A Thermal-Driven Floorplanning Algorithm for 3D Network-on-Chip (3D NoC) Systems

ABSTRACT

3D Network-on-Chip (NoC) has emerged as a cutting edge technology that provides better performance by combining features of NoC and die-stacking 3D Integrated Circuit (IC) technology. It is able to push the limits of Moore's law by increasing the density of components in a chip resulting in higher functionality. The increasing packing density and power consumption of Systems-on-Chip (SoC) have made thermal effects one of the most important concern of chip designers. Increase in temperature degrades the performance , life time, reliability and increases the maintenance cost. Addition of more layers in the z dimension has increased the length of heat conduction path and power density per unit area; besides, the cooling capabilities of the 3D stack are much less because of the adiabatic inter-layer materials. To ensure thermal safety, we propose a novel algorithm that uses multi-objective Genetic Algorithm to minimize the peak temperature of each layer in 3D NoC. Three different architectures have been considered in our experimental work and the results show a decrease in the peak temperature . Experimental results of three different architectures shows decrease in peak temperature of respective layers when compared to the original floor-plan.

Keywords

Routing and layout;Genetic algorithms;Parameter tuning;Multi-objective optimization ; Performance measures

1. INTRODUCTION AND RELATED WORK

With the need to integrate an increasing variety of Intellectual Property (IP) blocks in a single System-on-Chip (SoC), communication management becomes critical when highly diversified functions with varying latency and bandwidth requirements must be supported.Three dimensional Network-on-Chip ($3DNoC$) is emerging as a promising technology to improve the performance in terms of throughput, latency, energy, area overhead and communication cost [2]. Two or more dies Separate chips are stacked one above another or side by side known as Chip stack, System-in-Package (SiP) or Multi-Chip Module (MCM). This can result into a reduction of interconnect wire-length and chip area in comparison to its 2D counterpart. With this approach, the system modules communicates with each other by sending packets to each other over the network. The well controlled electrical parameters eliminates the timing iterations and the high performance circuits alleviates the latency and increases the bandwidth.

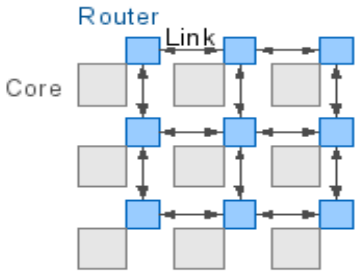

Figure 1: ***2D Homogenous Mesh Network-on-Chip***

NoC has borrowed ideas and concepts from computer networks and applies them to the embedded SoC domain. it is gaining popularity because of scalable communication, modular design and efficient use of communication links. NoCs use packets to route data from the source IP to the destination IP by using a network fabric that consists of network interfaces, routers and interconnection links. Interconnection links are also referred as channels or wire bundles. Network-on-Chips can be broadly divided into two types, homogenous and heterogeneous. Figure 1 shows a homogenous 2D Mesh NoC. Here each tile is a simple processor. Tiles can be replicated easily, so this structure is highly scalable and predictable. Due to homogenous integration, it gives less performance in comparison to heterogeneous one and network resource utilization is low. In case of heterogeneous NoC, the IPs can be memory,digital, analog, MEMS, RF, FPGA, core , general purpose or DSP processors. This architecture fits better to application domain. Most modern systems are heterogeneous in nature. It needs specialized routing and topology synthesis in more difficult.

The 2d chip fabrication has given rise to various 2D NoC structures in the recent years. Conventional 2D NoCs are facing lots of challenges due to exponential increase of PEs in a single layer. It leads to increase in number of hops to to from source to destination, ultimately increasing the latency to transmit the messages. Implementation of heterogeneous

ACM ISBN 978-1-4503-2138-9.
DOI: 10.1145/1235

technologies like high level scientific applications becomes critical in a single layer NoC . The increase in wire delay and power consumption made the chip designers to think adding IPs in the z-direction.

Multiple silicon layers are stacked vertically in 3D NoC. It has emerged as an effective mechanism to overcome the physical constraints and communication delays found in 2D NoCs. 3D NoCs saves space as it results in smaller footprints in each layer. The dies can be placed few millimeters apart vertically. Vertical placement of dies reduces interconnection length notably . The reduced interconnection length results in low latency and increase in performance. Data can be moved in both directions i.e. horizontally as well as vertically boosting the performance by 30% to 40%. More number of vertical vias between layers increases the bandwidth. 3D NoCs consume 40-50% less power than its 2D counterpart. As 3D NoC supports heterogeneous integration , the dies placed one above the other need not to be similar ones leading to optimization of chip components according to their diverse function.

