the on-chip wires connecting the TSV to the input/output (I/O) driver. The models developed in this paper are used to optimize the I/O driver size, the number of on-chip wires connecting the TSV to the driver, and the data-rate to maximize the aggregate bandwidth per unit energy. Furthermore, in order to maximize the aggregate bandwidth of a 3-D link, it is shown that splitting the TSV array into smaller subarrays and placing the I/O drivers closer to the TSVs is better compared with having large TSV arrays.**

***Index Terms*— 3-D integration, bandwidth, electromigration, interconnect, Joule heating, modeling, optimization, reliability, technology scaling, through silicon via (TSV).**

I. Introduction

THE economics of Moore's law has driven the semiconductor industry for the last five decades, with improved performance and power efficiency in every technology generation [1], [2]. However, with the device sizes shrinking well below 100 nm, the effort required to squeeze the required performance out of every new technology generation is rising exponentially. The need to connect a few billion transistors on a single chip has driven the industry toward adopting several novel interconnect technologies, such as double/quad patterning, low-k dielectrics, and so on. However, improvements in chip performance alone cannot guarantee an improvement in the overall system performance. For example, irrespective of how fast a microprocessor carries out its computations, it has to wait for data to be fetched from the different levels of cache or the main memory. Thus, if the off-chip bandwidth from the main memory to the processor cannot keep up with the microprocessor speed, the overall system performance will not improve. Microprocessor input/output (I/O) bandwidth demands approximately double every two years [3]. Hence, the conventional chip-to-chip interconnects, which suffer from significant conductor and dielectric losses at higher frequencies have become major bottlenecks in high performance nanoelectronic systems [4]. Without a considerable improvement in the performance and power of chip-to-chip interconnects, the boost in performance at the chip level cannot be translated to system-level improvements. Therefore, many alternative technologies, including optical interconnects [5], 3-D ICs [6], silicon interposers [7], and airgap interconnects [8], are being investigated.

Three dimensional integration aims to minimize the physical distance between the communicating ICs by stacking them on top of each other using through silicon vias (TSVs). Although there have been a significant number of papers on the modeling and characterization of isolated TSVs or TSV arrays [6], [9]–[12], it is essential to combine these TSV models with the models of I/O drivers, receivers, and on-chip interconnects to accurately estimate the performance of a 3-D IC. Katti *et al.* [9] and Bandyopadhyay *et al.* [10] focused mainly on estimating the depletion capacitance of TSVs as a function of TSV bias and frequency. In [6], the focus is primarily on estimating the circuit parameters such as resistance and capacitance of various components (redistribution layer, microbumps, and pads). In [11], curve fitting is used to obtain capacitance models for TSV arrays. High-frequency models are developed for TSVs in [12]. Although a system-level analysis of 3-D ICs is performed in [13], the impact of on-chip wires on the performance is ignored.

In this paper, we aim to identify the key bottlenecks that limit the performance of a 3-D IC and understand the impact of technology scaling. The role of maximum current density and delay in determining the performance of a 3-D IC are highlighted. Furthermore, based on the circuit models developed in this paper, optimal placement of the I/O drivers with respect to the TSVs is determined. In part II of this paper, the models developed in this paper are validated

Manuscript received February 28, 2016; revised April 16, 2016; accepted April 16, 2016. Date of publication May 6, 2016; date of current version May 19, 2016. This work was supported by Semiconductor Research Corporation within the Global Research Collaboration under Grant 2254.001. The review of this paper was arranged by Editor A. M. Ionescu.

V. Kumar was with the Georgia Institute of Technology, Atlanta, GA 30332 USA. He is now with Logic Technology Development, Intel Corporation, Hillsboro, OR 97124 USA (e-mail: vachkumar@gmail.com).

H. Oh, X. Zhang, L. Zheng, M. S. Bakir, and A. Naeemi are with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA (e-mail: hanju.oh@gatech.edu; xzhang96@gatech.edu; lizheng@gatech.edu; muhannad.bakir@mirc.gatech.edu; azad@gatech.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TED.2016.2556709

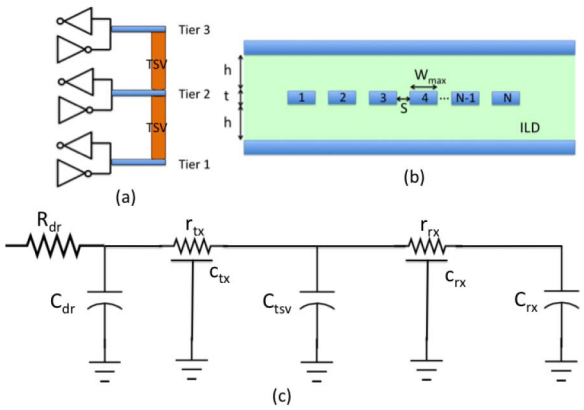


Fig. 1. (a) Schematic of a 3-D IC showing drivers and receivers (I/O circuits), TSVs, and on-chip interconnects that connect the I/O circuits to the TSVs (b) Cross section of on-chip interconnect connecting the I/O to the TSV (c) The circuit model used to predict the Elmore delay of a 3-D IC link comprising of lumped circuit models for I/O circuits and TSVs, and distributed RC models for on-chip interconnect.

