Malaysian Simulation SOCIETY

http://www.ams-mss.org

APPLICATIONS OF
MODELLING AND SIMULATION

eISSN 2600-8084

VOL 1, NO. 1, 2017, 36-41

Runtime Network-on-Chip Thermal and Power Balancing

M. S. Rusli*, M. N. Marsono and N. S. Husin

[1]Faculty of Electrical Engineering, Universiti Teknologi Malaysia, 81310 UTM Skudai, Johor, Malaysia.

*Corresponding author: shahrizal@fke.utm.my

Abstract: In Network-on-Chip (NoC), most thermal and peak power balancing methods are monitored by centralized power/thermal managers. This increases inter-core communication latency and imbalanced thermal distribution. These factors directly affect the hot spot formation caused by high power densities developed with increasing per-core transistor number. As a result, reliability decreases along with static power dissipation. This proposal aims to introduce hierarchical agents for balancing power and thermal distribution by manipulating system's parameters such as power, thermal, voltage below the optimal values. As some level of control is applied, this proposal also targets to achieve network scalability by implementing some level of independencies; self-organize and self-optimize distributed agent in overcoming core-level homogeneous processing element (PE) thermal and power variations at runtime. The aims of this work are to significantly contribute to achieving runtime thermal and power balancing, power and thermal management and reducing thermal hot spot formation in NoC.

Keywords: Network-on-Chip (NoC); thermal balancing; power and thermal management; homogeneous MPSoC; runtime

1. INTRODUCTION

According to Semiconductor Industry Association (SIA) [1], future deep submicron (DSM) technologies, operating frequency, transistor density and design complexity of multiprocessor system-on-chip (MPSoC) will continue to increase making latency, power and temperature dissipation a major concern. High power density per-core resulted from increasing number of transistors contributes to hot spot formation at respective cores which accelerates static power dissipation and mean-time-to-failure (MTTF) problem in System-on-Chip (SoC) leading to reliability issues [2]. Interconnect delay increases by 5% for every 10°C core temperature [3] rise. Some NoC prototypes significantly show NoCs taking substantial portion of system power e.g ~40% in RAW chip [4] and ~30% in Intel 80 core teraflop chip [5].

Related current solution approaches employed centralized and localized monitoring method and to concern on modular parameters at entire NoC chip without concerning capability of respective PEs (homogeneous vs heterogeneous) at runtime. The techniques applied will be discussed in the next section of this proposal. Different PEs generate different thermal and peak power consumption properties. A block of MPSoC consists of either multiple heterogeneous or homogeneous PEs for application computation. Homogeneous PEs have an average power and thermal characteristics compared to unique variations in heterogeneous PEs due to different PE types and processing modules. Hence, a distributed monitoring unit (agent) is needed to strategically provide some level of autonomous decisions in controlling PEs parameter variations.

An agent is an intelligent and independent identity which controls the communication activities in the part or the whole system. A hierarchical agent based method is known to be among the best method to allow hierarchical control independences in large scale NoC. By allowing low level agents to implement cell-level parametric control, algorithm complexity is decreased and can reduce communication latency of the entire chip. Based on these facts and evaluation, a hierarchical/distributed control strategies for thermal and power management is proposed.

The goal of this paper is to propose a runtime NoC thermal and power balancing by employing hierarchical agent strategy. The objectives of this research are to propose hierarchical agents for runtime NoC thermal and power parameter balancing and to enable the network to self-organize and self-optimize its own thermal distribution and power consumption to achieve scalability.

2. RELATED WORK

Actual chip power consumption is the summation of dynamic power and static power (leakage current). Static power is associated with leakage current at component level. Dynamic power can be controlled by manipulating voltage [6] and frequency [7][8] of the system. Most researches in centralized and distributed monitoring of both thermal and power consumption are preferably runtime based rather than design-time based.