3D NoC exacerbates the thermal problem in the inner layers. Clock frequency of micro-processors nearly doubles in each generation. The scaling in supply voltage is unable to remunerate the resultant increase in power [6]. If this will continue, it has estimated the value of power density to reach $10,000 W/cm^2$ by the end of 2016 [3].Increase in temperature affects the performance and reliability of the chip. With increase in temperature , 4% driving capability of transistors and 5% increase in the interconnection delay for each 10 degree Celsius rise in temperature [3]. has been noticed in case of More than 50% of the electronic failures arise due to increase in temperature and formation of hotspots [12].Lifetime of the chip decreases exponentially with increase of temperature. Higher temperature has adverse effects such as slow down of device, increase in leakage current and reduction in performance because of changing value of metal resistivity. The length of heat conduction path increases as more dies are stacked vertically [11].

2D Mesh is one of the well known 2D NoC architectures. This architecture consists of $r \times s$ routers interconnecting the IP blocked placed with them. This architecture is a popular one because of its regular structure and short interswitch wires. 3D Mesh based NoC can be formed this 2D NoC by simply adding the layers one above the other and ultra-wide buses are used to connect the layers in the third dimension.performance of 3D IC has been increased by integration of processor and memory in a stack [10]. By addition of a stack, the distance between processor and memory decreases leading to considerable decrease in memory access time.

Genetic algorithms proposed for latency aware mapping to reduce latency under congestion and no congestion [15].[1] is one of the first works to address the thermal-aware mapping for 3D NoC. Three ILP-based thermal-aware mapping algorithms for 3D NoC are proposed to explore the thermal constraints and their effects on temperature and performance in [9]. Another important fact in thermal issues is the distribution of temperature, which is not included. In this paper, we propose a thermal-aware floor-planning using multi-objective genetic algorithm which places the heat sources and heat sinks closer so that uniform distribution of temperature on the chip is possible.

2. THERMAL MODEL

In order to inspect the heat transfer domain, the thermodynamics system is converted into the thermal RC circuit. Some assumptions has been taken to model this. The heat flow is referred as current and the temperature difference is analogous to voltage. These equivalent circuits are called as thermal model in the thermal-design community. If the model includes thermal capacitors, it is then know as dynamic thermal model. In case of a 3D NoC system, multiple stacked layers are there . The corresponding thermal RC model of a 3D NoC is shown in table 1.

Table 1: Parameters used for temperature calculation in 3D NoC

Symbol	Description
$T_{x,y,z}$	Temperature of the node at the location (x, y, z)
$P_{x,y,z}$	Power consumption of the node at the location (x, y, z)
R_{inter}	Thermal resistance between vertical logic layer
R_{intra}	Thermal resistance between each horizontal logic layer
C_{inter}	Thermal capacitance between each vertical logic layer
R_{hs}	Thermal resistance for the heat spreader
C_{hs}	Thermal capacitance of the heat spreader
R_{hs}	Connective resistance of the heat sink

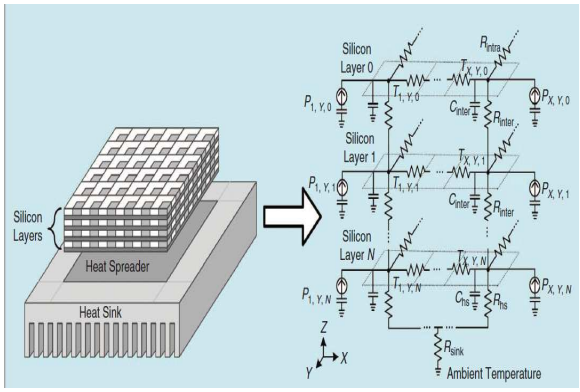

Figure 2: ***Thermal model of 3D NoC***

As shown in figure 2, the heat sink is attached in one side of the 3D NoC package. R_{sink} is used as the connective resistance of the heat sink in order to model the heat transfer between heat sink and ambiance. The transient temperature of each node has to be calculated in order to find out the total temperature of the chip. The change of temperature in a particular time interval can be formulated using FourierâĂŹs law. FourierâĂŹs law states that rate of heat flow is proportional to the gradient of temperature difference. Heat flow is the difference in temperature between system and surroundings. FourierâĂŹs law of heat transfer can be stated as follows, as follows,

$$\frac{\mathrm{d}T(t)}{\mathrm{d}t} = \frac{P(t)}{C} - \frac{T(t)}{RC} \tag{1}$$

