AN ANALYTICAL AND NUMERICAL MODEL FOR DESIGN OF NON-UNIFORMLY POWERED SILICON CHIPS BASED ON BOTH THERMAL AND DEVICE CLOCK PERFORMANCE

by

Abhijit D. Kaisare

Presented to the Faculty of the Graduate School of

The University of Texas at Arlington in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT ARLINGTON

DECEMBER 2007

This dissertation is dedicated to my beloved

Uncle late Mr. Raghunath G. Kaisare (Aba Kaka)

ACKNOWLEDGEMENTS

I would like to express my appreciation and gratitude to my advisor, Prof. Dereje Agonafer for his constant guidance, patience and motivation. I really thank him for supporting me during my entire course of PhD program. His constant support, expert advice and belief in me made this journey an unforgettable one.

I would also like to appreciate my Co-advisor Prof. A. Haji-Sheikh for his thoughtful insight, and valuable input during my program of work. It would have been impossible without his expert advice and sound knowledge to develop an analytical model which is the integral part of this dissertation.

I sincerely thank committee members Prof. Kent Lawrence, Prof. Seiichi Nomura and Prof. Zeynep Celik-Butler for being on my committee. I thank them for their thoughtful suggestions and comments during various stages of this dissertation.

This research is being done in conjunction with support of Intel Corp. Special thanks to Dr. Greg Chrysler and Dr. Ravi Mahajan, from Intel Corp, Chandler, AZ for their support and expert advice during this research. I would like to thank Dr. Ravi Mahajan for being on my committee, giving me valuable inputs and insight during this research project.

My graduate studies would not have been the same without the social and academic challenges and diversions provided by all my student-colleagues, faculty and staff in the Department of Mechanical Engineering, particular to Electronics, MEMS

and Nanosystems Packaging Center at University of Texas at Arlington. I am particularly thankful to my colleagues Veerendra Mulay, Saket Karajgikar and Nikhil Lakhkar for their encouragement and support.

I would also like to thank Ms. Sally Thompson of the Electronics, MEMS and Nanosystems Packaging Center for her constant support and encouragement throughout the course of this program.

Finally, I thank my parents Mr. Dattaram G. Kaisare and Mrs. Anuradha D. Kaisare and my wife Janhavi for their love, patience and support during my PhD Program.

November 26, 2007

ABSTRACT

AN ANALYTICAL AND NUMERICAL MODEL FOR DESIGN OF NON-UNIFORMLY POWERED SILICON CHIPS BASED ON BOTH THERMAL AND DEVICE CLOCK PERFORMANCE

Publication No. _____

Abhijit D. Kaisare, PhD.

The University of Texas at Arlington, 2007

Supervising Professors: Dereje Agonafer and A. Haji-Sheikh

Silicon chips continue to grow in capabilities, complexity and performance. Silicon chips typically integrate functional components such as logic and level two (L2) cache memory in their architecture. This functional integration of logic and memory results in improved performance of the chip. However, the integration also introduces a layer of complexity in the thermal design and management of silicon chips. As a direct result of functional integration, the power map on a silicon chip is typically highly non-uniform and the assumption of a uniform heat flux across the chip surface has been shown to be invalid post Pentium II architecture.

The active side of the silicon chip is divided into several functional blocks with distinct power assigned to each functional block. Current research is focusing on microprocessor as a leading example as this work applies equally to other applications where silicon chips demonstrate non-uniformity. For instance graphic chips can exhibit non-uniformity so can ASICs.

Initially, a numerical model is developed to minimize the on-die temperature of the package by optimizing the distribution of the non-uniformly powered functional blocks with different power matrices. In order to model the non-uniformly power dissipation on the silicon chip, the chip surface area is divided into different cases such as 3 x 3, 4 x 4, 8 x 8, 10 x 10 etc. of power matrix with a matrix space representing a distinct functional block with a constant heat flux. Using a FEM code, an optimization of the positioning of the functional blocks relative to each other is carried out in order to minimize the junction temperature Tj (temperature on die). The best possible Tjmax reduction could thus be found. Based on the data derived from numerical model, for different power matrices, trend for the data is evaluated which helps to draw design guideline for a typical die configuration (i.e. 30 x 30).

This is followed with a development of an analytical approach to temperature distribution of a first level package with a non-uniformly powered die which is carried for the first time. Previously developed analytical model [18] for two layer bodies is extended to this typical package which is a multilayer body. The solution is to begin by designating each surface heat flux as a volumetric heat source. An inverse methodology is applied to solve the equations for various surfaces to calculate maximum junction

temperature (Tjmax) for a given multilayer body. Finally validation of the analytical solution is carried out using developed numerical model.

In the end, a multi-objective optimization is carried out embedding developed numerical model into silicon floor plan to developed compact model satisfying both electrical and thermal performance. Recommendations are provided for an architecture design regarding maximum separation of functional block with minimum on-die temperature. The commercial finite element code ANSYS® is used in this study.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

CHAPTER 1

BACKGROUND & MOTIVATION

1.1 What is Non-uniform Power?

As the performance of microprocessor increases, the need to supply data and instructions to the core processor increases accordingly. In the past, these information, which normally reside in main memories are channeled to microprocessor via the bus, a parallel set of interconnects running between the microprocessor and the memory. In the recent Intel product generations, improved integrated circuit (IC) density makes it practical to integrate the Level Two (L2) cache into the microprocessor architecture. This integrated L2 cache stores the frequently used data thereby reducing the frequency of accessing information from the external main memories.

Consequently, this speeds up data execution and enhances microprocessor performance. Although the integration of L2 cache into microprocessor architecture renders definite design advantages, it also poses other challenges in the thermal design of next generation processors, especially for those with significantly large integrated L2 cache. One of the challenges is to design a thermal solution that meets the thermal performance of microprocessor with non-uniform power density within the silicon die. The non-uniform power distribution occurs when the central processing unit (CPU) or core processor region of the die dissipates a significant fraction of the total power while

the other cache regions of the die dissipate little or no power. This non-uniformity of power distribution results in a large die temperature gradient, with localized hot spots that are expected to affect the processor performance, product reliability as well as yield. This phenomenon [1] was reported in the design of Pentium® III XeonTM processor for 4-way and 8-way server systems where significantly large L2 cache was integrated to the silicon design. Figure 1.1 shows the 90 nm Intel Pentium Processor [2]. This is one of the example of Intel processor showing different functional blocks with non-uniform power. Figure 1.2 shows a plot of the overall cooling capability of different thermal solutions as a function of the density factor (DF), as a measure of the impact of power non-uniformity [2]. Figure 1.3 (a) shows the traditional assumption of uniform power while (b) shows a reality which is a highly non uniform power over a surface of die.

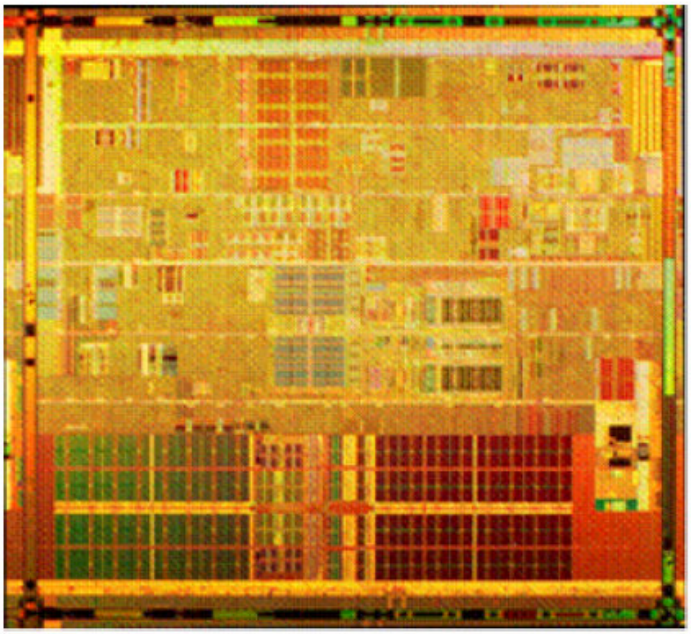


Figure 1.1: 90 nm Intel Pentium Processor [2]

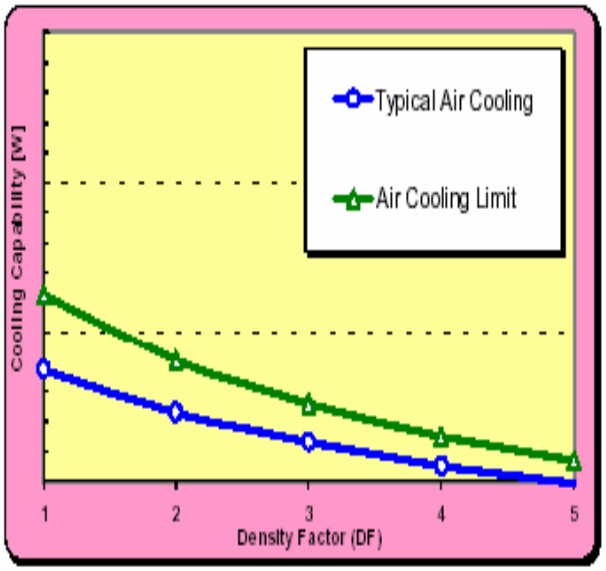
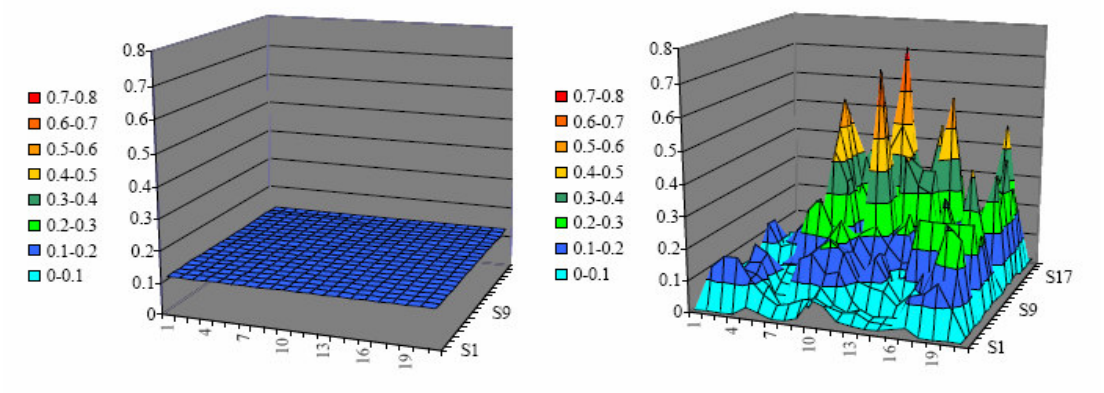


Figure 1.2: Impact of die power non-uniformity on cooling capability [2]



(a) (b)

Figure 1.3: Type of power (a) Uniform power which is a traditional assumption(b) Non uniform power [2]

1.2 Need for architecture level thermal management and thermal modeling

In recent years, power density in microprocessors has doubled every three years [3, 4], and this rate is expected to increase within one to two generations as feature sizes and frequencies scale faster than operating voltages [5]. As the total power increases,

the power density of the silicon chip is increasing as well. In order to keep transistors within their allowed operating temperature range, the generated heat has to be dissipated from its source in a cost-effective manner. These constraints limit the processor's peak power consumption. Peak power consumption limits apply both to desktops and mobile computers. Increased power dissipation results in higher operating temperature, more expensive cooling mechanisms and reduced reliability.

The integration level and the clock speed of high-performance microprocessors have increased so rapidly that the average power density of the advanced microprocessor chips is now more than 10 W/cm^2. Many chips are now running at peak power of >30 W even at V_{dd} =3.3V. Furthermore, the average power density is expected to increase in scaling. [40]

Power-aware design alone has failed to solve thermal management problem, requiring temperature-aware design at all system levels, including the processor architecture. Many low power techniques do not reduce power density in hot spots with having little impact on operating temperature. Temperature-aware design will make use of power-management techniques. In sub-100nm technologies, temperature plays an important role in performance and reliability. Localized heating occurs much faster than chip-wide heating; since power dissipation is spatially non-uniform across the chip as shown in figure 1.4, this leads to "hot spots" and spatial gradients that can cause timing errors or even physical damage. This means that power-management techniques, in order to be used for thermal management, must directly target the spatial and temporal behavior of operating temperature. Temperature-specific design techniques to date have

mostly focused on the thermal package (heat sink, fan, etc.). If the package is designed for worst-case power dissipation, they must be designed for the most severe hot spot that could arise, which is prohibitively expensive. To reduce packaging cost without unnecessarily limiting performance, it has been suggested that the package should be designed for the worst typical application [7,8,9]. It is therefore very important to create thermal compact model depends on architecture, power maps and floor plans [4] to solve thermal management problem.

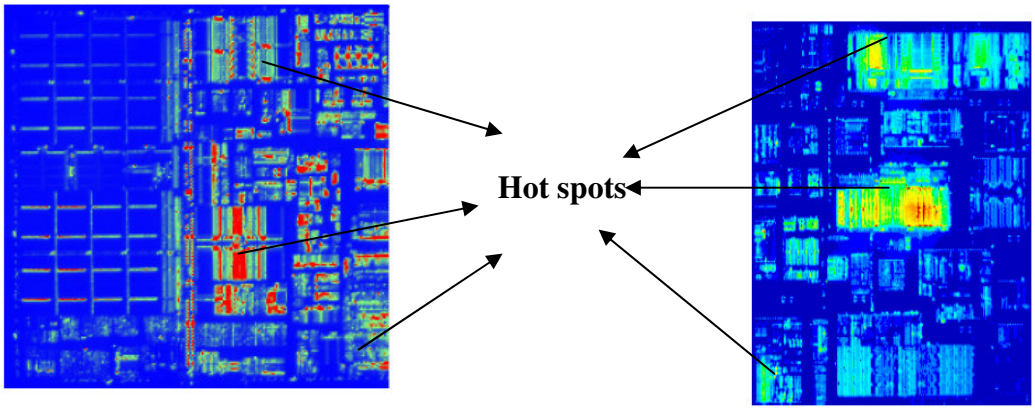


Figure 1.4: Hot spots as a effect of Non uniform power (a) Intel Pentium ® III Processor and (b) Intel Itanium ® Processor [9]

CHAPTER 2

LITERATURE STUDY

Non uniform power phenomenon [1] was reported in the design of Pentium® III XeonTM processor for 4-way and 8-way server systems where significantly large L2 cache was integrated to the silicon design. Previous work has been done in understanding the thermal solutions for the non-uniform power distribution within the silicon die. Some of the previous work includes Goh et al. [11] using genetic algorithm approach, Getkin et al. [12] using simplified fin modeling approach, Sikka et al. [13] using analytical temperature distribution method, June et al. [14] using cap-integral standoffs, Yu et al. [15] using fast placement dependent full chip thermal simulations .

