ABSTRACT

New technologies such as 3D integration are becoming a new force that is keeping Moore’s law in effect in today’s nano era. By adding a third dimension in current 2D circuits, we can greatly increase integration density, reduce interconnection length, and enable heterogeneous systems within one package. In order to exploit the advantages of 3D integration, layout designers and tool developers need to be fully aware of this rapid development. This paper gives an overview of recent 3D integration technologies, such as 3D packages and 3D integrated circuits. We then analyze and compare 3D data structures in order to draw conclusions about their future potential. Finally, the impact of 3D technologies on interconnect prediction is discussed.

Categories and Subject Descriptors

B.7.2 [Integrated Circuits]: Design Aids—Layout; B.7.1 [Integrated Circuits]: Types and Design Styles—Advanced technologies

General Terms

Algorithms, Design

Keywords

Three-dimensional Circuits, 3D Integration, 3D Floorplanning, Data Structures, Interconnect Prediction

1. INTRODUCTION

As predicted by Moore’s law, modern technologies allow high density integrated circuits with hundreds of millions of transistors. This remarkable progress has been principally achieved by reducing feature sizes. However, ongoing reduction of the lithographic features is increasingly expensive. Another enabler of Moore’s law has been the introduction of new technologies (e.g., copper interconnects, SOI, strained silicon). Currently, it is becoming more and more obvious that new technologies and methodologies rather than reduction in feature size are the key to further performance enhancements. The employment of 3D integration technologies, in which the active devices are placed in multiple layers, is one promising possibility to achieve a performance boost. Although discussed for some decades, 3D integration has only recently gained practical importance.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Figure 1: The increasing interconnect delay for various technology nodes according to the ITRS roadmap [1].

The power consumption and performance of a chip are increasingly influenced by interconnects (Fig. 1). Global wires which do not scale well need a growing number of repeaters. 3D integration allows to shorten interconnects which results in an improvement of performance and lower power consumption. Higher data bandwidths are possible and enable the design of highly efficient caches for microprocessor architectures. Smaller footprint sizes increase the yield of fabrication and improve the usage for mobile devices with tightened weight requirements. Overall, 3D integration reduces total wire length, signal delay, buffer count and power consumption.

Furthermore, a new set of devices, such as vertical transistors, could dramatically change the way integrated circuits are designed.
If there are so many advantages, why isn’t 3D integration widely used yet? One reason is that until now the scaling of the transistors was easier to achieve than investing into a new methodology. Furthermore, 3D integration faces enormous challenges in both technology and physical design. Typical problems in 3D integration technologies are, among others, reliability, alignment accuracy, testability and thermal issues. The physical design process for 3D technologies is even more complex than its 2D counterpart. Here, the third dimension dramatically increases the solution space. Besides the higher complexity, additional constraints, such as stringent thermal requirements, must be addressed.

In order to handle these complex problems in physical design, a feasible modeling of the layout problem is necessary. Specifically, all design steps need an efficient data structure to represent the real geometries of modules in the algorithms being used. Recently developed 3D data structures have improved this mapping and thus allow the use of efficient optimization approaches.

This paper gives an overview of currently applied 3D integration technologies and the 3D data structures that are required in order to model these technologies during design automation. One contribution of this paper is to categorize these data structures and describe their use in modeling the geometries of modules used in designs – a prerequisite for efficient layout optimization. We also discuss the consequences of 3D integration on interconnect prediction with emphasis on modeling interconnect resources and density. Layout designers and tool developers may find this information helpful in assessing new developments in this rapidly changing field.

In the first part of this paper we provide an overview of the different 3D integration technologies that are in use today. In the second part, modern data structures used for 3D integration are presented with a special emphasis on 3D floorplanning problems. In the third part, we discuss implications of 3D integration methodologies on interconnect prediction.

2. 3D INTEGRATION TECHNOLOGIES

A traditional integrated circuit is comprised of one active device layer covered with several metal layers. The transistors are integrated in the active layer and the metal layers are used to realize the interconnections between them. This so-called two dimensional (2D) approach is currently reaching its limits because, despite shrinking feature sizes, the rapidly increasing wire lengths pose a severe problem to electrical signal properties. For example, the total interconnect length of a chip fabricated with 32 nm and six metal layers is predicted to reach 3125 m/cm² in 2013 [1]. The maximum circuit area is also limited by the given signal propagation. Hence, the recent trend towards multifunctional mixed-signal circuits is hardly reachable with conventional 2D technology.