$T(t)$ is the temperature of a particular node at time t and $P(t)$ is the total power consumption of a particular node at time t. R and C are used as effective thermal resistance and effective thermal capacitance towards the ambiance. To

simplify the above equation, it can be rewritten as,

$$\frac{\mathrm{d}T(t)}{\mathrm{d}t} = \alpha \cdot P(t) - \beta \cdot T(t) \quad (2)$$

In the above equation, $\alpha = \frac{1}{C}$ and $\beta = \frac{1}{RC}$. α and β are known as physic constant. These constants depends upon the material of the employed technology. To solve the linear differential equation, we can set the boundary condition, where the initial temperature at t_0 is set as T_0. Above equation can be solved as,

$$T(t) = \int_{\gamma=t_0}^{\gamma=t} \alpha P(\gamma)\, e^{-\beta}(t-\gamma) + T_0 \cdot T_0 \cdot e^{-\beta}(t-\gamma) \quad (3)$$

3. PROPOSED ALGORITHM

3.1 *Genetic Algorithms*

Introduction of genetic algorithms were as a computational analogy of biological systems and they are adaptive by nature. They are suitable for solving problems with computational intractable solution space. The solution is usually encoded into a binary string called *chromosome*. The search begins with a random set of chromosomes instead of a single one. The random of chromosomes is called *initial − population*. A *fitness − score* is assigned to each chromosome which is directly linked to the objective function of the optimization problem [8].

In order to optimize the thermal deviation between nodes, we propose a novel thermal-aware floor-planning algorithm for 3D stacked Mesh based NoCs. The algorithm consists of four phases. The population of chromosomes is modified to a new generation by applying three operators *reproduction*, crossover and mutation. Reproduction takes good chromosomes based on the fitness function of each individual and replicates them. Crossover picks two chromosomes randomly and chromosomes are exchanged with a probability of p_{ch}. The mutation operator swaps 1 to 0 and vice-versa with a probability of p_{mu} . Genetic algorithm applies the above three operators in every generation until a stopping criteria is met. This algorithm can search a large solution space while ingoing the unsuitable regions. This methods leads to a very time consuming evaluation.

In our experimental flow, we have modified MFA proposed by David Cuesta et al.[4]. Our algorithm performs by an incremental floor-planning by placing the functional units at its correct location. Our algorithm is named as MFA_{noc}. It is a MOEA based on genetic algorithm. The floor-planner arranges the coded solutions that are eventually improved in the optimization process to provide configurations optimized both in performance and thermal response for the target architecture. To this end, MFA_{noc} simultaneously minimizes the following three objectives:

- Objective1: After the algorithm is going to test all the x, y, z values for the best place to fix the component, now it will check two constraints. (1) Whether the component is outside the die or not. (2) Overlapping with other components.
- Objective2: The peak temperature of the chip. The computation of this objective depends upon the thermal model chosen. When n number of functional units are present, the maximum temperature of two processing elements PE_i and PE_j is given by the cross product of the respective power densities divided by their euclidean distance.

$$O_2 = \sum_{i<j\epsilon 1\ldots n} \frac{PE_i \cdot PE_j}{\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2}} \quad (4)$$

By minimizing O_2 , the algorithm tries to place the hottest blocks as far as possible. By executing this process, the peak temperature of the layers is reduced.

Thermal comparator function is called by the sort algorithm. Thermal comparator is used to sort the list of components in decreasing order of power density, i.e. in decreasing order of temperature. In that way, the first individual in the population is a permutation where the components are sorted in decreasing power density. This is very useful to perform the placement algorithm since if we place first the hottest elements, we can reduce the maximum temperature pretty much. However, the trade off is that the wire-length is increased.

