

Interconnect-Based Design Methodologies for Three-Dimensional Integrated Circuits

Vertical integration is a novel communications paradigm where interconnect design is a primary focus.

By VASILIS F. PAVLIDIS, *Student Member IEEE*, AND EBY G. FRIEDMAN, *Fellow IEEE*

ABSTRACT | Design techniques for three-dimensional (3-D) ICs considerably lag the significant strides achieved in 3-D manufacturing technologies. Advanced design methodologies for two-dimensional circuits are not sufficient to manage the added complexity caused by the third dimension. Consequently, design methodologies that efficiently handle the added complexity and inherent heterogeneity of 3-D circuits are necessary. These 3-D design methodologies should support robust and reliable 3-D circuits while considering different forms of vertical integration, such as system-in-package and 3-D ICs with fine grain vertical interconnections. Global signaling issues, such as clock and power distribution networks, are further exacerbated in vertical integration due to the limited number of package pins, the distance of these pins from other planes within the 3-D system, and the impedance characteristics of the through silicon vias (TSVs). In addition to these dedicated networks, global signaling techniques that incorporate the diverse traits of complex 3-D systems are required. One possible approach, potentially significantly reducing the complexity of interconnect issues in 3-D circuits, is 3-D networks-on-chip (NoC). Design methodologies that exploit the diversity of 3-D structures to further enhance the performance of multiplane integrated systems are necessary. The longest interconnects within a 3-D

circuit are those interconnects comprising several TSVs and traversing multiple physical planes. Consequently, minimizing the delay of the interplane nets is of great importance. By considering the nonuniform impedance characteristics of the interplane interconnects while placing the TSVs, the delay of these nets is decreased. In addition, the difference in electrical behavior between the horizontal and vertical interconnects suggests that asymmetric structures can be useful candidates for distributing the clock signal within a 3-D circuit. A 3-D test circuit fabricated with a 180 nm silicon-on-insulator (SOI) technology, manufactured by MIT Lincoln Laboratories, exploring several clock distribution topologies is described. Correct operation at 1 GHz has been demonstrated. Several 3-D NoC topologies incorporating dissimilar 3-D interconnect structures are reviewed as a promising solution for communication limited systems-on-chip (SoC). Appropriate performance models are described to evaluate these topologies. Several forms of vertical integration, such as system-in-package and different candidate technologies for 3-D circuits, such as SOI, are considered. The techniques described in this paper address fundamental interconnect structures in the 3-D design process. Several interesting research problems in the design of 3-D circuits are also discussed.

KEYWORDS | Interconnect design methodologies; physical design; vertical integration; 3-D ICs; 3-D integration

Manuscript received January 31, 2008; revised July 2, 2008. Current version published February 27, 2009. This work was supported in part by the National Science Foundation under Contract CCF-0541206, by the New York State Office of Science, Technology and Academic Research under a grant to the Center for Advanced Technology in Electronic Imaging Systems, by Intel Corporation, Eastman Kodak Company, and Freescale Semiconductor Corporation, and foundry support from MIT Lincoln Laboratories. The authors are with the Department of Electrical and Computer Engineering, University of Rochester, Rochester, NY 14627 USA (e-mail: pavlidis@ece.rochester.edu; friedman@ece.rochester.edu).

Digital Object Identifier: 10.1109/JPROC.2008.2007473

I. INTRODUCTION

The majority of manufacturing technologies for three-dimensional (3-D) integration include die or wafer bonding, resulting in dense polyolithic systems where standard field-effect transistors are utilized to implement

logic functions [1]–[9]. Consequently, the intrinsic speed of a logic gate in 3-D circuits remains constant, while the interconnect performance can be significantly improved by vertically stacking the connected planes as compared to traditional two-dimensional (2-D) circuits. Vertical integration is therefore a novel communications paradigm where interconnect design is a primary focus. Since the interconnect segments are vertical, design methodologies that incorporate the signal propagation characteristics across the third dimension are necessary.

The inherent advantage of 3-D integration is the drastic decrease in interconnect length, particularly the long global interconnects, which directly results in increased speed [10]–[14]. The interconnect power is also reduced as the capacitance of the wires decreases [15], [16]. Additionally, the total power dissipated by an interconnect system is further decreased as the number of repeaters inserted along the interconnect is reduced [17]. Finally, coupling among intraplane adjacent interconnects is lower due to decreased length, improving signal integrity.

Another characteristic of 3-D ICs of even greater importance than the decrease in the interconnect length is the ability of these systems to include disparate technologies [18], greatly extending the capabilities of modern systems-on-chip (SoC). This defining feature of 3-D ICs offers unique opportunities for highly heterogeneous and sophisticated systems [19], [20]. A vast pool of applications such as medical, wireless communications, military, and low-cost consumer products, exists for vertical integration, as the proximity of the system components caused by the third dimension is suitable for either the high performance or low power ends of the SoC application space [114]. This heterogeneity, however, greatly complicates the interconnect design process within a multiplane system, as potential design methodologies need to manage the diverse interconnect impedance characteristics and process variations caused by the different fabrication processes and technologies employed in the different physical planes.

Consider, for example, a signal that traverses several digital and analog circuits on separate planes. In this case, the signal can behave both as an aggressor and as a victim. Consequently, to reliably propagate a signal across multiple physical planes at high speed is a difficult task. Another important global signaling issue is distributing the clock signal to the sequential elements on each plane of the 3-D stack. Delivering sufficient current to each circuit in a 3-D system is also a challenging problem. The number of package pins also limits the clock signal and power supply networks. Additional primary challenges in 3-D circuits include the development of methodologies at the front end of the design process [21], [22], multiplane functional testing [23], thermal management techniques [68], and maturing manufacturing technologies [24]–[28].

A system-in-package (SiP) can be described as an assemblage of either bare or packaged dice along the third dimension, where the interconnections through the z-axis are primarily implemented with solder balls, wire bonds, and vertical interconnects that penetrate through the silicon substrate to connect the circuits. The latter type of interconnects is typically called a through silicon via (TSV) [29]–[35]. This term is generalized in this paper to represent any type of interconnect that penetrates through the substrate to connect the circuits within a 3-D system. A system-in-package primarily includes standalone components or subsystems that are vertically integrated to comprise a complex multiplane system. The predominant benefits of SiP are the increased packaging efficiency as compared to 2-D integrated systems and shorter off-chip interconnects. The deleterious effects of the long on-chip interconnects, however, are not mitigated. These issues are effectively resolved by another form of vertical integration, 3-D ICs.

Three-dimensional circuits can be conceptualized as the bonding of multiple wafers or bare dice. The distinctive difference between an SiP and a 3-D IC is the granularity of the vertical interconnects. Different bonding styles between the planes within a 3-D system are also possible: front-to-front, back-to-front, and back-to-back [3], [36]. Examples of SiP structures and various bonding styles for 3-D circuits are illustrated in Fig. 1. Each of these bonding styles is likely to include through silicon vias (or interplane vias) with different physical dimensions. Consequently, the density of the vertical interconnects can vary not only among different 3-D circuits but also among the physical planes within a 3-D circuit. Another approach for vertical communication within 3-D circuits is through electromagnetic coupling. Both capacitive and inductive communication techniques have been developed for contactless 3-D circuits [37]–[39]. Although these techniques reduce or eliminate the need for galvanic connections among the planes of a 3-D circuit, issues such as dc power supply, interference from adjacent signals, and the physical area of

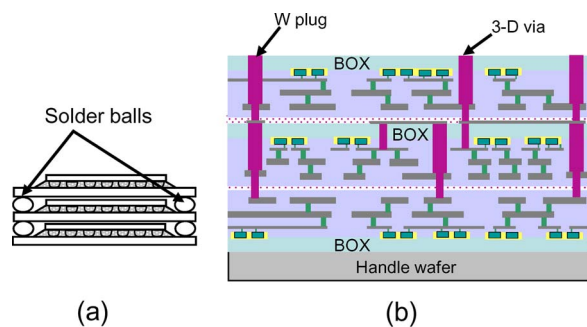


Fig. 1. Different forms of 3-D integration (not to scale). (a) System-in-package [32] and (b) a 3-D circuit with dense through silicon vias [8], [9]. Two different bonding styles, front-to-front and back-to-front, are illustrated.

the inductors and capacitors limit the applicability of these techniques [38].

Monolithic 3-D ICs include layers of planar devices successively grown on a conventional complementary metal-oxide-semiconductor (CMOS) or SOI plane [40]. Monolithic 3-D circuits support transistor-level integration. Consequently, the devices within a logic gate can be placed on different layers and, more importantly, implemented with different technologies such as CMOS or SOI. Independent of the technology utilized for the first device layer, these transistors are fabricated with conventional and mature processes. For the devices on the upper planes, however, different fabrication methods are required. Several techniques, based on laser recrystallization [41] or seed crystallization [42], are used to produce CMOS or SOI devices on the upper planes.

The major drawback of these circuits, however, is the inability to produce high-quality devices, specifically on the upper planes [41], [43]. In addition, the growth of the devices on the upper planes usually requires high temperatures. Insulator layers are used to protect the transistors on the first plane [40]. These insulators, however, further encumber the heat removal process with negative consequences for a monolithic 3-D circuit. Consequently, design methodologies for this type of 3-D circuit are not considered in this paper.

Physical and interconnect design techniques for 3-D circuits are the main focus of this paper. A variety of issues related to 3-D circuits, such as the complexity of these systems and novel interconnect structures, are discussed in Section II. The characteristics and related technological traits of the TSVs are also discussed in this section. Floorplanning techniques that consider the characteristics of 3-D circuits are presented in Section III. Different approaches have been proposed to floorplan these circuits and are discussed in this section. The placement of 3-D circuits targeting various performance objectives is reviewed in Section IV. The complex task of routing the multitude of individual cells within a 3-D structure is discussed in Section V. The disparity among the vertical interconnect impedance characteristics that can exist within a 3-D circuit can be exploited to further improve the performance of these circuits. Careful placement of the vertical interconnects is required, as highlighted in Section VI. The important and unexplored task of synchronization in 3-D circuits is discussed in Section VII. A 3-D test circuit investigating several different clock distribution networks is also described. The concept of 3-D networks-on-chip (NoC) for improving the communication throughput within a system-on-chip while reducing interconnect design complexity is presented in Section VIII. Several topologies and related improvements in the speed and power consumed by these global interconnects are described in this section. A brief description of several open research issues related to the 3-D interconnect design process along with some conclusions are offered in Section IX.

II. PHYSICAL DESIGN ISSUES IN 3-D CIRCUITS

The introduction of the third dimension has significantly increased the complexity of the integrated circuit design process. The characteristics of the vertical interconnects and the constraints that this type of interconnect poses on the physical design process are outlined in Section II-A. Various physical design approaches to manage thermal issues are highlighted in Section II-B. A discussion of complexity in the design of 3-D circuits is presented in Section II-C.

A. Vertical Interconnects

Increasing the number of planes that can be integrated into a single 3-D system is a primary objective of three-dimensional integration. A 3-D system with high-density vertical interconnects is therefore indispensable. Vertical interconnects implemented as TSVs produce the highest interconnect bandwidth within a 3-D system, as compared to wire bonding, peripheral vertical interconnects, and solder-ball arrays. Alternatively, the density of this type of interconnect dictates the granularity of the interconnected planes of the system, directly affecting the interplane communication bandwidth. Other important criteria should also be satisfied by the TSV fabrication process. A fabrication process for vertical interconnects should produce reliable and inexpensive TSVs. A high TSV aspect ratio, the ratio of the diameter of the top edge to the length of the via, may also be required for certain types of 3-D circuits. The effect of forming the TSVs on the performance and reliability of neighboring active devices should also be negligible.

