A Study of 3DIC Kogge-Stone Circuits

Christopher Erb, Nicholas Kelly

Introduction

Component delays have traditionally been the focus for performance optimization of logic circuits. As component sizes and delays shrink, wire delays become more significant. Awareness and planning of component placement to optimize wire length then becomes more valuable. The addition of another dimension through die stacking creates more opportunities for shortening this length by shrinking the absolute distance between components. Through Surface Vias (TSVs) act as the electrical connection between die layers. We've found that TSVs have RC characteristics that create a communication delay that is less than the delay created by the length of wire needed to separate modules in a two dimensional space, meaning the incurred delay can be lessened by separating the modules by die layer. While TSVs have beneficial delay characteristics, they have an area requirement far more expensive than their wire counterparts, which may make them hard to place in the design, especially if many signals need to transfer layers. To demonstrate the effectiveness of this method we choose a Kogge-Stone adder because of it notoriety for having many long wires.

Methodology

Partitioning Strategies

In order to minimize the amount of TSVs and maximize the possible gain from using 3D Integrated Circuit (3DIC) technology, an efficient partitioning method of a Kogge Stone adder must be selected. As discussed in [1], there are a few partitioning methods which have been used (see Figure 1). A conventional 2D Kogge-Stone adder has long interconnects, mainly in the later stages (lower levels) of the adder, especially for large widths. Additionally, routing such wires is difficult due to conflicts with other wires and delay constraints. Thus, one design which is proposed is to simply put each level on its own 3DIC tier (Level-Assigned, LA). However, since TSVs only travel in the z direction, this method will not alleviate the issue of interconnect length in the x and y directions. Furthermore, the method adds additional area and capacitance. The merits of this design are that it can reduce the interconnect congestion and each input or output is on the same tier (not so for other designs). The number of vias across layers is equal to: $n_t = 2(2\ log_2(N-1) - N - 2)$ for a *N*-wide adder (assuming one layer per level).

