

Exploring Compromises among Timing, Power and Temperature in Three-Dimensional Integrated Circuits

Hao Hua, Chris Mineo, Kory Schoenflies, Ambarish Sule, Samson Melamed,
Ravi Jenkal, and W. Rhett Davis
North Carolina State University, Raleigh, NC
{hhua, rhett_davis}@ncsu.edu

ABSTRACT

Three-dimensional integrated circuits (3DICs) have the potential to reduce interconnect lengths and improve digital system performance. However, heat removal is more difficult in 3DICs, and the higher temperatures increase delay and leakage power, potentially negating the performance improvement. Thermal vias can help to remove heat, but they create routing congestion, which also leads to longer interconnects. It is therefore very difficult to tell whether or not a particular system may benefit from 3D integration. In order to help understand this trade-off, physical design experiments were performed on a low-power and a high-performance design in an existing 3DIC technology. Each design was partitioned and routed with varying numbers of tiers and thermal-via densities. A thermal-analysis methodology is developed to predict the final performance. Results show that the lowest energy per operation and delay are achieved with 4 or 5 tiers. These results show a reduction in energy and delay of up to 27% and 20% compared to a traditional 2DIC approach. In addition, it is shown that thermal-vias offer no performance benefit for the low-power system and only marginal benefit for the high-performance system.

Categories and Subject Descriptors

17. Beyond Die-Integration and Packaging

General Terms

Performance, Design, Experimentation

Keywords

3DIC, temperature dependency, design flow, trade off

1. INTRODUCTION

Aggressive device scaling is imperative to meet the needs of high performance VLSI systems. While this scaling reduces gate delay, wiring parasitics have a much more pronounced effect, and interconnect delay begins to dictate system performance. New design strategies are needed to alleviate the impact of interconnect. Three-dimensional integrated circuits (3DICs) are a possible solution for interconnect-driven design, because stacking silicon layers allows more cells to be placed close to one another,

thereby decreasing the average interconnect length [10]. The term “tier” is used to refer to each active layer and its associated metal and dielectric layers. A recent example of a two-tier system was fabricated and demonstrated 15% reductions in both delay and power over the traditional 2DIC case [17].

But how can we know if a system will benefit from 3D integration? Stochastic estimates based on Rent’s rule have been used to predict that total interconnect capacitance for a system will decrease as the number of tiers increases, up to a maximum of a 40% reduction for 6 tiers [15]. With more than 6 tiers, the vertical interconnects dominate and wire-capacitance increases. Layout methodologies have been proposed for custom designs [7] and standard cell based designs [16] that predict similar improvements. These studies are limited, however, in that they have not included a thermal analysis. An underlying difficulty in 3DIC design is heat removal. The stacking of active layers exacerbates the heat-removal problem, and the higher temperatures lead to increased delay and leakage power [2][3]. This performance degradation could potentially negate the benefit of the shorter wires achieved with 3D integration.

Models have been proposed to enable thermal analysis of 3D chips [4][12][13], which has led to a number of proposed methodologies to thermally optimize the physical design of a system. Cong, *et al* presented a thermally-driven floorplanner to minimize wire-length, inter-tier interconnects and temperature [9]. Goplen, *et al.* [8] have presented a thermally aware approach to placement. These works focus on the placement of the hottest blocks close to the heat-sink. Another approach is to increase the thermal conductivity of the stack by inserting “thermal-vias”. Goplen, *et al.* [5] and Cong, *et al.* [6] have shown that thermal-via insertion can effectively reduce intra-tier and inter-tier temperature variation. Their work has demonstrated that the temperature gradient can be controlled in 3DIC.

However, previous researches have overlooked the fact that thermal vias increase routing congestion, which can lead to longer interconnects and increased dynamic power that in turn leads to higher temperatures and increased leakage power. In this work, we explore the trade-off between dynamic-power and leakage-power on two designs by varying tier count and thermal via density.

The design-flow used for this exploration is shown in Figure 1. This flow is based on a three-tier, three-metal-per-tier, 180nm SOI technology from MIT Lincoln Labs [21] that is currently being used to fabricate a Fast-Fourier Transform (FFT) test-chip [19]. As shown in the figure, the flow begins with the netlist result from standard-cell synthesis and partitions the design for minimum inter-tier cuts using k-METIS [18]. The design is then floorplanned, using the commercial 2DIC tool *Encounter* from

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, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

DAC 2006, July 24–28, 2006, San Francisco, California, USA.
Copyright 2004 ACM 1-59593-381-6/06/0007...\$5.00.