The electrical characteristics of the TSVs are of primary importance in 3-D ICs and are considerably different from the horizontal interconnect segments [44], as described by recent electrical models [45]. This situation is due to the structure of these interconnects and the diverse technologies, such as CMOS and SOI, that can exist in a 3-D system. Producing low resistance and capacitance TSVs is a fundamental objective of manufacturing technologies. Finally, not properly characterizing the contribution of the TSVs to the delay of the critical interplane interconnect can result in significant inaccuracy in the performance of a 3-D system [46]. Consequently, these structures must be carefully considered during the 3-D physical design process. Examples of a TSV used in CMOS and SOI circuits are illustrated in Fig. 2(a) and (b), respectively. The impedance and physical characteristics of these structures are listed in Table 1. A pitch equal to twice the diameter of the TSV is assumed to determine the density of the TSV for these processes where this dimension is not provided.

The thermal traits of the TSVs are also significant, as these vias can affect the thermal behavior of a 3-D IC. TSVs can be used to provide high thermal conductivity paths to facilitate the flow of heat from the upper planes to the

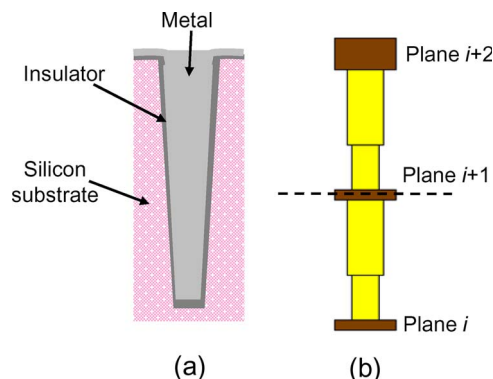


Fig. 2. Examples of through silicon vias (not to scale) used in (a) SiP and 3-D CMOS technologies [28], [47] and (b) 3-D SOI processes [9].

plane attached to the heat sink, maintaining the temperature of a 3-D circuit within acceptable levels. Materials with low thermal resistance, such as copper, are therefore preferred.

B. Thermal Issues in 3-D ICs

A 3-D system consists of disparate materials with considerably different thermal properties including semiconductor, metal, dielectric, and possibly polymer layers used for plane bonding. Although the power consumption of these circuits is expected to decrease due to the considerably shorter interconnects, the power density increases since there is a greater number of devices per unit volume as compared to a 2-D circuit. As the power density increases, the temperature of the planes nonadjacent to the heat sink of the package can rise, resulting in degraded performance or thermal gradients that can accelerate wear out mechanisms [49]–[51]. Design methodologies at various stages of the IC design flow, such as synthesis, floorplanning, and placement and routing, which maintain the temperature of a circuit within specified limits or alleviate thermal gradients among the planes of the 3-D circuit, are therefore necessary.

Two key elements are required to establish a successful thermal management strategy: a thermal model, to characterize the thermal behavior of a circuit, and design techniques that alleviate thermal gradients among the physical planes of a 3-D stack while maintaining the operating temperature within acceptable levels. The primary

requirements of a thermal model are high accuracy and low complexity [52]–[54], while thermal design techniques should produce high-quality circuits without incurring long computational design time [55]. To reduce the complexity of the modeling process, standard methods to analyze heat transfer, such as finite difference, finite element, and boundary element methods, have been adopted to evaluate the temperature of a 3-D circuit. Simpler analytic expressions have also been developed to characterize the temperature within a 3-D system.

Thermal design techniques can be classified into two categories: thermal strategies that improve the thermal profile of a 3-D circuit without requiring any redundant interconnect resources for thermal management and those methodologies that are an integral part of a more aggressive thermal policy that utilize thermal TSVs, sacrificing other design objective(s). These TSVs are typically called thermal or dummy vias [18] to emphasize the objective of conveying heat rather than providing signal communication for circuits located on different physical planes. Thermal wires can also be employed to transfer heat [56]. Thermal wires correspond to those horizontal wires that connect regions with different thermal via densities through thermal interplane vias.

C. Complexity of 3-D Physical Design Process

The solution space for classical physical design methodologies increases significantly in 3-D systems, as the physical distance of two circuit cells is reduced not only by placing these cells near each other on the same plane but also by placing the cells in vertically adjacent locations. This situation results in a formidable increase in the number of solutions that can be explored, resulting in an exponential growth in the computational time. The increase in the number of metal layers yields similar computational issues for the routing task [57]. Computationally efficient heuristic algorithms are the primary tool to manage the dramatic increase in the solution space for 3-D circuits. Methods such as simulated annealing (SA) and genetic algorithms complete the mosaic of the 3-D physical design process [58], [59]. Another important issue is the efficient representation of the cell locations within a 3-D circuit to maintain sufficiently low storage requirements.

In addition, traditional objectives, such as wire length and area, are insufficient for 3-D circuits, particularly

Table 1 Impedance and Physical Characteristics of TSVs

Process	Depth [μm]	Diameter [μm]	Total resistance [$\text{m}\Omega$]	Density [$1/\text{mm}^2$]
[47]	25	4	140	$\sim 1.6 \times 10^4$
[22]	30	1.2	<350	$\sim 1.7 \times 10^5$
[28]	80	5/15	9.4/2.6	$\sim 1 \times 10^4 / 1.1 \times 10^3$
[28]	150	5/15	2.7/1.9	$\sim 1 \times 10^4 / 1.1 \times 10^3$
[45]	90	75	2.4	~ 44
[9], [48]	~ 12	1.75	148	$\sim 8.2 \times 10^4$

heterogeneous multiplane integrated systems. Since these systems can combine disparate technologies, such as radio-frequency (RF), analog, and digital circuits, other objectives, such as noise and signal integrity, need to be simultaneously considered in addition to conventional objectives.

These objectives require the synergistic development of design methodologies, which previously were individually developed for each type of circuit. Furthermore, design aids that employ these novel techniques will be necessary. To date, the lack of these methods has considerably limited the evolution of heterogeneous 3-D systems and deprived three-dimensional circuits from exploiting important manufacturing improvements that have recently been achieved.

III. FLOORPLANNING FOR 3-D CIRCUITS

The predominant design objective for floorplanning a circuit has traditionally been to achieve the minimum area or, alternatively, the maximum packing density while interconnecting these blocks with minimum length wires. Most floorplanning algorithms can be classified as either slicing [60] or nonslicing [61], [62]. Floorplanning techniques belonging to both of these categories have been proposed for 3-D circuits [63]–[66]. An efficient floorplanning technique for 3-D circuits should adequately handle two important issues: representation of the third dimension and the related increase in the solution space. Floorplanning techniques for 3-D circuits that address these issues are discussed in this section. Multiobjective techniques are also reviewed.

Notating the location (i.e., the x , y , and z coordinates) and dimensions (i.e., width, length, and height) of the circuit cells in a volumetric system typically requires a considerable amount of storage. A 3-D circuit, however, consists of a limited number of planes. Consequently, such a system can be described as an array of two-dimensional planes, where circuit cells are treated as rectangles that can be placed on any of the planes within a 3-D system [65]–[68]. The second challenge for 3-D floorplanning is to effectively explore the solution space, where a hierarchical approach can often be more efficient for floorplanning 3-D circuits than a flat approach.

In nonhierarchical floorplanning algorithms, the floorplanning process proceeds by assigning the cells to the planes of the stack followed by simultaneous intraplane and interplane cell swapping, potentially exploring the entire solution space. Interplane moves, however, result in a formidable increase in the solution space, directly affecting the computational time of a flat floorplanning algorithm. For example, assuming N (e.g., 100) cells of a 3-D system consisting of n (e.g., three) planes and applying a low overhead technique to denote the 3-D arrangement of the blocks, a flat floorplanning approach increases the number of candidate solutions by

$N^{n-1}/(n-1)!$ (e.g., 5000) times as compared to a 2-D circuit consisting of the same number of cells [65].

Alternatively, a hierarchical approach can be used to significantly reduce the number of candidate solutions, where a two-step solution to the floorplanning problem is followed. Initially, the circuit cells are assigned to the physical planes. In the second step, a simulated annealing based engine simultaneously generates the floorplan of each of the planes by only permitting intraplane moves, considerably decreasing the search space for the optimal floorplan [65]. An example of the increase in the solution space due to the third dimension is illustrated in Fig. 3(a) and (b). By adopting a two-step solution, the computational complexity and time can be considerably decreased. Similar hierarchical approaches can also be applied to other stages of the 3-D physical design process.

The partitioning scheme adopted in the initial step of the hierarchical approach plays a crucial role in determining the compactness of a particular floorplan, as interplane moves are not allowed when floorplanning the planes. Different partitions correspond to different subsets of the solution space which may exclude the optimal solution(s). The criterion for partitioning should therefore be carefully selected. Partitioning can, for example, be based on minimizing the estimated total wirelength of the system [69] and/or the number of vertical interconnects [70].

During floorplanning, the starting point for a simulated annealing engine is generated by randomly assigning the blocks to the planes of the system to balance the area of the individual planes [71]. The SA process progresses by swapping blocks between planes or changing the location

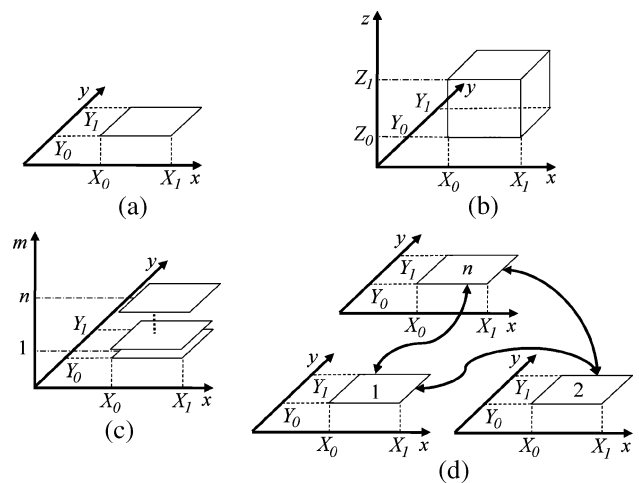


Fig. 3. Example of physical design solution space for floorplanning 2-D and 3-D circuits: (a) available area for floorplanning a planar circuit, (b) available volume for floorplanning a 3-D circuit, (c) a finite number of planes is considered to reduce the solution space, and (d) the floorplan of the planes is generated after the circuit cells are assigned to each plane. The arrows represent global constraints among planes that guide the floorplan of a 3-D system.

of the blocks within one plane. The expected wire length and number of vertical vias are reevaluated after each modification of the partition, where the algorithm progresses until the target solution is achieved at the desired low temperature of the SA algorithm. After the partition is complete, a candidate solution is perturbed by selecting a plane within the 3-D stack and applying a variety of intraplane moves [71]. Application of a hierarchical approach to the MCNC and GSRC benchmark suites¹ demonstrates a small reduction, on the order of 3%, in the number of vertical vias and a significant 14% reduction in wire length, as compared to nonhierarchical 3-D floorplanning techniques [68]–[71], [72].

The complexity of three-dimensional integration requires several dissimilar metrics for producing efficient floorplans for 3-D circuits beyond the use of traditional area and wire length metrics. These metrics can consider, for example, the communication throughput among the circuit blocks [36] or the number of interplane vias [68]. Techniques that include a thermal objective have also been developed [68]. The thermal objective typically aims at producing a uniform temperature distribution across each plane while peak temperatures are maintained sufficiently low. A multiobjective cost function inevitably increases the total computational runtime. A significant portion of this time is attributed to thermal profiling the 3-D circuit each time a candidate floorplan is generated. To reduce this time, simple thermal models are utilized, slightly degrading the quality of the solution [68].

IV. PLACEMENT FOR 3-D CIRCUITS

Placement algorithms have traditionally targeted minimizing the area of a circuit and the interconnect length among the cells, while reserving space for routing the interconnect. In vertical 3-D integration, a “placement dilemma” arises in deciding whether two circuit cells sharing a large number of interconnects can be more closely placed within the same plane or placed on adjacent physical planes, decreasing the interconnection length. Placing the circuit blocks on adjacent planes can often produce a line with the shortest wirelength to connect these blocks. An exception is the case of small blocks within an SiP where the length of the interplane vias is greater than 100 μm [24], [74]. Placement methodologies have also been discussed where other objectives, such as thermal gradients among the physical planes and the temperature of the planes [75], are considered.