One solution to this problem has been the increased implementation of so-called 3D circuits based on 3D integration technologies [2, 3]. Here, the active devices are not limited to one layer, but are also placed on top of each other. One famous example for the usage of this technology is the stacking of memory dies to realize high capacity memory cards. Through silicon vias (TSVs) are, among others, an enabling technology for this kind of integration. These vertical electrical connections that are passing completely through a silicon wafer or die require diameters and pitches significantly larger than that of conventional (signal) vias. Currently, TSVs need contact pitches between 2.6 and 3.8 μm and diameters of 1.3 - 1.9 μm [1].

A distinguishing feature of 3D integration technologies is the hierarchy level of the vertical integration. 3D integration on package level such as System on Package (SoP) and System in Package (SiP) are often classified as 3D packages (see Section 2.1). 3D integration on wafer level is another promising 3D integration technology that leads to so-called 3D integrated circuits (see Section 2.2). Here, multiple device layers are integrated directly on wafer level, e.g. by using donor wafer bonding or wafer stacking.

Figure 2 shows an overview and classification of different 3D integration technologies.

2.1 3D Packages

3D packages are separate chips or bare dies stacked vertically in order to build a multiple device layer circuit. In contrast to 3D integrated circuits which use through silicon vias (TSVs), the vertical connections between these chips or dies are realized using their external pins/pads. This allows the combination of chips with a wide range of technologies to be combined into one package.

Important enabling technologies for 3D packages are wire bonding and C4 which stands for Controlled Collapsed Chip Connection, i.e., a flip-chip related technology. Interconnects between the various modules in the package comprises all types of signals such as digital, analog and optical.

As shown in Fig. 2, the best known representatives of 3D packages are: 3D System on Chip (SoC), 3D Multi-Chip Module (MCM), Package on Package (PoP), System in Package (SiP), and Systems on Package (SoP). While the SoC technology combines all modules on one silicon chip, MCMs consist of numerous chips (bare dies) mounted on a common (ceramic) substrate and interconnected on that substrate (Fig. 3 b). SiP evolved out of both SoC and MCM technologies by vertically stacking multiple chips (bare dies) within one package (Fig. 3 a, c). Package on Package stacks several packed circuits for higher integration (Fig. 3 d).

One of the most enhanced 3D packaging method is the System on Package (SoP, Fig. 3 e) [4]. It combines the ad-
3D Integrated Circuits

Figure 4: Classification of 3D ICs with regard to their technology. While layer growth (a) and donor wafer bonding (b) require a sequential processing of the various device layers, wafer level stacking (c) enables a parallel fabrication of these layers.

Figure 5: Fabrication of a Silicon on Insulator (SoI) wafer [6, 7]: (a) Circuit layer 1 is attached to a glass carrier, etching process at the backside, buried oxide layer as etch stop, (b) alignment and bonding of circuit layer 1 onto layer 2, (c) removal of top glass carrier and fabrication of through silicon vias for interlayer interconnects.

In comparison to 3D packages, 3D integrated circuits achieve a higher level of integration. Key technology enablers are through silicon vias (TSVs), wafer thinning, high accuracy alignment and viable wafer bonding methods.

One example of the layer growth technology is metal-induced lateral crystallization (Fig. 4 a). Here, after the deposition of amorphous silicon, a recrystallization process is induced by a partial deposited metal layer at higher temperatures (> 400 °C). While this technology enables a high via density, the temperature of the growth process can have negative effects on the lower device layers.

Donor wafer bonding is based on a sequential fabrication of multiple circuit layers, sequentially from bottom to top (Fig. 4 b). This method avoids temperature-related drawbacks on the lower layers due to the absence of a recrystallization process.

Various prefabricated wafers (circuit layers) are combined either face-to-face (F2F) or face-to-back (F2B) using the wafer level stacking technology (Fig. 4 c). Even though the process is called wafer level stacking, so-called die-to-wafer methods (D2W) can be used to decrease the influence of defects on single wafers. This increased fabrication yield is achieved by stacking single, pre-tested dies (rather than an entire wafer) onto the wafer.

The IBM 3D IC fabrication process (described in [6, 7]) is a promising example for 3D integration on wafer level. This method is shown in Fig. 5. In order to achieve thin circuit layers required for short interlayer vias, Silicon on Insulator (SoI) wafers with a buried oxide (BOX) layer are used. The bonding of one circuit layer onto another is realized by either oxide fusion, eutectic bonding or the use of adhesives. Top circuit layers are often supported by a glass carrier.
Due to their high level of mutual integration, the various modules of 3D circuits require a top-down hierarchical design with a significant degree of interdependencies. Outlines of physical design of 3D ICs are given in [8–10].