using TSV capacitance and 3-D link measurements. This paper has been organized as follows. The circuit models, limit to the data-rate, and optimization of a 3-D IC are presented in Section II. In Section III, system-level models are developed to evaluate the trade-off between on-chip wire length and TSV density. The results and discussions are presented in Section IV. Important conclusions from this paper are summarized in Section V.

II. Modeling and Design of a 3-D IC

In this section, circuit models are developed for a 3-D link and used to estimate the maximum data-rate that the 3-D link can support. The data-rate is primarily limited by the maximum current density that the on-chip interconnects can handle and the delay of the 3-D link. Furthermore, based on the maximum data-rate of operation and the energy per bit, the optimal number of wires connecting the TSV to the I/Os and the driver size are chosen.

A. Circuit Model

Circuit-level models are developed for 3-D links consisting of I/O drivers and receivers, transmitter side on-chip interconnects, TSVs, and receiver side on-chip interconnects, as shown in Fig. 1(a). The on-chip interconnect between the TSV and the I/O circuit is comprised of multiple wires, as shown in Fig. 1(b). In modern process technologies, the upper limit on the local density of wires imposed by chemical mechanical polishing determines the maximum width of each of the wires. In this analysis, the maximum width of each of the wires is assumed to be $3\times$ the minimum width allowed for the metal level. The equivalent circuit model for the 3-D link is shown in Fig. 1(c). The I/O driver is modeled as a resistance connected to the driver capacitance, the receiver is modeled as a load capacitance, and the on-chip interconnects are modeled as distributed RC networks. The wires connecting TSVs to the I/O are connected in parallel by shorting at the I/O and the TSV; as a result, the major component of wire capacitance is from the parallel plate capacitance to the layers above and below. The resistance of the wires is estimated using the effective resistivity model presented in [14]. TSVs are modeled as capacitances, with their capacitance given by the maximum depletion capacitance developed in [9]. However, at high frequencies, and when the substrate in the vicinity of the TSVs is not connected to ground through substrate taps, the effective capacitance is reduced, and determined by the coupling capacitances to the neighboring TSVs, as shown in Part II of this paper. Important technology parameters such as supply voltage, driver resistance, gate capacitance, wire dimensions, and dielectric constant of inter-layer dielectric are obtained from the International Technology Roadmap for Semiconductors (ITRS) roadmap [15].

B. Data-Rate Limited by Reliability

Because the TSV is a large capacitive load on the driver circuit, it draws/dumps a large current during charging/discharging. As a result, the on-chip interconnect connecting the I/O driver and the TSV should be designed to handle large currents. The high current densities in the wires connecting the TSVs to the I/O driver can cause severe failures due to electromigration and Joule heating [16]. The rms current flowing through the on-chip interconnect connecting the TSV to the driver is given by [17]

$$I_{\text{rms}} = \sqrt{\frac{V_{\text{dd}}^2 T_{RF} F_{\text{max}}}{3R_{\text{dr}}^2}} \tag{1}$$

where V_{dd} is the supply voltage of the I/O driver, T_{RF} is the maximum rise/fall time at the output of the I/O driver, R_{dr} is the driver resistance, and F_{max} is the maximum frequency at which the link can toggle. The projected maximum current density requirement for several technology generations is obtained from ITRS projections [15]. In this analysis, it is assumed that advanced interconnect technologies to combat electromigration/Joule heating will be developed by the semiconductor industry to achieve the target set by the ITRS projections. Assuming $10\% - 90\%$ rise/fall time requirement at the output of the I/O driver, the minimum size of the I/O driver can be computed as

$$\begin{aligned} T_{RF} &= 2.2\frac{R_{\text{min}}}{S_{\text{min}}}\left((C_{\text{tsv}} + c_{tx}L_{tx} + c_{rx}L_{rx}) + 2C_{\text{min}}S_{\text{min}}\right) \\ \Rightarrow S_{\text{min}} &= \frac{2.2R_{\text{min}}(C_{\text{tsv}} + c_{tx}L_{tx} + c_{rx}L_{rx})}{T_{RF} - 4.4R_{\text{min}}C_{\text{min}}} \end{aligned} \tag{2}$$