Cadence. This floorplan step optimizes for wirelength, critical path delay and maximum temperature. Next, thermal-design is performed, which involves insertion of a pre-determined number of thermal-vias. These vias can be moved later to eliminate hot spots. The cells are then placed in each tier independently using *Encounter*, followed by an inter-tier via alignment step to ensure consistent via positions on each tier. The clock-tree is then inserted, the design is routed, and RC parasitics are extracted for each tier. These parasitics are then merged into a single SPEF file, inserting a value for each inter-tier via determined from 3D field-solver simulations. The delays and power are then analyzed using the cell-based analysis tools *PrimeTime* and *PowerCompiler* from *Synopsys* and switching-activity annotations from Verilog simulation. This flow is described more completely in [19]. This work builds on the approach from the test-chip to support place-and-route for up to 10 tiers and 5 layers of metal per tier.

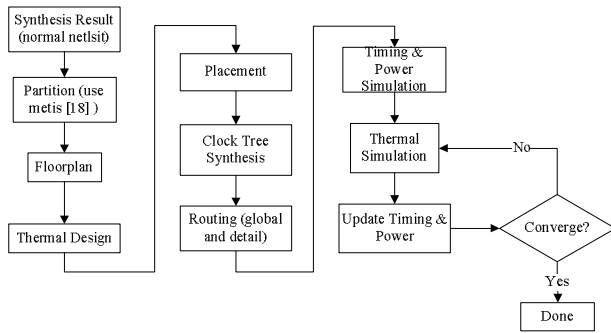


Figure 1: 3DIC design flow (modified from [20])

A severe limitation of our original flow was that the cell-based delay and power estimates assume a single temperature and are therefore inaccurate. Because the power and temperature cannot be determined independently, we extended the cell-based approach as illustrated in Figure 1 to iteratively determine the power and temperature and converge on a final solution. Section 2 of this paper introduces our thermal model, Section 3 discusses how delay and leakage power are dependent on temperature, and Section 4 describes the iterative solver.

Ultimately, we would like to have a simple way to determine if a system will benefit from 3D integration. We begin to explore this complex issue by presenting a design-study of two different systems: a low-power and a high-performance design as shown in Table 1. The low-power design is based on the 8-point FFT presented in [19], but has been pipelined to increase throughput. The high-performance design is based on the OpenRISC Platform System-on-Chip (ORPSOC) [22], which includes a 32-bit OpenRISC micro-processor, memory controllers, and 40KB of embedded memory. We have modified the ORPSOC to include a second OpenRISC core that communicates with the same memories through a bus arbitration unit. Section 5 of this paper presents and explains experimental results, quantifying the design trade-offs in terms of path delay, power consumption and maximum temperature.

Table 1: Design summary

Design	Path delay	Power	# of std cells	Die area
FFT	26.1 ns	0.809 W	158 K	11.6 mm ²
ORPSOC	17.8 ns	3.298 W	120 K	18.8 mm ²

2. Flow Implementation and Thermal Model

We vary the number of tiers and density of thermal-vias to explore the trade-off between dynamic and leakage power in 3DIC designs. Thermal-vias are implemented with a standard-cell that contains two inter-tier vias. The “density” in our study refers to the ratio of via-cell area to total area, rather than actual via-area, which is only 26% of the via-cell area. Thermal via pattern generation consists of two phases: global and detail. In global phase, we uniformly distribute thermal-via cells within one silicon tier (Figure 2a) and assume all tiers to have the same die area and thermal-via density. A prototype was created to estimate performance and average temperature, which were set as targets for the detail phase. In the detail phase, these via cells are re-located after clock tree synthesis (Figure 2b) to achieve a more uniform temperature profile, which aims to avoid timing and/or leakage violation caused by high temperature. In a case, the detail phase only need to be executed once when the optimal combination of tier count and thermal via density are obtained. For simplicity, we use the average temperature over one tier to represent the uniform temperature matrix obtained in detail phase. These thermal-vias are considered placement and routing blockages. We restrict our study to thermal-via cell densities below 20% of the total chip area, because trial-routes have shown that congestion makes our designs unroutable at higher densities.