2.3 Challenges and Comparison of 3D Technologies

3D integration is the logical consequence of the increased bottleneck of 2D integration with regard to interconnect problems that prevent further improvements in performance and functionality. However, these new technologies pose severe challenges that include technological problems, like high accuracy alignment or reliability issues, and physical design issues such as considering new thermal constraints. A comparison of 3D integration technologies with regard to their main characteristics is given in Table 1.

### Table 1: Comparison of 3D integration technologies.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>3D Packages</th>
<th>3D Integrated Circuits</th>
</tr>
</thead>
<tbody>
<tr>
<td>SiP/SoP/PoP</td>
<td>Low</td>
<td>Very high</td>
</tr>
<tr>
<td>Fabrication</td>
<td>Parallel</td>
<td>Sequential</td>
</tr>
<tr>
<td>Distance between layers</td>
<td>High</td>
<td>Very low</td>
</tr>
<tr>
<td>Heterogeneous configuration</td>
<td>Yes</td>
<td>Difficult</td>
</tr>
<tr>
<td>Thermal interaction between layers</td>
<td>Low</td>
<td>Very high</td>
</tr>
<tr>
<td>Testability</td>
<td>Chip level</td>
<td>Difficult</td>
</tr>
</tbody>
</table>

Sequence Pair \( (acbed, abcdfe) \) \Rightarrow Sequence Triple/Quintuple \( (bdefca, dfabcd, abdef, fabced, cabfdef) \)

Figure 6: Classical 2D floorplanning data structures (left) and their 3D counterparts (right, see also Table 2).

3. 3D DATA STRUCTURES

3.1 Overview

In electronic design automation, specifically during physical design, data structures are used to store information about various layout elements and their properties. Hence, data structures are an abstract model of the corresponding design problem.

Efficient data structures integrate additional helpful properties of the layout elements, such as direct access to neighbors. They build up a solution space which then can be investigated with an optimization method like simulated annealing. This solution space should be non-redundant, as small as possible and include the best solutions. Furthermore, an efficient implementation of a data structure must allow a fast execution of operations like rotation, exchange and the transformation from the abstract representation to the real geometry.

Early 3D physical design flows utilized classical, proven data structures, such as one slicing tree for one active device layer. However, these so-called 2.5D approaches have the disadvantage that tight linking between layers has been neglected, thus, preventing for example a successful thermal-driven design.

It is obvious that modern 3D data structures should fully support 3D layouts. As illustrated in Fig. 6, the majority of the currently developed 3D data structures are extensions of their well known 2D predecessors. One example is the T-Tree presented in [11]. Here, the efficient properties of its 2D version (B*-Tree) could be maintained by implementing a ternary tree (instead of a binary tree) that enables the representation of the third dimension. However, some of these 3D extensions of data structures, such as sequence triple [12], lack those properties which make them so efficient in the 2D case. New data structures specifically designed for the 3D integration – which are truly able to fulfill the new demand determined by 3D integration technologies – are required. Next, we outline the progress that has been achieved so far in the area of 3D floorplanning.

3.2 3D Data Structures in Floorplanning

During floorplanning, the shapes and positions of various modules (circuit partitions) are determined. While in many 3D applications placement and routing within the individual modules are performed separately and thus, can maintain their overall 2D data structure, the top-level floorplanning of these modules requires a 3D approach. As such, floorplanning algorithms were among the first to tackle the challenges posed by 3D integration.

Conventional floorplanning assumes a single two-dimensional layer on which several modules have to be arranged. A wide variety of different algorithmic approaches have been...
4. 3D INTERCONNECT PREDICTION

Predicting characteristics of interconnect systems has been essential for any technology in order to evaluate interim solutions during the IC design development. But what distinguishes interconnect prediction for 3D circuit technologies from that of previous technologies?

4.1 Modeling Interconnect Resources

Interconnect resources of an IC can be generally modeled by a rectilinear graph with weighted edges as shown in Fig. 9(a). It consists of nodes that correspond to regions in the design, and weighted edges that correspond to interconnect resources in those regions. Each of the nodes is assigned a design coordinate (x, y, z). x and y are continuous coordinates within a layer, while z is discrete and specifies the interconnect layer. In most practical applications, however, x and y are artificially restricted to discrete values. Furthermore, preferred directions for interconnect are often assumed for layers. These two measures simplify the design process by limiting the number of nodes and reducing the number of edges of the graph (Fig. 9b).