where R_{min} is the resistance of a minimum sized driver in the technology, C_{min} is the capacitance at the output node of an I/O driver formed of minimum sized transistors, C_{tsv} is the TSV capacitance, C_{dr} is the capacitance at the I/O node, and c_{tx}, c_{rx}, L_{tx}, and L_{rx} are the wire capacitances and wire lengths on the transmitter and receiver sides, respectively. Combining (1) and (2), the maximum current density J_{max} is given by

$$J_{\text{max}} = \sqrt{\frac{V_{\text{dd}}^2 T_{RF} F_{\text{max}} S_{\text{min}}^2}{3W^2T^2R_{\text{min}}^2}} \tag{3}$$

where $W = N_w W_{\max}$ is the effective wire width, $W_{\max} = 3 \times W_{\min}$ is the maximum wire width for a given metal layer, N_w is the number of wires, $W_{\min}$ is the minimum wire width for the metal layer, and T is the wire thickness. Therefore, the maximum data-rate limited by reliability that the on-chip interconnect can handle $F_{\mathrm{dr,rel}}$ is given by

$$F_{\mathrm{dr,rel}} = \frac{F_{\max}}{2} = \frac{3}{T_{RF}} \left(\frac{J_{\max} W_{\max} N_w T R_{\min}}{V_{\mathrm{dd}} S_{\min}} \right)^2 \quad (4)$$

where N_w is the number of wires with width $W_{\max}$. From (4), it is clear that the maximum data-rate increases quadratically with N_w at small widths. However, at larger values of N_w, the driver size $S_{\min}$ increases linearly due to the increased top and bottom capacitances with an increase in the value of N_w; as a result, the maximum data-rate limited by reliability saturates at larger widths.

C. Data-Rate Limited by Delay

The circuit model developed earlier is used to estimate the $50\% - 50\%$ delay of the 3-D link, given by Elmore delay model [18]

$$\begin{aligned} t_d(L_{tx}, L_{rx}) = {} & 0.69(R_{\mathrm{dr}} + r_{tx} L_{tx}) C_{\mathrm{tsv}} \\ & + 0.69 R_{\mathrm{dr}} (c_{tx} L_{tx} + c_{rx} L_{rx} + C_{rx} + C_{\mathrm{dr}}) \\ & + 0.69 r_{tx} c_{rx} L_{tx} L_{rx} \\ & + 0.38 (r_{tx} c_{tx} L_{tx}^2 + r_{rx} c_{rx} L_{rx}^2) \end{aligned} \quad (5)$$

where R_{dr} and C_{dr} are the driver resistance and capacitance, respectively, C_{tsv} is the capacitance of the TSV, C_{rx} is the load capacitance at the receiver, r_{tx}, c_{tx}, r_{rx}, and c_{rx} are the resistances and capacitances per unit length of the on-chip interconnects on the transmitter and receiver side, respectively, and L_{tx} and L_{rx} are the wire lengths on the transmitter and receiver side. The advantage of using Elmore delay model is that it can be used to estimate the delay contributions from different components, and to identify the most dominant component of delay. From our simulations, the driver resistance and driver side wire resistance, TSV capacitance, and driver side wire capacitance dominate the delay of the 3-D link. Although the rise/fall time at the receiver input is a more accurate estimate of the maximum data-rate of the link, the total link delay t_d is a good indicator of the rise/fall time at the receiver. As a result, the maximum data-rate limited by delay is given by $F_{\mathrm{dr,del}} = (1/t_d)$.

D. Optimal Design of On-Chip Interconnect

Because the maximum data-rate of the 3-D link is limited either by reliability or by the delay of the 3-D link, the effective limit on the data-rate of the 3-D link is given by the smaller of the two data-rates

$$F_{\mathrm{dr,eff}} = \min(F_{\mathrm{dr,rel}}, F_{\mathrm{dr,del}}). \quad (6)$$

Due to the high current density at smaller N_w (the number of on-chip wires connecting the TSV to the I/O driver), the data-rate is determined by reliability concerns, and improves rapidly with an increase in the value of N_w. However, beyond a certain value of N_w, the data-rate is limited by delay. Furthermore, the total capacitance increases with an increase in the value of N_w. As a result, the energy required to transmit one bit over the channel increases linearly with N_w. To get the best performance per unit energy supplied to the link, we maximize the compound metric ($F_{\mathrm{dr,eff}}$/EPB) [19], [20], where EPB is the energy per bit given by (7). By choosing the optimal number of wires N_w that maximizes the compound metric, we not only avoid severely limiting the data-rate due to reliability concerns, but also avoid using a large number of wires for a small improvement in data-rate

$$\mathrm{EPB} = 0.5 C_{\mathrm{tot}} V_{\mathrm{dd}}^2 \quad (7)$$

$$C_{\mathrm{tot}} = C_{\mathrm{dr}} + C_{rx} + C_{\mathrm{tsv}} + c_{tx} L_{tx} + c_{rx} L_{rx}. \quad (8)$$