Skardon et al. [7,8] developed a compact model as shown in figure 2.1 also known as “hot spots” which deals with temperature aware micro-architecture. This work describes a thermal-modeling approach that is easy to use and computationally efficient for modeling thermal effects and thermal-management techniques at the processor architecture level. This approach is based on modeling thermal behavior of the microprocessor die and its package as a circuit of thermal resistances and capacitances that correspond to functional blocks at the architecture level. This yields a simple compact model, yet heat dissipation within all major functional blocks and the heat flow among blocks and through the package are accounted for. The model is parameterized, boundary and initial conditions independent, and is derived by a

structure assembly approach. It is based on a simple network of thermal resistances and capacitances, which are derived using area and floorplan information for the major functional blocks at the architecture level, and estimates of thermal resistance and capacitance for the microprocessor's thermal package.

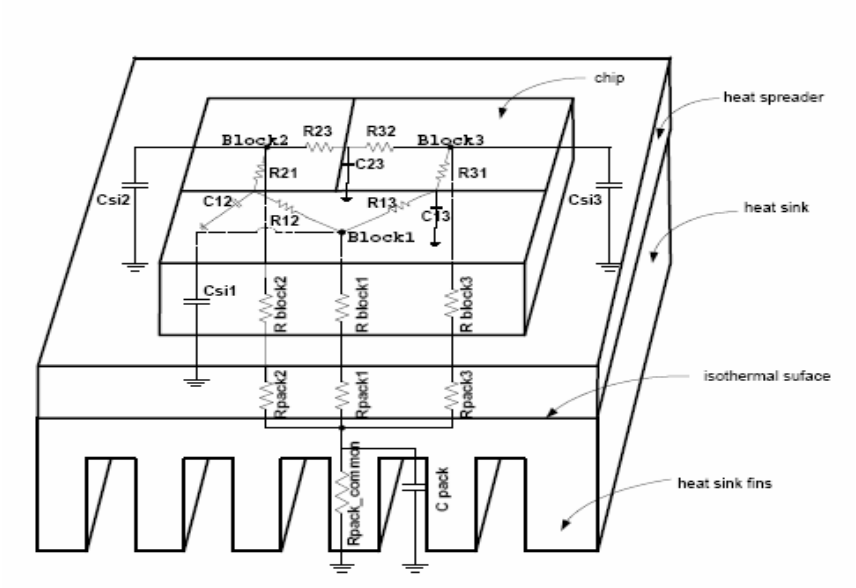


Figure 2.1: Skardon's Model (Overall RC model, including both vertical and lateral components) [7]

Hot spot uses the duality between electrical and thermal resistances and capacitances to derive a dynamic compact model of the heat flow among the different architecture-level blocks within a microprocessor. [7]

Wu et al. [21] mathematically estimated the active power of functional units in modern microprocessors. The method leverages the phasic behavior in power consumption of programs, and capture as many power phases as possible to form a linear system equations such that functional unit power can be solved. Suwa and Hamid [16] presented a multidisciplinary design and optimization methodologies in electronic

packaging which includes integrated multidisciplinary CAD environment and automated design and optimization techniques. Skardon et al. [7,8] shows that there is significant peak temperature reduction potential in managing lateral heat spreading through floorplanning. It was also suggested that potential warrants considerations of the temperature performance trade off early in the design stage at the microarchitectural level using floorplanning.

There has not been a design guideline or compact model developed for thermal based optimization of functional block distribution in non-uniformly powered microprocessor. In addition there is no close form solution (Analytical model) available for temperature distribution for a first level package with non-uniformly powered die. It is also very important to develop a best known method to develop a compact model based on both thermal and device clock performance for a non-uniformly powered microprocessor.

CHAPTER 3

OBJECTIVE

(I) Development of numerical model for thermal optimization of non-uniformly powered microprocessors.

(i) Thermal based optimization of functional block distributions in a non-uniformly powered microprocessors.

(ii) Design rule for minimizing on-die temperature in a non–uniformly powered microprocessors.

(II) Development of an analytical model for non-uniformly powered microprocessors.

(i) Analytical approach to temperature distribution of a first level package with uniformly powered microprocessors.

(ii) Analytical approach to temperature distribution of a first level package with non-uniformly powered microprocessors.

(III) Development of numerical model using multi-objective optimization based on both thermal and device clock performance for non-uniformly powered microprocessors.

3.1 Description Of A Package

A schematic diagram of a typical package shown in figure 3.1 is used in this research. It has a heat sink of 64 x 64 x 6.35 mm attached to the spreader of 31 x 31 x 1.8 mm to the silicon die with the help of TIM I and TIM II with a thickness of 0.025

mm. as shown in table 3.1. The various power maps used in this research are provided by Intel corp.

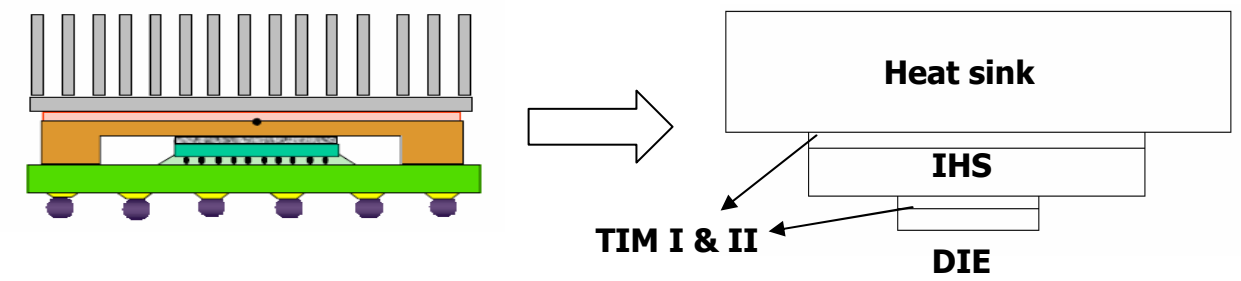


Figure 3.1: Model of a Flip chip package used in this study

Table 3.1: Dimensions and material properties data

Type	Dimensions (mm)	Thickness (mm)	Conductivity (W/m°K)
Die	10.17 x 10.62	0.75	120
TIM I	10.17 x 10.62	0.025	1
IHS	31 x 31	1.8	390
TIM II	31 x 31	0.025	1
Heat sink	64 x 64	6.35	390

CHAPTER 4

DESIGN APPROACH AND MODELING METHODOLOGY

4.1 Development of numerical model for thermal optimization of non-uniformly powered microprocessors

The objective of this study is to minimize the on-die temperature of the described package by optimizing the distribution of the non-uniformly powered functional blocks on the microprocessor. The different specific cases of power maps that will be used for the optimization are 4 x 4, 5 x 5, 6x 6, 8 x 8, 10 x 10, 12 x 12 and 15 x 15. In order to model the non-uniform power dissipation on the microprocessor, the die surface area is divided into a given matrix with a matrix space representing a distinct functional block with a constant heat flux. Using a FEM code, an optimization of the positioning of the functional blocks relative to each other is carried out in order to minimize the junction temperature Tj. A Schematic diagram of a typical attachment shown in figure 3.1 is used for the analysis. It has a heat sink of 64 x 64 x 6.35 mm attached to the spreader of 31 x 31 x 1.8 mm to the silicon die with the help of TIM I and TIM II with a thickness of 0.025 mm. Geometry is created using given dimensions for die, spreader, heat sink and TIMs as shown in table 3.1. Die is divided into a matrix of above mentioned power maps to study the optimization of given power maps as described. A steady state thermal analysis is carried out to calculate the maximum junction temperature (Tj) for given power maps for both the test cases such as 4 x 4 and

6 x 6. Mesh sensitivity analysis is carried out with temperature varies in the range of +/- 1 ° C; coarse meshing is used for the analysis. Figure 4.1 shows the finite element model of the package considered for this study.

Parameters such as thicknesses of TIM I and TIM II, thermal conductivities of TIMs and convective transfer coefficient are calculated to satisfy the given package conditions. All the material properties are applied as per the material property data. Total power of 115 W is applied over a die as per the given power maps. A (h) of 1200 W/m^2 °C is used on the top of package for the analysis. This (h) is an effective heat transfer coefficient applied per unit area of the heat sink base (equal to the fin average h times the fin/base area ratio). Once baseline cases for given non-uniform power maps are carried out, an optimization study is carried out to see the effect of non-uniform power on the maximum junction temperature. Percentage variation in the optimized temperature is plotted against baseline temperature. Design guidelines are then suggested for any number of power maps. Tool Ansys Workbench is used for the analysis and optimization study.

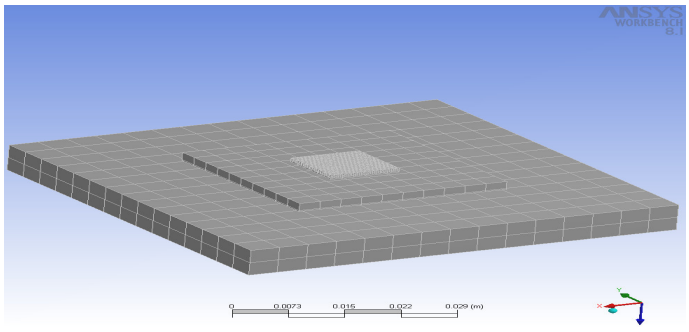
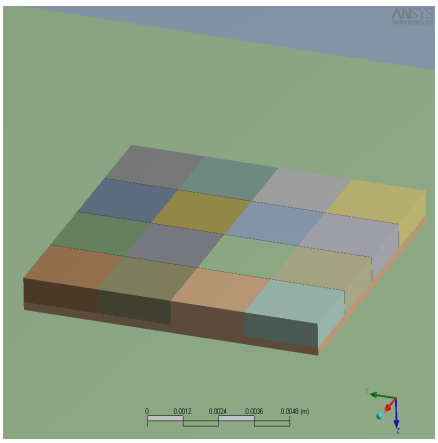
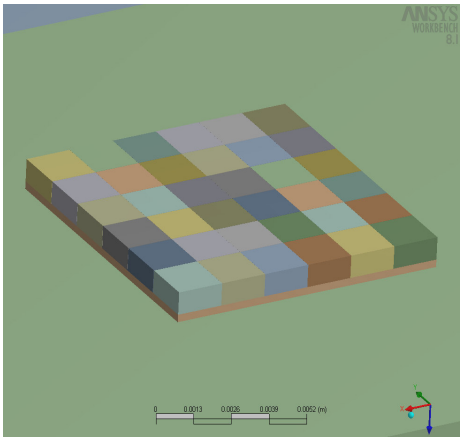


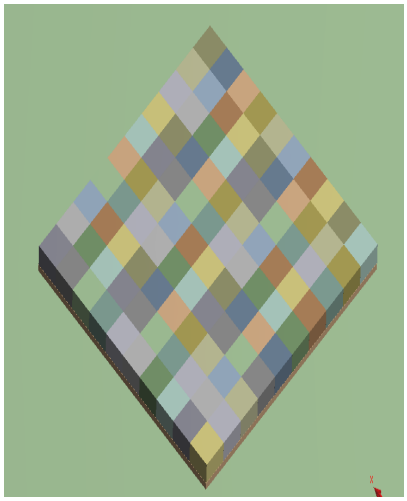
Figure 4.1: FEM Model of the package with number of elements = 4574 and number of nodes = 29085



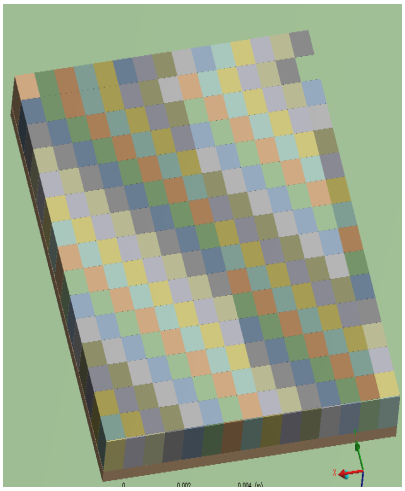
(a)



(b)



(c)



(d)

Figure 4.2: Divided die for various different cases of power maps
(a) 4 x 4 case and (b) 6 x 6 case (c) 10 x 10 case and (b) 15 x 15 case

4.1.1 Modeling Methodology

Maximum junction temperature is simulated for the different power maps as mentioned in above section as baseline case. Initially optimization study is carried out for test cases such as 4 x 4 and 6 x 6 power maps to define the number of functional blocks for any number of power maps. Figure 4.3 gives a flow chart which is used to develop this numerical model for thermal optimization of non-uniformly powered microprocessor. For optimization study a functional blocks of 4, 6 and 12 different individual powers are used as a design variable for 6 x 6 power map. Blocks of 16 design variables are used for the case of 4 x 4 power map. Design Xplorer which is a module in ANSYS is used to analyze the optimization study. Functional block is a group of number of design variables i.e. different individual powers in a given power map. This analysis has no constraints placed on the redistribution of functional blocks. Figure 4.4 shows a solid model of divided die for four different cases. Figure 4.5 and 4.6 gives the different cases considered for optimization for two different power maps such as 6 x 6 and 4 x 4. A functional block of groups of 4,6 and 12 design variables can be seen on a 6 x 6 power map and 16 design variable in case of 4 x 4 power map as a test cases for optimization study. This study is then extended for all other power maps to get an optimized junction temperature. The best possible Tjmax reduction could thus be found. In reality constraints must be placed regarding the maximum separation of any 2 (or more) functional blocks to satisfy electrical timing and compute performance requirements.

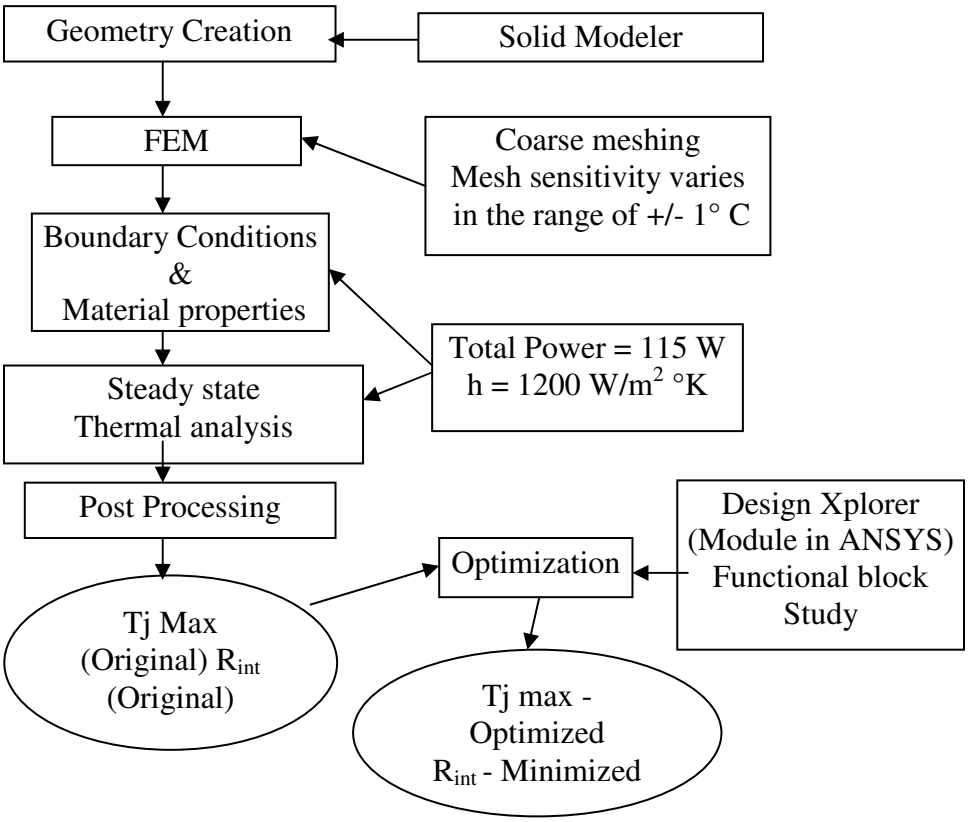


Figure 4.3: Flowchart for developing numerical model

4.1.2 Optimization Details

For the Design Xplorer design-of-experiment method (DOE), which is a module in ANSYS [17] , the locations of the automatic design points are determined according to a DOE that is called a central composite design with a fractional factorial design. The number of generated automatic design points is calculated using the number of input parameters selected in the parameters view. Design Xplorer automatically generates the response surfaces as part of the execution process. The upper and lower levels of the DOE points depend on whether the input variable is a design variable (optimization or

robust design) or an uncertainty variable (Six Sigma Analysis). Generally, a response surface will be more accurate when closer to the DOE points. Therefore, the points should be close to the areas of the input space that are critical to the effects being examined [17].