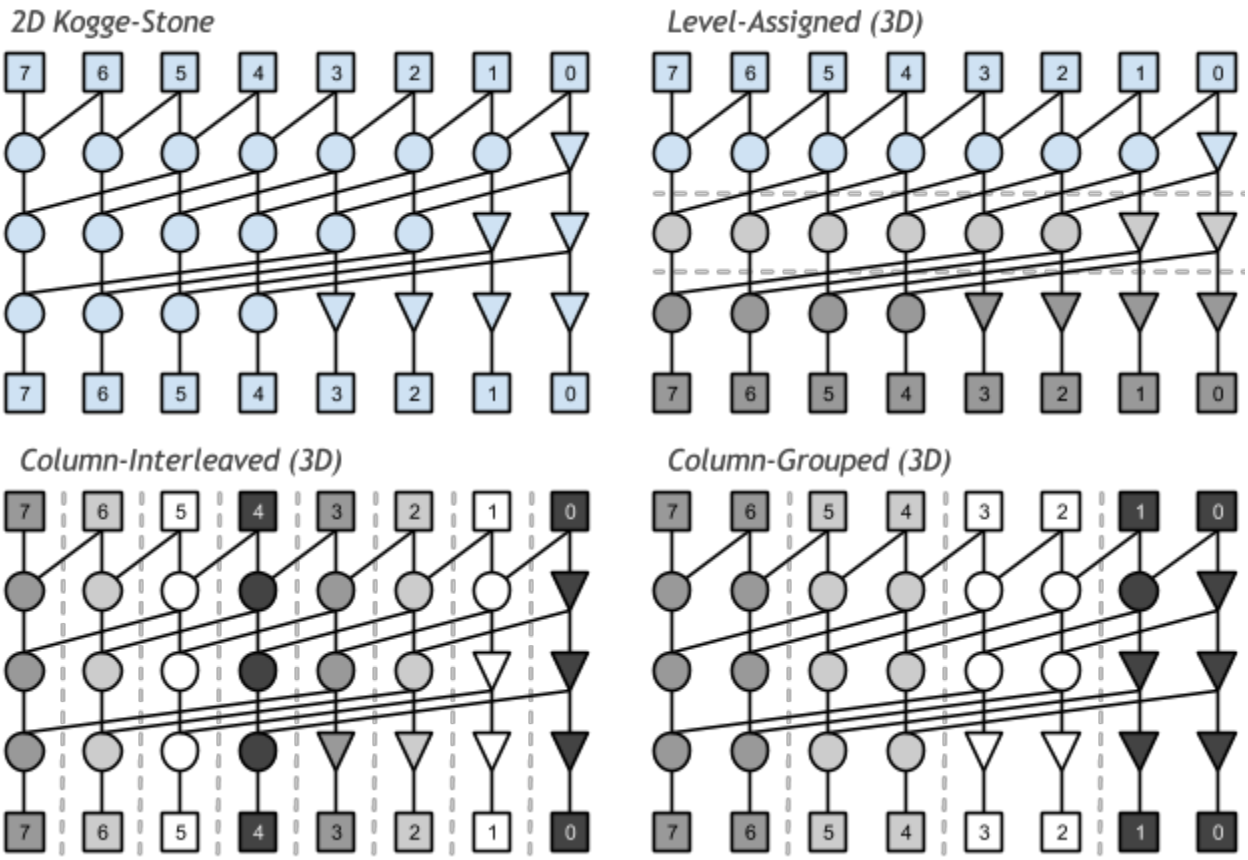

Figure 1: Kogge-Stone partitioning methods (color corresponds to tier)

Another 3DIC strategy is to interleave the columns (Column-Interleaved, CI) to different layers. The advantage of this method presents itself in later stages (depending on the amount of layers used) where previously long interconnects are now close by (e.g. 7 and 3). However, previously close interconnects are now farther away and require TSVs. Thus the distribution of TSVs is heavily in the early stages and partially in the later stages depending on the amount of tiers used. An additional drawback is the fact that the input and output are on different layers (requiring additional TSVs). The amount of TSVs used in the CI strategy is proportional to: $n_t = 2N\ log_2(N)$ for a *N*-wide adder (number of tiers still affects number of TSVs). The final partitioning strategy is to group columns by layer (Column-Grouped, CG). This strategy is nearly the equivalent of having many smaller adders on multiple tiers. Interconnect length can be at most the length of one tier (depending on layout). Compared to the CI layout, the CG method has a fairly even distribution of TSVs, though there are more in the later stages. The amount of TSVs used in the CG strategy is proportional to: $n_t = 2N\ log_2(N)$; however, it is still dependent on the number of tiers. While there are possibly other partitioning methods, only those which partition across columns (i.e. split column-wise) will lead to significant improvement in interconnect length and delay. We looked into other methods (variants of CI and CG), but none seemed efficient in terms of layout.

While the number of TSVs certainly matters in terms of area, the length of each TSV has some effect as well. As [2] discusses, the length of TSVs does not have impact on the overall capacitance or resistance of the TSV; however, long TSVs can cause more congestion for dies it intersects. Since TSVs can be comparatively large (>5μm) and go through all metal and substrate layers, they can constrict placement of MOSFETs and interconnects. For a circuit with many possible TSVs (such as a Kogge-Stone adder), this can lead to much more area than desired.

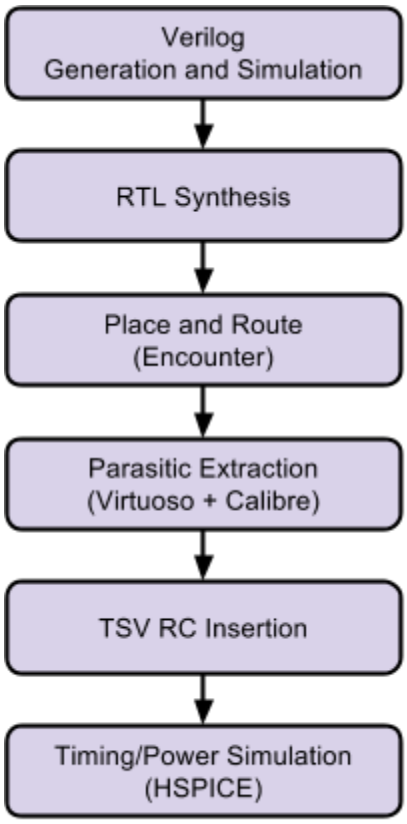

Figure 2: 3DIC workflow

3DIC Workflow

Since 3DIC technology and design tools are in the very early stages and current (or available) tools do not support multiple tiers or TSVs, we had to use a custom workflow. Our workflow is described in Figure 2 and later sections. Overall, the flow consists of generating functional (and structural) models in verilog. Then, Place and Route is used to create layouts for each tier. Using the layouts, the parasitic RC values are extracted into SPICE netlists. Finally, the netlists are combined and simulated in HSPICE with stimuli to yield power and timing values.