III. System-Level Modeling

In Section II, it was shown that on-chip interconnects can be critical in determining the performance of 3-D links. In this section, the trade-off between on-chip interconnect length and TSV density is studied by comparing the two structures shown in Fig. 2. The structure in Fig. 2(a) (Structure 1) aims to pack as many TSVs as possible in a large TSV array, but suffers due to longer on-chip wires. On the other hand, the structure in Fig. 2(b) (Structure 2) aims to reduce the on-chip wire length at the expense of smaller area available for TSVs. Although Structure 1 is an unlikely structure to be implemented in a real chip, it gives us an upper limit to how densely TSVs can be packed. Mathematically, the number of TSVs and the worst-case on-chip wire lengths for Structures 1 and 2 are given by

$$N_{s1} = \left\lfloor \frac{X_{\max} - K_{\mathrm{oz}}}{D+S} \right\rfloor \left\lfloor \frac{Y_{\max}}{D+S} \right\rfloor \quad (9)$$

$$N_{s2} = M \left\lfloor \frac{X_{\max}}{D+S} \right\rfloor \left\lfloor \frac{Y_{\max}}{H_{I/O} + MD + (M-1)S + 2K_{\mathrm{oz}}} \right\rfloor \quad (10)$$

$$L_{s1} = K_{\mathrm{oz}} + \left(\left\lfloor \frac{X_{\max} - K_{\mathrm{oz}}}{D+S} \right\rfloor - 1 \right) (D+S) \quad (11)$$

$$L_{s2} = K_{\mathrm{oz}} + \left\lfloor \frac{M-1}{2} \right\rfloor (D+S) \quad (12)$$

where $\lfloor . \rfloor$ is the floor function, N_{s1} is the number of TSVs in Structure 1, $X_{\max} Y_{\max}$ represents the maximum area available for TSVs, K_{oz} is the keep-out zone, D and S are the TSV diameter and spacing, respectively, M is the maximum number of rows in the TSV array of Structure 2, and $H_{\mathrm{I/O}}$ is the height of I/O cell for a given technology. In this analysis, because the driver and the receiver are assumed to be inverters, it is assumed that the I/O cell height is the same as the standard cell height for the technology node. The aggregate bandwidth of the two structures is given by

$$BW_{s1,2} = N_{s1,2} F_{\mathrm{dr,eff}}(L_{s1,2}, L_{s1,2}). \quad (13)$$

IV. Results and Discussion

In this section, the models developed in sections II and III are used to compute the upper limits to the performance of 3-D ICs. Two values of TSV diameter are considered for

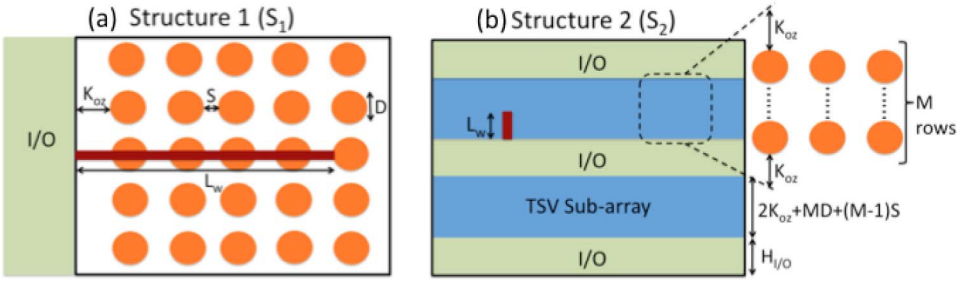


Fig. 2. (a) Schematic showing the top view of Structure 1, where TSVs are packed tightly, but on-chip wires are long. (b) Schematic showing the top view of Structure 2, where the available area is divided into multiple rectangular TSV arrays, with a few rows of standard cells between them for I/O placement.

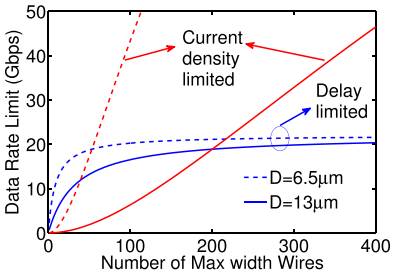


Fig. 3. Maximum data-rate of a 3-D link limited by the maximum current density allowed ($F_{\mathrm{dr,rel}}$) and by the delay of the link ($F_{\mathrm{dr,del}}$), as a function of the number of maximum width wires. TSV diameters of 13 and 6.5 μm are considered, local on-chip interconnect length of 100 gate pitches is assumed, and the circuit/interconnect parameters are obtained from the ITRS projections for the 7.5-nm node.