For example, for goal driven optimization, the DOE points should be located close to where the optimum design is determined to be. For a Six Sigma Analysis, the DOE points should be close to the area where failure is most likely to occur. In both cases, the location of the DOE points depends upon the outcome of the analysis. Not having that knowledge at the start of the analysis, you can determine the location of the points as follows,

For a design variable, the upper and lower levels of the DOE range coincide with the bounds specified for the input parameter. It often happens in optimization that the optimum point is at one end of the range specified for one or more input parameters. For an uncertainty variable, the upper and lower levels of the DOE range are the quantile values corresponding to a probability of 0.1% and 99.9%, respectively.

This is the standard procedure whether the input parameter follows a bounded (e.g., uniform) or unbounded (e.g., Gaussian) distribution. Because the probability that the input variable value will exactly coincide with the upper or lower bound (for a bounded distribution) is exactly zero. That is, failure can never occur when the value of the input variable is equal to the upper or lower bound. Failure typically occurs in the tails of a distribution, so the DOE points should be located there, but not at the very end of the distribution.

DOE Using a Central Composite Design [17]

The location of the automatic design points for the Design Xplorer DOE method is based on a central composite design. If *N* is the number of input parameters, then a central composite design consists of one center point, 2 x *N* axis point located at the -α and +α position on each axis of the selected input parameters and $2^{(N-f)}$ factorial points located at the -1 and +1 positions along the diagonals of the input parameter space. The fraction *f* of the factorial design and the resulting number of automatic design points are given in the table 4.1

Design parameters such as input, derived and response parameters are defined to start the optimization. Different powers (individual power levels) are the input parameters and maximum junction temperature is the response parameter. Using the DOE method the number of design points is directly related to the number of selected input Parameters. Number of design variables is usability variables, which depend on number of input parameters. Automatic design points are created using central composite design theory. Once it runs for the entire design points goal driven optimization is carried out to extract the desired results. In this case, goal is to have summation of all the individual powers in an assigned functional block should be same with keeping the individual power distinct to lower down maximum junction temperature. Numbers of design variables are restricted to 18, which allow to perform two different types of study on 6 x 6 and 4 x 4 power maps. These test cases are then applied to various cases such as 3 x 3, 5 x 5 etc to get trend of results. Once this study is carried out, a graph of percentage variation change in optimized temperature is plotted

against baseline case. From this, guidelines are set for distribution of power which will help to get comparable values of optimized temperature for a realistic power map i.e. 30 x 30. Following is a study shown for two test cases considered in this study.

Case II) 4 x 4 power map

As shown in figure 4.4, all the 16 different individual powers are grouped in a single functional block to optimize the design. This attempt comes up with design in a single attempt.

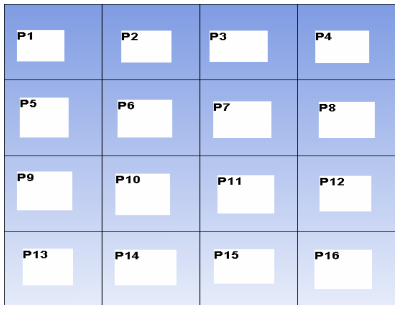


Figure 4.4: A test case with functional block with 16 design variables for 4 x 4 power map

Case I) 6 x 6 Power map

As shown in figure 4.5, functional blocks of groups of 4, 6 and 12 are created to study its effect on the overall junction temperature. A combination of theses 3 methods is used to optimize the maximum junction temperature. These different cases are chosen randomly to account for variations of individual powers. Once data for individual cases of different functional blocks is extracted, a combination for best cases is selected to come up with optimized power map for 6 x 6.

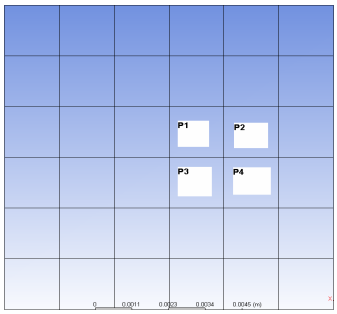
Table 4.1: Number of automatic design points as a function of the number of input parameters [17]

Number of input parameters	Factorial number f	Number of automatic design points
1	0	5
2	0	9
3	0	15
4	0	25
5	1	27
6	1	45
7	1	79
8	2	81
9	2	147
10	3	149
11	4	151
12	4	281
13	5	283
14	6	285
15	7	287
16	8	289
17	9	291
18	9	549
19	10	551
20	11	553

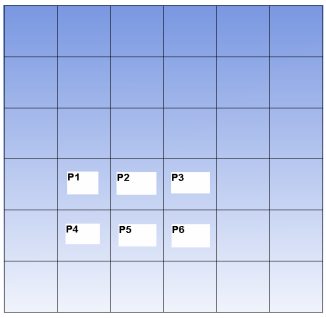
4.2 Development of analytical model for non- uniformly powered microprocessors

4.2.1 Methodology (Analytical Model) – Uniformly powered die

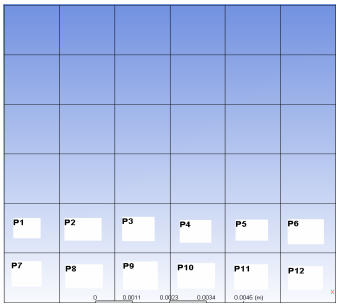
Analytical model as shown in figure 4.6 for two layer bodies developed by Haji-Sheikh et al [18] is used to calculate (Tjmax) maximum junction temperature for a attachment as shown. Figure 4.7 shows dimensions for two layer bodies.



(a)



(b)



(c)

Figure 4.5: Different test cases with groups of functional blocks of (a) 4 design variables (b) 6 design variables (c)12 design variables (b) for 6 x 6 power map

The dimensions in the x, y, and z directions are a, b, and c, respectively. L1, L2, W1 and W2 are the dimensions of the heating site or power generation on top of layer 2. The numerical examples are for two-layer bodies and they include boundary conditions of the first, second, and third kind. Power distribution can be located at any constant x plane (0 or finite thickness). Contact Resistance can be included at interface. Boundary condition of 1st, 2nd or 3rd kind on constant x planes. Boundary condition of 1st and 2nd on constant x and z planes.

The two layer model is extended to the described package within 3 steps. Uniform heat flux is applied to the base of TIM II and a convective coefficient (h) of

1200 W/m^2 °C is used on top of heat sink. This (h) is an effective heat transfer coefficient applied per unit area of the heat sink base (equal to the fin average h times the fin/base area ratio).

Using the temperature solutions given by equation 4.1 and 4.2, temperature distribution over a TIM II base is calculated. Thus using that as a boundary condition which is uniform temperature on top of IHS which is of same dimension to TIM II and uniform heat flux on TIM I, maximum temperature is calculated for those two layers. Finally following the same procedure, maximum temperature over die surface is calculated for uniform heat flux. Validation of the analytical model with numerical model is carried using previously developed numerical model.

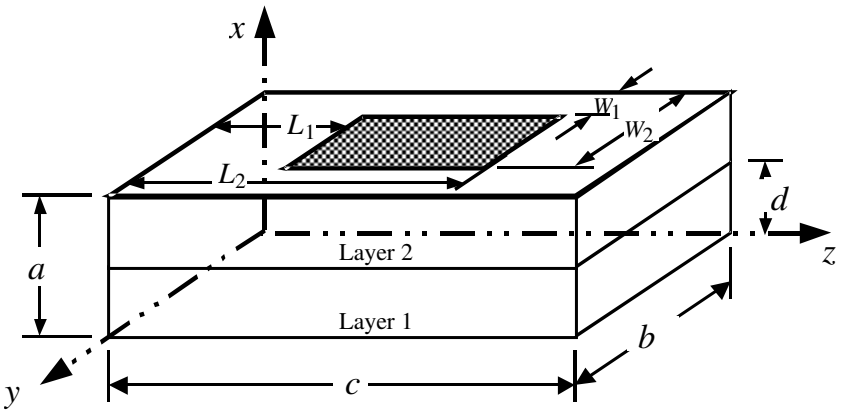


Figure 4.6: Schematic of a two-layer body [18]

The temperature solution [18] having boundary conditions of the first kind at y=0 and y=c are,

$$T_1(x,y,z) = -\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \frac{X_m(\beta_m x)Z_n(\nu_n z)}{k_2 N_{x,m} N_{z,n}} \frac{Y_{1,mn}(\gamma_{mn} y)}{\eta_{mn} Y'_{2,mn}(\eta_{mn} c)} \quad \text{---------- (4.1)}$$
$$\times \int_{z=0}^{d}\int_{x=0}^{a} q_0(x',z') X_m(\beta_m x') Z_n(\nu_n z') dx' dz'$$

$$T_2(x,y,z) = -\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \frac{X_m(\beta_m x)Z_n(\nu_n z)}{k_2 N_{x,m} N_{z,n}} \frac{Y_{2,mn}(\eta_{mn} y)}{\eta_{mn} Y'_{2,mn}(\eta_{mn} c)} \quad \text{---------- (4.2)}$$

$$\times \int_{z=0}^{d}\int_{x=0}^{a} q_0(x',z') X_m(\beta_m x') Z_n(\nu_n z') dx' dz'$$

Where T1 is in layer 1 and T2 is in layer 2

4.2.2 Methodology (Analytical Model) – Non-Uniformly powered die

An exact analytical solution to this problem is a difficult task if not impossible. To date, there is no analytical solution available to such a package with non-uniformly powered microprocessor. However, it is important to seek an alternative analytical solution that can also serve as a tool for verification of any numerical undertaking. The system under consideration, as shown in Figure 4.7, can be decomposed into three different solution sets to be designated as T_I, T_{II}, and T_{III}. The temperature T_I is a solution for the combined regions 1 and 2 with the boundary conditions of the second kind on all surfaces while there is a variable heat over the $x = x_1$ surface of the region 1. The symbol T_I stands for two solutions: T_1 in region 1 and T_2 in region 2. Similarly, the temperature T_{II} is the analytical solution for the combined regions 3 and 4 with the boundary conditions of the second kind on all sides and it stands for T_3 in region 3 and T_4 in region 4.

Finally, the plate number 5 has a boundary condition of the third kind of surface 6 while all other surfaces are having the boundary conditions of the second kind. Assuming the heat transfer coefficient over the surface 6 is specified, the temperature $T_{III} = T_5$ has a standard analytical solution except the heat flux over the surface 5 is an unknown, to be determined. Also, the solutions for each of the two-layer regions are

readily available [18]. As discussed for T_{III}, the heat flux over the surface 3 is also an unknown. Therefore, the heat flux values over the surfaces 3 and 5 must be computed a priori by employing an inverse methodology.

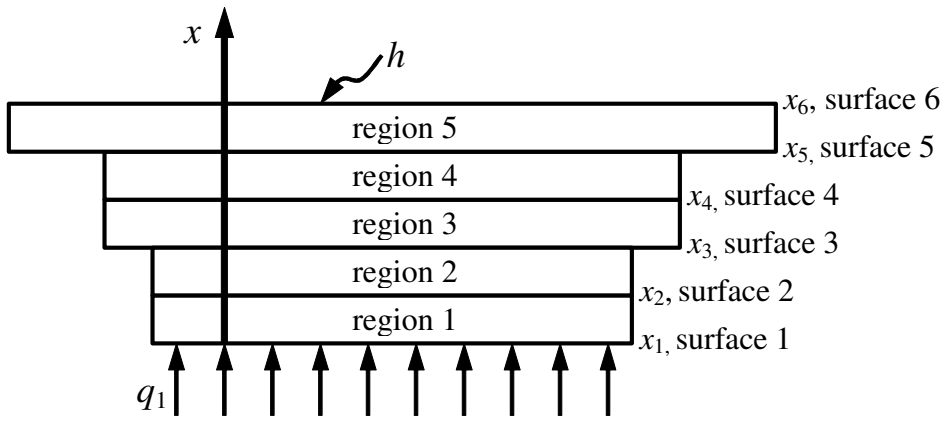


Figure 4.7: Schematic of the regions being studied

The solution is to begin by designating each surface heat flux as a volumetric heat source

$$g_i = q_i \, \delta(x - x_i) \qquad \text{---------- (4.3)}$$

For the surface between regions 2 and 3, $i = 3$ and for the surface between regions 4 and 5, $i = 5$. The symbol $\delta(x - x_i)$ designates the Dirac delta function and x_i is the coordinate of the surface i while the symbol q_i stands for the corresponding surface heat flux component in the x-direction. Once T_I, T_{II}, and T_{III} solutions are in hand, the contact conditions over surfaces 3 and 5 are

$$q_3 = (T_2 - T_3)\big|_{x=x_3} / R_3 \qquad \text{---------- (4.4)}$$

along the contact area on surface 3 and

$$q_5 = (T_4 - T_5)\big|_{x=x_5} / R_5 \qquad \text{---------- (4.5)}$$

along the contact area on surface 5. The parameter R_2 is the contact resistance between regions 2 and 3 at $x = x_3$ and R_5 is the contact resistance between regions 4 and 5 at $x = x_5$. In the case of perfect contact between the plates,

$$T_2\big|_{x=x_3} = T_3\big|_{x=x_3} \quad \text{---------- (5.6)}$$

on surface 3 when the contact resistance $R_3 = 0$ and

$$T_4\big|_{x=x_5} = T_5\big|_{x=x_5} \quad \text{---------- (5.7)}$$

on surface 5 when the contact resistance $R_5 = 0$.

Since the exact solutions for T_2 and T_3 contain the unknown heat flux q_3, while the exact solution for T_4 and T_5 contain the unknown heat flux q_5 this system represents two equations to be solved in order to determine q_3 and q_5. However, the temperature solutions would indicate that these two equations are two integral equations to be solved simultaneously using an inverse methodology. Although, the formulation of temperature solutions is exact, an approximate procedure is needed to determine the numerical values for q_3 and q_5.