Heat sources are those components of a circuit, where electrical current flows and the electrical current is converted into thermal energy. The components stores the thermal energy depending on the type of their materials and thermal capacitance. Our floor-planner helps to keep the heat sources as far as possible and it generally places them at the border of the chip. This technique helps in reducing the on chip temperature due to diffusion. Vertical heat spread is taken into account by not placing the heat sources one above the other. From the optimized floor-plans , it is clearly visible that most of the IPs are placed either in the first layer or in the last layer. Routers are placed at the border lines of the chip which yields to less heat.

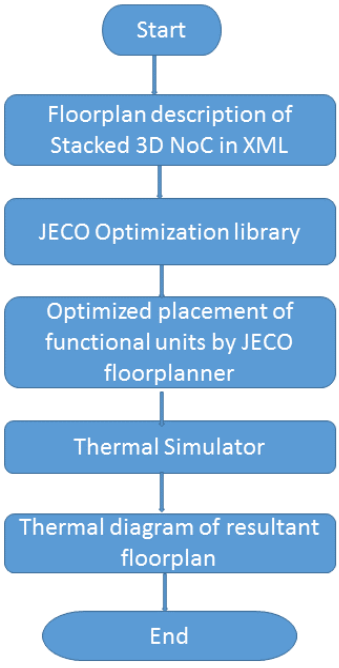

Figure 3: *Experimental Flow*

The Network-on-Chip is divided into small blocks called IPs and switches. Each IP is connected to its corresponding switch. Each processing element PE_i is characterized by a width PE_{wi} , a height PE_{hi} and a length PE_{li}. Similarly, the width, height and length of switch SW_i are denoted as SW_{wi}, SW_{hi} and SW_{li}. The outer volume has a maximum length L, maximum width W and maximum length L. The geometrical location of Processing element PE_i is (x_i, y_i, z_i) and switch SW_i is (a_i, b_i, c_i), where $0 \leq a_i \leq x_i \leq L - PE_{li}$, $0 \leq b_i \leq y_i \leq L - PE_{hi}$, $0 \leq c_i \leq z_i \leq L - PE_{wi}$. The switch is always having less dimension in comparison to its corresponding processing element, so the above geometrical constraints hold good. The co-ordinate (x_i, y_i, z_i) is used to denote the left-bottom-back co-ordinate of block PE_i and (a_i, b_i, c_i) for the attached switch.

In our evaluation process, we have used permutation encoding [13], every chromosome takes string of values characterizing different functional units of the target architecture. These records collects information related to the functional unit, namely layer number, width, length. Further freedom is assigned to the optimization process by managing length and width and layer number like rotation. Power densities of individual units are also taken into account with its connections with the other units by the algorithm. As these information are common to all functional units, our algorithm is not codifying these in to chromosomes.

All the chromosomes are having size N, where N is the number of functional units to be placed. The cardinality of the reseach solution space is N!.

- Selection: Binary tournament selection strategy is implemented. The best pair of couples are selected from the random couple of chromosomes to participate in crossover.

- Crossover: Cycle crossover is used to produce the offsprings. The crossover operator takes into account that all the components should appear once and only once in the chromosomes. In this way, we are avoiding redundancy.

- Mutation: Position of two blocks in the chromosomes are swapped , producing in a change of the placement sequence of the mutated chromosome. Another effect is the rotation of the individual.

3.2 Thermal profiling framework

Our thermal estimation framework is a compact resistive model based on the 3D ICE [14]. To calculate the temperature , we have given the average power consumption of each IP and each on-chip router. , location of each functional units , connections, physical dimensions and the number of layers in the 3-D design. Our Matlab simulator returns peak temperature of each layer based on the given values.

3.3 Communication framework

NoC traffic includes the data packets traveling from IP to IP via Switches. Our algorithm uses a 3-D stacked mesh based NoC structure, employing distance order routing algorithm. Our communication pattern is modeled on the basis of counting the number of hops from source router to destination router. The total communication cost can be obtained by the sum of all hops for each and every packet traveling from source IP to destination IP.

The pseudo code presentation for our experimental flow is shown . The framework is initialized with IP and router power values, traffic, connectivity and physical properties such as physical co-ordinates of all the functional units and number of layers. Original floor-plan is then generated. An initial number of solutions is produced. For each generation mapping and placement is done. Communication cost and worst-case temperature of each layer is calculated.The solutions are ranked based on the cost function. Crossover and mutation are applied on those solutions to generate new chromosomes. At the end of number of generations, the best solution is selected.