Several approaches have been adopted for placing circuit cells within a volume [76]–[80]. Different types of circuit cells for various 3-D technologies have been investigated in [81]. Layout algorithms for these cells have also been devised, demonstrating the benefits of 3-D integration. Since TSVs consume silicon area, possibly

increasing the length of some interconnects, an upper bound on this type of interconnect resource is necessary. Alternatively, sparse utilization of the vertical interconnects can result in insignificant savings in wire length. To consider the effect of the vertical interconnects, a weighting factor has been used to increase the distance in the vertical direction, controlling the decision as to where to insert the interplane vias [78]. This weight essentially behaves as a controlling parameter that favors the placement of highly interconnected cells within the same or adjacent physical planes.

Alternatively, TSVs are treated as circuit cells since these interconnects occupy silicon area [82] and are included in the individual cell placement process within each plane. Since this approach can result in two different locations for placing a TSV, as illustrated in Fig. 4, a weighted average distance between these two locations can be utilized to place a TSV [82]. Although these approaches consider the location of the TSV, the fundamental objective is to decrease the interconnect length. The maximum achievable reduction in the interconnect length for the longest on-chip interconnect is proportional to \sqrt{n} , where n is the number of planes constituting a 3-D circuit [15]. Any further improvement in the performance of the interplane interconnects can be obtained by considering the electrical characteristics of the TSV. Such a TSV placement methodology is discussed in Section VI.

As with floorplanning, multiobjective placement techniques for 3-D circuits are necessary. Additional objectives that affect both the cell placement and wire length are simultaneously considered. The force directed method is a well-known technique used for cell placement [83], where repulsive or attractive forces are placed on the cells as if these cells are connected through a system of springs. The force directed method has been extended to incorporate the thermal objective during the placement process [84]. In this approach, repulsive forces are applied to those blocks that exhibit high temperatures (i.e., “hot blocks”) to ensure that the high-temperature blocks are placed at a greater distance from each other. The efficiency of this

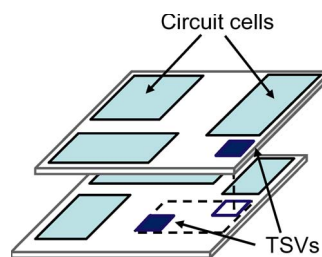


Fig. 4. Treating the TSVs as circuit cells on different planes can result in two different locations for placing a TSV. These locations define a region in which the TSV can be placed to satisfy different design objectives.

¹<http://www.cse.ucsc.edu/research/surf/GSRC/progress.html>.

force directed placement technique has been evaluated on the MCNC² and IBM-PLACE benchmarks,³ demonstrating a 1.3% decrease in the average temperature, a 12% reduction in the maximum temperature, and a 17% reduction in the average thermal gradient. The total wire length, however, increases by 5.5%. As demonstrated by these results, this technique primarily achieves a uniform temperature distribution across each plane, resulting in a significant decrease in thermal gradients as well as the maximum temperature. The average temperature throughout a 3-D IC, however, is only slightly decreased.

As a more practical example that demonstrates the need to include the thermal objective in 3-D physical design techniques, consider an Intel Pentium 4 processor, which has been redesigned in two planes [85]. The increased power density due to stacking can increase the peak temperature within the 3-D processor by approximately 26 °C, as compared to the original 2-D system if thermal issues are ignored [85]. This increase can significantly degrade the performance and reliability of the processor. If the thermal objective is incorporated during the placement process, a negligible 2 °C increase is observed [85].

Alternatively, additional TSVs that do not function as a signal path can be utilized to further enhance the heat transfer process. The design objective is to identify those regions where thermal vias are most needed (the hot spots) and place thermal vias within those regions at the appropriate density. Such an assignment, however, is mainly restricted by two factors; the routing blockage caused by these vias and the size of the unoccupied regions or white space that exist within each plane. Although thermal via insertion can be applied as a postplacement step, integrated techniques produce a more efficient distribution of the thermal TSVs for the same temperature constraint [69]. The integrated technique requires 16% fewer thermal vias for the same temperature constraint, with a 21% increase in computational time and an almost 3% reduction in total area.

V. ROUTING FOR 3-D CIRCUITS

Routing is the most complex and least developed of the physical design techniques used in 3-D circuits. The multiple metal layers available for routing on each physical plane exacerbate the difficulty in routing a net connecting several circuit cells located on different planes. As these interconnects also compete with the transistors for silicon area, routing is a formidable task for 3-D circuits. An early paper on routing 3-D circuits demonstrated several issues related to this physical design task [86]. Consequently, several heuristics have been developed that address routing in the third dimension [87], [88].

An effective approach for routing 3-D circuits is to convert the routing interplane interconnect problem into a 2-D channel routing task, as the 2-D channel routing problem has been efficiently solved [89], [90]. A number of methods can be applied to transform the problem of routing the interplane interconnects into a 2-D routing task, which requires utilizing a portion of the available routing resources for interplane routing (usually the top metal layers). Interplane interconnect routing can be implemented in five major stages including interplane channel definition, pseudoterminal allocation, interplane channel creation (channel alignment), detailed routing, and final channel alignment [87]. Additional stages route the 2-D channels, both the interplane and intraplane interconnects, and perform channel ordering to determine the wire routing order for the 2-D channels.

Alternatively, multilevel algorithmic techniques [91] have been applied to route 3-D circuits. The advantages of multilevel routing are the lower computational time and higher completion rates as compared to flat and hierarchical routers. Multilevel routing can be treated as a three-stage process, as illustrated in Fig. 5: a coarsening phase, an initial solution generation at the coarsest level (level p) of the grid, and a subsequent refinement process until the finest level of the grid is reached. Before the coarsening phase is initiated, the routing resources in each unit block of the grid are determined by a weighted area sum model. The routing resources are allocated during each coarsening step. The resources for the local nets within a block are transferred at each coarsening step. At the coarsest level, an initial routing tree is generated. This initial routing task commences with a minimum spanning tree for each multiterminal net. A Steiner tree heuristic and a maze searching algorithm generate a 3-D Steiner tree for each of these interconnects. Additionally, the TSVs are estimated for each block. During the last phase, the initial routing tree is refined until the finest level is reached. In this refinement phase, the signal (and thermal) TSVs are successively assigned and distributed within each block. The routing of the wires follows the refinement of the TSVs. At the finest level, a detailed router completes the routing of the circuit [91].

Although this technique offers a routing solution for standard cell and gate array circuits, alternative techniques that support different forms of vertical integration, for example, system-on-package (SoP), are also required. In an SoP, the routing problem can be described as connecting the I/O terminals of the blocks located on the planes of the SoP through interconnect and pin layers. For systems where the routing resources, such as the number of pin distribution layers, is limited, multiobjective routing is required to achieve a sufficiently small form factor. Other issues, such as integrating passive and active components, further enhance the demand for multiobjective routing approaches. A multiobjective approach can consider, for example, wire length, crosstalk, congestion, and/or area [88].

²<http://er.cs.ucla.edu/benchmarks/ibm-place>.

³http://www.cbl.ncsu.edu/pub/Benchmark_dirs/LayoutSynth92.

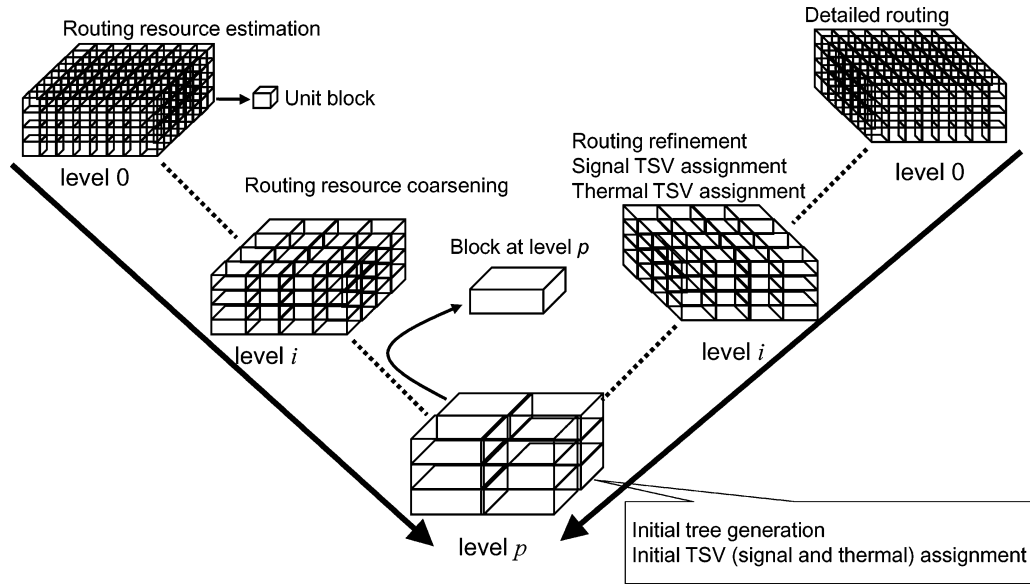


Fig. 5. Multilevel routing for 3-D circuits. The technique can be adapted to include multiple objectives for routing a 3-D circuit [91].

Multilevel routing for 3-D ICs has been extended to include the thermal objective [92], [93]. In addition to routing resources, the power density within each block of the grid is determined at each coarsening step. An initial TSV assignment to each block is implemented during the coarser step along with generation of an initial routing tree. The TSV assignment includes both signal and thermal TSVs, with priority given to the signal TSVs. Alternatively, thermal TSVs are assigned to a block after insertion of the signal TSVs without exceeding the maximum TSV capacity of the block. In addition to the benefits that the added thermal vias produce, thermal wires can also be utilized to enhance the heat transfer process. These thermal wires are treated as routing channels wherever there are available tracks [56].

VI. TIMING OPTIMIZATION OF INTERPLANE INTERCONNECTS

Three-dimensional integration demonstrates many opportunities for heterogeneous SoCs [20]. Integrating circuits from diverse fabrication processes into a single multiplane system can result in substantially different interconnect impedance characteristics of each physical plane within a 3-D circuit. By considering the disparate interconnect impedance characteristics of 3-D circuits, the performance of the interplane interconnects can be significantly improved. An efficient technique to decrease the delay of interplane interconnects by optimally placing the TSVs is discussed in Section VI-A. The problem of placing TSVs to decrease the delay of interplane trees is discussed in Section VI-B.

A. Two-Terminal Interplane Nets

The impedance characteristics of the interconnect layers belonging to different physical planes of a 3-D circuit can vary significantly. The interplane interconnects are therefore modeled as an assembly of horizontal interconnect segments with different impedance characteristics connected by interplane vias. In this section, a heuristic for near-optimal interplane via placement of two-terminal nets that include several TSVs is described.