As shown in the flow chart given in figure 4.11, analytical model is developed using different steps which are explained as follows. The two layer model developed by Haji-Sheikh et al [18] is used to develop the model using developed procedure as explained in flow chart. As shown in figure 4.8, step 1 includes TIM II and spreader while figure 4.9, step 2 includes TIM I and die. Initial boundary conditions are applied to these 2 different two layer bodies as shown in figure 4.9 and 4.10. In step 1, effective h of 5117 W/m^2 ° C is calculated using area average of heat sink base and TIM II. Non

uniform heat flux is applied on bottom of die as well as bottom of spreader. Once initial boundary conditions are applied, temperature at the interface of heat spreader (IHS) and TIM I is compared with modifying reference temperature. Figure 4.10 shows temperature locations T_I and T_{II} at spreader and die surface respectively. Use new T_{ref} to calculate T_{II} at same locations. There is a perfect contact i.e. zero contact resistance between two layer bodies considered for this study.

Calculate q at different locations on interfaces of IHS and TIM I using following equations

$$q_{i,k+1}=[q_i]_{,k} + \epsilon(T_{I,i}-T_{II,j})_k \quad \text{----------} \quad (4.8)$$

Use steps 1-7 as mentioned below to find ϵ and S. Once comparison of S and ϵ is carried out, recalculate (q)new at that minimum value to get final T_I and T_{II} which will give junction temperature (Tj)

$$S^2 = \frac{\Sigma(T_{I,i} - T_{II,i})^2}{N} \quad \text{----------} \quad (4.9)$$

Step 1 : Use Initial guess for ϵ

Step 2: Calculate $(q)_{new}$ i.e. q_{k+1} using equation (4.8)

Step3 : Calculate S using equation (4.9) for ϵ

Step 3 :Calculate $\sum(q)_{new}$ and $\sum(q)_{old}$

Step 4 : Calculate $\delta = \sum(q)_{new} / \sum(q)_{old}$

Step 5: Update $\epsilon_{new} = \epsilon / \delta$

Step 6: Calculate S equation (4.9) for ϵ_{new}

Step 7: Compare S for ϵ and ϵ_{new} (If S = very small value, Stop)

If value of S_{new} is lower than S , the solution is converging

If value of S_{new} is greater than S , the solution is not converging, use different value of Є as a Initial guess

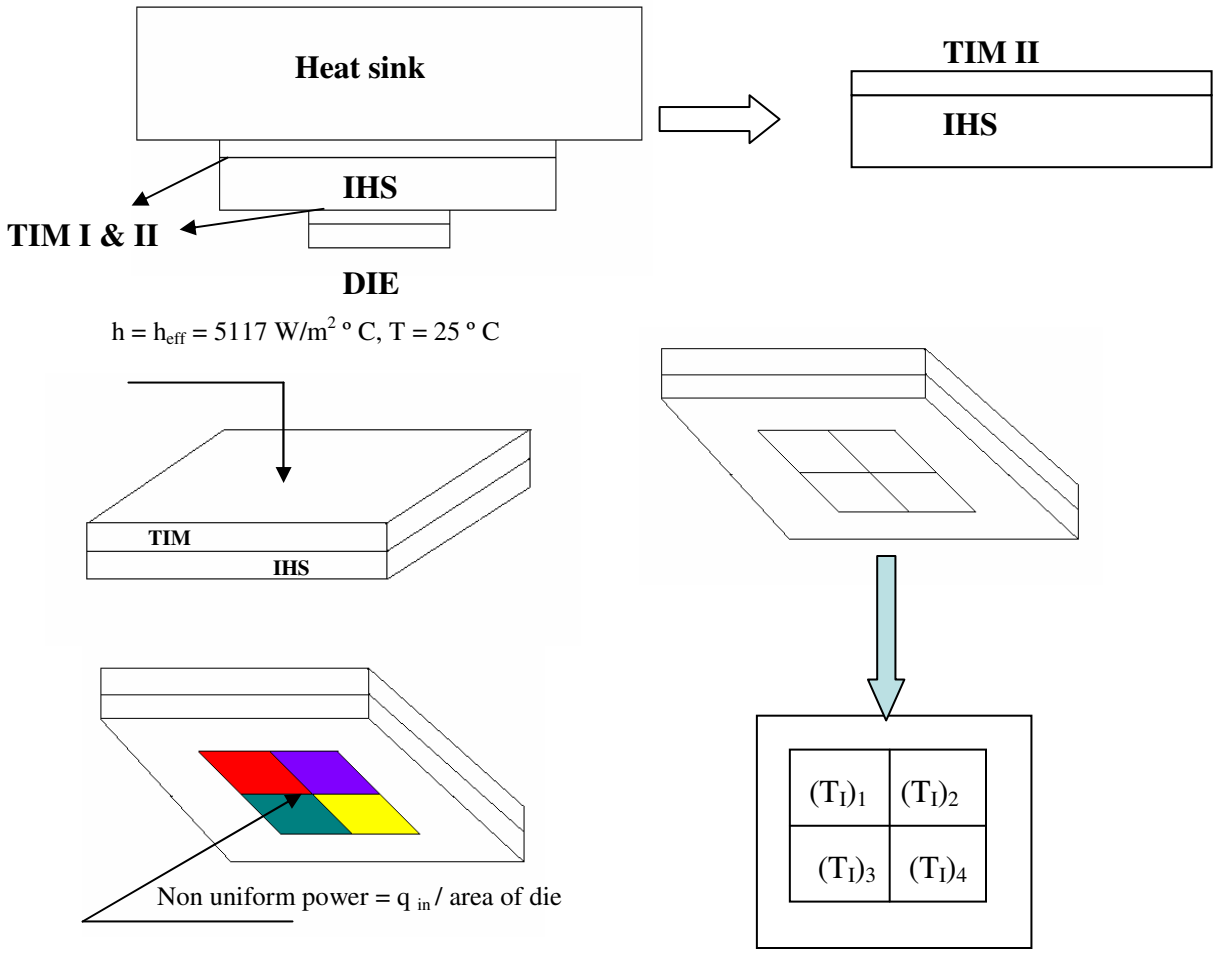


Figure 4.8: Step 1 (TIM II and Heat spreader)

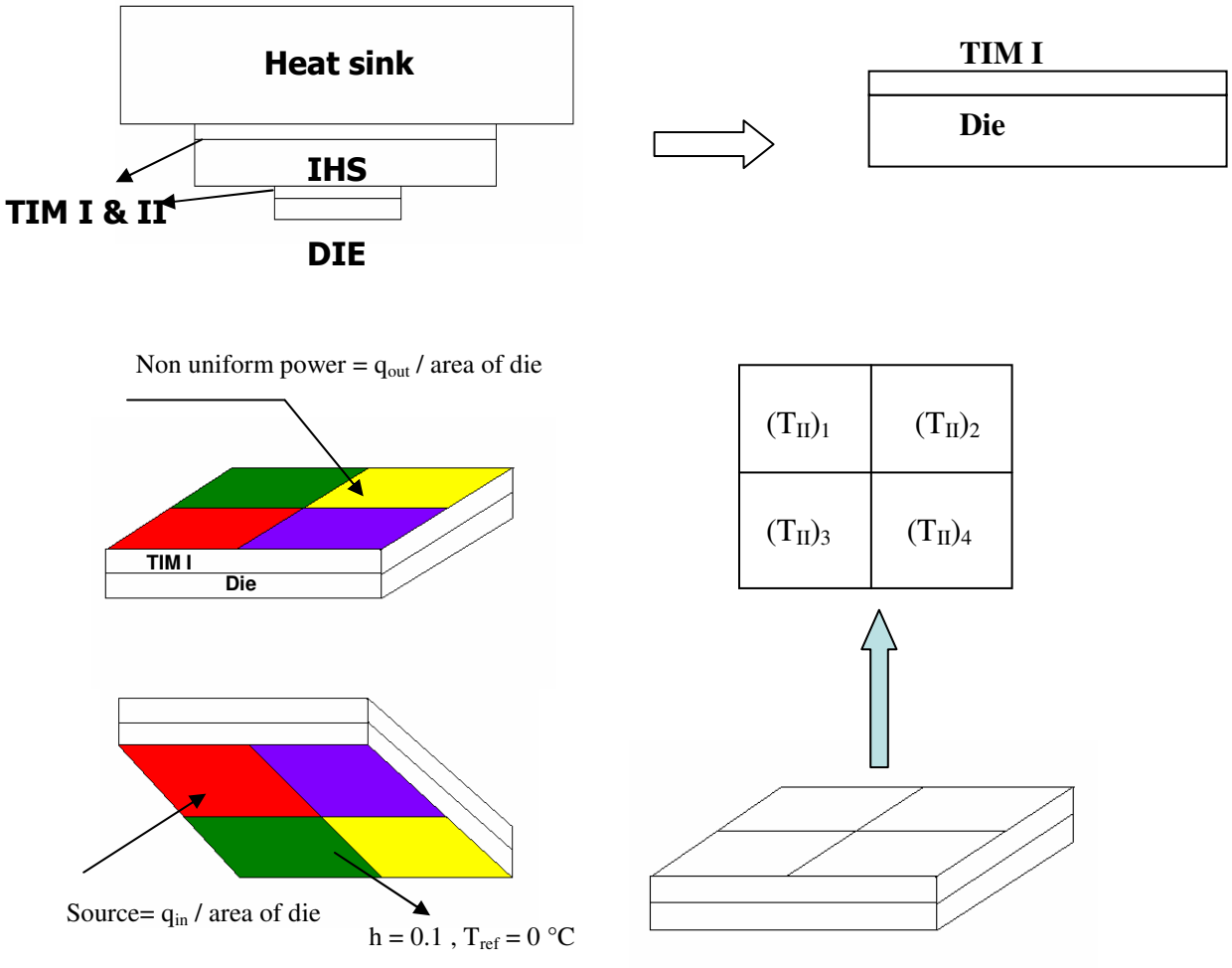


Figure 4.9: Step 2 (TIM I and Die)

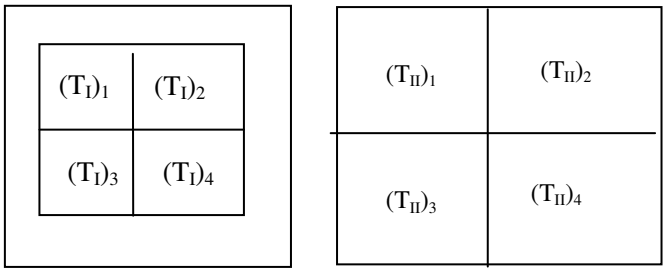


Figure 4.10: Values of T_I and T_{II} at interfaces from step 1 and step 2

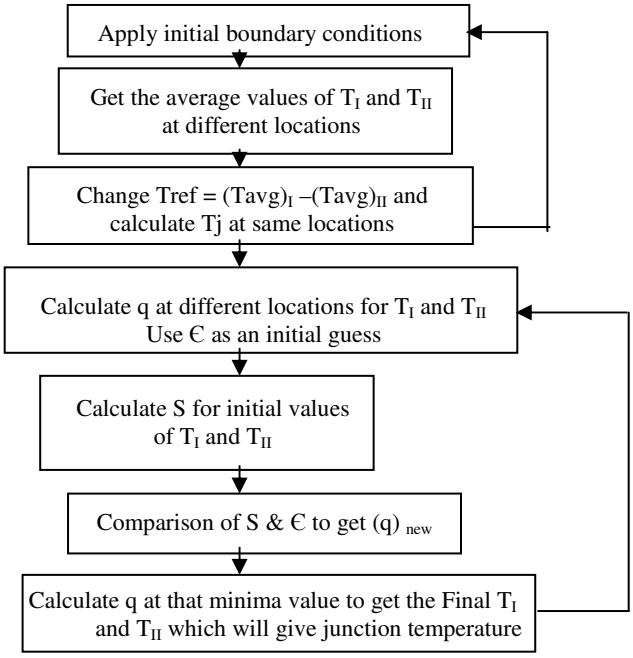


Figure 4.11: Flow chart for developing an analytical model for non-uniformly powered microprocessor

4.3 Development of numerical model using multi-objective optimization based on both thermal and device clock performance for non-uniformly powered microprocessors

4.3.1 Modeling Methodology

In this study, numerical model is developed that utilizes multi-objective optimization consists of redistribution of functional blocks to improve device performance and minimize on-die temperature. The same attachment consists of base of heat sink, TIM I and TIM II and heat spreader with the different die architecture are used to develop the numerical model. Developed design guideline using numerical model for thermal optimization is used as a base line case. This baseline case is

embedded into silicon floor plan to developed compact model satisfying both electrical and thermal performance. Non-uniformity is a common trend as Compaq and Pentium processors are two case studies considered in this study for a more universal problem. As shown in figure 4.12 and 4.16 floor plan of Compaq alpha 21364 and Pentium IV architecture are two cases considered to develop numerical model to satisfy thermal and electrical performance.

Constraints for the electrical optimization have been regarding the maximum separation of any 2 (or more) functional blocks. Once positioning of the functional blocks is carried out, thermal optimization of these non-uniformly powered functional blocks is carried out to minimize on-die temperature. Finally recommendations are provided for an architecture that optimizes based on both minimization of the length between critical functional blocks and minimization of on-die temperature.

Case I: Compaq alpha 21364

Geometry is created using given dimensions for Die which is a Compaq alpha processor of size 10 x 10 mm and core size of 4 x 4 mm, spreader, heat sink and TIMs. Die core is divided into various functional blocks as shown in figure 4.13. A steady state thermal analysis is carried out to calculate the maximum junction temperature (Tj) for given power distribution for various functional blocks as shown in table. Mesh sensitivity analysis is carried out with temperature varies in the range of +/- 1 ° C; coarse meshing is used for the analysis.

Parameters such as thicknesses of TIM I and TIM II, thermal conductivities of TIMs and convective transfer coefficient are calculated to satisfy the

given package conditions. All the material properties are applied as per the material property data. A h of 1200 W/m2 °C is used on the top of package for the analysis. This (h) is an effective heat transfer coefficient applied per unit area of the heat sink base (equal to the fin average h times the fin/base area ratio).

Figure 4.14 shows the critical functional blocks in a die core. These critical functional blocks are identified from the design guideline developed from the numerical model. These are the functional blocks which are dissipating the highest power. Once the baseline case is modeled, an electrical optimization is carried out for the same critical functional blocks. Different planes are selected as shown in figure 4.15 to optimize the floor plan. As shown in table, range of the planes is specified to come up with the floor plan (Best case and Worst case) that gives maximum junction temperature. Once the optimization is carried out, recommendations are provided for the optimum floor plan and critical functional blocks are identified within the Compaq alpha processor.