Centralized monitoring is able to coordinate and balance function among resources as the manager can fully observe the chip's resource properties. At centralized monitoring unit (CMU), core parameters such as temperature, power and latency are optimized for all components in a system. Researches in centralized low power and thermal monitoring agent NoC are links shut down based on power usage [7] and distributed clock frequencies based on globally asynchronous locally synchronous (GALS) concept [8][9]. Besides that, adaptive routing is implemented where it is a packet routing scheme to avoid congestion during its transmission in heavy on-chip traffic. Examples of its implementation are power utilization of NoC component based adaptive routing scheme [10] and topology- and deadlock free, strongly connected component - based routing scheme [11]. NoC core placement problem has also been studied by [12][13].

Dynamic voltage scaling (DVS) or dynamic voltage and frequency scaling (DVFS) techniques are proposed by [6], [8] and [14]. Adjustment of voltage and/or frequency at processor (genetic algorithm) is applied depending on the power utilization information of network components i.e routers and network adapters to reduce overall energy consumption. Enhancement in centralized monitoring agent is proposed by Dennis et al [15] by implementing Diagnostic Adaptivity Processing (DAP) at geographical centre of a system. The agent is able to vary frequency and voltage of cores based on demand and conduct offline parametric core test. However, this technique increase communication latency in time-division method (TDM) caused by flits buffering. Real-time algorithm complexity also increases with the increase in simultaneous parametric monitoring activity.

Distributed or hierarchical agent however, is able to overcome communication latency problem and balance load at resource level. It also reduces algorithm complexity at CMU, hence parametric control can be hierarchically distributed to a few levels. This design is inspired by the biological coordination of the human body, whereby the nerve system consists of the central nervous system (the brain and the spinal cord) and the peripheral nervous system that is scattered throughout the body. The central nervous system controls all activities in the peripheral nervous system and decides the action taken by the nerve cells. Distributed monitoring reduces latency in respective cell and interconnects. [16] proposes hierarchical global (GA) and cluster agents (CA). This heuristic approach invokes handshaking between agents for packet transmission. [17] improves the previous design by embedding thermal sensor in another layer of agent, tile agent (TA) in each cell unit. [18] introduces hardware based and software based thermal agents to tackle hot spot formation in agent units. This approach is different with what this proposal aims at doing in the sense that, we are giving some level of independencies to the low level hierarchy in making its decision to self-organize and self-optimize its own parametric variations.

3. PROPOSED ARCHITECTURE

In this section, we propose a hierarchical agent for NoC platform using software implementation. The proposed architecture is a 8x8 mesh topology NoC and is scalable for improvement. Each tile consists of a router, R, a network interface, N and processing element/block, P. Figure 1 shows an example of common 4x4 mesh topology NoC.

We propose a three-level distributed agent architecture; tile agent, cluster agent and global agent. Each tile contains a tile agent - the smallest functional unit that monitors its own communication unit. It monitors properties of the tile i.e power and temperature of the component and reports to the cluster agent each time requested to execute task and when performance problems arise.

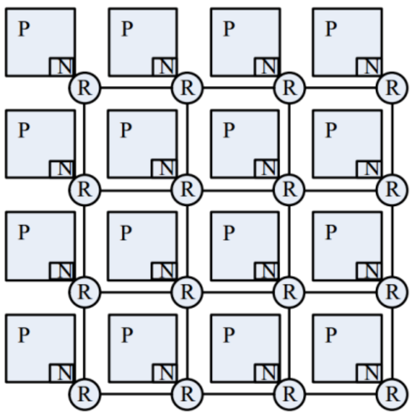


Figure 1. Common 4×4 mesh topology [19]

Cluster agent is placed at a strategic geographical radius of its cluster and monitors power and temperature reports of a specific group number of tiles as shown in Figure 2. It is able to identify the number of tiles needed for task computation assigned to it. It also identifies suitable tiles to be grouped into its cluster and computes task assigned by global agent. Besides, it forecasts runtime thermal energy consumption pattern of tiles under its cluster for dynamic thermal management and reports the state to global agent.