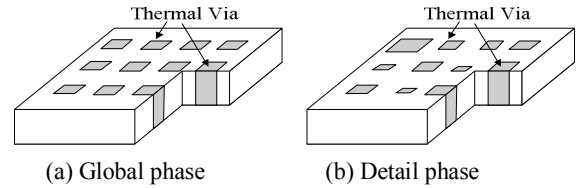


Figure 2: Thermal via formation.

Thermal via insertion in the FFT and ORPSOC is demonstrated in Figure 3. They differ because the ORPSOC contains blocks of memory (IP) into which we cannot insert thermal vias. Because of this, memories are in approximately the same locations on each tier because the area above and below memories has no room for thermal via too. As described in [17], this practice of stacking memories rather than logic and memory between tiers helps to maximize performance.

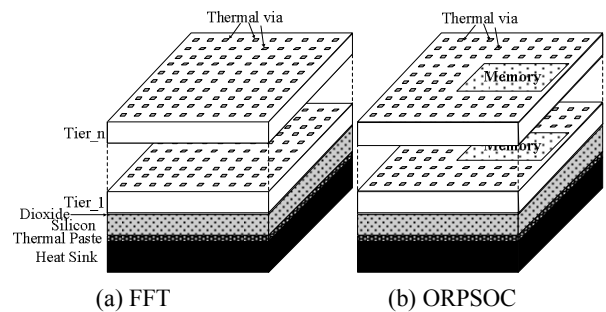


Figure 3: Thermal via insertion in (a) FFT, (b) ORPSOC

Assuming adiabatic boundary conditions on the sides of the die not connected to the heat-sink, the 3D system fits the thermal model shown in Figure 4, where RI_{bottom} is the equivalent thermal resistance between tier 1 active layer and the heat sink, RI_2 is the equivalent thermal resistance between tier 1 and tier 2 active layers, etc. The steady-state power consumption on each

tier ($P_{tier1,2,3}$ etc) is modeled as a constant current source in this model. The ambient temperature (heat sink temperature) is represented as a constant voltage source. The node voltages represent the average temperature on each tier.

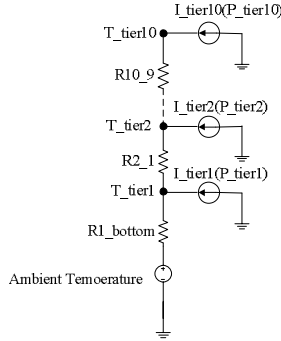


Figure 4: Thermal model

The calculation of equivalent thermal resistance is based on the equivalent thermal conductivity. In this work, we use the worst case (smallest) equivalent thermal conductivity by assuming that only the thermal vias, inter-tier signal vias and dielectric are conducting heat (we ignore lateral metallization, metal to metal via, etc). The equivalent thermal conductivity for above materials conducting heat in parallel can be expressed as (1), where λ is the thermal conductivity and A is the area. The equivalent thermal resistance is calculated using (3), where t is the thickness between two adjacent active layers. This model tends to under-estimate the temperature at any one point because it assumes perfect heat spreading on each tier, but it provides a good first-order estimate of the average temperature.

$$\lambda_{equ} = \frac{\lambda_{SiO_2} A_{SiO_2} + \lambda_{thermal-via} A_{thermal-via} + \lambda_{signal-via} A_{signal-via}}{A_{die}} \quad (1)$$

$$A_{die} = A_{SiO_2} + A_{thermal-via} + A_{signal-via} \quad (2)$$

$$R_{equ} = \frac{\lambda_{equ} t}{A_{die}} \quad (3)$$

3. Performance/Temperature Dependence

3DIC performance degrades, and reliability is compromised when operating at high temperatures. Though dynamic power is independent of temperature [3], logic gate delay and leakage power will increase with temperature. Longer delay has negative effect on temperature due to lower frequency and smaller dynamic power, but leakage power has positive effect on temperature because higher temperature increases leakage power considerably. In evaluating the performance of 3DIC, we must consider this relationship until the point at which temperature and performance converge. In practical timing or power analyses, it is impossible to include tens of libraries at different operating temperatures. Power and timing calculations become very difficult and computationally intensive when accounting for temperature variation. Many timing arcs span multiple silicon tiers, therefore single timing paths operate with their transistors at significantly different temperatures. Short of the feasible, naïve brute-force approach making use of tens of separate libraries, it is very difficult for current tools (such as *PrimeTime* and *PrimePower*) to perform a fast yet accurate analysis of system performance. In this section, we develop two models to address the delay-temperature and the leakage-temperature dependencies. The resistance-temperature

dependency is ignored in this work because it is relatively small compared to transistor output (holding) resistance.