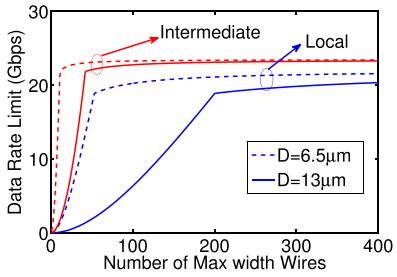


Fig. 4. Effective maximum data-rate of a 3-D link ($F_{\mathrm{dr,eff}}$), derived from Fig. 3 for local interconnects, and similarly for intermediate interconnects. TSV diameters of 13 and 6.5 μm are considered, and the circuit/interconnect parameters are obtained from the ITRS projections for the 7.5-nm node.

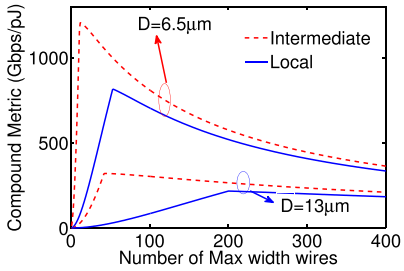


Fig. 5. Compound metric ($F_{\mathrm{dr,eff}}$/EPB) as a function of the number of maximum width wires. TSV diameters of 13 and 6.5 μm are considered, and the circuit/interconnect parameters are obtained from the ITRS projections for the 7.5-nm node.

the analysis, 13 and 6.5 μm. The TSV pitch is assumed to be 2$\times$ the TSV diameter, the aspect ratio is assumed to be 10, the liner thickness is assumed to be 0.38 μm, and the keep-out zone is assumed to be half the diameter. The TSV diameter of 13 μm and the liner thickness of 0.38 μm correspond to the dimensions demonstrated in part II of this paper. The technology scaling parameters are derived from ITRS projections [15]. The on-chip interconnect spacing is assumed to be equal to the width. The minimum allowed width of the intermediate interconnect is assumed to be twice the minimum allowed width of the local interconnect. The I/O driver size is determined by the rise/fall time requirement at the driver output, assumed to be 135 ps.

A. Data-Rate Limit of a 3-D IC

The upper bound of the data-rate of a 3-D IC link limited by reliability concerns and delay, as a function of effective interconnect width (or equivalently the number of maximum width wires N_w) is shown in Fig. 3. At smaller N_w, the data-rate limited by the maximum current density is smaller compared with the data-rate limited by delay. This is because, at smaller N_w, the maximum current density limits how fast the signals can toggle without causing the physical breakdown of the wires. However, the current density limited data-rate increases rapidly with N_w, thus making the delay limited data-rate dominates at larger widths. At smaller N_w, the product of driver side wire resistance and TSV capacitance dominates the delay of the 3-D link; hence, the delay improves significantly with an increase in the value of N_w. However, beyond a certain width, the improvement in resistance is nullified by the increase in the wire capacitance; as a result, maximum data-rate saturates with an increase in the value of N_w, as shown in Fig. 3. The effective data-rate limit as a function of N_w is shown in Fig. 4. When the TSV diameter is larger, large current is drawn to charge the large TSV capacitance; as a result, a large number of wires are needed to carry the large currents and meet the current density requirement. Thus, the crossover between current density limited and delay limited data-rate occurs at larger value of N_w, as shown in Fig. 3. Furthermore, because the local interconnects have smaller thickness compared with the intermediate interconnects, larger number of wires are needed to support the same current density; hence, the crossover width is smaller for the case where on-chip wires are routed in an

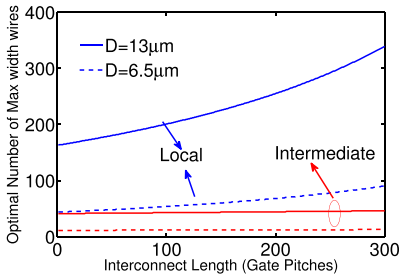


Fig. 6. Optimal number of maximum width wires as a function of interconnect length in gate pitches. TSV diameters of 13 and 6.5 μm are considered, and the circuit/interconnect parameters are obtained from the ITRS projections for the 7.5-nm node.