Verilog Models

Verilog models were created for two 32-bit Kogge-Stone adders. One model is a reference design for a standard 2D implementation and the second is our multi-module design for 3D die stacking using the CG strategy. The 2D implementation is written in structural verilog and organized into initialization phase, the 5 generate/propagate levels, and the addition phase. For the 3D implementation, a custom verilog generator was created which allows for Kogge-Stone adders to be generated from arbitrary tiers and input widths. Due to the logarithmic nature of the adder, construction is relatively simple. The generator conforms to the CG strategy (though could easily be expanded to other strategies) and groups tiers into verilog modules where each input/output is effectively a TSV. In this manner, each module can be put into Place and Route separately, though a top level module is created for reference and simulation purposes. The CG strategy was selected due to its simplicity, even distribution of TSVs, and interconnect benefits. However, an implementation and analysis of the different methods is considered future work.

Layout & Extraction

For our interactions with Cadence we use the 45nm libraries provided by NCSU in the FreePDK45 <cite?>. The RTL is imported into Encounter and synthesized. From this we get a gate level verilog file, from which we can import our designs into Encounter. The 2D design is imported all at the same time. It is interesting to note that the layout that Encounter creates for the 2D design places the generate/propagate (GP) levels in concentric circles with the first level being the outermost and each level moving closer towards the center. Connections between GP cells in Kogge-Stone adders cross twice as many bit positions at each level compared to the previous level, so it is beneficial that later levels be closer together. A layout is generated for each tier in the 3D model individually. Initially we tried separating modules/tiers in a 2D design space, but it would have not modeled TSV delay accurately. This is because the wire length equivalent to model the delay of a TSV would place the modules on top of each other. Module/Tier inputs and outputs (which signify TSVs) are routed to the edge of the die (I/O) which can signify the required routing to the TSV (see discussion about TSV placement). After layout, each layer is imported into Virtuoso for parasitic RC extraction. Calibre PEX is used in order to perform LVS and generate a SPICE netlist for the tier, complete with distributed-RC wire models.

TSV Modeling & Simulation

After SPICE netlists have been created for each tier in the design, they need to be stitched together. This stage takes each SPICE subcircuit and connects them together while also inserting TSV models at the inputs/outputs (except for initial summation input/output). The TSVs are modeled as a distributed-RC model depending on their depth (see discussion for more information) [2]. Additionally, depending on TSV placement (elaborated in discussion), RC models for the routes could be removed or altered in order to simulate TSVs which are closer to the signal's point of origin. Thus, the result is a single SPICE netlist which effectively emulates a 3DIC process. Finally, the combined SPICE netlist is simulated in HSPICE with various stimuli. Each voltage stimulus is modeled as a step-function in order to adequately record dynamic and static power. Timing is measured based upon when the sum bits switch (input dependent).

Discussion & Results

Theoretical 2D/3D Analysis

Modern tools are capable of efficiently organizing components in order to minimize wire delays experienced by components on a two dimensional grid, however, they do not address three dimensional spaces. We have calculated that delay can be shortened by using TSVs and moving components to adjacent layers. A 5μm diameter via has capacitance between 8 fF and 70 fF for heights between 5μm and 100μm [2]. The delay for a 5μm wide, 20μm tall, copper via is as follows: A 5μm wide via that is 20μm in height has a capacitance of 18.3 fF. The resistance equation for a via is the height of the via divided by the conductivity of the metal times its area:

$$R_{via} = \frac{l_v}{\sigma \pi r_v^2}$$

At 20°C copper has a conductivity of 59.59 M $\Omega^{-1} \cdot m^{-1}$, so $R_{via} = 17.09\ m\Omega$. Hence the $(RC)_{via}$ delay is 0.313 fs. If we compare this to metal4 - metal6 delay (from NCSU FreePDK45):

$$(RC)_{wire} / \mu m^2 = 0.0751 \text{ f s}/\mu m^2,\ (RC)_{wire} / \mu m^2 = (RC)_{via} / x^2$$

Where R = 0.439 Ω/μm, C = 0.171 fF/μm, then x = 2.04 μm. This via has the equivalent delay to a metal(4-6) trace that is 2 μm long. There are many instances in a 32-bit Kogge-Stone adder where wires need to travel much farther than 2 μm, and using a die stacked model we can reduce the delays experienced by a portion of these longer wires to that experienced by a wire of length 2 μm.