3.4 DESIGN FLOW

Our proposed method is of two fold: (1) to place the intellectual property cores (IPs) , routers at their respective positions in the decreasing order of their power density and (2) generation of thermal diagram of the 3D NoC architecture. To this end, we have proposed a novel thermal-aware floor-planner for 3D NoC , that includes various phases as shown in Figure 1.

4. EXPERIMENTAL SET-UP

The 3D NoC studied in our experimental work is based on the Mesh based architecture , with IPs fabricated in 90nm technology. The original architecture is discussed in [7] . We have taken the 3D stacked mesh based NoC as our baseline architecture. Some changes has been made in the original architecture i.e. more numbers of layers are added for inclusion of more number of processing elements and routers. Each IP core is connected to a router in the 2D layers. Routers in the same layer are connected with each other using horizontal links and in the different layers are connected by vertical links or up/down ports respectively.

The floor-planner will place functional units in different layers that compose the 3-D NoC architecture minimizing the peak temperature of each layer.The area of the chip is kept constant from the beginning of the optimization as we are not aiming towards reduction of chip area. The thermal results obtained by our floor-planner will be compared with the original peak temperature.

Our experiment work will concentrate on the analysis of the thermal optimization obtained by the floor-planner in three different scenarios. The first scenario resembles 3x3x3 3D Stacked Mesh NoC, where it is having 9 PEs and 9 corresponding Switches. The same structured is repeated in the other layers. IN case of 4 layer and 5 layer architecture, we have simply added one layer each having same structure .

On the other side, the genetic algorithms are configured with various parameters. The algorithm which is in charge of placing the components, crossover probability is set to 0.90 and the mutation probability is set to 1/ number of blocks as discussed in [5]. The algorithm which optimizes the placement of the blocks is configured with maximum pupulation of 100 and a maximum number of 300 generations. The probability of mutation is the inverse of available points in the plane. Then , cycle cross over is set with a probability of 0.9 and the tournament selection method following the guidelines given in [5]

The architectural parameters taken in our experiment is taken from [7] . As shown in Table 2, the inter-layer distance between two layers is 20um. Stacked Mesh combines multiple layers of two dimensional Mesh networks and connects

Table 2: Architectural parameters of 3D Stacked Mesh based NoC

Inter-layer wire-length(um)	20
Vertical bus wire-legth(um)	60
Horizontal bus wire-length(mm)	2.5
Port count	6 (+ bus arbitration)
Switch area (mm2)	0.1225
Switch static energy (pJ)	81.3
Longest wire-delay (ps)	219

Table 3: Physical properties of on-chip components

PARAMETER	VALUE
Die thickness	0.15 mm
Inter-layer material thickness	0.02 mm
Die capacitance	$1.63\times10^{6} J/(m^3 - K)$
Die resistance	0.076 mK/W
Inter-layer material resistance	0.25 mK/W
Ambient temperature	250 degree Kelvin

them vertically with a bus. This architecture is a suitable choice for adding layers in the z-direction because of the negligible vertical interconnect distance. Each router contains six number of ports : one to the bus, one to the IP and the remaining to the four cardinal directions. Wider buses [10] can also be added for higher bandwidth and cost effective communication among routers present in different layers.

Table 4: Average hop and link count for 3D stacked mesh with addition of layers

No. of layers	Links present	Average no. of hops
3	54	2.6666
4	75	2.75
5	96	2.8

5. EXPERIMENTAL RESULTS

This section presents the thermal results obtained in the scenarios described in Section 3. All the thermal values have been calculated using the thermal model described in Section 2. All the thermal diagrams presented in this section have been computed with MFA_{noc}. In addition to peak temperature (inK) , we have also calculated the number of hops with increase in number of layers and increase in number of links as these two parameters also determine increase in on-chip temperature. The values can be obtained from Table 4.

Fig 4 depicts the thermal distribution of our baseline scenario, where all processing elements are labeled as $IPid$ and routers as $SWid$, where id is an identifier. As can be seen in the figure, hot spots appear in the topmost layer as the heat conduction path is longer in topmost layers. This is an threat to the SoC. Similarly, figure 6 and figure 8depicts the original thermal diagram of 4 layer and 5 layer stack based 3-D NoC respectively. From the Table number 5,6 AND 7, it is clear that as we are adding more number of layers the peak temperature increases to a great extent.