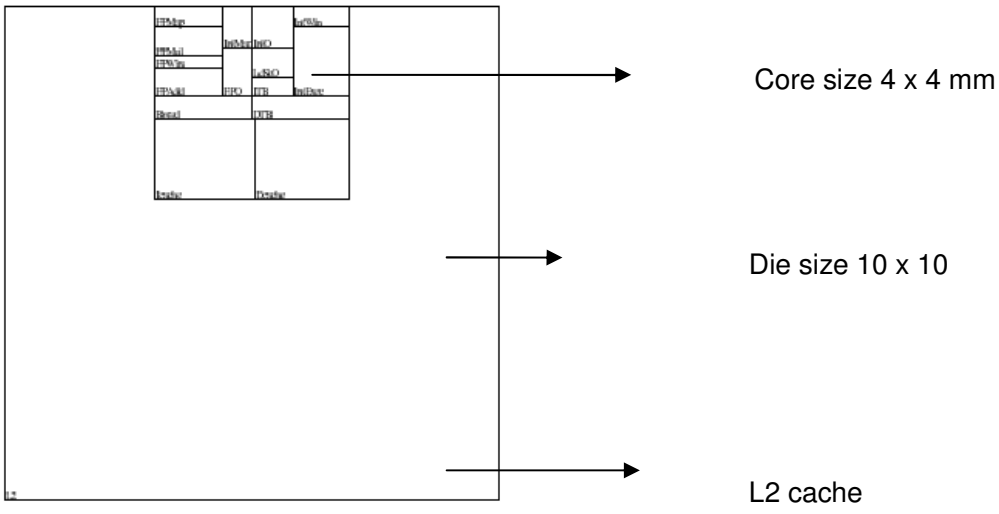


Figure 4.12: Die photo of Compaq Alpha 21364 [7,19]

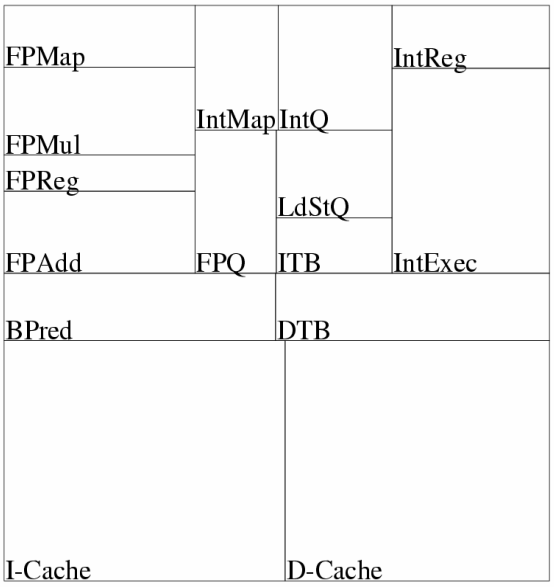


Figure 4.13: Floor plan – Expanded view
of Compaq Alpha 21364 with CPU core [7,19]

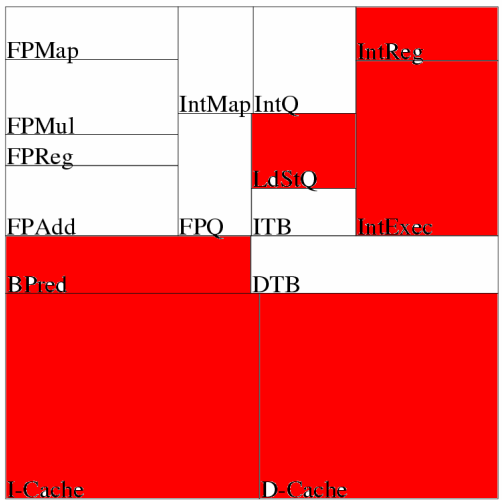


Figure 4.14: Functional blocks in a Compaq
processor core with highest power

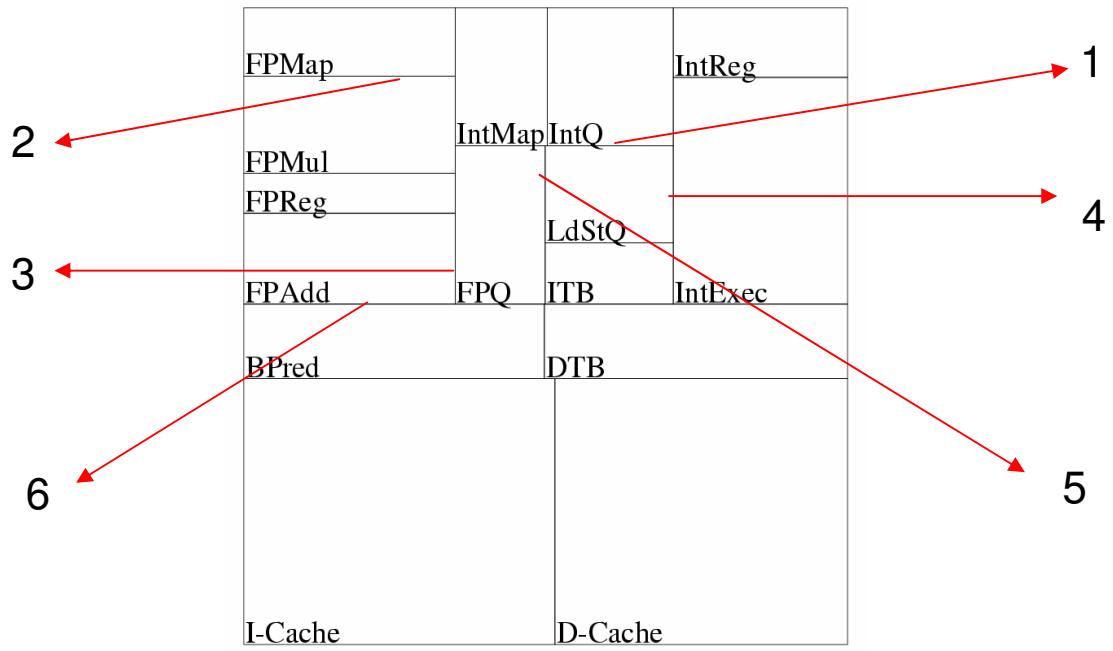


Figure 4.15: Planes selected in the floor plan of the Compaq core processor for optimization

Table 4.2: Range of values of different planes related to different functional blocks used in the optimization study

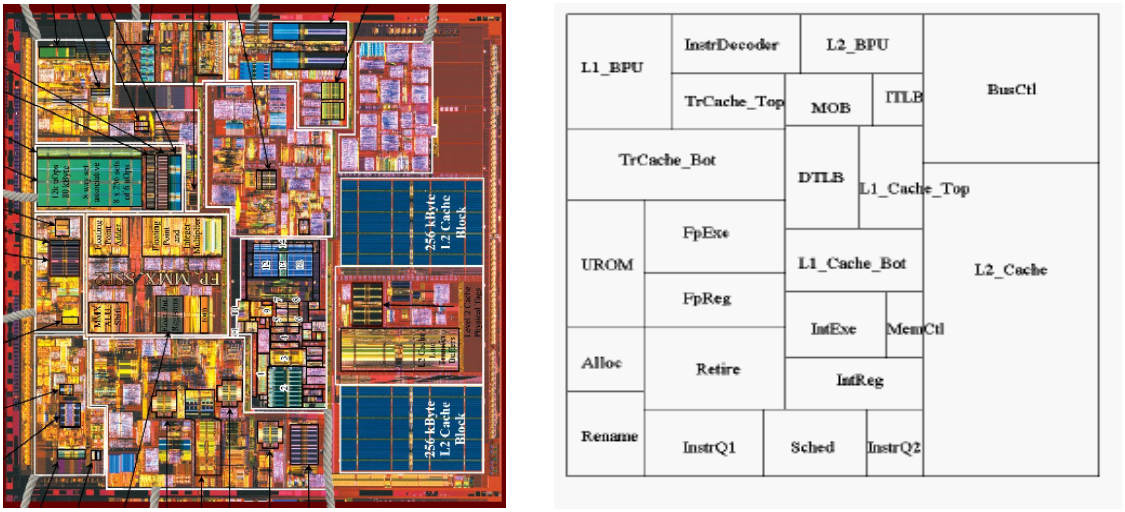
Number	Type of functional block	Range (Lower) mm	Range (Upper) mm
Plane 1	Intqmap	0.4	0.85
Plane 2	Fpmap	0.2	0.5
Plane 3	Intmap	1.2	1.6
Plane 4	Intexereg	1.2	1.8
Plane 5	Intqdtb	1.5	2.1
Plane 6	Dtbpred	1	2

Case II : Pentium IV Architecture

Figure 4.16 (a) shows a Pentium IV architecture [20]. It consists of the core and Cache of the microprocessor. The architecture shows the different functional blocks of the typical non-uniformly powered microprocessor. Figure 4.16 (b) shows a floor plan

layout of the same architecture. These functional blocks have specific arrangements on the die according to the transformation of the information between the blocks. The proximity of certain block to another depends on the function it does and the criticality of the information.

There are 24 functional blocks in this corresponding floor plan for Pentium IV architecture. In this study functional blocks with similar functionality are clubbed together to come up with total of 16 functional blocks as shown in figure 4.17.



(a) (b)

Figure 4.16: Pentium IV floor plan (a) Pictorial view (b) functional block distribution [20]

Geometry is created using given dimensions for Die architecture as shown in figure 4.17 which is a Pentium IV processor of size 10 x 10 including 16 functional blocks , Spreader, Heat sink and TIMs. Die is divided into various functional blocks as shown in Table 4.3. Pro-E is used to model the given die architecture. In the Pro-E

model, the each functional block may contain one or few blocks of the power map depending on the size of the functional block. Assembly is created at the end for various parts of the die (i. e. individual functional block is a part) . Each functional blocks in the model is parameterized which have two three parameter depending upon the shape of the block. This model then exported to Ansys Workbench taking all the parameters which is a key to carry out the multi-objective optimization. Table shows parameterization details for the various functional blocks in the Pentium IV architecture

A steady state thermal analysis is carried out to calculate the maximum junction temperature (Tj) for given power distribution for various functional blocks as shown in table 4.3. Mesh sensitivity analysis is carried out with temperature varies in the range of +/- 1 ° C; coarse meshing is used for the analysis.

Parameters such as thicknesses of TIM I and TIM II, thermal conductivities of TIMs and convective transfer coefficient are calculated to satisfy the given package conditions. All the material properties are applied as per the material property data. A (h) of 1200 W/m2 °C is used on the top of package for the analysis. This (h) is an effective heat transfer coefficient applied per unit area of the heat sink base (equal to the fin average h times the fin/base area ratio).

Once the baseline case is analyzed, different scenarios are considered by changing the location of the critical functional blocks to come up with the recommendation for the functional block distribution in the given Pentium IV architecture. This change in the location is determined using previously developed design guideline. Electrical constraints are defined as a minimum wire distance which

can be interpreted through the various parameters generated for an individual functional block.

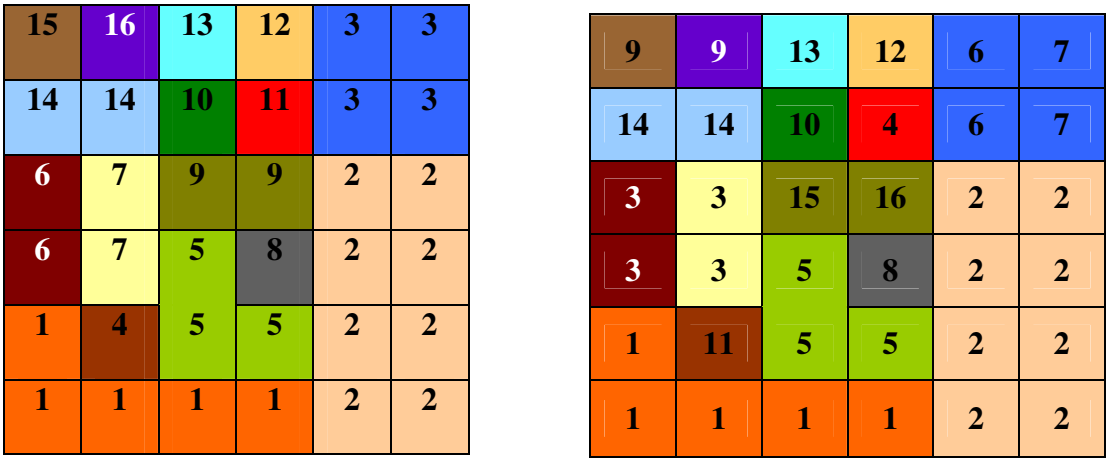
Figure 4.17: Baseline case of the Pentium IV architecture floor plan with clubbing different functional blocks from the original floor plan

Changing the location of functional blocks

In this case, the functional block locations are interchanged. Once the location is changed steady state thermal simulations is carry out to calculate maximum junction temperature (Tj) for this arrangement. As shown in figure 4.18, the functional block 3(Bus control) is interchanged with 6 & 7 (Floating point), the block 4(Retire) with 11 (MOB) and 9 (L1cache) with 15 & 16 (L1 Bpu & Decoder). This is done using the previously developed design guideline.

Table 4.3: Functional block distribution with their dimensions and power consumed

Block No.	Functional Unit	Dimensions (mm)	Power (Wu et al) W
1	Allocation+Rename+Instruction	6.66 x 3.33 x 1.667	7
2	L2 Cache	3.33 x 6.66	12.9
3	Bus Control	3.33 x 3.33	11
4	Retire	1.667 x 1.667	2.5
5	Integer Register + Integer Execution	3.33 x 3.33 x 1.667	0.5
6	Urom	1.667 x 3.33	3
7	Fp register + Fp execution	1.667 x 3.33	3.6
8	Memory Control	1.667 x 1.667	0.5
9	L1 cache	3.33 x 1.667	2.6
10	Dtlb	1.667 x 1.667	6.1
11	MOB	1.667 x 1.667	13.5
12	IlTB	1.667 x 1.667	3.5
13	L2 Bpu	1.667 x 1.667	7.7
14	Tr Cache	3.33 x 1.667	2.5
15	L1 Bpu	1.667 x 1.667	6.8
16	Instruction decoder	1.667 x 1.667	10



(a) (b)

Figure 4.18: Interchanging locations of functional blocks (a) Original floor plan (b) Modified floor plan

CHAPTER 5

RESULTS AND DISCUSSIONS

5.1 Development of numerical model for thermal optimization of non-uniformly powered microprocessors

Table 5.1 gives the various values of maximum junction temperature for overall study. Different cases are defined as per the functional block used such as 4-1, 6-1 etc. Various results within the different functional blocks are product of different cases of the positions on the given power maps. These cases are more effective for functional blocks of 4 and 6 as compare to 12 in achieving an optimized temperature. Combination study of functional blocks for 6 x 6 case is more effective than 4 x 4 case for the initial cases. Figure 5.1 shows the graph of minimum and maximum junction temperatures on different power maps. This also includes effect on temperature gradient for different power maps. Minimum junction temperature is comparatively stable as compared to maximum temperature as power matrix increases. In other cases, maximum temperature goes on increasing as the power matrix increases. This is possible due to the required uniform distribution of temperature for the higher power map matrices. This clearly shows that, non-uniform power distribution has significant effect on maximum junction temperature and therefore the gradient as power map matrix number increases. It is noted here that the average junction temperature from Tmax and Tmin is almost constant for all the various power matrices.