TSV Placement

TSV placement is complicated by its large footprint, 5 μm in diameter. The area footprint of a TSV is quite large compared to metal traces and transistors. This will quickly consume potential component area and block other signal traces, so careful placement and signal choices are essential to avoid increasing inter-component distances. Currently, in our models we assume the TSVs are placed on the outer edges of each tier. However, since the interconnect has to be routed to that location, it causes diminishing returns on performance/power gains. Instead, TSVs could be dispersed throughout the circuit, ideally near the point of origin (difficulties arise with area constraints and standard cell placement).

Results

We focused on a 32-bit Kogge-Stone adder since it has been widely used in many architectures and [1] has focused on larger (greater than 512-bit) adders. As previously stated, we created a custom flow to model the effects of TSVs in a 3DIC process (using conventional 2D techniques). We followed our flow for the results shown in Table 1, for various different tier counts for a 32-bit Kogge-Stone adder. As initially expected, the area usage decreases at least by a factor of two upon doubling the amount of tiers each time. Though, it is important to note that these numbers do not incorporate the TSV area. If the TSV area was incorporated the numbers would be either very similar or in favor of 2D.

However, delay and energy actually seems to increase with increased tiers. Since previous authors have published some research indicating improvement (often minor), we can conclude that there might be some inaccuracies in our modeling of 3DICs. The main inaccuracy within our system is there are still large routing paths from the point of origin to the edge of the die (see Figure 3). However, as stated previously, this depends on the TSV placement, which is not currently modeled within our Place and Route system. We make the assumption that the TSVs are on the outer edges of the die, thus the need for routing. If instead the TSVs were placed in a grid-like fashion or strategically (based upon timing), we would probably expect to see similar results as previous publications.

# Tiers	Area (μm^2)*	Delay (*ns*)	Power (*mW*)	Energy (*pJ*)
1	2596	0.70	3.13	2.16
2	665 (+ 63 TSVs)	1.10	2.24	2.45
4	250 (+ 125 TSVs)	1.15	3.04	3.47

Table 1: Initial results (* TSV area not factored in)

However, these results show how dependent delay improvement for is for TSVs. In terms of power, the results were similar or better than the 2D model. This might be attributed to the decreased resistance of TSV compared to metal layers. Overall, these results show that more work needs to be done for our model, and that TSV performance is highly dependent on TSV placement, the amount of TSVs, and the TSV technology.

Related Work

3D Stacked Adders

Voicu, et al. [1] conducted a case study on 3D stacked adders, focusing primarily on wide-input, prefix adders which are important for cryptographic operations (e.g. RSA). In addition to proposing a new partition method (CG), they analyze Brent-Kung and Kogge-Stone adders using various partition methods. They compared the the different configurations spanning across different amounts of tiers (2, 4, 8, and 16) and different input widths (512, 1024, 2048, and 4096 bit) in terms of delay, footprint, and cost. Since cost was an important factor, they developed a 3D hybrid adder which contained a carry-select adder stage on each tier where each adder was a Kogge-Stone or Brent-Kung. The advantage of this design is that only one TSV is required between layers (carry), where many TSVs are required for 3D Kogge-Stone adders. Additionally, each tier was identical (not so for other designs), which when coupled with less TSVs amounted to a floorplan and cost reduction.

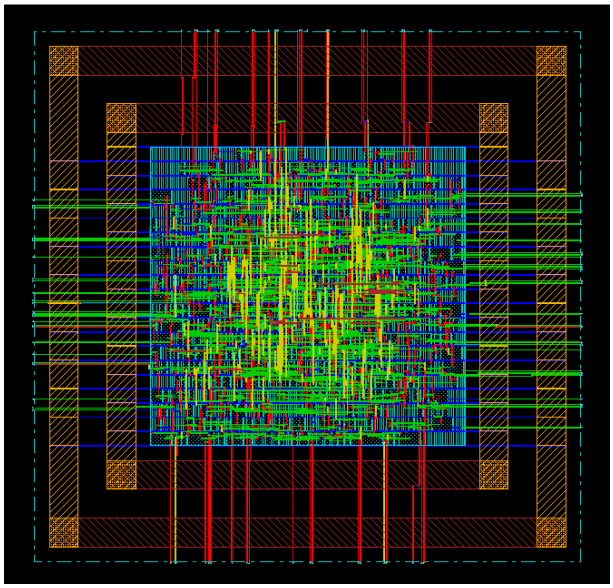

Figure 3: Place and Route

Additionally, Ouyang, et al. [5] proposed a Kogge-Stone adder design in 180nm technology. Their Kogge-Stone design contained higher-radix cells on early levels, and they used the CI partitioning method. Additionally, they explored 3D Wallace multipliers which split different counting stages onto different tiers. Their final design was implemented in 1.3mm x 1.3mm and contained a 36-bit

Kogge-Stone adder and a 32 x 32 Wallace multiplier. In their initial results, they found that the 3D Kogge-Stone scales better in terms of delay, where the 2D has exponential delay for large input widths (as seen in [1]).