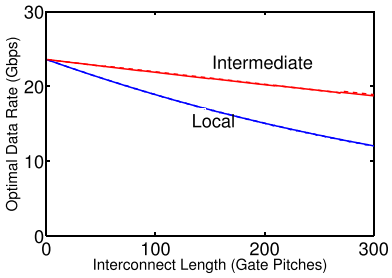


Fig. 7. Optimal data-rate as a function of interconnect length in gate pitches. The circuit/interconnect parameters are obtained from the ITRS projections for the 7.5-nm node.

intermediate layer. Because the maximum data-rate saturates with the number of wires, increasing the number of wires beyond a certain value only increases the total capacitance and the energy per bit; as a result, a compound metric that combines data-rate and energy per bit is maximized to obtain the optimal number of wires, as shown in Fig. 5. The optimal number of wires as a function of interconnect length in gate pitches is shown in Fig. 6. With an increase in the interconnect length, the total capacitance and the driver size increase; hence, a larger number of wires is necessary to carry the larger current. However, for intermediate-level interconnects, the increase in the current density is not significant; as a result, the optimal number of wires does not increase significantly. The optimal data-rate decreases with an increase in the interconnect length, as shown in Fig. 7. This is due to the increases in the total capacitance and the driver size, which increase the current density.

B. Impact of Technology Scaling

The optimal driver size and the data-rate as a function of ITRS technology node for an on-chip interconnect length of 100 gate pitches are shown in Fig. 8. The improvement in the ITRS projected current carrying capacity and the lower gate pitch with technology scaling ensure that local interconnects need fewer wires and smaller drivers at advanced technology nodes. However, because intermediate wires have lower interconnect capacitance due to larger dielectric height and smaller optimum number of wires, the impact of technology scaling on the optimal number of wires and

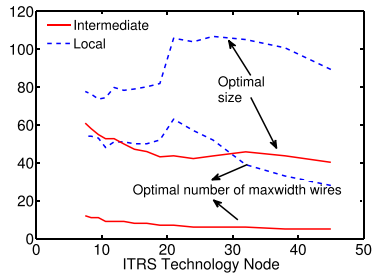


Fig. 8. Optimal number of maximum width wires and optimal driver size as a function of ITRS technology node. TSV diameters of 13 and 6.5 μm are considered, and the interconnect length is 100 gate pitches.

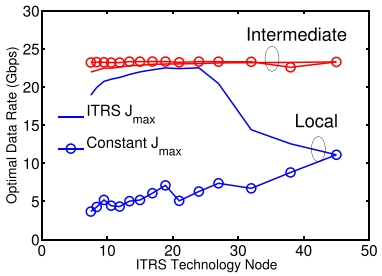


Fig. 9. Optimal data-rate as a function of ITRS technology node, computed assuming constant $J_{\max}$ and $J_{\max}$ targets projected by ITRS. The interconnect length is 100 gate pitches.

optimal driver size is smaller. The optimal data-rate as a function of technology node for an on-chip interconnect length of 100 gate pitches is shown in Fig. 9. If the ITRS projections for current density with technology scaling are met, the optimal data-rate of a 3-D link using local interconnect improves with technology scaling. However, if the industry is unable to keep up with the ITRS projections for maximum current density, the optimal data-rate decreases with technology scaling. As a result, current carrying capacity of on-chip wires is a key factor in determining the performance advantage of technology scaling in 3-D IC links.

C. System-Level Analysis

The models developed in section III are applied here to study the impact of on-chip wires on the performance of a 3-D IC. The key trade-off between the two structures shown in Fig. 2 is that S_1 has long on-chip wires, but S_2 has fewer TSVs due to the overhead of keep-out zone on the top and bottom of the TSV array. In spite of the smaller number of TSVs, S_2 achieves a larger bandwidth density compared with S_1, as shown in Fig. 10. Furthermore, it is clear that increasing the number of rows in the TSV array (M) adversely impacts the performance due to the increase in the on-chip wire length, as shown in Fig. 10. Furthermore, the worst-case wire length is the same for $M = 2n - 1$ and $M = 2n$, because the I/O cells can be placed above or below the TSV array. However, the overhead of keep-out zones is smaller for $M = 2n$; hence, the bandwidth density is slightly higher for $M = 2n$. The optimal number of rows in the TSV array is 2, as shown in Fig. 2(b). Because the total capacitance and the on-chip interconnect length for the

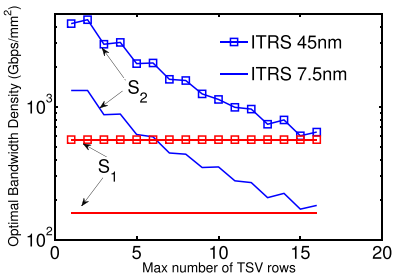


Fig. 10. Bandwidth density as a function of number of rows in the TSV array of Structure 2 (M).

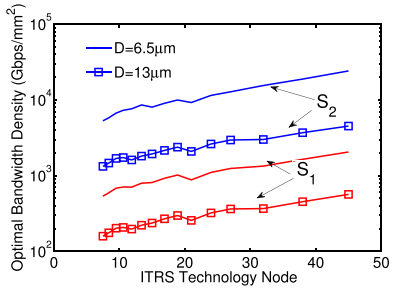


Fig. 11. Bandwidth density as a function of ITRS technology node for Structures 1 and 2.

two structures decrease considerably with TSV scaling, the bandwidth density improves significantly, as shown in Fig. 11. However, if TSV dimensions do not scale with technology, bandwidth density degrades with technology scaling. As a result, scaling TSV dimensions to reduce the TSV capacitance and the on-chip interconnect length is essential to improve the overall performance of a 3-D link.