Our floor-planner spaces heat sources abd heat sinks as far as possible, specially it tries to place all the blocks having high temperature value near the border of the chip, leading to cooling down of the chip. Vertical heat spread is also taken into account by our floor-planner , and each optimized layer is having a different layout. From the figures, it is clearly evident that our floor-planner places the routers in the lower layer in order to cool down the peak temperature. The reduction determines a more homogenous thermal distribution, which is translated into lower leakage current and reduced reliability risk.

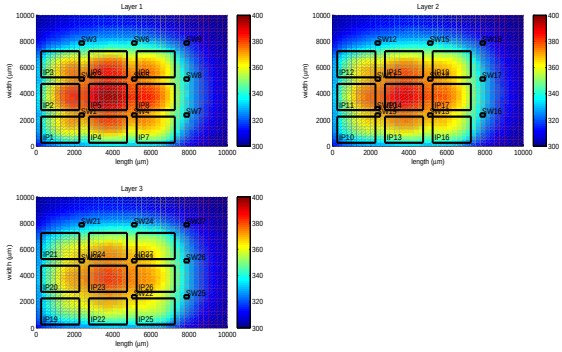

Figure 4: ***Thermal diagram of 3X3X3 3D stacked Mesh NoC***

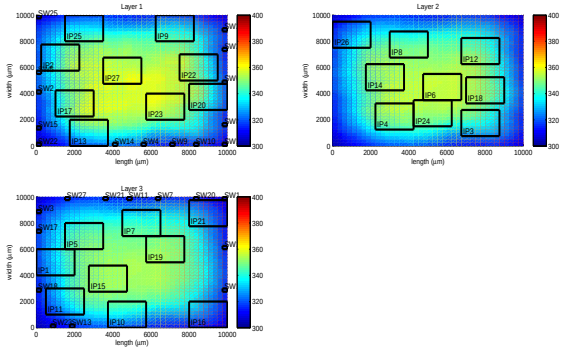

Figure 5: ***Optimized thermal diagram 3X3X3 3D stacked Mesh NoC after applying GA***

The Multi-objective Genetic floor-planner is able to reduce 27.1342K, 29.8688K and 32.6588K for layer0, layer1 and layer2 respectively. In case of layer3, the differences are better appreciated. Similarly, in case of 4 layer architecture, 32.3946K, 35.7064K, 39.6877Kand 44.3604K reduction of worst-case temperature is noticed for layer0, layer1, layer2 and layer3 respectively. The floor-planner is able to reduce more temperature in the higher layers resulting in cooler chips. The experimental results coincides with the optimized thermal diagrams obtained. For layer5, thermal reduction for layer0, layer1, layer2, layer3 and layer4 are 51.3943K, 56.O28K, 61.8737K, 69.1706K and 75.1354K respectively. The number of hops and number of links increases as we are adding more number of layers.

6. CONCLUSION

This work has presented a novel and efficient Multi-objective Genetic algorithm to cope with the problem of thermal-aware floor-planning in 3-D stacked Mesh NoC. Our floor-

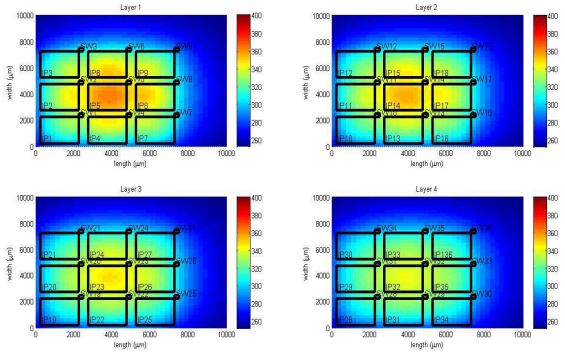

Figure 6: ***Thermal diagram of 3D stacked Mesh NoC***

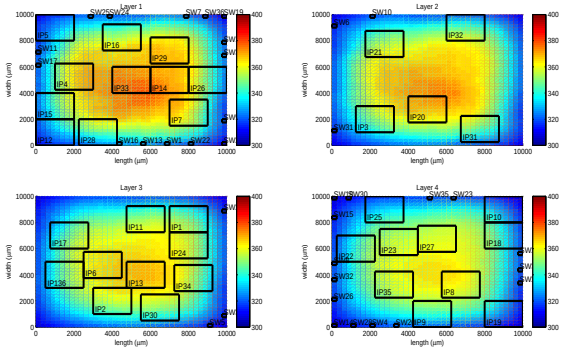

Figure 7: ***Optimized thermal diagram 4X4X4 3D stacked Mesh NoC after applying GA***