3.1 Transistor-Delay/Temperature Dependence

The transistor delay-temperature dependence can be expressed as (4) [11][14]. In this case I_D is the drain current, T is the absolute temperature in kelvin, α is the velocity saturation index, and μ is the mobility. The typical value of α is 1.5, but it is actually smaller for a short channel MOSFET [14]. The typical β value is 1.5 too. The temperature dependence of threshold voltage V_{TH} and mobility μ can be expressed as (5) and (6) [11] where T_0 is the reference temperature. As shown in [1], the threshold voltage temperature coefficient, k , is also a weak function of temperature. For fully depleted SOI, the variation in k due to temperature is relatively small [1]; for simplicity, we may treat it as a linear function of temperature (7).

$$Delay \propto \frac{CV_{DD}}{I_D} \propto \frac{CV_{DD}}{\mu(T)(V_{DD} - V_{TH}(T))^\alpha} \quad (4)$$

$$V_{TH}(T) = V_{TH}(T_0) - k(T - T_0) \quad (5)$$

$$\mu(T) = \mu(T_0) \left(\frac{T}{T_0}\right)^{-\beta} \quad (6)$$

$$k = k_0 + \gamma(T - T_0) \quad (7)$$

Substitute (5) and (6) into (4), we derive

$$delay(T) \propto \frac{CV_{DD}}{\mu(T_0)(T/T_0)^{-\beta}(V_{DD} - V_{TH}(T_0) + k(T - T_0))^\alpha} \quad (8)$$

$$delay(T) = \frac{delay(T_0)(V_{DD} - V_{TH}(T_0))^\alpha T^\beta}{T_0^\beta (V_{DD} - V_{TH}(T_0) + k(T - T_0))^\alpha} \quad (9)$$

We have varied α , γ , and k_0 values for a 1X buffer driving 50fF and 150fF loads, between 0°C and 250°C, to obtain results in line with SPICE simulations. It is necessary to use slightly different k_0 values for falling and rising delays. Table 2 shows simulated delay values from SPICE, versus those obtained using equation (9). The table gives the maximum error and the corresponding temperature when the maximum error occurs (column 6 and 7). Figure 5 gives a clear picture of how the rise delay of a 1X buffer depends on temperature. As can be seen in this figure, our model is fairly close to SPICE simulation for this simple case.

Table 2: Comparison between SPICE simulation and our model for 1x buffer delay with temperature as variable

	k_0 (mV/K)	γ (mV/K)	β	α	Max Err	Max Err T
Rise	1.2	0.003	1.5	1.3	6.1%	0 °C
Fall	1	0.003	1.5	1.4	5.2%	170 °C

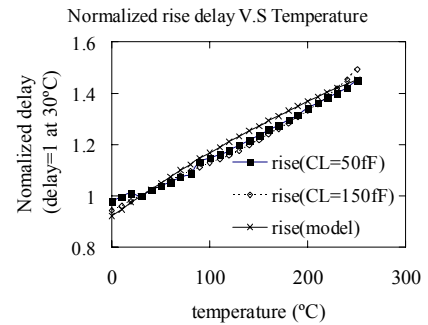


Figure 5: SPICE simulation vs model on 1x buffer delay

The delay model in equation (9) is the basis for our timing analysis. For our standard cell library, across the range temperatures shown and the range of loads common in our designs, the largest disparity between SPICE and our model is an XOR gate with a delay error of 11.6%.

3.2 Leakage-Power/Temperature Dependence

Leakage power drastically varies with operating temperature. Though leakage power in SOI devices is considerably smaller than that of bulk devices, it can still comprise a large portion of the total power when the temperature is very high. To create an accurate leakage model with respect to temperature, we have run extensive SPICE simulations using the BSIMSOI model to determine the temperature-leakage relationship for the standard cells in our library. The leakage temperature dependency is super-linear, and can be modeled as a polynomial function [2]. For the temperature range between 0°C and 250°C, we found that a third-order polynomial can describe the dependencies very well, with a maximum error of 5%. The model is of the form:

$$\frac{I_{leakage}(T)}{I_{leakage}(T_0)} = 1 + a_1 \cdot (T - T_0) + a_2 \cdot (T - T_0)^2 + a_3 \cdot (T - T_0)^3 \quad (10)$$

Using a curve fitting technique, we found values for the coefficients a_1 , a_2 and a_3 in (10). Because the coefficients are slightly different for each standard cell, we used the average values given in Table 3.