V. Conclusion

The importance of on-chip interconnects on the performance of 3-D ICs with TSVs is quantified in this paper. It is shown that the performance of a 3-D link is limited not only by the on-chip interconnect RC, driver resistance and the TSV capacitance, but also by the current carrying capacity of the on-chip wires connecting the TSV to the I/O driver. This is because the large TSV capacitance results in a huge charging/discharging current, which has to be supported by the on-chip interconnects connecting the driver to the TSV. Because a large current is drawn by the TSVs, it is shown that a significant number of on-chip wires are required to reliably connect the I/O drivers to the TSVs and achieve optimal data-rates. Furthermore, this paper presents a method to co-optimize the driver size, on-chip interconnect, and data-rate to meet the reliability constraints and maximize the aggregate bandwidth per unit energy supplied to the link. Based on the models developed, it is shown that smaller TSV arrays with distributed I/O drivers are better compared with large TSV arrays with clustered I/O drivers because of the shorter on-chip wires connecting the TSV to the driver. Finally, if on-chip wires in the local metal levels are used to connect the TSV to the driver, the performance of 3-D ICs is shown to degrade significantly with technology scaling unless the current carrying capacity of the wires is improved with every technology generation.

References

[1] G. E. Moore, "Cramming more components onto integrated circuits," *Electronics*, vol. 38, no. 8, pp. 114–117, Apr. 1965.
[2] R. H. Dennard, F. H. Gaensslen, H.-N. Yu, V. L. Rideovt, E. Bassous, and A. R. Leblanc, "Design of ion-implanted MOSFET's with very small physical dimensions," *IEEE Solid-State Circuits Soc. Newslett.*, vol. 12, no. 1, pp. 38–50, Oct. 1974.
[3] G. Balamurugan *et al.*, "A scalable 5–15 Gbps, 14–75 mW low-power I/O transceiver in 65 nm CMOS," *IEEE J. Solid-State Circuits*, vol. 43, no. 4, pp. 1010–1019, Apr. 2008.
[4] H. Cho, P. Kapur, and K. C. Saraswat, "Power comparison between high-speed electrical and optical interconnects for interchip communication," *J. Lightw. Technol.*, vol. 22, no. 9, pp. 2021–2033, Sep. 2004.
[5] D. A. B. Miller, "Rationale and challenges for optical interconnects to electronic chips," *Proc. IEEE*, vol. 88, no. 6, pp. 728–749, Jun. 2000.
[6] J. Kim *et al.*, "High-frequency scalable electrical model and analysis of a through silicon via (TSV)," *IEEE Trans. Compon., Packag., Manuf. Technol.*, vol. 1, no. 2, pp. 181–195, Feb. 2011.
[7] T. O. Dickson *et al.*, "An 8 × 10-Gb/s source-synchronous I/O system based on high-density silicon carrier interconnects," *IEEE J. Solid-State Circuits*, vol. 47, no. 4, pp. 884–896, Apr. 2012.
[8] T. J. Spencer, P. J. Joseph, T. H. Kim, M. Swaminathan, and P. A. Kohl, "Air-gap transmission lines on organic substrates for low-loss interconnects," *IEEE Trans. Microw. Theory Techn.*, vol. 55, no. 9, pp. 1919–1925, Sep. 2007.
[9] G. Katti, M. Stucchi, K. De Meyer, and W. Dehaene, "Electrical modeling and characterization of through silicon via for three-dimensional ICs," *IEEE Trans. Electron Devices*, vol. 57, no. 1, pp. 256–262, Jan. 2010.
[10] T. Bandyopadhyay, R. Chatterjee, D. Chung, M. Swaminathan, and R. Tummala, "Electrical modeling of through silicon and package vias," in *Proc. IEEE Int. Conf. 3D Syst. Integr. (3DIC)*, Sep. 2009, pp. 1–8.
[11] E. Beyne *et al.*, "Through-silicon via and die stacking technologies for microsystems-integration," in *Proc. IEEE Int. Electron Devices Meeting (IEDM)*, Dec. 2008, pp. 1–4.
[12] I. Ndip *et al.*, "High-frequency modeling of TSVs for 3-D chip integration and silicon interposers considering skin-effect, dielectric quasi-TEM and slow-wave modes," *IEEE Trans. Compon., Packag., Manuf. Technol.*, vol. 1, no. 10, pp. 1627–1641, Oct. 2011.
[13] M. A. Karim, P. D. Franzon, and A. Kumar, "Power comparison of 2D, 3D and 2.5D interconnect solutions and power optimization of interposer interconnects," in *Proc. IEEE 63rd Electron. Compon. Technol. Conf. (ECTC)*, May 2013, pp. 860–866.
[14] G. Lopez, J. Davis, and J. Meindl, "A new physical model and experimental measurements of copper interconnect resistivity considering size effects and line-edge roughness (LER)," in *Proc. IEEE Int. Interconnect Technol. Conf.*, Jun. 2009, pp. 231–234.
[15] *International Technology Roadmap for Semiconductors-2010 Update, PIDS2 Table*. [Online]. Available: http://www.itrs.net/
[16] H. S. Kamineni *et al.*, "Challenges to via middle TSV integration at sub-28 nm nodes," in *Proc. IEEE Int. Interconnect Technol. Conf.*, May 2014, pp. 199–202.
[17] A. M. Martinez, "Quick estimation of transient currents in CMOS integrated circuits," *IEEE J. Solid-State Circuits*, vol. 24, no. 2, pp. 520–531, Apr. 1989.
[18] J. Rubinstein, P. Penfield, and M. A. Horowitz, "Signal delay in RC tree networks," *IEEE Trans. Comput.-Aided Design Integr.*, vol. 2, no. 3, pp. 202–211, Jul. 1983.
[19] V. Kumar, R. Bashirullah, and A. Naeemi, "Modeling, optimization and benchmarking of chip-to-chip electrical interconnects with low loss air-clad dielectrics," in *Proc. IEEE 61st Electron. Compon. Technol. Conf. (ECTC)*, May/Jun. 2011, pp. 2084–2090.
[20] V. Kumar *et al.*, "Airgap interconnects: Modeling, optimization, and benchmarking for backplane, PCB, and interposer applications," *IEEE Trans. Compon., Packag., Manuf. Technol.*, vol. 4, no. 8, pp. 1335–1346, Aug. 2014.