A schematic of an interplane interconnect connecting two circuits located n planes apart is illustrated in Fig. 6. The horizontal segments of the line are connected through the vias, which can traverse more than one plane where each via is placed within a certain physical interval. The via placement is constrained

$$0 \leq x_j \leq \Delta x_j, \quad (1)$$

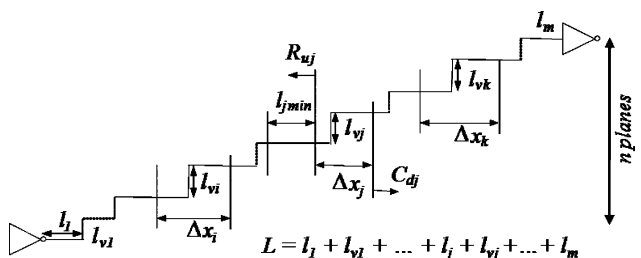


Fig. 6. Interplane interconnect connecting two circuits located n planes apart.

where Δx_j is the length of the interval where the via connecting planes j and $j + 1$ can be placed. This interval length is called the “allowed interval” here for clarity. x_j is the distance of the via location from the edge of the allowed interval. Due to the nonuniformity of the interconnects, each segment is modeled as a distributed RC line with different impedance characteristics. In order to analyze the delay of a line, the distributed Elmore delay model has been adopted due to the simplicity and high fidelity of this model [94]. The accuracy of the model can be further improved as discussed in [95]. However, unlike a single plane, more than one set of fitting coefficients is required in a 3-D system. Alternatively, higher order models with higher accuracy as compared to the Elmore delay model can be utilized to characterize the delay of the interplane nets. Due to the particular traits of interplane nets in 3-D circuits, however, the optimization problem can be nonconvex even for the simple Elmore delay model. Employing higher order delay models further exacerbates the difficulty of optimizing the interconnect delay as the convexity of these timing models cannot be easily proved. Consequently, any solutions based on these models can produce local minima, possibly creating inferior solutions than that produced by the less accurate Elmore delay model. An increase in the computational time should also be considered as a natural tradeoff for greater accuracy when utilizing these models.

Based on this delay model, the key concept in the heuristic is that the optimum via placement depends primarily upon the size of the allowed interval (that is estimated or known after an initial placement) rather than the exact location of the via. Consider the interplane interconnect shown in Fig. 6. The optimum via location x_j^* for via j is a monotonic function of R_{uj} and C_{dj}

$$x_j^* = f(R_{uj}, C_{dj}) \quad (2)$$

where R_{uj} and C_{dj} are the upstream resistance and downstream capacitance, respectively, of the allowed interval for via j . As the size of the allowed intervals for all of the vias is constrained by (1), the minimum and maximum value of R_{uj} and C_{dj} can be readily determined, permitting the values of x_j^* for these extrema $x_{j\min}^*$ and $x_{j\max}^*$ to be evaluated. Due to the monotonic dependence of x_j on R_{uj} and C_{dj} , the optimum location for via j , x_j^* lies within the range delimited by $x_{j\min}^*$ and $x_{j\max}^*$. By iteratively decreasing the range of values for x_j^* , the optimal location for via j can be determined.

This heuristic has been used to implement an algorithm that exhibits an optimal or near-optimal TSV placement for two-terminal interplane interconnects in 3-D ICs, with significantly lower computational time as compared to general optimization engines. Two-terminal interplane interconnects for different numbers of physical planes

have been analyzed. The impedance characteristics of the horizontal segments and vias are extracted for several interconnect structures using a commercial impedance extraction tool [96]. Copper interconnect has been assumed with an effective resistivity of $2.2 \mu\Omega\text{-cm}$. Based on the extracted impedances, the resistance and capacitance of the horizontal segments range from 25 to 125 Ω/mm and 100 to 300 fF/mm , respectively, for a 90 nm technology node⁴ [97]. The cross-section of the vias is $1 \times 1 \mu\text{m}$, with $1 \mu\text{m}$ spacing from the surrounding horizontal metal layers, assuming an SOI process as described in [8]. For all of the interconnect structures, the total and minimum length of each horizontal segment is randomly generated. For simplicity, all of the vias connect the segments of two adjacent physical planes.

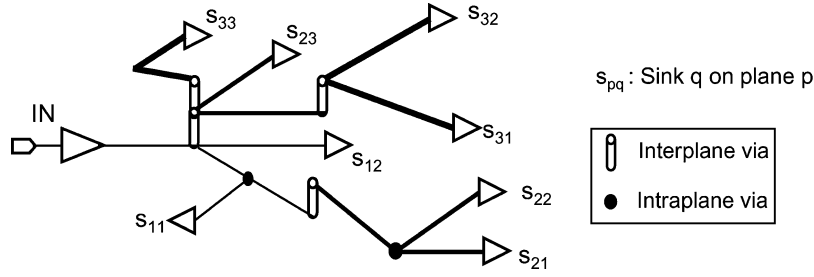
The via locations or, equivalently, the length of the horizontal segments, are determined from the via placement heuristic for relatively short interconnects ($< 2 \text{ mm}$). SPICE delay simulations demonstrate an average improvement of 8.9% as compared to the case where the vias are placed at the center of the line and 14.1% as compared to random via placement, respectively. The two-terminal via placement algorithm is also compared in terms of both optimality and efficiency to an optimization solver, YALMIP [98]. The algorithm exhibits high accuracy as compared to YALMIP independent of the number of planes that comprise the 3-D interconnect, demonstrating that optimum solutions are obtained for most interconnect configurations. In addition, for those cases where some of the vias are not optimally placed, the loss of optimality is insignificant ($< 0.01\%$). Furthermore, the algorithm is approximately two orders of magnitude faster than YALMIP while the complexity of the algorithm exhibits an almost linear dependence on the number of interplane vias.

From these results, exploiting the nonuniform impedance characteristics of the interplane interconnects when placing the vias can improve the delay of multiplane lines. This improvement in delay can decrease the number of repeaters required to drive a global line or eliminate the need for repeaters in semiglobal (medium length) lines. In addition, wire sizing can be avoided, thereby saving significant power. Decreasing the number of repeaters and avoiding wide lines reduces the overall power consumption, which is a particularly important issue in 3-D circuits.

B. Multiterminal Interplane Nets

Multiterminal tree-like interconnects constitute a significant portion of the interconnects in an integrated circuit. Improving the performance of these nets in 3-D circuits is a challenging task, as the leaves of these interconnects can be located on different physical planes. A technique for placing the vias to decrease the delay of an

⁴See <http://www.eas.asu.edu/~ptm>.


Fig. 7. An example of an interplane interconnect tree.

interplane tree is described in this section. The application of this technique to various interplane trees targeting both 3-D ICs and SiP is also discussed.

A simple interplane interconnect tree (also called an interconnect tree for simplicity) is illustrated in Fig. 7. The leaves of the tree are located on different physical planes within a 3-D stack. Subtrees not directly connected to the interplane vias that do not contain any interplane vias (i.e., intraplane trees) are also shown. The weighted summation of the distributed Elmore delay of the branches of an interconnect tree is considered as the objective function

$$T_w = \sum_{\forall s_{pq}} w_{s_{pq}} T_{s_{pq}} \quad (3)$$

where $w_{s_{pq}}$ and $T_{s_{pq}}$ are the weight and distributed Elmore delay of sink s_{pq} , respectively. Weights are assigned to the sinks according to the relative criticality of the sinks. The constrained optimization problem for placing a via within an interplane interconnect tree can be described as

$$(P1) \quad \text{minimize } T_w, \text{ subject to } (1), \forall \text{ via } v_j. \quad (4)$$

The heuristic for two-terminal nets has been extended to address the task of placing TSVs to decrease the delay of an interplane tree. TSV placement algorithms based on this heuristic have been applied to several interplane interconnect tree examples. Trees for different numbers of

planes and sinks are analyzed. The impedance characteristics of the horizontal segments are similar to those utilized for two-terminal nets, as discussed in Section VI-A. These trees are optimized for two different 3-D technologies, a 3-D IC technology based on [8], where the TSV length is $l_{v3-D} = 10 \mu\text{m}$ and an SiP technology where the TSV length is $l_{vSiP} = 70 \mu\text{m}$ [44], [45]. The impedance characteristics of the TSVs are $r_{v3-D} = 22 \Omega/\text{mm}$ and $c_{v3-D} = 210 \text{ fF}/\text{mm}$ and $r_{vSiP} = 22 \Omega/\text{mm}$ and $c_{vSiP} = 6 \text{ pF}/\text{mm}$ for the 3-D IC and SiP technology, respectively. The savings in delay achieved by optimally placing the vias is listed in Table 2 for different via placement scenarios.

The improvement in delay of the interconnect trees is listed in columns 6 through 9 of Table 2. The results are compared to the case where the vias are initially placed at the center of the allowed interval (i.e., $x_i = \Delta x_i/2$) and the case where the vias are placed at the lower edge of the allowed interval (i.e., $x_i = 0$). The improvement in delay depends upon the length of the allowed interval. This dependence, however, is weak as compared to two-terminal nets. In addition, the improvement in delay is lower than point-to-point nets for the same allowed length intervals. The reason for this behavior is that any modifications to the routing tree are strictly confined within the allowed interval that least affects the routing tree. If this constraint is relaxed, the length of the interconnect segments can be further reduced, resulting in a considerably greater improvement in speed. Note that the improvement in delay achieved by optimally placing

Table 2 Delay of Various Interplane Interconnect Trees for Different Number of Sinks, Physical Planes n , and 3-D Technologies

n	Technology	Number of sinks	Avg. branch length [μm]	Δx_i 's [μm]	Delay improvement [%]				Instances
					$x_i^* = \Delta x_i/2$		$x_i^* = 0$		
					Avg	Max	Avg	Max	
3	3-D IC	4	216	50	1.31	7.11	5.33	13.00	10000
4	3-D IC	8	407	50	1.47	6.88	6.83	13.22	10212
3	3-D IC	4	815	150	1.15	5.74	4.42	10.02	11000
4	3-D IC	8	909	150	1.29	4.98	5.70	9.48	10219
3	SiP	4	216	50	1.21	4.99	1.78	5.58	10000
4	SiP	8	407	50	0.90	3.54	1.98	5.72	10212
3	SiP	4	815	150	1.31	4.10	1.98	5.68	11000
4	SiP	8	909	150	1.04	3.28	2.34	5.71	10219

the TSVs in a 3-D IC is substantially greater than the improvement for an SiP technology. This difference is due to the significantly longer length and larger impedance characteristics of the TSVs utilized in an SiP. Manufacturing processes that provide short vertical interconnects with low parasitic impedances are therefore necessary; otherwise, the performance benefits due to the reduction in interconnect length will decrease since the TSVs contribute significantly to the overall interconnect delay.

VII. SYNCHRONIZATION IN 3-D CIRCUITS

An omnipresent and challenging issue for synchronous digital circuits is the reliable distribution of the clock signal to the many thousands of sequential elements distributed throughout a synchronous circuit [99], [100]. The complexity is further increased in 3-D ICs as sequential elements belonging to the same clock domain (i.e., synchronized by the same clock signal) can be located on different planes. Another important issue in the design of clock distribution networks is low power consumption, since the clock network dissipates a significant portion of the total power consumed by a synchronous circuit [101], [102]. This demand is stricter for 3-D ICs due to the increased power density and related thermal limitations.

In 2-D circuits, symmetric interconnect structures, such as H- and X-trees, are widely utilized to distribute the clock signal across a circuit [100]. The symmetry of these structures permits the clock signal to arrive at the leaves of the tree at the same time, resulting in synchronous data processing. Maintaining this symmetry within a 3-D circuit, however, is a difficult task. Consequently, asymmetric structures are useful candidates for distributing the clock signal within a 3-D circuit. Issues related to the distribution of the clock signal within a 3-D system are discussed in this section. Experimental results of a 3-D test circuit manufactured by MIT Lincoln Laboratories composed of several different 3-D clock network architectures are also described.

To evaluate the specific requirements of a 3-D clock network, consider a traditional H-tree topology. At each branch point of an H-tree, two branches emanate with the same length. An extension of an H-tree to three dimensions does not guarantee equidistant interconnect paths from the root to the leaves of the tree. The clock signal propagates through interplane vias from the output of the clock driver to the center of the H-tree on the other planes. The impedance of these vias can increase the time for the clock signal to arrive at the leaves of the tree on these planes as compared to the time for the clock signal to arrive at the leaves of the tree located on the same plane as the clock driver. Furthermore, in a multiplane 3-D circuit, three or four branches can emanate at each branch point. The third and fourth branches propagate the clock signal to the other planes of the 3-D circuit. Similar to a design

methodology for a 2-D H-tree topology, the width of each branch is reduced to a third (or more) of the segment preceding the branch point in order to match the impedance at that branch point. This requirement, however, is difficult to achieve as the third and fourth branches are implemented by an interplane via. Note that the vertical interconnects are of significantly different length as compared to the horizontal branches and exhibit different impedance characteristics.