Table 3: Coefficients of leakage model (10)

coefficient	a_1	a_2	a_3
value	0.0226	0.00033	1.77E-6

We verified this model by re-characterizing the standard-cell library for a range of temperatures and using *PowerCompiler* to predict the power on two benchmark circuits. The benchmark circuits are a 4-bit adder with 19 gates and a 32-bit adder with 1,381 gates. Figure 6 shows how our model compares to the simulation. We note that the prediction in equation (10) becomes more accurate for large circuits because the total leakage approaches the average among all cells.

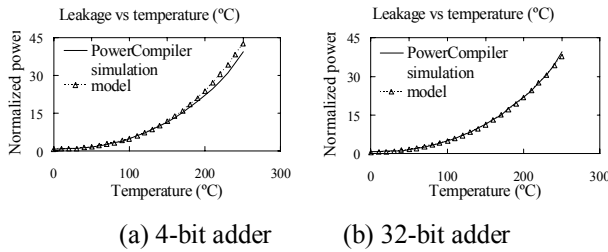


Figure 6: Leakage model compared to simulation

4. Iterative Calculation of Timing, Power and Temperature

Power density in 3D integration increases drastically as the number of silicon tiers increases. Circuits are often designed for the worst case temperature of 125°C, but studies have shown that SOI devices are still theoretically useful up to temperatures of 250°C [1]. Because the temperature, delay, and power are co-dependent variables we determine these values for the final circuit using the iterative approach illustrated in Figure 7. For simplicity, we assume dynamic power is simply dependent on the

clock frequency f , as shown in (11). With our assumption of boundary conditions, we can write the temperature-power dependency in (12) [12], where ΔT_i is the temperature difference between tier i and $i-1$, R_i is the equivalent thermal resistance between tier i and $i-1$, and P_j is the power consumption of tier j . When R_i is determined, the average temperature is proportional to power consumption. This iteration is easily coded into most scripting environments, and our experiments show that the temperature, delay, and power tend to converge with less than 0.1% error in 4 iterations or less.

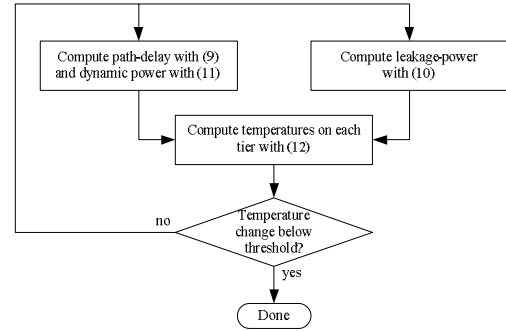


Figure 7: Iterative solution of timing, power and temperature

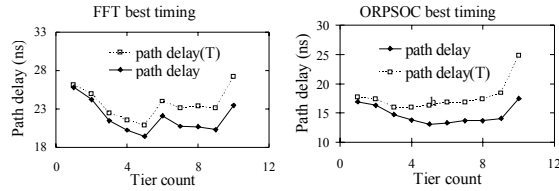
$$P_{dynamic}(f) = P_{dynamic}(f_0) \cdot \frac{f}{f_0} \quad (11)$$

$$\Delta T_i = R_i \sum_{j=i}^{10} P_j \quad (12)$$

5. Experimental Results

In this section we present our experiments with the FFT and ORPSOC circuits. All results are based on the flow illustrated in Figure 1, which was applied on designs with 1 to 10 tiers and thermal-via densities ranging from 0% to 20%, using an 83-cell standard-cell library that was characterized using *Cadence SignalStorm* and *Synopsys HSPICE* with nominal device parameters. The memories used in the ORPSOC system were not completed in time for this study, however, and so an estimate based on an 8KB SRAM block is used. A total of five of these SRAM blocks are used, each having an estimated read/write delay of 4.85ns, dynamic power of 41.6 pJ/cycle, and leakage power of 4.9 mW. Since the memory delay is relatively small compared to the path delay, we place these memories in the upper tiers so that the most power-hungry and timing-critical blocks can be placed close to the heat-sink.