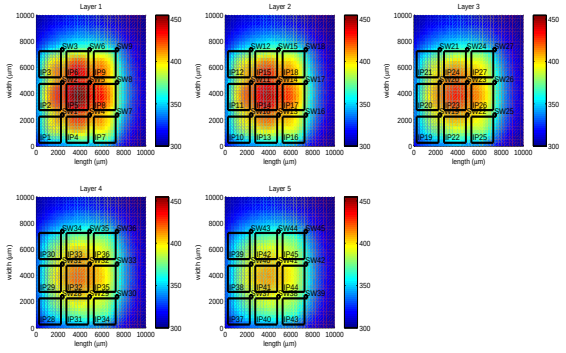

Figure 8: ***Thermal diagram of 5X5X5 3D stacked Mesh NoC***

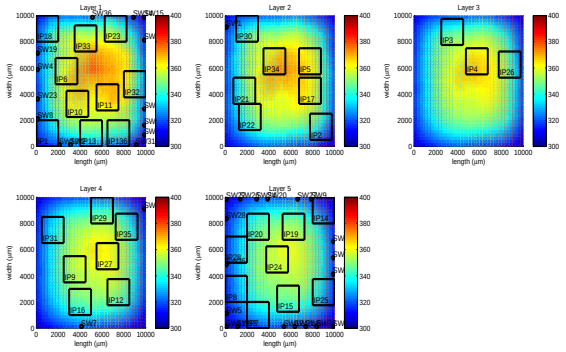

Figure 9: ***Optimized thermal diagram 5X5X5 3D stacked Mesh NoC after applying GA***

planner has resulted in optimization of the floor-plan and it is attached with an accurate thermal model, which has produced promising results in the minimization of peak temperature of different layers in case of 3 layer, 4 layer and 5 layer Mesh NoC. Our results outperform previous experimental results obtained by traditional thermal-aware floor-planners which use MILP ,Simulated Annealing or PSO.

7. REFERENCES

[1] C. Addo-Quaye. Thermal-aware mapping and placement for 3-d noc designs. In *SOC Conference, 2005. Proceedings. IEEE International*, pages 25–28. IEEE, 2005.

[2] A. Ahmed and A. Abdallah. Low-overhead routing algorithm for 3d network-on-chip. In *Networking and Computing (ICNC), 2012 Third International Conference on*, pages 23–32. IEEE, 2012.

[3] D. Chatterjee and T. Manikas. Power-density aware floorplanning for reducing maximum on-chip temperature. In *18th IASTED Int. Conf. on Modelling and Simulation (ICMS)*, pages 319–324, 2007.

[4] D. Cuesta, J. L. Risco-Martin, J. L. Ayala, and J. I. Hidalgo. 3d thermal-aware floorplanner using a moea approximation. *Integration, the VLSI Journal*, 46(1):10–21, 2013.

[5] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan. A fast and elitist multiobjective genetic algorithm: Nsga-ii. *Evolutionary Computation, IEEE Transactions on*, 6(2):182–197, 2002.

[6] J. ESIA and T. KSIA. Sia. international technology roadmap for semiconductors. Technical report, Technical report, ESIA and JEITIA and KSIA and TSIA and SIA, 2005. http://public. itrs. net, 2012.

[7] B. S. Feero and P. P. Pande. Networks-on-chip in a three-dimensional environment: A performance evaluation. *Computers, IEEE Transactions on*, 58(1):32–45, 2009.

[8] D. E. Goldberg and J. H. Holland. Genetic algorithms and machine learning. *Machine learning*, 3(2):95–99, 1988.

[9] P. K. Hamedani, S. Hessabi, H. Sarbazi-Azad, and N. E. Jerger. Exploration of temperature constraints for thermal aware mapping of 3d networks on chip. In *Parallel, Distributed and Network-Based Processing (PDP), 2012 20th Euromicro International Conference on*, pages 499–506. IEEE, 2012.