A test circuit exploring four different clock network topologies for 3-D circuits has been designed, manufactured, and measured. The test circuit is based on a 3-D fully depleted (FD) SOI fabrication technology recently developed by MIT Lincoln Laboratories (MITLL) [8]. The MITLL process is a wafer-level 3-D integration technology with up to three FDSOI wafers bonded to form a 3-D circuit. The minimum feature size of the devices is 180 nm, with one polysilicon layer and three metal layers interconnecting the devices on each wafer. A backside metal layer also exists on the upper two planes, providing the starting and landing pads for the TSVs, and the I/O, power supply, and ground pads for the entire 3-D circuit. An attractive feature of this process is the high density 3-D vias. The dimensions of these vias are $1.75 \times 1.75 \mu\text{m}$, much smaller than the size of the through silicon vias in many existing 3-D technologies [24], [25].

The test circuit consists of four blocks. Each block contains the same logic circuit but different clock distribution networks. The total area of the test circuit is $3 \times 3 \text{ mm}^2$. All of the blocks share the same logic circuitry to emulate the variety of switching patterns in a synchronous digital circuit. In each of the circuit blocks, the clock driver for the overall clock distribution network is located on the second plane. The location of the clock driver on that plane is chosen to ensure that the clock signal propagates through identical vertical interconnect paths to the first and third plane, resulting in the same delay for the clock signal to arrive at the registers located on the first and third planes. The off-chip clock signal is received by the clock driver through an RF pad located at the middle of each block. Additional RF pads are placed at different locations on the topmost plane of each block for probing. The fabricated test circuit is depicted in Fig. 8, where the RF and dc pads on the back side metal layer of the third plane are shown.

The clock distribution networks combine commonly used networks such as H-trees, meshes, and rings. These clock network topologies range from highly symmetric topologies, such as H-trees, as the block shown in Fig. 9(a), to fully asymmetric topologies, such as a trunk-based topology. Normal operation has been demonstrated during preliminary testing of the fabricated test circuit. The clock input is a 1.5 V peak-to-peak sinusoidal signal with 0.75 V dc offset. The clock driver is implemented with a traditional chain of tapered buffers [103]–[105], which produces a square waveform at the root of the clock

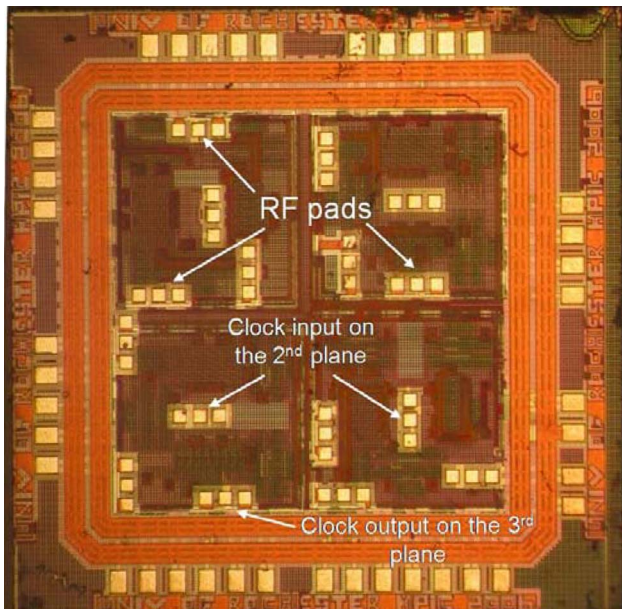


Fig. 8. Fabricated 3-D test circuit. The total area is $3 \times 3 \text{ mm}^2$. There are four different blocks, with one input and three output RF pads for each block. The area of each block is approximately 1 mm^2 .

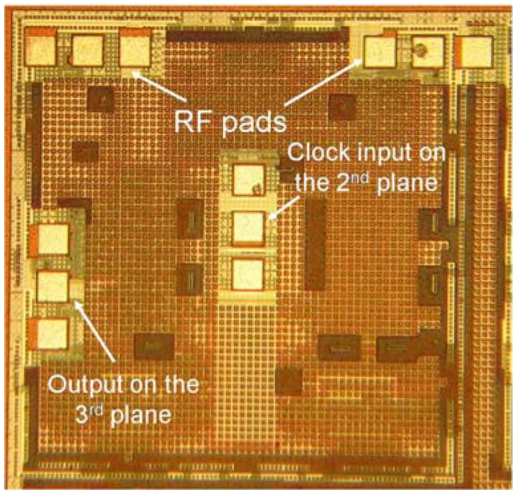
distribution network. The clock distribution network of the block illustrated in Fig. 9(a) contains a four level H-tree (i.e., equivalent to 16 leaves) with identical interconnect characteristics in each plane. All of the H-trees are connected through a group of interplane vias. Note that the H-tree on the second plane is rotated by 90° with respect to the H-trees on the other two planes. This rotation effectively eliminates inductive coupling between

the H-trees. The second plane is front-to-front bonded with the first plane and both of the H-trees are implemented on the third metal layer. The vertical distance between these clock networks is approximately $2 \mu\text{m}$. All of the H-trees are shielded with two parallel lines connected to ground. The waveform shown in Fig. 9(b) is the clock signal at a leaf of the H-tree on the third plane, demonstrating operation of the circuit at 1 GHz. Experiments demonstrate that a clock distribution network that combines an H-tree on the second plane and meshes on the other two planes exhibits moderate skew, within 10% of the clock period, and the lowest power consumption. The superior performance of this topology is due to the symmetry of the H-tree and the balancing characteristic of the meshes.

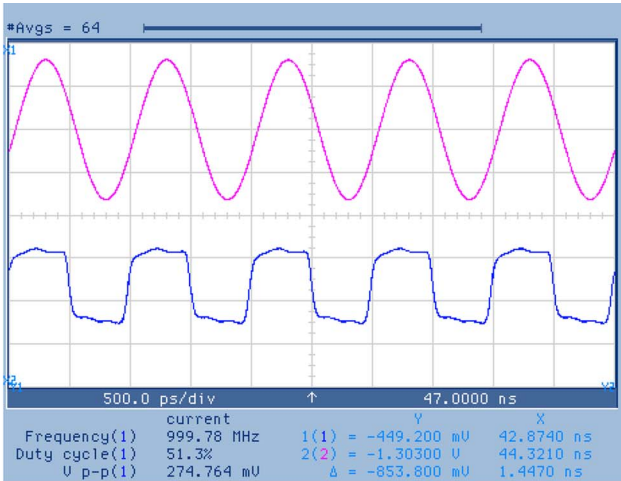
VIII. COMMUNICATION CENTRIC 3-D ARCHITECTURES

A promising design paradigm to appease foreseen interconnect problems is networks-on-chip [106], where information is communicated among circuits within packets in an internet-like fashion. The synergy between these two design paradigms, NoCs and 3-D ICs, can be exploited to significantly improve the performance and decrease the power consumption of communications limited systems. Several interesting topologies that emerge by incorporating the third dimension in networks-on-chip are discussed in this section.

On-chip networks differ from traditional interconnection networks in that communication among the network elements is implemented through the on-chip routing layers rather than the metal tracks of the package or printed circuit board. NoCs offer high flexibility and



(a)



(b)

Fig. 9. Experimental results of the fabricated 3-D circuit: (a) tested circuit block and (b) clock signal waveform from the H-tree on the third plane operating at 1 GHz.

regularity, supporting simpler interconnect models and greater fault tolerance. The canonical interconnect backbone of the network combined with appropriate communication protocols enhance the flexibility of these systems [107]. NoCs provide communication among a variety of processing elements (PEs), such as processor and digital signal-processing cores, memory blocks, field-programmable gate arrays, and dedicated hardware [108]–[110]. Furthermore, the length of the communication channel is primarily determined by the area of the PE, which is typically unaffected by the network structure. Mesh structures have been a popular network topology for conventional 2-D NoCs [106], [111], as illustrated in Fig. 10(a), where each PE is connected to the network through a router [106].

Integration in the third dimension introduces a variety of topological choices for NoCs. For a 3-D NoC, as shown in Fig. 10(b), the total number of nodes is $N = n_1 \times n_2 \times n_3$, where n_1 , n_2 , and n_3 is the number of network nodes in the x , y , and z direction, respectively. In this topology, each PE is on a single yet possibly different physical plane (2-D IC/3-D NoC). In other words, a PE can be implemented on only one of the n_3 physical planes of the system and, therefore, the 3-D system contains $n_1 \times n_2$ PEs on each of the n_3 physical planes, where the total number of nodes is N [59], [112]. A 3-D topology is illustrated in Fig. 10(c), where the interconnect network is contained within one physical plane (i.e., $n_3 = 1$), while each PE is integrated on multiple planes, notated as n_p (3-D IC/2-D NoC). Finally, a hybrid 3-D NoC based on the two previous topologies is depicted in Fig. 10(d). In this NoC topology, both the interconnect network and the PEs can span more than one physical plane of the stack (3-D IC/3-D NoC).

Analytic models of the zero-load latency and power consumption with delay constraints of these networks capturing the effects of the topology on the performance of 3-D NoCs have been developed. The overall zero-load network latency for a 3-D NoC is [113]

$$T_{\text{network}} = \text{hops}(t_a + t_s) + \text{hops}_{2-D}t_h + \text{hops}_{3-D}t_v + \frac{L_p}{w_c}t_h \quad (5)$$

where t_a , t_s , t_v , and t_h are the delay of the arbiter, crossbar switch, and vertical and horizontal channels, respectively. hops_{2-D} and hops_{3-D} denote the average number of hops within the two dimensions n_1 and n_2 and within the third dimension n_3 , respectively (see Fig. 10). hops is equal to the summation of hops_{2-D} and hops_{3-D} . L_p and w_c denote, respectively, the size of a data packet and the width of the interconnect buss connecting adjacent network routers. L_v denotes the length of the vertical buss, which is equal to one or more TSV lengths.

These models do not incorporate the effects of the routing scheme and traffic load. Since minimum distance paths and no contention are implicitly assumed in these expressions, nonminimal path routing schemes and heavy traffic loads will increase both the latency and power consumption of the network. These models can therefore be treated as lower bounds for both the latency and the power consumption of the network. Alternatively, these expressions provide the maximum improvement in the performance of a conventional NoC that can be achieved with vertical integration.

The resulting decrease in network latency as compared to a standard 2-D/IC 2-D NoC is illustrated in Fig. 11(a) for increasing network size, where the area of each PE is

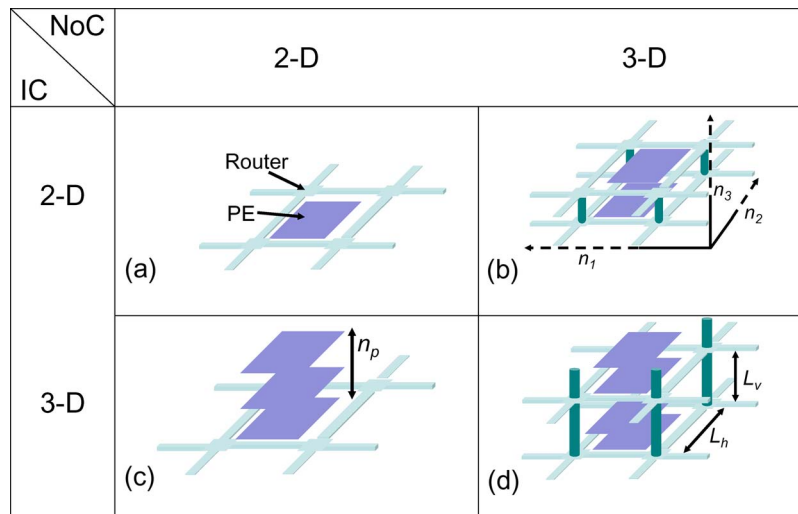


Fig. 10. Various NoC topologies (not to scale): (a) 2-D IC/2-D NoC, (b) 2-D IC/3-D NoC, (c) 3-D IC/2-D NoC, and (d) 3-D IC/3-D NoC.