Figure 5.2 (a) and (b) shows the temperature distribution for original power map and optimized one for 6 x 6 case. Original power map clearly shows the hot spots in the center with non-uniform temperature all over the area of the die. With optimized case, temperature is distributed more evenly over the area of the die. Figure 5.3 (a) and (b) shows the graph of power distribution before and after optimization study for 6 x 6 case. Figure 5.4 shows the baseline cases (data before optimization) for 5 x 5 , 8 x 8 and 15 x 15. It can be seen that the hot spots are right in the center for all these cases. Figure 5.5 shows the temperature gradient (ΔT) for two different cases with values of $\Delta T = 47$ ° C for 5 x 5 case while values of $\Delta T = 53$ ° C for 12 x 12. Table 5.2 gives the various values of ΔT and maximum junction temperature (baseline and optimized) with a percentage variation for all different power maps. Figure 5.6 (a), (b), (c) and (d) shows temperature distribution for optimized one for different cases such as 5 x 5, 8 x 8, 10 x 10 and 15 x 15. Figure 5.7 and 5.8 is the graphical representation of the data given in Table 5.2. It can be seen that temperature distribution shows huge temperature drop within the die after optimization. Base line (original) temperature goes on increasing as the power map matrix number increases. This is possible due to the required uniform distribution of temperature for the higher power map matrices. It shows effect on percentage variation change in optimized temperature on optimized temperature compared to baseline case. With the optimized temperature, range of 5 % to 13 % drop in temperature is achieved compared to original temperature for different cases of power maps. As the number of design variables increases, a significant drop in temperature can be possible for higher functional blocks. Increase in the percentage

variation change in optimized temperature for ascending power maps is a key example of this fact.

Table 5.1: Results of various cases for two different power maps of 4 x 4 and 6 x 6

	Type of Case	Max. Junction temperature. (° C)
Case I: 6 x 6		
	Base line case (Actual power map)	124.8
	Functional block of 4 design variables	
	4-1	124.7
	4-2	124.7
	4-3	124.5
	4-4	124.2
	4-5	120.8
	4-6	122.0
	4-7	121.5
	4-8	122.4
	4-9	124.7
	Functional block of 6 design variables	
	6-1	125.3
	6-2	120.6
	6-3	124.2
	6-4	121.7
	6-5	124.8
	6-6	120.4
	Functional block of 12 design variables	
	12-1	124.4
	12-2	128.0
	12-3	124.0
	Modified Case (Optimized power map)	114.3
Case II: 4 x 4		
	Base line case (With actual power map)	121.8
	Functional block of 16 design variables	115.0
	Modified Case (Optimized power map)	115.0

Table 5.2: Junction temperature (maximum and minimum) and gradient over a given die surface for different power maps

	Tj (max)	Tj (min)	Gradient
	(ºC)	(ºC)	(ºC)
4 x 4 case	121	79	42
5 x 5 case	124	77	47
6 x 6 case	124	77	47
8 x 8 case	126	76	50
10 x10 case	128	76	52
12 x 12 case	129	76	53
15 x 15 case	132	76	56

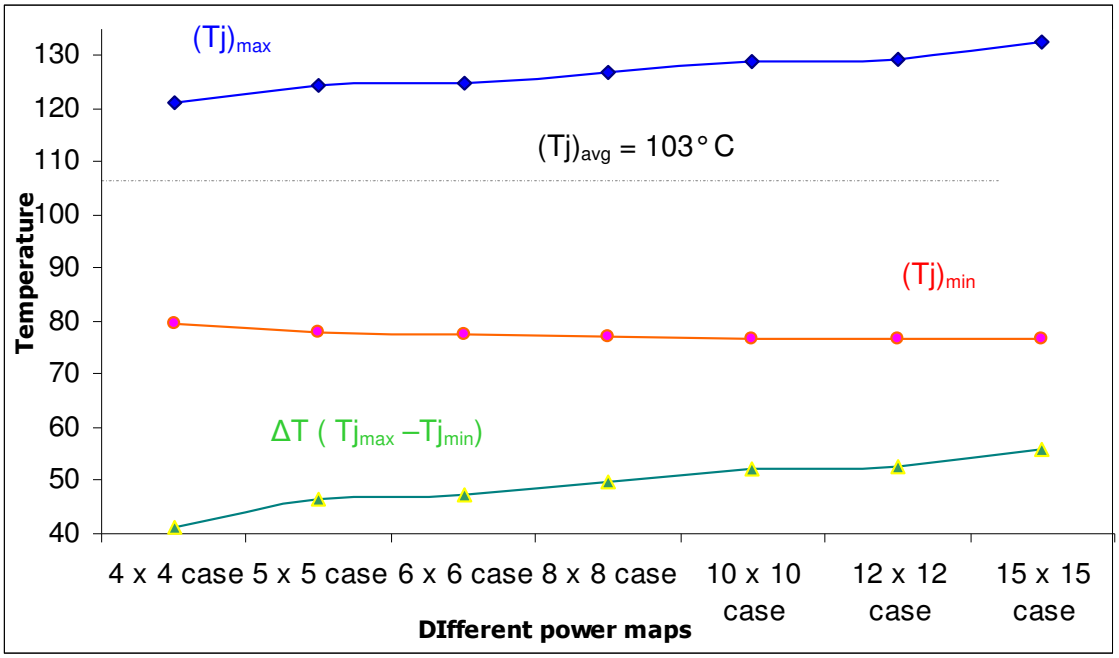
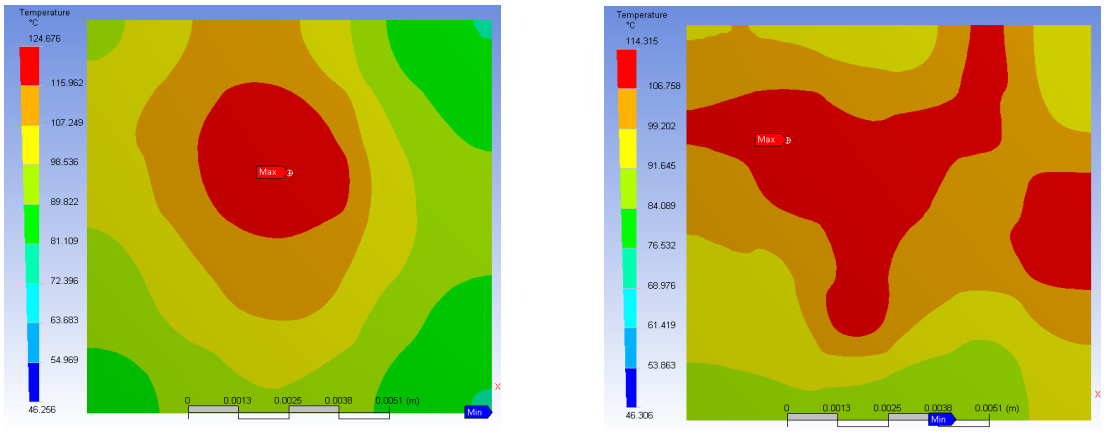
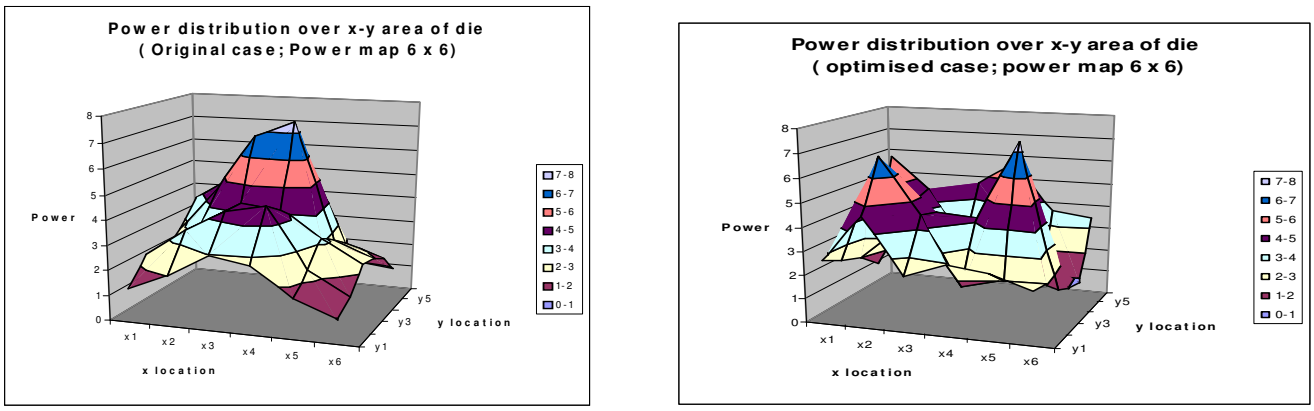


Figure 5.1: Graph of minimum and maximum junction temperatures on different power maps



(a) (b)

Figure 5.2: Temperature distribution for 6 x 6 power map (a) Original power map (b) Optimized power map



(a) (b)

Figure 5.3: Graph showing power distribution for 6 x 6 power map (a) Original power map (b) Optimized power map

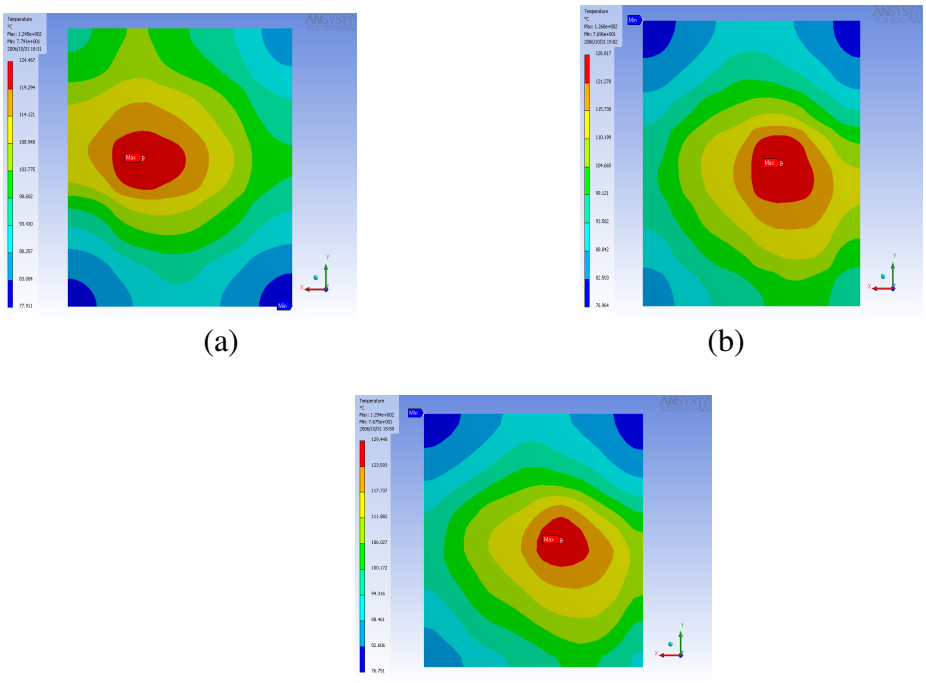


Figure 5.4: Temperature distributions for different power maps
(Baseline before optimization (a) 5 x 5 case (b) 8x 8 case (c) 15 x 15 case

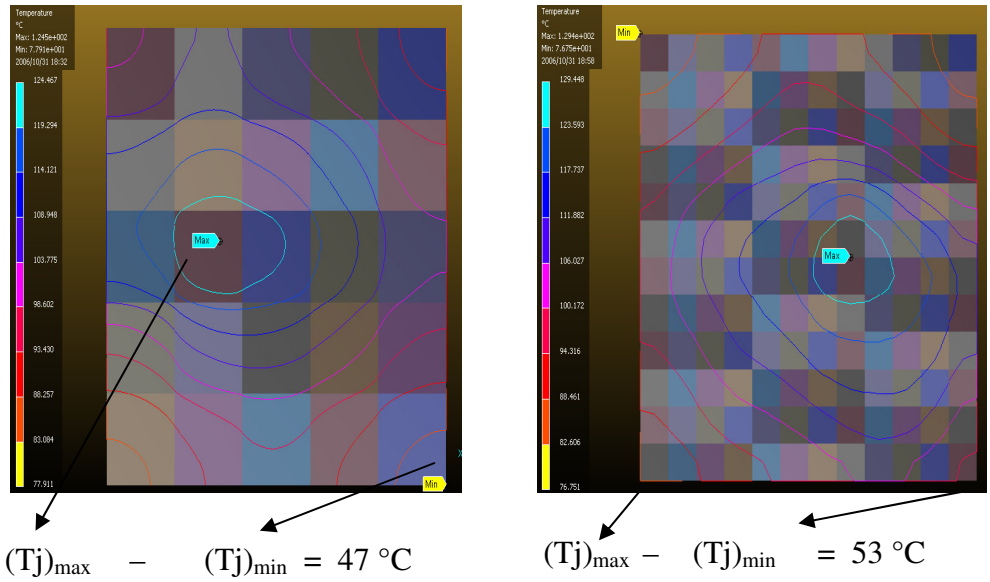
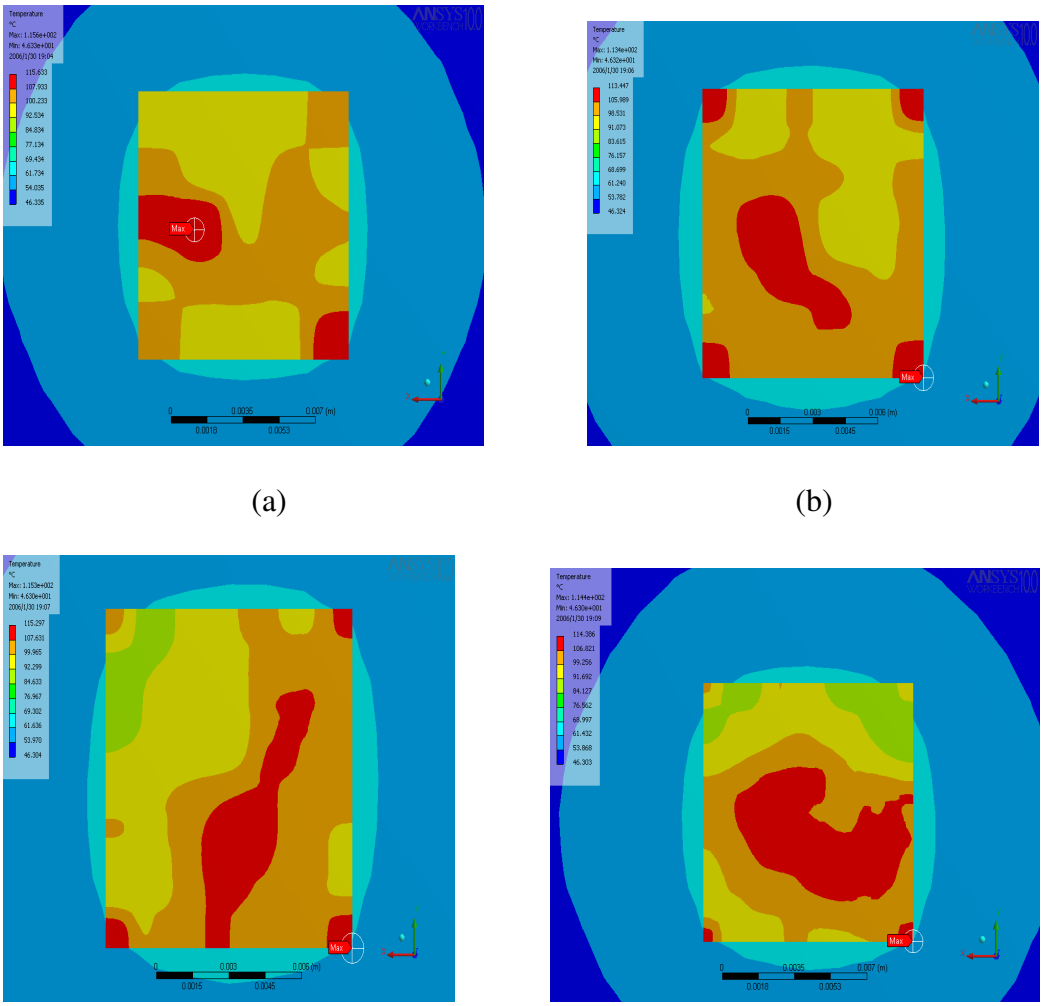