Our first comparison shows the impact of our iterative timing/power/temperature calculation methodology. Figure 8 shows the delay values for the FFT and ORPSOC both considering and neglecting the temperature effect. This figure shows that delay-temperature dependence is less critical in the low-power application, because the temperature-corrected delay curve more closely matches the original estimate for the FFT than for the ORPSOC. Note also the increase in delay for the FFT with 6 tiers. This is due to a difficulty in partitioning the design into 6 tiers. As Figure 8 (b) shows, the delays of the ORPSOC increase rapidly when the tier count exceeds eight, because the excessively high temperature begins to dominate.



(a) FFT timing (b) ORPSOC timing

Figure 8: Best timing value with/without delay temperature dependence

Our next comparison was to see how much improvement in energy-per-operation and path delays could be achieved with 3D integration compared to a traditional 2D approach. Table 4 shows the results with the corresponding number of tiers and thermal via density. The FFT design achieved the most improvement in both energy and delay (27% and 20% respectively) when using 5 tiers and no thermal-vias. The ORPSOC design showed the most improvement with 4 tiers, and different thermal-via densities were shown to minimize energy and delay.

Table 4: Best energy/cycle and timing of 2D/3D integration

Design	F=FFT O=ORPSOC	2D	3D	tiers	via den	improve
F	E/cyc (nJ)	21.2	15.5	5	0%	26.9%
	Delay (ns)	26.1	20.9	5	0%	19.9%
O	E/cyc (nJ)	58.7	47.9	4	2%	18.4%
	Delay (ns)	17.8	14.8	4	5%	16.9%

Figure 9 gives a more detailed picture of the design-space with plots showing how maximum-temperature, total wire-length, path delays, and power vary with the two variables (tier number and thermal via density). The darker areas represent smaller values, and the brighter areas represent larger values. The energy-per-cycle numbers in Table 4 are determined by the product of power and delay. Figure 9 (a) and (b) give the maximum temperature trend of the FFT and ORPSOC. As can be seen from the figures, the maximum temperature is monotonic as tier number increases and thermal via density decreases. This is because temperature is directly proportional to power density, which increases with the number of tiers, and directly proportional to thermal-resistance, which increases as thermal-via density decreases. However, as thermal via area increases, the temperature drop in the ORPSOC is more pronounced than in the FFT. This can be explained by the polynomial leakage power-temperature dependency. Because the ORPSOC design is hotter, the leakage power is more sensitive to an increase in thermal-via density. Though higher temperature usually causes longer path delays and in turn reduces dynamic power, it had a different effect in low power versus high performance applications. Figure 9 (c) and (d) gives the total wire-length of each combination of tier count and thermal via density. With same thermal via density, wire-length decreases as tier count increases due to the proximity of blocks. Wire-length increases with thermal-via density, however, due to the increased routing congestion. Note that the figure is somewhat misleading, because it does not include vertical length, which is difficult to quantify in a meaningful way. Figure 9 (e) and (f) show the power trend for the FFT and ORPSOC. In the FFT design, the upper left portion of the graph is where the most power is consumed, where tier count is low and thermal via density is high. This is because the

design consumes relatively little power, and the total temperature rise above ambient is small. Therefore, the power is mainly determined by interconnect wire-length (dynamic-power), rather than leakage. Therefore, the total power consumption of the FFT