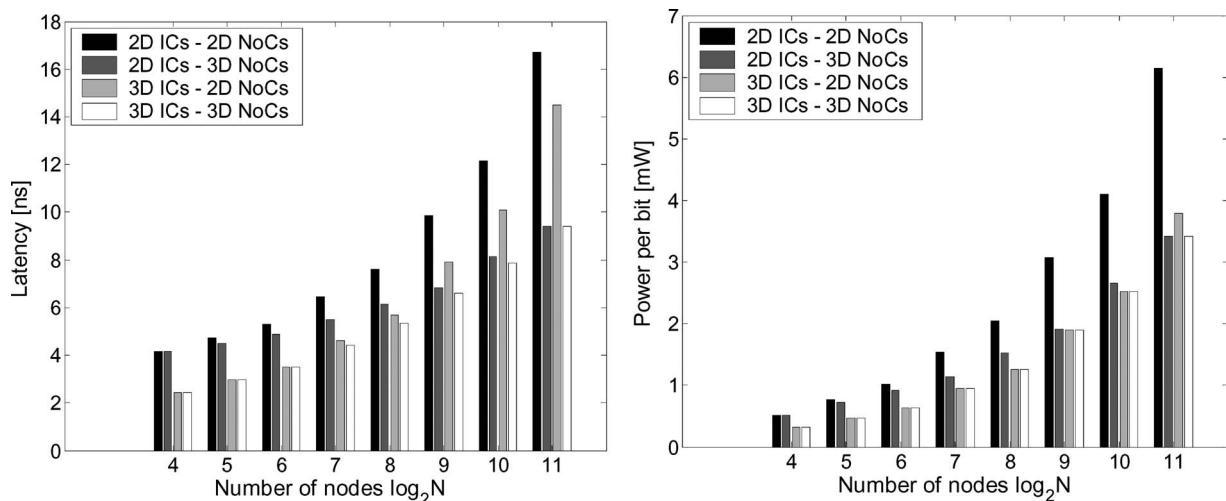


Fig. 11. Performance of 3-D NoC topologies for a range of network sizes where $A_{PE} = 1 \text{ mm}^2$: (a) zero-load latency and (b) power consumption with delay constraints.

denoted by A_{PE} . The 2-D IC/3-D NoC topology decreases the number of hops while the interconnect buss delay remains constant. With a 3-D IC/2-D NoC, the buss delay is smaller but the number of hops remains unchanged. With a 3-D IC/3-D NoC, all of the latency components can be decreased by assigning a portion of the available physical planes to the network while the remaining planes of the stack are used for the PEs. A decrease in latency of 31.5% and 29.7% can be observed for $N = 128$ and $N = 256$ nodes, respectively, with $A_{PE} = 1 \text{ mm}^2$. Note that the 3-D IC/3-D NoC topology achieves the greatest savings in latency by optimally balancing n_3 with n_p . Consequently, the tradeoff between the number of hops and the buss length for various 3-D topologies can be exploited to improve the performance of a network-on-chip.

As with the zero-load latency, each topology affects the power consumption of a network in a different way. The power consumption can be reduced by either decreasing the number of hops that a packet travels or by decreasing the buss length. Note that by reducing the buss length, not only is the interconnect capacitance reduced but also the number and size of the repeaters required to drive the lines are decreased, resulting in a greater savings in power. In Fig. 11(b), the power consumption of a 2-D NoC topology is compared to the three-dimensional topologies previously discussed. A power savings of 38.4% is achieved for $N = 128$ with $A_{PE} = 1 \text{ mm}^2$. Allowing the available physical planes to be utilized either for the third dimension of the network or for the PEs, the 3-D IC/3-D NoC scheme achieves the greatest savings in power in addition to the minimum delay. For each topology and network size, the distribution of the network nodes n_1 , n_2 , and n_3 in the three physical dimensions is chosen to ensure that the target objective is minimized. This assignment can be different

for the various topologies, network size, and latency or power objective.

Note that these topologies emphasize the latency and power consumption of a network, neglecting the performance requirements of the individual PEs. If the performance of the individual PEs is important, only one 3-D topology may be available; however, despite this constraint, a significant savings in latency and power can be achieved since in almost every case the network latency and power consumption are lower than for the 2-D IC/2-D NoC topology.

IX. CONCLUSIONS AND FUTURE WORK

Developing a design flow for 3-D ICs is a complicated task with many ramifications. Design methodologies at the front end and mature manufacturing processes at the back end are required to effectively provide large scale 3-D systems. Physical design techniques at different stages of a developmental design flow for 3-D circuits have been discussed in this paper, emphasizing the effect of the 3-D nature on each design stage. A variety of floorplanning, placement, and routing techniques and algorithms for 3-D circuits have been described that consider the unique characteristics of 3-D circuits. In these techniques, the discrete nature of the third dimension is exploited to decrease the number of candidate solutions and, consequently, the computational time required to design a 3-D circuit.

The objective function of 3-D layout techniques has been extended to include routing congestion, power supply noise, and decoupling capacitance allocation in addition to traditional objectives, such as wire length and area. Due to increased power densities and greater distances between the circuits on the upper planes and

the heat sink, physical design techniques that embody a thermal objective can be a useful mechanism to manage thermal issues in 3-D ICs. Design techniques can reduce thermal gradients and temperatures in 3-D circuits by redistributing the blocks among and within the planes of a 3-D circuit. Alternatively, thermal vias can be utilized in 3-D circuits to convey heat to the heat sink. Thermal wires in the horizontal direction are similar in function to thermal vias and can also be utilized to lower thermal gradients within 3-D circuits.

Significant performance improvements can be achieved by optimally placing interplane vias in 3-D circuits. Algorithms for determining the minimum delay of the interplane interconnects are an integral element of the physical design process for 3-D circuits. Interplane interconnect impedances of 3-D circuits, however, vary considerably from 2-D interconnect impedances. This difference is due to several reasons, such as the heterogeneity of 3-D circuits, diverse fabrication technologies, and the variety of bonding styles.

Another requirement for maximizing the speed of 3-D circuits is to reliably distribute the clock signal within these circuits. A 3-D clock distribution network, however, cannot be directly extended from a 2-D circuit due to the asymmetry of a multiplane 3-D circuit and the effect of the interplane via impedances. Several clock distribution networks have been developed to investigate synchronization issues in 3-D systems. These network topologies have been included in a 3-D test circuit manufactured in the 3-D FDSOI fabrication technology developed at MITLL. The circuit is composed of four independent blocks, where each block is a three-plane 3-D circuit with a different clock distribution network. This circuit constitutes the first effort to investigate the critical design issue of synchronization in vertical integration. Successful high-speed operation of the test circuit has been demonstrated.

In addition to higher performance, 3-D integration offers significant opportunities for designing highly diverse and complex systems. On-chip networks can be a useful solution to provide sufficient communication throughput among the components of these 3-D systems. Three-dimensional NoCs are a natural evolution of 2-D NoCs,

exhibiting superior performance. Several novel 3-D NoC topologies are discussed in this paper. These topologies decrease the latency and power consumption by reducing both the number of hops per packet and the length of the communications channels. These 3-D topologies demonstrate the tradeoff between the number of planes required to implement a network and those planes required to implement the PEs. Consequently and not surprisingly, the 3-D IC/3-D NoC topology achieves the greatest improvement in latency and power consumption by most effectively exploiting the third dimension.

Research on the design of 3-D ICs has only recently begun to emerge. Many challenges remain unsolved and significant effort is required to provide effective solutions to the problems encountered in the design of 3-D ICs. Two important challenges related to global interconnect issues in 3-D circuits are robust clock and power distribution networks. Heterogeneous 3-D circuits, for instance, pose further limitations on the design of 3-D clock distribution networks in addition to clock skew, power dissipation, and delay uncertainty. In these systems, the clock signal can behave both as the victim and the aggressor when propagating through noisy digital and sensitive analog planes. Techniques that enhance signal integrity in 3-D structures are therefore necessary.

Furthermore, distributing power to the planes of the stack located far from the power/ground pads is another fundamental issue in 3-D ICs. As the power/ground pads are typically located along the edges of the plane, providing sufficient current while satisfying target voltage levels for every transistor within a 3-D IC requires innovative power distribution networks. Decoupling capacitance allocation strategies are also required for 3-D ICs, as the decoupling capacitors can be placed closer to the transistors, for instance, directly above or below the high current loads. Addressing these important design issues will considerably accelerate the development of commercial 3-D integrated systems. The material described in this paper is intended to shed light on those areas related to the design of 3-D integrated systems in an effort to develop large-scale multifunctional multiplane systems to continue the micro-electronics revolution. ■

REFERENCES

- [1] A. Fan, A. Rahman, and R. Reif, "Copper wafer bonding," *Electrochem. Solid-State Lett.*, vol. 2, no. 10, pp. 534–536, Oct. 1999.
- [2] R. Reif, A. Fan, K. N. Chen, and S. Das, "Fabrication technologies for three-dimensional integrated circuits," in *Proc. IEEE Int. Symp. Quality Electron. Design*, Mar. 2002, pp. 33–37.
- [3] R. J. Gutmann *et al.*, "Three-dimensional (3D) ICs: A technology platform for integrated systems and opportunities for new polymeric adhesives," in *Proc. IEEE Int. Conf. Polymers Adhesives Microelectron. Photon.*, Oct. 2001, pp. 173–180.
- [4] J.-Q. Lu *et al.*, "Stacked chip-to-chip interconnections using wafer bonding technology with dielectric bonding glues," in *Proc. IEEE Int. Interconnect Technol. Conf.*, Jun. 2001, pp. 219–221.
- [5] A. Klumpp, R. Merkel, R. Wieland, and P. Ramm, "Chip-to-wafer stacking technology for 3D system integration," in *Proc. IEEE Int. Electron. Compon. Technol. Conf.*, May 2003, pp. 1080–1083.
- [6] T. Fukushima, Y. Yamada, H. Kikuchi, and M. Koyanagi, "New three-dimensional integration using self-assembly technique," in *Proc. IEEE Int. Electron Devices Meeting*, Dec. 2005, pp. 348–351.
- [7] S. Tiwari *et al.*, "Three-dimensional integration for silicon electronics," in *Proc. IEEE Lester Eastman Conf. High Perform. Devices*, Aug. 2002, pp. 24–33.
- [8] *FDSOI Design Guide*, MIT Lincoln Laboratories, Cambridge, MA, 2006.
- [9] J. A. Burns *et al.*, "A wafer-scale 3-D circuit integration technology," *IEEE Trans. Electron Devices*, vol. 53, pp. 2507–2515, Oct. 2006.
- [10] J. W. Joyner *et al.*, "Impact of three-dimensional architectures on interconnects in gigascale integration," *IEEE*