Figure 5.5: Temperature difference (Gradient) on overall die
before optimization (a) 5 x 5 case (b) 12 x 12 case



(a) (b)

(c) (d)

Figure 5.6: Temperature distributions for different power maps after optimization (a) 5 x 5 case (b) 8 x 8 case (c) 10 x 10 case And (d) 15 x 15 case

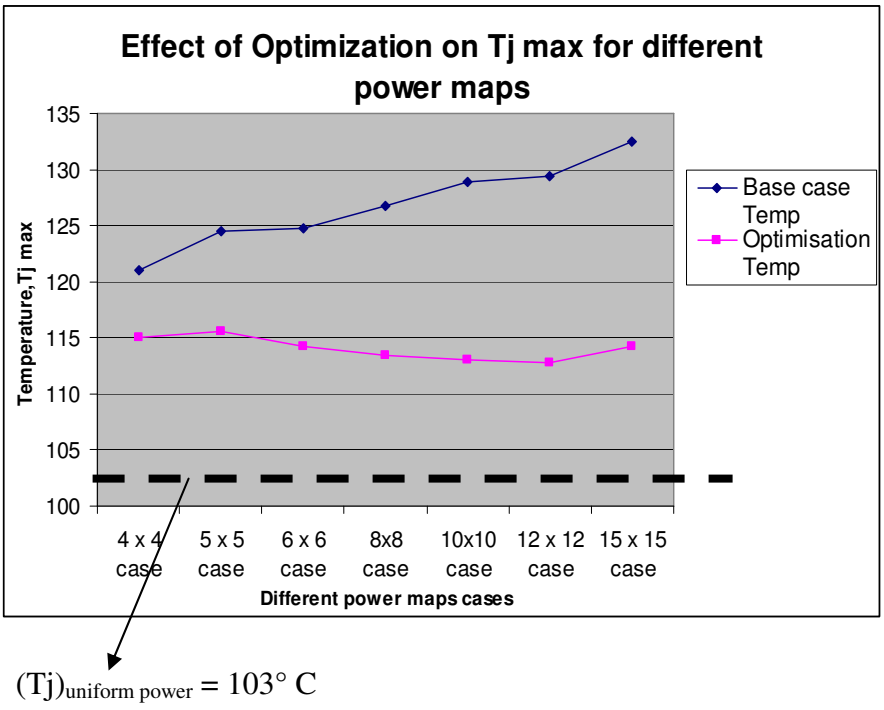


Figure 5.7: Maximum Junction temperatures (Baseline and Optimized) for different power maps

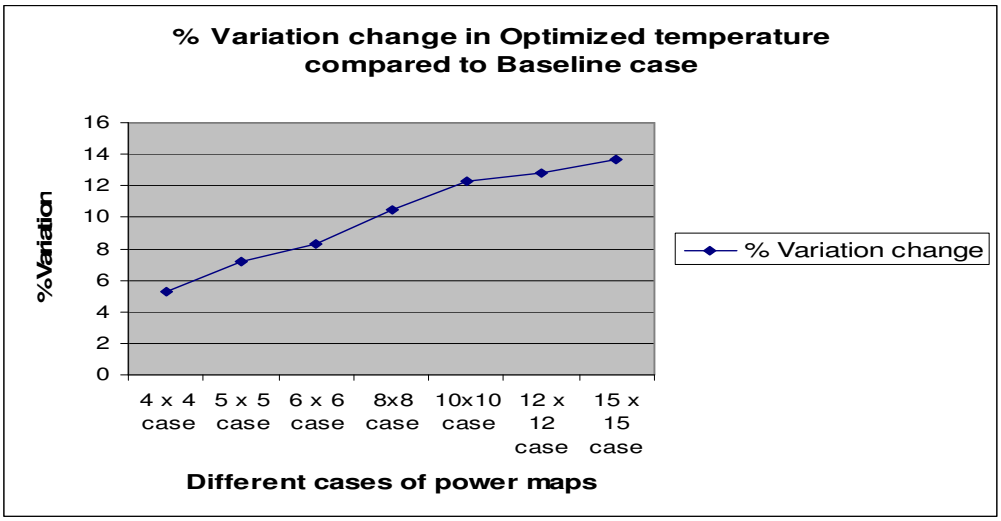


Figure 5.8: Percentage variation changes in optimized temperature for different cases of power maps

5.1.1 Design Guideline

Based on the results from the thermal optimization for many power maps, trend for the data is evaluated which helps to draw design guideline for a typical die configuration i.e. 30 x 30. This guideline is focuses on nature and distribution of power over a given die area with functional blocks to achieve minimum junction temperature.

Optimal power map is realized when the low individual power levels are placed in the periphery as compared to high power. Functional blocks of wide range of powers (combination of lower and higher power levels) are more effective as compared to blocks of same range. It is very important to consider the location of very high and very low power levels distributed over a given power map. Interchanging theses locations is very sensitive for optimization. Groups of high powers, which are located near the center, are more responsible for optimization within the die. As power map matrix increases control over optimization enhances because of different levels of individual powers within that power map. By following these design guidelines, further optimization studies can provide even better enhancement for a realistic power map case such as 30 x 30 with 900 functional blocks or more.

5.2 Development of analytical model for uniformly and non- uniformly powered microprocessors

Temperature distribution of a typical package consisting of a non-uniformly and uniformly powered die, heat spreader, TIM I & II and the base of the heat sink is calculated using an analytical model. The results are then compared numerically for validation. This uniform heat flux approach gives temperature of around 100 ºC for analytical model and 103.3 °C for numerical model. Figure 5.9 shows the temperature

distribution for uniform heat flux. This uniform temperature value can be ideal case for optimization of different functional blocks distribution within die for a given power matrix. The non uniformly powered model gives temperature around 118 °C and 121 °C for developed analytical model while 121°C and 125 °C for numerical model for power map cases of 4 x 4 and 5 x 5 respectively. The results of analytical model are within 5% of numerical model as shown in figure 5.10

This model is very robust, which will be ideally suited for upstream design which can be easily incorporated in Best Known Method for simplicity and optimization procedures. Furthermore, it can handle an arbitrary number of functional blocks i.e. non-uniformly powered test cases (Realistic case is 30 x 30) and not limited as computational tools are. Subsequently, the analytical model can be coupled with an optimization tool such as Ansys Workbench, providing a robust solution.

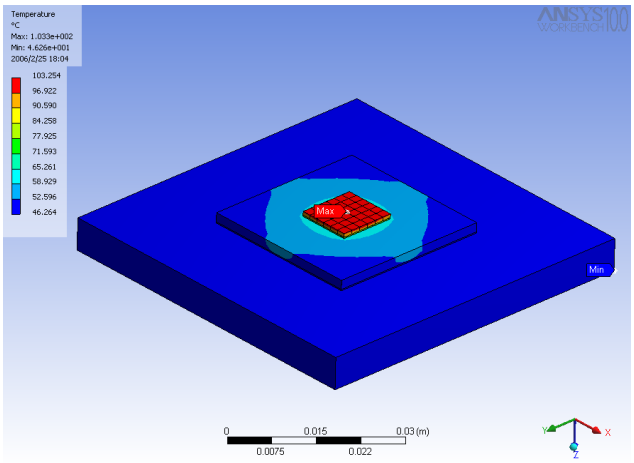


Figure 5.9: Temperature distribution for distribution of uniform heat flux over a die surface

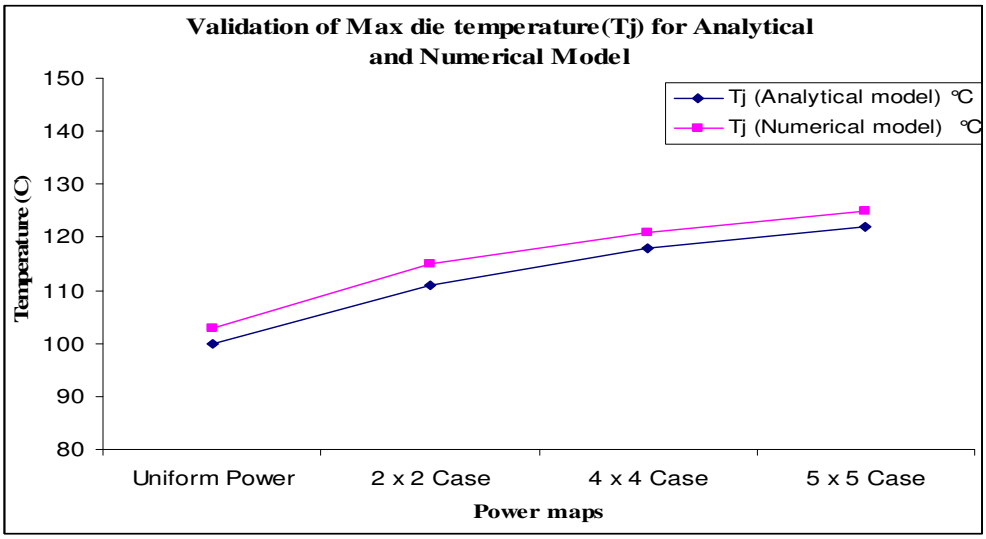


Figure 5.10: Validation of junction temperature (Tj) for die for numerical and analytical model for different cases

5.3 Development of numerical model using multi-objective optimization based on both thermal and device clock performance for non-uniformly powered microprocessors.

Case I : Compaq alpha Processor

Figure 5.11 shows temperature distribution in a original floor plan of Compaq Alpha 21364 CPU core before optimization which is considered as baseline case. Figure 5.12 shows a floor map of CPU core with 2 different architecture arrangements (Best case and Worst case) after optimization study. Temperature plots show junction temperature of 76°C for best case while 96°C for worst case architecture arrangements as shown in figure 5.10. Figure 5.13 shows the identification of critical functional blocks in a given floor plan after thermal optimization study. Functional blocks such as Integer units (Intmap, IntQ, IntTreg, Intexe), Bpred, Dtb are the critical functional blocks and most responsible for optimization of junction temperature (Tjmax). Table

5.3 shows the values of the planes selected for critical functional blocks after optimization study.

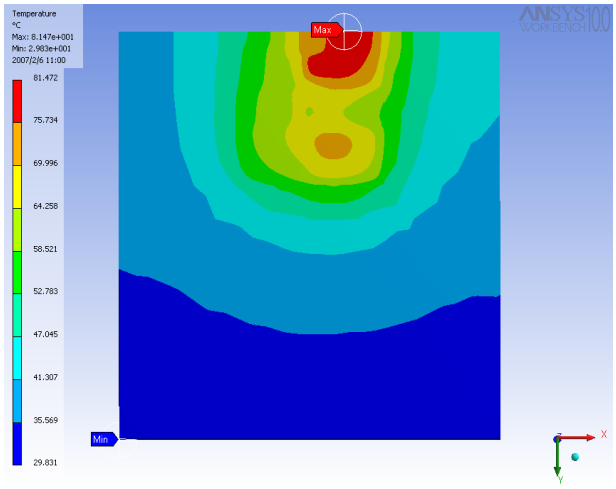
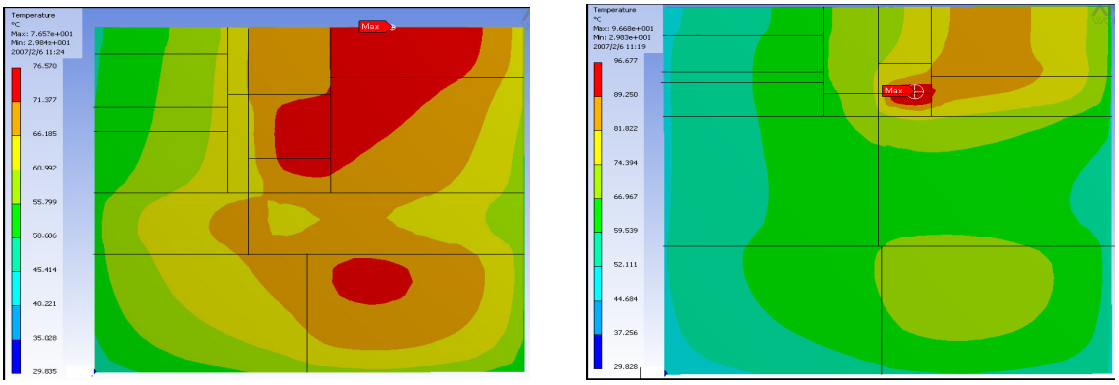


Figure 5.11: Temperature distribution (Original Floor plan for Compaq alpha Processor)



(a) (b)

Figure 5.12: A floor map of Compaq Alpha core with 2 different architecture arrangements (a) Best case (b) Worst case after optimization study

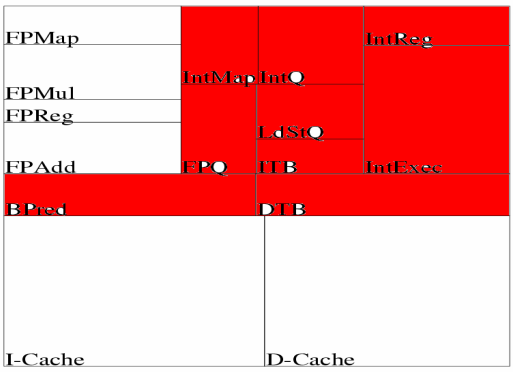


Figure 5.13: Identification of critical functional blocks in a given floor plan

Table 5.3: Temperature values for the best case and worst case floor plans with specific values of the planes selected for optimization

Number	Baseline Case	Best Case	Worst Case
Plane 1	0.6	0.82	0.84
Plane 2	0.35	0,22	0.3
Plane 3	1.4	1.4	1.52
Plane 4	1.5	1.76	1.58
Plane 5	1.85	1.58	2.06
Plane 6	1.5	1.95	1.08
Tj (ºC)	**81**	**76**	**96**

Case I : Pentium IV processor

Figure 5.14 shows the baseline case for a given Pentium IV architecture. It gives a maximum junction temperature (Tj) of 138 ° C. After changing the locations of the functional blocks the junction temperature (Tj) decreases to 129 ° C as shown in figure 5.15. Keeping the same wiring length, it is very important to keep the critical functional blocks relative to each other to give minimum junction temperature. Functional blocks such as Bus Control, MOB, Retire, L1 cache, L1 Bpu and Instruction decoder are the critical functional blocks. Junction temperature (Tj) is

reduce by 10 °C with redistribution of the critical functional blocks; although wiring distance remains almost the same as compare to original floor plan.