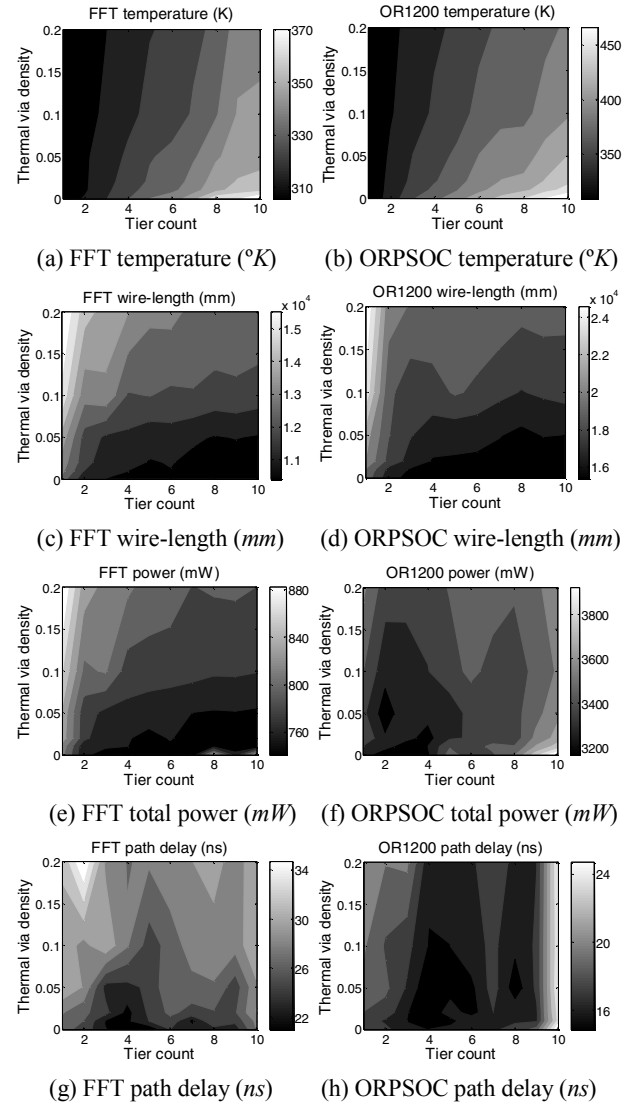


Figure 9: Max temperature, power and timing with different silicon tier number and thermal via area

varies inversely with the number of tiers. However, when tier count is large, the total wire-length no longer decreases, and so the power reduction is less pronounced. The case with the ORPSOC is different. Due to the relatively high temperature, the leakage power tends to dominate, and the highest-power is in the lower-right of the graph. Therefore, the optimum energy-per-operation occurs with fewer tiers. For each tier count, as thermal via density increases, the total power consumption first decreases and then increases. This is because the leakage-power dominates for low via-densities, but dynamic-power dominates for higher densities. In our experiments, the leakage power of ORPSOC varies from 2% to 30% of the total power. We are faced with the trade-off between dynamic-power and leakage-power to achieve optimal system performance. One reason for the severe leakage-power effect in the ORPSOC design is that the memories

were placed on the top-tier, which has the highest temperature. Without using any thermal-vias, these memories would consume approximately 25X more leakage power and run 60% slower with 10 tiers. If our critical path delay is close to the memory latency, we must be very careful to find the ideal location for the memories. In such a case, we would not opt for a design with a large number of tiers, because of the additional difficulty associated with the heat removal. Figure 9 (g) and (h) shows the timing trend of the FFT and ORPSOC. The optimal timing region of FFT is when the tier number is equal to 5 and the thermal via density is equal to 0, while the optimal timing region of the ORPSOC is when the number of tiers is approximately equal to 4, and the thermal via density equal to 5%. Timing is mainly restricted by interconnect wire-length in a low power design. In a high performance design, both wire-length and temperature have a large impact on timing.

6. Conclusion

Thermal-vias in 3DICs can be used to remove heat, but they create routing congestion, which leads to a trade-off between leakage-power and dynamic-power. We have proposed a methodology to explore this trade-off and have applied it to two case-studies, representing both low-power and high-performance applications. Contrary to some researches that emphasis on temperature control in 3DIC, we have pointed out that the overuse of thermal via cannot benefit 3DIC system performance due to the increase in wire-length. In low-power designs, the temperature gradient across tiers is not significant. For these designs, thermal-vias do not help, and increasing the number of tiers beyond 5 does not bring any improvement in energy-per-operation or delay. The interaction among timing, power, and temperature is more pronounced in high-performance applications. For these systems, careful thermal via placement can effectively reduce the temperature gradient within the 3D chip, and its effect on leakage power reduction helps to reduce the maximum temperature.

This work has shown that 3DICs are an attractive way to improve system performance. For the low-power system, a reduction in energy-per-operation of 27% was achieved along with a 20% improvement in speed. However, if no revolutionary heat removal or packaging methods become available, this work suggests that there may be no benefit in fabricating more than 5 tiers.

7. ACKNOWLEDGMENTS

The authors would like to thank DARPA for supporting this work. We would also like to thank Cadence, Synopsys, PTC, and Ansoft for generously providing the CAD tools. Thanks to MIT Lincoln Labs for providing access to their FD-SOI library and for their aid in developing our design kit. Lastly, thanks to James Stine at the Illinois Institute of Technology for generously providing access to the IIT-SoC standard-cell characterization scripts.