[10] P. Jacob, O. Erdogan, A. Zia, P. M. Belemjian, R. P. Kraft, and J. F. McDonald. Predicting the performance of a 3d processor-memory chip stack. *Design & Test of Computers, IEEE*, 22(6):540–547, 2005.

[11] K. Puttaswamy and G. H. Loh. Thermal herding: Microarchitecture techniques for controlling hotspots in high-performance 3d-integrated processors. In *High Performance Computer Architecture, 2007. HPCA*

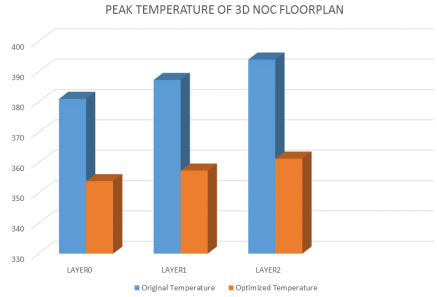

Figure 10: ***Original and optimized peak temperature of 3X3X3 layer 3D NoC***

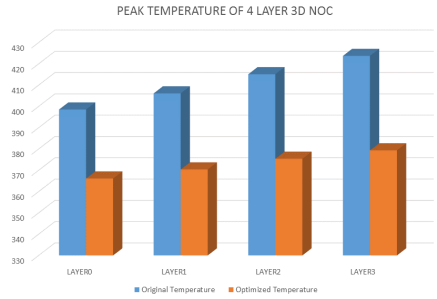

Figure 11: ***Original and optimized peak temperature of 4X4X4 layer 3D NoC***

2007. IEEE 13th International Symposium on, pages 193–204. IEEE, 2007.

[12] L. Shang, L.-S. Peh, A. Kumar, and N. K. Jha. Thermal modeling, characterization and management of on-chip networks. In *Proceedings of the 37th annual IEEE/ACM International Symposium on Microarchitecture*, pages 67–78. IEEE Computer Society, 2004.

[13] S. Sivanandam and S. Deepa. *Introduction to genetic algorithms*. Springer Science & Business Media, 2007.

[14] A. Sridhar, A. Vincenzi, M. Ruggiero, T. Brunschwiler, and D. Atienza. 3d-ice: Fast compact transient thermal modeling for 3d ics with inter-tier liquid cooling. In *Proceedings of the International Conference on Computer-Aided Design*, pages 463–470. IEEE Press, 2010.

[15] J. Wang, L. Li, H. Pan, S. He, and R. Zhang. Latency-aware mapping for 3d noc using rank-based multi-objective genetic algorithm. In *ASIC (ASICON), 2011 IEEE 9th International Conference on*, pages 413–416. IEEE, 2011.

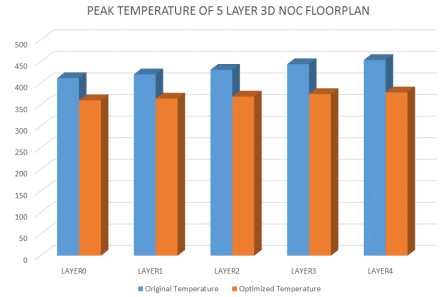

Figure 12: ***Original and optimized peak temperature of 5X5X5 layer 3D NoC***

Table 5: Peak and optimized temperature of 4 layer floorplan

Layer Number	Temp. of original floorplan	Temp. of optimized floorplan	Reduction of Temp.
Layer0	398.5079	366.1133	32.3946
Layer1	406.0838	370.3774	35.7064
Layer2	415.0838	375.3961	39.6877
Layer3	423.6995	379.3391	44.3604

Table 6: Peak and optimized temperature of 5 layer floorplan

Layer Number	Temp. of original loorplan	Temp. of optimized floorplan	Reduction of Temp.
Layer0	411.5411	360.1468	51.3943
Layer1	420.2010	364.1730	56.028
Layer2	430.7982	368.9245	61.8737
Layer3	443.8383	374.6677	69.1706
Layer4	453.5233	378.3879	75.1354

Table 7: Peak and optimized temperature of 3 layer floorplan

Layer Number	Temp. of original floorplan	Temp. of optimized floorplan	Reduction of Temp.
Layer0	381.013	353.8788	27.1342
Layer1	387.2342	357.3654	29.8688
Layer2	393.9507	361.2919	32.6588