3DIC and TSVs

While there are several commercial tools available for 3DIC development, many of them are still experimental and the area as a whole is developing. TSV placement is an area of research which is crucial for the utilization of TSVs. Tradeoffs exist for regular (grid) and non-regular placements in terms of timing and manufacturability. Additionally, TSVs cause coupling, layout, and stress issues which need to be accounted for. Lim's research group has explored various methods of EDA with TSVs, including clocking routing, microfluidics, and more [6]. In terms of clock routing, TSVs can cause clock tree fragments due to the large area used by TSVs.

Conclusion

TSVs can plausibly decrease perceived wire delays and increase the performance of designs with long interconnects. We have seen that TSVs are area expensive. This cost is alleviated if they can be moved out of the way, and further if the optimal interconnects are chosen. A strategy using TSVs is most effective when: the longest metal traces are shortened by moving one of the two connected components to an adjacent level, and this move doesn't result in an excessive increase in the length of another of the moved component's connections. However, in our initial results, we saw little or no improvement in area, delay, and power characteristics of a 32-bit Kogge-Stone adder. This is partially due to how we model the TSVs (as discussed earlier), though it also indicates that TSVs must be used correctly. A adder design, such as the Kogge-Stone, with many long interconnects is a poor fit for 3DIC. Any benefits in interconnect length reduction will be lost in the increase in area or conflicts due to TSVs. Voicu, et al. [1] make a better suggestion, were adders which are easily grouped (such as Carry-Select) are used on multiple tiers, allowing for short interconnects and decreased area without conflicts. Additionally, TSVs could be used for busses which will get higher utilization of the TSV (e.g. memory).

In general, 3DIC will bring large improvements to processor speed and density. However, 3DIC EDA and manufacturing details are a large topic of research. Many of the issues relate to how to properly incorporate TSVs into existing design methodologies. Although, if designs are selected for reduced TSV usage (or efficient usage), many benefits can come from 3DIC.

References

[1] Voicu, G.R.; Lefter, M.; Enachescu, M.; Cotofana, S.D., "3D stacked wide-operand adders: A case study," *Application-Specific Systems, Architectures and Processors (ASAP), 2013 IEEE 24th International Conference on*, vol., no., pp.133,141, 5-7 June 2013.

[2] Dae Hyun Kim; Mukhopadhyay, S.; Sung Kyu Lim, "Fast and Accurate Analytical Modeling of Through-Silicon-Via Capacitive Coupling," *Components, Packaging and Manufacturing Technology, IEEE Transactions on* , vol.1, no.2, pp.168,180, Feb. 2011.

[3] Weerasekera, R.; Grange, M.; Pamunuwa, D.; Tenhunen, H.; Li-Rong Zheng, "Compact modelling of Through-Silicon Vias (TSVs) in three-dimensional (3-D) integrated circuits," *3D System Integration, 2009. 3DIC 2009. IEEE International Conference on* , vol., no., pp.1,8, 28-30 Sept. 2009.

[4] D. H. Kim and S. K. Lim, “Through-silicon-via-aware delay and power prediction model for buffered interconnects in 3D ICs,” in Proc. ACM/IEEE Int. Workshop Syst. Level Interconnect Prediction, Jun. 2010, pp. 25–31.

[5] Ouyang, J.; Sun, G.; Chen, Y.; Duan, L.; Zhang, T.; Xie, Y.; Irwin, M.J., "Arithmetic unit design using 180nm TSV-based 3D stacking technology," *3D System Integration, 2009. 3DIC 2009. IEEE International Conference on* , vol., no., pp.1,4, 28-30 Sept. 2009.

[6] Lim, S.K., "3D Circuit Design with Through-Silicon-Via: Challenges and Opportunities," *Electronic Design Processes (EDP) 2010 Symposium Workshop*, Apr. 2010.