8. REFERENCES

- [1] G. Groeseneken, *et al* "Temperature Dependence of Threshold Voltage in Thin-Film SOI MOSFET's", *IEDL, Vol. 11, No. 8*, Aug. 1990.
- [2] H. Su, F. Liu, A. Devgan, E. Acar and S. Nassif, "Full Chip Leakage Estimation Considering Power Supply and Temperature Variations", *ISLPED*, Aug. 2003.
- [3] W. Liao, L. He and K. M. Lepak, "Temperature and Supply Voltage Aware Performance and Power Modeling at Microarchitecture Level", *TCAD*, Vol. 24, No. 7, Jul. 2005.
- [4] A. Rahman, A. Fan and R. Reif, "Thermal analysis of three-dimensional (3-D) integrated circuits (ICs)", *IITC*, June 2001.
- [5] B. Goplen and S. Sapatnekar, "Thermal Via Placement in 3D ICs", *ISPD*, 2005.
- [6] J. Cong and Y. Zhang, "Thermal via Planning for 3-D ICs", *ICCAD*, Nov. 2005.
- [7] S. M. Alam, D.E. Troxel and C.V. Thompson, "A comprehensive layout methodology and layout-specific circuit analyses for three-dimensional integrated circuits" *ISQED*, Mar. 2002.
- [8] B. Goplen and S. Sapatnekar, "Efficient Thermal Placement of Standard Cells in 3D ICs using a Force Directed Approach," *ICCAD*, Nov. 2003.
- [9] J. Cong, W. Jie and Z. Yan, "A Thermal-Driven Floorplanning Algorithm for 3D ICs", *ICCAD*, Nov. 2004.
- [10] K. Banerjee, S. J. Souri, P. Kapur and K. C. Saraswat, "3-D ICs: a novel chip design for improving deep-submicrometer interconnect performance and systems-on-chip integration", *Proceedings of the IEEE*, May 2001.
- [11] A. Bellaouar A., A. Fridi, M. J. Elmasry and K. Itoh, "Supply voltage scaling for temperature insensitive CMOS circuit operation". *TCAS II*, Mar 1998.
- [12] S. Im and Banerjee, "Full chip thermal analysis of planar (2-D) and vertically integrated (3-D) high performance ICs", *IEDM* 2000.
- [13] T. Y. Chiang, K. Banerjee and K. C. Saraswat, "Compact Modeling and SPICE-Based Simulation for Electrothermal Analysis of Multilevel ULSI Interconnects", *ICCAD*, Nov. 2001.
- [14] K. Kanda, K. Nose, H. Kawaguchi and T. Sakurai "Design impact of positive temperature dependence on drain current in sub-1-V CMOS VLSIs". *JSSCC*, Oct. 2001.
- [15] R. Zhang, K. Roy, Cheng-Kok Koh, D.B. Janes, "Power trends and performance characterization of 3-dimensional integration for future technology generations", *ISQED*, Mar. 2001.
- [16] S. Das, A. Chandrakasan and R. Reif, "Design tools for 3-D integrated circuits", *ASP-DAC*, Jan. 2003.
- [17] B. Black, D.W. Nelson, C. Webb and N. Samra, "3D processing technology and its impact on iA32 microprocessors", *ICCD*, Oct. 2004.
- [18] G. Karypis and V. Kumar, *The METIS Serial Graph Partitioning Tool*, available online at <http://www-users.cs.umn.edu/~karypis/metis>
- [19] W. R. Davis *et al*, "Demystifying 3D ICs: the pros and cons of going vertical," *IEEE Design & Test of Computers*, vol 22, no 6, Nov.-Dec. 2005.
- [20] H. Hua, C. Mineo, S. Melamed and W. R. Davis, "The 3DIC Phase 1 Place and Route Flow". *NCSU Design-Flow Database*, available online at <http://www.ece.ncsu.edu/muse/flowdb>
- [21] V. Suntharalingam *et al*, "Megapixel CMOS Image Sensor Fabrication in Three-Dimensional Integrated Circuit Technology", *ISSCC*, Feb. 2005.
- [22] *OpenRISC Reference Platform System-on-a-Chip and OpenRISC 1200 IP Core Specification*, available online at <http://www.opencores.org/projects.cgi/web/or1k/orpsoc>