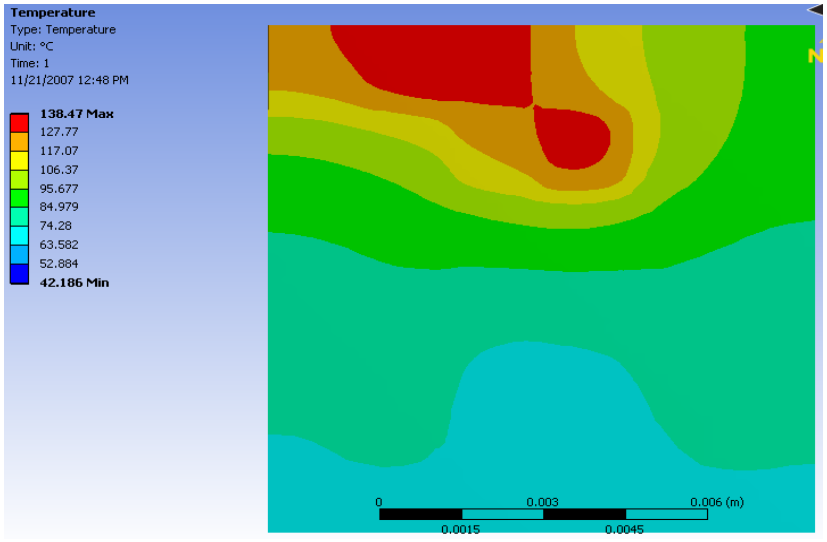


Figure 5.14: Maximum junction temperatures for the baseline case for a given Pentium IV architecture

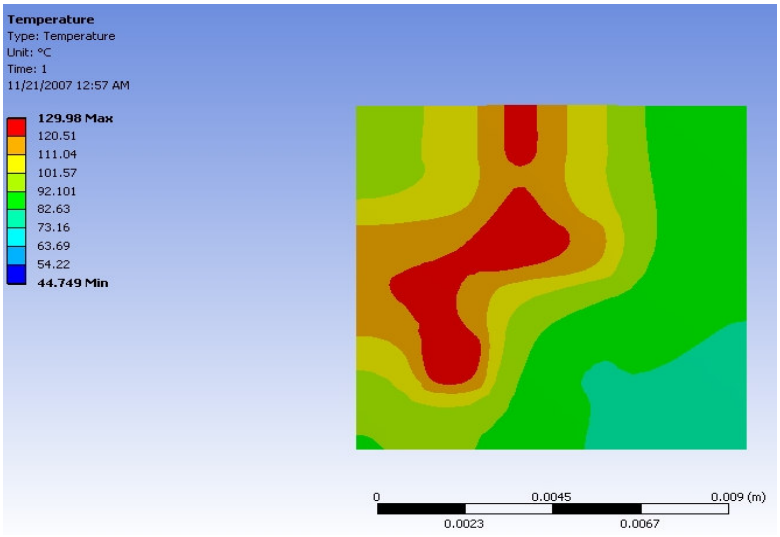


Figure 5.15: Maximum junction temperature for the modified floor plan (Changing the locations of functional blocks) for a given Pentium IV architecture

CHAPTER 6

CONCLUSION AND FUTURE WORK

- Development of numerical model for thermal optimization of non-uniformly powered microprocessors
    - By following design guidelines developed by numerical model, further optimization studies can provide even better enhancement for a realistic power map case such as 30 x 30 with 900 functional blocks or more.
- Development of an analytical model for uniformly powered and non-uniformly powered microprocessors.
    - This model is very robust, which will be ideally suited for upstream design which can be easily incorporated in Best Known Method for simplicity and optimization procedures. Furthermore, it can handle an arbitrary number of functional blocks and not limited as computational tools are.
- Development of numerical model using multi-objective optimization non- uniformly powered microprocessors.
    - Summarizing, given an initial floor plan in terms of functional blocks and corresponding power maps, the numerical model results in a new "floor plan" that optimizes the performance based on both thermal and device clock performance.

- This study can be extended to multi-core system which can also include various issues such as core hopping, leakage current to further improve the optimization.

APPENDIX A

NOMENCLATURE

h	Convective heat transfer coefficient, $W/m^2 \cdot K$
k	Thermal Conductivity, $W/m \cdot K$
N	Number of power map matrix
P	Total dissipated power, W
q	heat flux (W/m^2)
R	Resistance
T	Temperature, °C
T_I	Temperature at bottom of heat spreader
T_{II}	Temperature at top of TIM I
a	Ambient
j	Junction
max	maximum
a,b,c,d	dimensions in Figure 4.6, *cm*
A,B	constants, in Region 1
A_{mn}	Fourier coefficients
Bi_1	$h_1 b / k_{1y}$ in Region 1
Bi_2	$h_2 (c-b) / k_{2y}$ in Region 2
Bi_b	$1/R_b$
C,D	constants, in Region 2
F	a specified function
g_i	volumetric heat source in region *i*, W/cm^3
i, j	indices
k_{ix}	thermal conductivity in Region *i* along *x*, W/cmK
k_{iy}	thermal conductivity in Region *i* along *y*, W/cmK
k_{iz}	thermal conductivity in Region *i*, along z, W/cmK

m, n, p indices

$N_{x,m}$ norms for x-direction

$N_{y,mnp}$ norms for y-direction

$N_{z,n}$ norms for z-direction

q heat flux, W/cm^2

q_x x-component of heat flux vector, W/cm^2

q_y y-component of heat flux vector, W/cm^2

q_z z-component of heat flux vector, W/cm^2

r_i $(k_{ix}/k_{iy})^{1/2}$

R contact resistance, cm^2K/W

R_b Rk_{1y}/b

R_c $Rk_{2y}/(c-b)$

s_i $\cdot\sqrt{k_{iz}/k_{iy}}$

T_i temperature in Regions i, K

x, y, z coordinates, cm

X eigenfunction in x-direction

Y eigenfunction in y-direction

Z eigenfunction in z-direction

Greek

β_m eigenvalue for x-direction, cm^{-1}

γ_{mnp} eigenvalue for y-direction in Region 1, cm^{-1}

η_{mnp} eigenvalue for y-direction in Region 2, cm^{-1}

ν_n eigenvalue for z-direction, cm^{-1}

θ temperature function

Ψ eigenfunction,

REFRENCES

1. Teck Joo Goh, A.N. Amir, Chia-Pin Chiu and J. Torresola, "Novel thermal validation metrology based on non-uniform power distribution for Pentium(R) III XeonTM cartridge processor design with integrated level two cache," 51st Electronic Components and Technology Conference, 2001, pp. 1181-1186.
2. Ravi Mahajan, Raj Nair, V. Wakharkar, "Emerging Directions for packaging technologies," Intel Technology Journal, Q2,2002 http://download.intel.com/technology/itj/2002/volume06issue02/art07_emerging directions/vol6iss2_art07.pdf
3. S. Borkar, "Design challenges of technology scaling," *IEEE Micro*, pages 23–29, Jul.–Aug. 1999.
4. R. Mahajan, "Thermal management of CPUs: A perspective on trends, needs and opportunities," Oct. 2002. Keynote presentation at the 8th Int'l Workshop on Thermal Investigations of ICs and Systems
5. SIA. International Technology Roadmap for Semiconductors, 2005
6. S. Gunther, F. Binns, D. M. Carmean, and J. C. Hall, "Managing the impact of increasing microprocessor power consumption," In Intel Technology Journal, Q1 2001.

7. K. Skadron, M. R. Stan, W. Huang, S. Velusamy, K. Sankaranarayanan, and D. Tarjan. , "Temperature-Aware Microarchitecture," In Proceedings of the 30th International Symposium on Computer Architecture, pp. 2-13, June 2003.
8. K. Skadron, M.R. Stan, W. Huang, K. Sankaranarayanan, Z. Lu, and J. Lach., "The Need for a Computer-Architecture Approach to Thermal Management in Computer Systems," In Proceedings of the 5th IEEE International Conference on Thermal, Mechanical and Thermo-Mechanical Simulation and Experiments in Micro-electronics and Micro-systems (EuroSimE), pp. 415-22, May 2004
9. W. Huang, M. R. Stan, K. Skadron, K. Sankaranarayanan, S. Ghosh, and S. Velusamy, "Compact Thermal Modeling for Temperature-Aware Design," In Proceedings of the 41st Design Automation Conference, June 2004
10. Prakash, M., "Cooling Challenges for Silicon Integrated Circuits," The Ninth Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems, ITHERM '04, Volume 2, June 2004. *Pages 705-706*
11. Teck Joo Goh, "Thermal methodology for evaluating the performance of microelectronic devices with non-uniform power dissipation," 4th Electronics Packaging Technology Conference, 2002. pp. 312-317.
12. V. Gektin, Ron Zhang, M. Vogel, Guoping Xu and M. Lee, "Substantiation of numerical analysis methodology for CPU package with non-uniform heat dissipation and heat sink with simplified fin modeling," Thermal and Thermo-mechanical Phenomena in Electronic Systems, 2004, pp. 537-542 Vol.1

13. K.K. Sikka, “An analytical temperature prediction method for a chip power map,” Semiconductor Thermal Measurement and Management Symposium, 2005, pp. 161-166
14. M.S. June and K.K. Sikka, “Using cap-integral standoffs to reduce chip hot-spot temperatures in electronic packages,” Thermal and Thermo-mechanical Phenomena in Electronic Systems, 2002, pp. 173-178.
15. Zhiping Yu, D. Yergeau, R.W. Dutton, S. Nakagawa and J. Deeney, “Fast placement-dependent full chip thermal simulation,” VLSI Technology, Systems, and Applications, 2001 , pp. 249-252.
16. Tohru Suwa, Hamid Hadim, “Integrated Multidisciplinary Design And Optimization Methodologies In Electronics Packaging: State-Of-The Art Review,” ASME International Mechanical Engineering Congress and Exposition,2006
17. Ansys Workbench (10.0) reference manual , ANSYS, Inc. Canonsburg, PA,USA
18. Haji-Sheikh, A; Beck, J.V.; Agonafer, D, “Steady-state heat conduction in multi-layer bodies,” International Journal of Heat and Mass Transfer, v 46, n 13, June, 2003, p 2363-2379
19. Compaq 21364 die photo. From website: CPU Info Center. http://bwrc.eecs.berkeley.edu/CIC/die photos.
20. www.chip-architect.com

21. Wei Wu; Lingling Jin; Jim Yang; Pu Liu; Tan, S.X.-D., "A systematic method for functional unit power estimation in microprocessors," Design Automation Conference, 2006 43rd ACM/IEEE , vol., no., pp. 554-557, 24-28 July 2006
22. Peng Li, Yangdong Deng, Lawrence T. Pileggi, "Temperature-Dependent Optimization of Cache Leakage Power Dissipation," Computer Design: VLSI in Computers and Processors, 2005. ICCD 2005 Pp. 7- 12
23. Kiran Puttaswamy, Gabriel H. Loh, "Thermal Analysis of a 3D Die Stacked High Performance Microprocessor," Proceedings of the 16th ACM Great Lakes symposium on VLSI, 2006 Pp. 19 - 24
24. Yongkui Han, "Temperature Aware Techniques for Design, Simulation and Measurement in Microprocessors," Doctor of Philosophy Dissertation, February 2007
25. Wei Huangy, Eric Humenayz, Kevin Skadronz, and Mircea R. Stany, Charles L. Brown, "The Need for a FullChip and Package Thermal Model for Thermally Optimized IC Designs," ISLPED'05, San Diego, California, USA, August 8–10, 2005.
26. Pedro Chaparro, José González and Antonio González, "Thermal-Aware Clustered Microarchitectures," Computer Design: VLSI in Computers and Processors, 2004. ICCD 2004 Pp. 48- 53
27. Zuoyuan Li, Xianlong Hong, Qiang Zhou,Jinian Bian, Hannah H. Yang, Vijay Pitchumani, "Efficient Thermal-Oriented 3D Floorplanning and Thermal Via

Planning for Two-Stacked-Die Integration," ACM Transactions on Design Automation of Electronic Systems, Vol. 11, No. 2, April 2006, Pp. 325–345

28. David Brooks, Margaret Martonosi, "Dynamic Thermal Management for High-Performance Microprocessors," Proceedings of the 7th International Symposium on High-Performance Computer Architecture, Monterrey, Mexico, January 2001
29. Soraya Ghiasi, Dirk Grunwald, "Thermal Management with Asymmetric Dual Core Designs," University of Colorado Technical Report CU-CS-965-03. , June 2004
30. http://tbp.berkeley.edu/~jdonald/research/cmp/ghiasi_asymmetric.pdf
31. Jeremy W. Sheaffer, Kevin Skadron, David P. Luebke, "Studying Thermal Management for Graphics-Processor Architectures," Proceedings of IEEE International Symposium on Performance Analysis of Systems and Software, 2005
32. Kevin Skadron, Mircea R. Stan, Wei Huang, Sivakumar Velusamy, Karthik Sankaranarayanan, David Tarjan, "Temperature-Aware Computer Systems: Opportunities And Challenges," Published by the IEEE Computer Society, 2003 0272-1732/03
33. Kevin Skadron, "The Importance Of Computer Architecture In Microprocessor Thermal Design," Inter Society Conference on Thermal Phenomena, 2004
34. Jeremy W. Sheaffer, Kevin Skadron, and David P. Luebke, "Temperature-Aware GPU Design," International Conference on Computer Graphics and Interactive Techniques, 2004, Pp.100

35. W. Huang, M. R. Stan, and K. Skadron, “Physically-based compact thermal modeling—achieving parametrization and boundary condition independence,” in Proc. 10th Int. Workshop THERMal Investigations of ICs Syst., Oct. 2004, pp. 287–292.
36. http://www.cs.virginia.edu/~skadron/Papers/hotspot_therminic04.pdf
37. Mircea R. Stan , Kevin Skadron, Marco Barcella , Wei Huang , Karthik Sankaranarayanan ,Sivakumar Velusamy, “HotSpot: a Dynamic Compact Thermal Model at the Processor-Architecture Level,” Special issue on Therminic 2002 as paper T02MEJ 14
38. Karthik Sankaranarayanan, Sivakumar Velusamy, Kevin Skadron, Mircea Stan, “Microarchitectural Floorplanning for Thermal Management: A Technical Report,” Virginia University Charlottesville Dept Of Computer Science Technical Report CS-2005-08, May 20, 2005
39. Wei Wu, Lingling Jin, Jun Yang, Pu Liu, Sheldon X.-D. Tan, “A Systematic Method For Functional Unit Power Estimation in Microprocessors,” Design Automation Conference, July 24.28, 2006, San Francisco, CA, USA
40. Ning, T.H., “A CMOS technology roadmap for the next fifteen years,” Microelectronics and VLSI, 1995. TENCON '95., IEEE Region 10 International Conference on , vol., no., pp.1-4, 6-10 Nov 1995

BIOGRAPHICAL INFORMATION

Abhijit D. Kaisare received his B.S. degree in mechanical engineering from University of Pune in 1997 and M.S. degree in mechanical engineering from University of Texas at Arlington in 2002. He then pursued his Ph.D. degree in mechanical engineering. He worked as Graduate Research Assistant in the Electronics, MEMS and Nanoelectronics and Systems Packaging Center (EMNSPC). He also worked as Graduate Teaching Assistant in Department of Mechanical Engineering. His research interests are in the area of thermal management and characterization for electronic packaging and MEMS applications. He received his PhD in Mechanical Engineering from the University of Texas at Arlington in December 2007.