The general model of interconnect resources described above is valid for 3D technologies as well as for previous (2D) technologies. On a first sight, this observation would imply that interconnect routing and prediction are not more complicated for 3D ICs. However, during 3D interconnect prediction the number of layers that has to be taken into account is larger, design objectives are different compared to previous technologies, and additional constraints must be considered.

Figure 7: Illustration of 3D Slicing Tree operations to permute a given 3D floorplan. A rotation alters an inner node (representing a cut through the normal plane) resulting in a physical rotation of modules contained in the subtrees of that node. An exchange swaps two subtrees resulting in a physical exchange of modules contained in these subtrees.

Figure 8: 3D block swapping example using new 3D data structures. As illustrated at the top, block a is exchanged with blocks (b, c).
Table 2: Most common 3D data structures applied to 3D floorplanning problems with regard to their publication date, worst-case complexity of the operations, size of the solution space and main characteristics \((n = \text{number of modules, } n.g. = \text{not given})\). Except the last two data structures, all are directly derived from 2D predecessors (see also Fig. 6).

<table>
<thead>
<tr>
<th>Data Structure</th>
<th>Year</th>
<th>Complexity</th>
<th>Solution Space</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequence Triple/Quintuple</td>
<td>2000</td>
<td>(O(n^2))</td>
<td>(O((n!)^3)/O((n!)^5))</td>
<td>Sequence: Three or five sequences (locii)</td>
</tr>
<tr>
<td>3D Sub-Transitive Closure Graph</td>
<td>2004</td>
<td>(O(n^2))</td>
<td>(O((n!)^3))</td>
<td>Graph: Three transitive graphs</td>
</tr>
<tr>
<td>T-Tree</td>
<td>2004</td>
<td>(O(n))</td>
<td>(O(n! \cdot 3^{(n-1)} / 2^{2n} \cdot n^{1.5}))</td>
<td>Graph: Ternary tree, nodes: modules, branches: neighbor information</td>
</tr>
<tr>
<td>3D Bounded-Sliceplane Grid</td>
<td>2005</td>
<td>(O(n^2))</td>
<td>depends on grid</td>
<td>Grid: Three dimensional grid</td>
</tr>
<tr>
<td>3D Corner Block List</td>
<td>2005</td>
<td>(O(n))</td>
<td>(O(n! \cdot 3^{(n-1)} \cdot 2^{4n-4}))</td>
<td>Sequence/List: Sequence of modules, list of orientations, list of tri-branches</td>
</tr>
<tr>
<td>3D Slicing Tree</td>
<td>2005</td>
<td>(O(n))</td>
<td>(O(n! \cdot 3^{(n)} / n^{1.5}))</td>
<td>Graph: Binary tree, inner nodes: slices, leaves: modules</td>
</tr>
<tr>
<td>O-Sequence</td>
<td>2006</td>
<td>(O(n))</td>
<td>n.g.</td>
<td>Sequence: Sequence of modules and symbols</td>
</tr>
<tr>
<td>Double Tree and Sequence</td>
<td>2007</td>
<td>(O(n^2))</td>
<td>(O(n! \cdot n^{2(n-1)}))</td>
<td>Sequence/Graph: Two rooted trees (x-tree, y-tree), sequence (z-order)</td>
</tr>
<tr>
<td>Labeled Tree and Dual Sequences</td>
<td>2008</td>
<td>(O(n^{4/3} \log n))</td>
<td>(O((n!)^2 \cdot n^{n-1}))</td>
<td>Sequence/Graph: Sequence of modules, number sequence and tree</td>
</tr>
</tbody>
</table>

Figure 9: A general graph modeling interconnect resources (a) and a graph with a reduced number of edges by assuming preferred routing directions (b).

- Interconnect in 3D ICs topologically differs from interconnect in previous (2D) technologies such that it potentially connects elements that reside in different active layers. Hence, the endpoints of a net may be located on different layers that are vertically “far apart”.

- The impact of through silicon vias (TSVs) on electrical interconnect properties is much larger than the impact of conventional signal vias. For example, TSVs are often significantly longer than regular signal vias which connect layers in the same tier.