- Trans. Very Large Scale (VLSI) Syst., vol. 9, pp. 922–928, Dec. 2001.
- [11] J. W. Joyner, P. Zarkesh-Ha, and J. D. Meindl, “A stochastic global net-length distribution for a three-dimensional system-on-a-chip (3D-SOC),” in *Proc. IEEE Int. ASIC/SOC Conf.*, Sep. 2001, pp. 147–151.
- [12] A. Rahman, A. Fan, J. Chung, and R. Reif, “Wire-length distribution of three-dimensional integrated circuits,” in *Proc. IEEE Int. Interconnect Technol. Conf.*, May 1999, pp. 233–235.
- [13] A. Rahman and R. Reif, “System level performance evaluation of three-dimensional integrated circuits,” *IEEE Trans. Very Large Scale (VLSI) Syst.*, vol. 8, pp. 671–678, Dec. 2000.
- [14] D. Stroobandt and J. Van Campenhout, “Accurate interconnection lengths in three-dimensional computer systems,” *IEICE Trans. Inf. Syst. (Special Issue on Physical Design in Deep Submicron)*, vol. 10, no. 1, pp. 99–105, Apr. 2000.
- [15] J. W. Joyner and J. D. Meindl, “Opportunities for reduced power distribution using three-dimensional integration,” in *Proc. IEEE Int. Interconnect Technol. Conf.*, Jun. 2002, pp. 148–150.
- [16] R. Zhang, K. Roy, C.-K. Koh, and D. B. Janes, “Stochastic interconnect modeling, power trends, and performance characterization of 3-D circuits,” *IEEE Trans. Electron Devices*, vol. 48, pp. 638–652, Apr. 2001.
- [17] B. S. Cherkauer and E. G. Friedman, “A unified design methodology for CMOS tapered buffers,” *IEEE Trans. Very Large Scale (VLSI) Syst.*, vol. 3, pp. 99–111, Mar. 1995.
- [18] K. Banerjee, S. K. Souri, P. Kapour, and K. C. Saraswat, “3-D ICs: A novel chip design paradigm for improving deep-submicrometer interconnect performance and systems-on-chip integration,” *Proc. IEEE*, vol. 89, pp. 602–633, May 2001.
- [19] M. Koyanagi et al., “Future system-on-silicon LSI chips,” *IEEE Micro*, vol. 18, pp. 17–22, Jul./Aug. 1998.
- [20] V. K. Jain, S. Bhanja, G. H. Chapman, and L. Doddannagari, “A highly reconfigurable computing array: DSP plane of a 3D heterogeneous SoC,” in *Proc. IEEE Int. Syst. Chip Conf.*, Sep. 2005, pp. 243–246.
- [21] K. Bernstein et al., “Interconnects in the third dimension: Design challenges for 3-D ICs,” in *Proc. IEEE/ACM Design Automation Conf.*, Jun. 2007, pp. 562–567.
- [22] R. S. Patti, “Three-dimensional integrated circuits and the future of system-on-chip designs,” *Proc. IEEE*, vol. 94, pp. 1214–1224, Jun. 2006.
- [23] D. L. Lewis and H.-H. S. Lee, “A scan-island based design enabling pre-bond testability in die-stacked microprocessors,” in *Proc. IEEE Int. Test Conf.*, Oct. 2007, pp. 1–8.
- [24] M. W. Newman et al., “Fabrication and electrical characterization of 3D vertical interconnects,” in *Proc. IEEE Int. Electron. Compon. Technol. Conf.*, Jun. 2006, pp. 394–398.
- [25] P. Dixit and J. Miao, “Fabrication of high aspect ratio 35 μ m pitch interconnects for next generation 3-D wafer level packaging by through-wafer copper electroplating,” in *Proc. IEEE Int. Electron. Compon. Technol. Conf.*, Jun. 2006, pp. 388–393.
- [26] S. X. Zhang, S.-W. R. Lee, L. T. Weng, and S. So, “Characterization of copper-to-silicon for the application of 3D packaging with through silicon vias,” in *Proc. IEEE Int. Conf. Electron. Packag. Technol.*, Sep. 2005, pp. 51–56.
- [27] N. Ranganathan et al., “High aspect ratio through-wafer interconnect for three-dimensional integrated circuits,” in *Proc. IEEE Int. Electron. Compon. Technol. Conf.*, Jun. 2005, pp. 343–348.
- [28] D. Henry et al., “Low electrical resistance silicon through vias: Technology and characterization,” in *Proc. IEEE Int. Electron. Compon. Technol. Conf.*, Jun. 2006, pp. 1360–1365.
- [29] S. F. Al-Sarawi, D. Abbott, and P. D. Franzon, “A review of 3-D packaging technology,” *IEEE Trans. Compon., Packag., Manuf. Technol. B*, vol. 21, pp. 2–14, Feb. 1998.
- [30] P. Garrou. (2005, Feb.). Future ICs go vertical. *Semiconductor Int.*
- [31] M. Karnezos, “3-D packaging: Where all technologies come together,” in *Proc. IEEE/SEMI Int. Electron. Manuf. Technol. Symp.*, Jul. 2004, pp. 64–67.
- [32] J. Miettinen, M. Mantysalo, K. Kaija, and E. O. Ristolainen, “System design issues for 3D System-in-Package (SiP),” in *Proc. IEEE Int. Electron. Compon. Technol. Conf.*, Jun. 2004, pp. 610–615.
- [33] E. Beyne, “The rise of the 3rd dimension for system integration,” in *Proc. IEEE Int. Interconnect Technol. Conf.*, Jun. 2006, pp. 1–5.
- [34] J. U. Knickerbocker et al., “3-D silicon integration and silicon packaging technology using silicon through-vias,” *IEEE J. Solid-State Circuits*, vol. 41, pp. 1718–1725, Aug. 2006.
- [35] W. J. Howell et al., “Area array solder interconnection technology for the three-dimensional silicon cube,” in *Proc. IEEE Int. Electron. Compon. Technol. Conf.*, May 1995, pp. 1174–1178.
- [36] M. Healy et al., “Multiobjective microarchitectural floorplanning for 2-D and 3-D ICs,” *IEEE Trans. Computer-Aided Design Integr. Circuits Syst.*, vol. 26, pp. 38–52, Jan. 2007.
- [37] E. Culurciello and A. G. Andreou, “Capacitive inter-chip data and power transfer for 3-D VLSI,” *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 53, pp. 1348–1352, Dec. 2006.
- [38] S. A. Kühn, M. B. Kleiner, R. Thewes, and W. Weber, “Vertical signal transmission in three-dimensional integrated circuits by capacitive coupling,” in *Proc. IEEE Int. Symp. Circuits Syst.*, May 1995, vol. 1, pp. 37–40.
- [39] J. Xu et al., “AC coupled interconnect for dense 3-D ICs,” *IEEE Trans. Nucl. Sci.*, vol. 51, pp. 2156–2160, Oct. 2004.
- [40] K. Sugahara et al., “SOI/SOI/bulk-Si triple-level structure for three-dimensional devices,” *IEEE Electron Device Lett.*, vol. EDL-7, pp. 193–194, Mar. 1986.
- [41] S. Akiyama et al., “Multilayer CMOS device fabricated on laser recrystallized silicon islands,” in *Proc. IEEE Int. Electron Devices Meeting*, Dec. 1983, pp. 352–355.
- [42] V. Subramania and K. C. Saraswat, “High-performance germanium-seeded laterally crystallized TFTs for vertical device integration,” *IEEE Trans. Electron Devices*, vol. 45, pp. 1934–1939, Sep. 1998.
- [43] S. Kawamura et al., “Three-dimensional CMOS ICs fabricated by using beam recrystallization,” *IEEE Electron Device Lett.*, vol. EDL-4, pp. 366–368, Oct. 1983.
- [44] C. Ryu et al., “High frequency electrical circuit model of chip-to-chip vertical via interconnection for 3-D chip stacking package,” in *Proc. IEEE Topical Meeting Electr. Perform. Electron. Packag.*, Oct. 2005, pp. 151–154.
- [45] D. M. Jang et al., “Development and evaluation of 3-D SiP with vertically interconnected through silicon vias (TSV),” in *Proc. IEEE Int. Electron. Compon. Technol. Conf.*, Jun. 2007, pp. 847–850.
- [46] V. F. Pavlidis and E. G. Friedman, “Interconnect delay minimization through interlayer via placement in 3-D ICs,” in *Proc. ACM Great Lakes Symp. VLSI*, Apr. 2005, pp. 20–25.
- [47] C. A. Bower et al., “High density vertical interconnect for 3-D integration of silicon integrated circuits,” in *Proc. IEEE Int. Electron. Compon. Technol. Conf.*, Jun. 2006, pp. 399–403.
- [48] I. Savidis and E. G. Friedman, “Electrical modeling and characterization of 3-D vias,” in *Proc. IEEE Int. Symp. Circuits Syst.*, May 2008, pp. 784–787.
- [49] S. Im and K. Banerjee, “Full chip thermal analysis of planar (2-D) and vertically integrated (3-D) high performance ICs,” in *Proc. IEEE Int. Electron Devices Meeting*, Dec. 2000, pp. 727–730.
- [50] T.-Y. Chiang, S. J. Souri, C. O. Chui, and K. C. Saraswat, “Thermal analysis of heterogeneous 3-D ICs with various integration scenarios,” in *Proc. IEEE Int. Electron Devices Meeting*, Dec. 2001, pp. 681–684.
- [51] C. C. Liu, J. Zhang, A. K. Datta, and S. Tiwari, “Heating effects of clock drivers in bulk, SOI, and 3-D CMOS,” *IEEE Trans. Electron Device Lett.*, vol. 23, no. 12, pp. 716–728, Dec. 2002.
- [52] Z. Tan, M. Furmanczyk, M. Turowski, and A. Przekwas, “CFD-micromesh: A fast geometrical modeling and mesh generation tool for 3D microsystem simulations,” in *Proc. Int. Conf. Model. Simul. Microsyst.*, Mar. 2000, pp. 712–715.
- [53] P. Wilkerson, M. Furmanczyk, and M. Turowski, “Compact thermal model analysis for 3-D integrated circuits,” in *Proc. Int. Conf. Mixed Design Integr. Circuits Syst.*, Jun. 2004, pp. 277–282.
- [54] G. Digele, S. Lindenkreuz, and E. Kasper, “Fully coupled dynamic electro-thermal simulation,” *IEEE Trans. Very Large Scale (VLSI) Syst.*, vol. 5, pp. 250–257, Sep. 1997.
- [55] M. B. Kleiner, S. A. Kühn, P. Ramm, and W. Weber, “Thermal analysis of vertically integrated circuits,” in *Proc. IEEE Int. Electron Devices Meeting*, Dec. 1995, pp. 487–490.
- [56] T. Zhang, Y. Zhang, and S. Sapatnekar, “Temperature-aware routing in 3-D ICs,” in *Proc. IEEE Asia South Pacific Design Autom. Conf.*, Jan. 2006, pp. 309–314.
- [57] S. Tayu and S. Ueno, “On the complexity of three-dimensional channel routing,” in *Proc. IEEE Int. Symp. Circuits Syst.*, May 2007, pp. 3399–3402.
- [58] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*. Reading, MA: Addison-Wesley, 1989.
- [59] C. Addo-Quaye, “Thermal-aware mapping and placement for 3-D NoC designs,” in *Proc. IEEE Int. SOC Conf.*, Sep. 2005, pp. 25–28.
- [60] R. H. J. M. Otten, “Automatic floorplan design,” in *Proc. IEEE/ACM Design Autom. Conf.*, Jun. 1982, pp. 261–267.