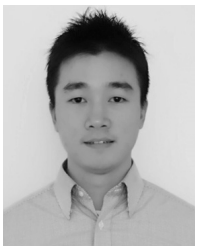
Vachan Kumar (S'10–M'14) received the B.E. degree in electronics and communication engineering from the National Institute of Technology Karnataka, Surathkal, India, in 2006, and the M.S. and Ph.D. degrees in electrical and computer engineering from the Georgia Institute of Technology, Atlanta, GA, USA, in 2012 and 2014, respectively.

He is currently with Logic Technology Development, Intel Corporation, Hillsboro, OR, USA, where he is involved in technology benchmarking.

Hanju Oh (S'14) received the B.S. degree in electrical and computer engineering from Hanyang University, Seoul, South Korea, in 2010, and the M.S. degree from the Georgia Institute of Technology, Atlanta, GA, USA, in 2014, where he is currently pursuing the Ph.D. degree in electrical and computer engineering.

Xuchen Zhang (S'15) received the B.S. degree in electronics engineering and the M.S. degree in microelectronics from Shanghai Jiao Tong University, Shanghai, China, in 2008 and 2011, respectively, and the M.S. degrees in electrical engineering from the Georgia Institute of Technology, Atlanta, GA, USA, in 2012, where he is currently pursuing the Ph.D. degree.

His current research interests include 3-D electronic system integration, advanced packaging, and embedded microfluidic cooling.



Li Zheng received the B.S. degree from Zhejiang University, Hangzhou, China, in 2006, the M.S. degree from Shanghai Jiao Tong University, Shanghai, China, in 2009, and the Ph.D. degree in electrical and computer engineering from the Georgia Institute of Technology, Atlanta, GA, USA, in 2015.

His current research interests include advanced 3-D packaging, numerical power distribution network modeling, and microfluidic cooling for high-performance computing systems.

Muhannad S. Bakir (SM'12) received the B.E.E. degree from Auburn University, Auburn, AL, USA, in 1999, and the M.S. and Ph.D. degrees in electrical and computer engineering from the Georgia Institute of Technology (Georgia Tech), Atlanta, GA, USA, in 2000 and 2003, respectively.

He is currently an Associate Professor with the School of Electrical and Computer Engineering, Georgia Tech.

Azad Naeemi (S'99–M'04–SM'04) received the B.S. degree in electrical engineering from Sharif University, Tehran, Iran, in 1994, and the M.S. and Ph.D. degrees in electrical and computer engineering from the Georgia Institute of Technology (Georgia Tech), Atlanta, GA, USA, in 2001 and 2003, respectively.

He joined as a Faculty Member with the School of Electrical and Computer Engineering, Georgia Tech, where he is currently an Associate Professor.