- The resources for vertical interconnect between tiers (inter-tier vias, subsequently labeled as TSVs) are more scarce/expensive compared to vias between interconnect layers within a tier. As such they require specific consideration during interconnect prediction. While the exact properties of TSVs are technology dependent, the TSV pitch is at least one magnitude larger than the signal via pitch for current and foreseeable 3D technologies (see Section 2). Additionally, reliability and tolerances of TSVs are an issue.

- In general, a \(k\)-tier 3D IC generates \(k\) times the heat of a 2D chip. Hence, heat distribution and heat dissipation are much more important in 3D ICs. As a consequence, thermal vias and other thermal elements have to be integrated into 3D ICs and require consideration during 3D interconnect prediction. Furthermore, the influence of heat on the electrical properties of interconnect becomes more important.

- Blockage avoidance and congestion management are more complex with the addition of a third dimension. For instance, a TSV or thermal via that spans two or more tiers constitutes a blockage that wires have to navigate around. Furthermore, these vias must traverse active layers, thus preventing any active component within that area. Often, an additional keepout area of the TSV is required because transistors cannot be placed in TSV vicinity due to the impact of TSV-related stress.

Interconnect prediction has been always a trade-off between accuracy and computation time. This is even more relevant for 3D ICs, as the problem size is generally larger. Depending on the desired accuracy, the size of the graph that models interconnect resources is reduced by applying different levels of simplification. Referring to 2D technologies, and in particular to the prediction of routing congestion and wire lengths, there are three means of simplification (illustrated in Fig. 10):

- Individual layers of interconnect are combined and, hence, not distinguished during prediction. This is practical if the interconnect in the combined layers has similar properties and if blockages are distributed evenly. The resulting interconnect prediction is not providing information about vias when using this simplification.
Interconnect layers are stored with the resulting tiles treated as atomic units that are not inspected in detail. This simplification is a basic concept of global routing. It reduces the x- and y-resolution of interconnect prediction by an arbitrary factor, which is adjustable to the requirements of specific designs.

- Modeling of interconnect can be disassociated from the layout topology. At this level of simplification, only global values, such as the average wire length of all nets in a design, can be estimated. Prediction methods that rely on few or no topological features are, for example, Rent’s Rule [21] and methods of complexity analysis [22].

As illustrated in Fig. 11, a new intermediate level of simplifying the graph of interconnect resources is needed for 3D technologies. While interconnect modeling to two dimensions has proven to be practical in many cases for previous (2D) designs, 3D technologies require specific consideration of inter-tier vias as they are an expensive resource and significantly influence interconnect properties. Hence, it is necessary to distinguish whether layers are connected by vias or by more expensive TSVs. Consequently, in an intermediate level of simplification, only layers within the same tier are combined, because at this level of simplification the properties of inter-tier vias are still to be included in the interconnect prediction. If further simplification is required, layers of different tiers are also combined (Fig. 11). However, this step removes all information about inter-tier via properties which limits its applicability.

4.2 Interconnect Density Prediction

Lou’s statistical model [23] provides a method to predict interconnect congestion in 2D while respecting blockages. It is based on the assumption that each net is routed along one of its shortest Manhattan paths. From each routing grid point within the Manhattan window, the net is routed in x-direction with the same probability as in y-direction. This model allows computing an expected 2D interconnect density for each routing region in the design (Fig. 12).

As suggested above, interconnect prediction models that consider inter-tier vias (TSVs) and other vertical/blockage information are necessary for 3D technologies. To convey Lou’s model to 3D technologies with several tiers, a probability for vertical routing segments between tiers (layers) must be added in order to include those requirements (see also [24]). Such an extended model respects blockages and vertical resources that naturally differ throughout the design area. The resulting interconnect prediction considers the expected interconnect density for each tier as well as an expected density of inter-tier vias (intermediate level, Fig. 13).

5. SUMMARY

3D integration is becoming a new force that is keeping Moore’s law in effect in today’s nano era. Numerous 3D data structures have been recently developed in order to make 3D technologies accessible to an efficient and automatic design. To fully exploit the advantages of the extra dimension in 3D integrated circuits, layout designers and tool developers need to be aware of this rapid development. We have presented a survey of the emerging 3D integration technologies and provided an up-to-date overview of the most efficient 3D data structures. While 3D integration technologies can be divided into two groups (3D packages and 3D integrated circuits), related data structures fall into three categories: sequence and/or list representations, grids, and graph-based
6. REFERENCES


Acknowledgments

This work is supported by the German Science Foundation (DFG, project 1401/1).