- [61] X. Hong *et al.*, "Corner block list: An effective and efficient topological representation of non-slicing floorplan," in *Proc. IEEE/ACM Int. Conf. Computer-Aided Design*, Nov. 2000, pp. 8–11.
- [62] E. F. Y. Yong, C. C. N. Chu, and C. S. Zion, "Twin binary sequences: A non-redundant representation for general non-slicing floorplan," *IEEE Trans. Computer-Aided Design Integr. Circuits Syst.*, vol. 22, pp. 457–469, Apr. 2003.
- [63] H. Yamazaki, K. Sakanushi, S. Nakatake, and Y. Kajitani, "The 3D-packing by meta data structure and packing heuristics," *IEICE Trans. Fundam. Electron., Commun. Comput. Sci.*, vol. E83-A, no. 4, pp. 639–645, Apr. 2000.
- [64] L. Cheng, L. Deng, and D. F. Wong, "Floorplanning for 3-D VLSI design," in *Proc. IEEE Int. Asia South Pacific Design Autom. Conf.*, Jan. 2005, pp. 405–411.
- [65] Z. Li *et al.*, "Hierarchical 3-D floorplanning algorithm for wirelength optimization," *IEEE Trans. Circuits Syst. I, Regular Papers*, vol. 53, no. 12, pp. 2637–2646, Dec. 2006.
- [66] Y. Deng and W. P. Maly, "Interconnect characteristics of 2.5-D system integration scheme," in *Proc. IEEE Int. Symp. Phys. Design*, Apr. 2001, pp. 341–345.
- [67] P. H. Shiu, R. Ravichandran, S. Easwar, and S. K. Lim, "Multi-layer floorplanning for reliable system-on-package," in *Proc. IEEE Int. Symp. Circuits Syst.*, May 2004, vol. V, pp. 69–72.
- [68] J. Cong, J. Wei, and Y. Zhang, "A thermal-driven floorplanning algorithm for 3-D ICs," in *Proc. IEEE/ACM Int. Conf. Computer-Aided Design*, Nov. 2004, pp. 306–313.
- [69] Z. Li *et al.*, "Efficient thermal via planning approach and its application in 3-D floorplanning," *IEEE Trans. Computer-Aided Design Integr. Circuits Syst.*, vol. 26, pp. 645–658, Apr. 2007.
- [70] T. Yan, Q. Dong, Y. Takashima, and Y. Kajitani, "How does partitioning matter for 3D floorplanning," in *Proc. ACM Int. Great Lakes Symp. VLSI*, Apr./May 2006, pp. 73–76.
- [71] X. Hong *et al.*, "Non-slicing floorplan and placement using corner block list topological representation," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 51, pp. 228–233, May 2004.
- [72] J. M. Lin and Y. W. Chang, "TCG: A transitive closure graph based representation for non-slicing floorplans," in *Proc. IEEE/ACM Design Automation Conf.*, Jun. 2001, pp. 764–769.
- [73] W.-L. Hung *et al.*, "Interconnect and thermal-aware floorplanning for 3-D microprocessors," in *Proc. IEEE Int. Symp. Quality Electron. Design*, Mar. 2006, pp. 98–103.
- [74] W.-C. Lo *et al.*, "An innovative chip-to-wafer and wafer-to-wafer stacking," in *Proc. IEEE Int. Electron. Compon. Technol. Conf.*, Jun. 2006, pp. 409–414.
- [75] B. Goplen and S. Sapatnekar, "Placement of thermal vias in 3-D ICs using various thermal objectives," *IEEE Trans. Computer-Aided Design Integr. Circuits Syst.*, vol. 25, pp. 692–709, Apr. 2006.
- [76] M. Ohmura, "An initial placement algorithm for 3-D VLSI," in *Proc. IEEE Int. Symp. Circuits Syst.*, May 1998, vol. IV, pp. 195–198.
- [77] T. Tanprasert, "An analytical 3-D placement that preserves routing space," in *Proc. IEEE Int. Symp. Circuits Syst.*, May 2000, vol. III, pp. 69–72.
- [78] I. Kaya, M. Olbrich, and E. Barke, "3-D placement considering vertical interconnects," in *Proc. IEEE Int. SOC Conf.*, Sep. 2003, pp. 257–258.
- [79] Y. Deng and W. P. Maly, "Interconnect characteristics of 2.5-D system integration scheme," in *Proc. ACM Int. Symp. Phys. Design*, Apr. 2001, pp. 171–175.
- [80] S. T. Obenaus and T. H. Szymanski, "Gravity: Fast placement for 3-D VLSI," *ACM Trans. Design Autom. Electron. Syst.*, vol. 8, no. 3, pp. 298–315, Jul. 2003.
- [81] A. Harter, *Three-Dimensional Integrated Circuit Layout*. Cambridge, U.K.: Cambridge Univ. Press, 1991.
- [82] W. R. Davis *et al.*, "Demystifying 3D ICs: The pros and cons of going vertical," *IEEE Design Test Comput.*, vol. 22, Nov./Dec. 2005.
- [83] H. Eisenmann and F. M. Johannes, "Generic global placement and floorplanning," in *Proc. IEEE/ACM Design Automation Conf.*, Jun. 1998, pp. 269–274.
- [84] B. Goplen and S. Sapatnekar, "Efficient thermal placement of standard cells in 3-D ICs using a force directed approach," in *Proc. IEEE/ACM Int. Conf. Computer-Aided Design*, Nov. 2003, pp. 86–89.
- [85] B. Black *et al.*, "Die stacking (3D) microarchitecture," in *Proc. IEEE/ACM Int. Symp. Microarchit.*, pp. 469–479, Dec. 2006.
- [86] R. J. Enbody, G. Lynn, and K. H. Tan, "Routing the 3-D chip," in *Proc. IEEE/ACM Design Autom. Conf.*, Jun. 1991, pp. 132–137.
- [87] C. C. Tong and C.-L. Wu, "Routing in a three-dimensional chip," *IEEE Trans. Comput.*, vol. 44, pp. 106–117, Jan. 1995.
- [88] J. Minz and S. K. Lim, "Block-level 3-D global routing with an application to 3-D packaging," *IEEE Trans. Computer-Aided Design Integr. Circuits Syst.*, vol. 25, pp. 2248–2257, Oct. 2006.
- [89] A. Hashimoto and J. Stevens, "Wire routing by optimizing channel assignment within large apertures," in *Proc. IEEE/ACM Design Autom. Conf.*, Jun. 1971, pp. 155–169.
- [90] T. Ohtsuki, *Advances in CAD for VLSI*. Amsterdam, The Netherlands: Elsevier, 1986.
- [91] J. Cong, M. Xie, and Y. Zhang, "An enhanced multilevel routing system," in *Proc. IEEE/ACM Int. Conf. Computer-Aided Design*, Nov. 2002, pp. 51–58.
- [92] J. Cong and Y. Zhang, "Thermal driven multilevel routing for 3-D ICs," in *Proc. IEEE Asia and South Pacific Design Autom. Conf.*, Jun. 2005, pp. 121–126.
- [93] J. Cong and Y. Zhang, "Thermal via planning for 3-D ICs," in *Proc. IEEE/ACM Int. Conf. Computer-Aided Design*, Nov. 2005, pp. 744–751.
- [94] K. D. Boese *et al.*, "Fidelity and near-optimality of elmore-based routing constructions," in *Proc. IEEE Int. Conf. Comput. Design*, Oct. 1993, pp. 81–84.
- [95] A. I. Abou-Seido, B. Nowak, and C. Chu, "Fitted elmore delay: A simple and accurate interconnect delay model," *IEEE Trans. Very Large Scale (VLSI) Syst.*, vol. 12, no. 7, pp. 691–696, Jul. 2004.
- [96] *Metal User's Guide*. [Online]. Available: www.oea.com
- [97] W. Zhao and Y. Cao, "New generation of predictive technology model for sub-45 nm design exploration," in *Proc. IEEE Int. Symp. Quality Electron. Design*, Mar. 2006, pp. 585–590.
- [98] J. Löfberg, "YALMIP: A toolbox for modeling and optimization in MATLAB," in *Proc. IEEE Int. Symp. Computer-Aided Control Syst. Design*, Sep. 2004, pp. 284–289.
- [99] E. G. Friedman, Ed., *Clock Distribution Networks in VLSI Circuits Syst.* Piscataway, NJ: IEEE Press, 1995.
- [100] E. G. Friedman, "Clock distribution networks in synchronous digital integrated circuits," *Proc. IEEE*, vol. 89, pp. 665–692, May 2001.
- [101] D. W. Bailey and B. J. Benschneider, "Clocking design and analysis for a 600-MHz alpha microprocessor," *IEEE J. Solid-State Circuits*, vol. 22, pp. 1627–1633, Nov. 1998.
- [102] T. Xanthopoulos *et al.*, "The design and analysis of the clock distribution network for a 1.2 GHz alpha microprocessor," in *Proc. IEEE Int. Solid-State Circuits Conf.*, Feb. 2001, pp. 402–402.
- [103] N. Hedenstierna and K. O. Jeppson, "CMOS circuit speed and buffer optimization," *IEEE Trans. Computer-Aided Design Integr. Circuits Syst.*, vol. CAD-6, pp. 270–281, Mar. 1987.
- [104] N. C. Li, G. L. Haviland, and A. A. Tuszynski, "CMOS tapered buffer," *IEEE J. Solid-State Circuits*, vol. 25, pp. 1005–1008, Aug. 1990.
- [105] C. Punty and L. Gal, "Optimum tapered buffer," *IEEE J. Solid-State Circuits*, vol. 27, pp. 1005–1008, Jan. 1992.
- [106] A. Jantsch and H. Tenhunen, *Networks on Chip*. Norwell, MA: Kluwer Academic, 2003.
- [107] L. Benini and G. De Micheli, "Networks on chip: A new SoC paradigm," *IEEE Computer*, vol. 31, pp. 70–78, Jan. 2002.
- [108] D. Bertozzi *et al.*, "NoC synthesis flow for customized domain specific multiprocessor systems-on-chip," *IEEE Trans. Parallel Distrib. Syst.*, vol. 16, pp. 113–129, Feb. 2005.
- [109] J. C. Koob *et al.*, "Design of a 3-D fully depleted SOI computational RAM," *IEEE Trans. Very Large Scale (VLSI) Syst.*, vol. 13, pp. 358–368, Mar. 2005.
- [110] S. Kumar *et al.*, "A network on chip architecture and design methodology," in *Proc. Int. IEEE Annu. Symp. VLSI*, Apr. 2002, pp. 105–112.
- [111] M. Millberg *et al.*, "The nostrum backbone—A communication protocol stack for networks on chip," in *Proc. IEEE Int. Conf. VLSI Design*, Jan. 2004, pp. 693–696.
- [112] F. Li *et al.*, "Design and management of 3D chip multiprocessors using network-in-memory," in *Proc. IEEE Int. Symp. Comput. Architect.*, Jun. 2006, pp. 130–142.
- [113] V. F. Pavlidis and E. G. Friedman, "3-D topologies for networks-on-chip," *IEEE Trans. Very Large Scale (VLSI) Syst.*, vol. 15, no. 10, pp. 1081–1090, Oct. 2007.
- [114] V. F. Pavlidis and E. G. Friedman, *Three-Dimensional Integrated Circuit Design*. Morgan Kaufmann Publishers, 2009.

ABOUT THE AUTHORS

Vasilis F. Pavlidis (Student Member, IEEE) received the B.S. and M.Eng. degrees in electrical and computer engineering from the Democritus University of Thrace, Xanthi, Greece, in 2000 and 2002, respectively. He received the M.Sc. and Ph.D. degrees from the Department of Electrical and Computer Engineering, University of Rochester, Rochester, NY in 2003 and 2008, respectively. He is currently with EPFL, Lausanne, Switzerland.



From 2000 to 2002, he was with INTRACOM S.A., Athens, Greece. In summer 2007, he was with Synopsys Inc, Mountain View, CA. His current research interests are in the area of interconnect modeling, 3-D integration, networks-on-chip, and related design issues in VLSI.

Eby G. Friedman (Fellow, IEEE) received the B.S. degree from Lafayette College, Easton, PA, and the M.S. and Ph.D. degrees from the University of California, Irvine, all in electrical engineering.



From 1979 to 1991, he was with Hughes Aircraft Company. He has been with the Department of Electrical and Computer Engineering, University of Rochester, Rochester, NY, since 1991, where he is a Distinguished Professor. He is also a Visiting Professor at the Technion—Israel Institute of Technology, Haifa. His research interests are in high performance synchronous digital and mixed-signal microelectronic design and analysis. He is the author of more than 300 papers and book chapters. He has received several patents. He is the author or editor of ten books in the fields of high-speed and low-power CMOS design techniques, high-speed interconnect, and the theory and application of synchronous clock and power distribution networks.

Dr. Friedman is a Senior Fulbright Fellow. He was Editor-in-Chief of the IEEE TRANSACTIONS ON VERY LARGE SCALE INTEGRATION (VLSI) SYSTEMS and a Member of the Editorial Board of the PROCEEDINGS OF THE IEEE. He received the University of Rochester Graduate Teaching Award and a College of Engineering Teaching Excellence Award.