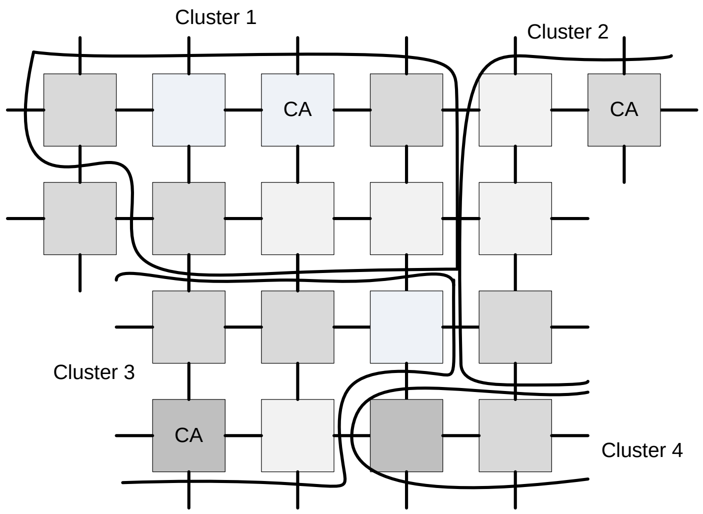


Figure 2. Cluster regions formation on NoC platform

Global agent identifies suitable cluster to perform tasks based on communication tasks graph (CTG), monitors power consumption and thermal dissipation of clusters. It collaborates with cluster agent to perform power- and thermal-aware workload balancing when requested by cluster agent. The decision output from this agent is transmitted to the cluster agent and tile agent. If the reconfiguration fails, the global agent will inform the requested client to wait until the resources are ready to execute new request.

The thermal pattern generated by cluster agent will be manipulated by global agent to determine thermal imbalance in each cluster. By comparing the pattern with predefined thermal threshold values, a thermal management optimization algorithm is computed to modify currently executing cluster's properties such as operating power, frequency, CTG remapping or reclustering. This is done if and only if the forecast value is less than the threshold value. The results of this algorithm is used for decision making process whether to reconfigure part or whole of the system. The decision involves whether the cluster needs to recluster, migrate tasks or global agent reconfigures the whole system The process is continuously done until task execution finishes. If the temperature pattern value is less than the threshold value, the system waits until new modification request appears. The flowchart of the proposed architecture is shown in Figure 3.

4. METHODOLOGY

This research is divided into two stages; Firstly, to study the existing hierarchical agent thermal and power in NoC model, issues and application. Secondly is the implementation in term of software codes which is further subdivided into five stages. Below are the stages of the research plan.

Firstly, assuming CTG and task scheduling have already been assigned, we will only consider software implementation of agents beginning from this part. The platform will be developed using Access Noxim simulator which combines Noxim and HotSpot simulator for real-time power and thermal measurement. We will implement the most commonly used NoC mesh topology with random routing traffic applied on the 8x8 tile NoC processors. The simulation is application specific on a Reed-Solomon code encoder with codeword format RS(32,28,8). The tool involved will be Access Noxim software, an improved version of Noxim simulator. It is developed using SystemC, a system description language based on C++. This software comes in a package with thermal simulator, Hotspot. Power and thermal energy dissipation at each interval during NoC simulation can be determined at runtime. The two properties will be used for thermal management decision making process in the next phase. Reed-Solomon code encoder traffic pattern will be applied to the simulation as real application example to benchmark with other NoC software. Agent module will be implemented for energy acquisition and thermal/power management process. However, other simulators are still under consideration.

Second methodology is the design of thermal pattern generation inculcating power profile information gathered from the platform. It will be computed based on identified NoC thermal model. Thirdly, an optimization algorithm will be designed to achieve power and thermal balancing for requesting cluster. Based on the input from previous stage, the global agent will compute necessary measurements before commending output that yields better thermal balancing techniques to be implemented by its low-level agents. Type of output will be determined later as the research continues.

Next, the decision making process is the most important part of this research. The output from the previous stage will be used by cluster agent to decide whether to accept the suggestion or continue running its current state besides depending on other parameters. Final decision made by cluster agent will be employed by tile agent in its region. Finally, analysis of results will be done with other 2D NoC platform for validation and verification.

Based on predefined NoC properties on cycle-accurate Access Noxim, thermal and power consumption values at NoC tile level can be collected at particular intervals. The simulator collects both thermal and power properties of router at each NoC tile during runtime. During runtime, change of workload affects both power and thermal dissipation at each tile. Access

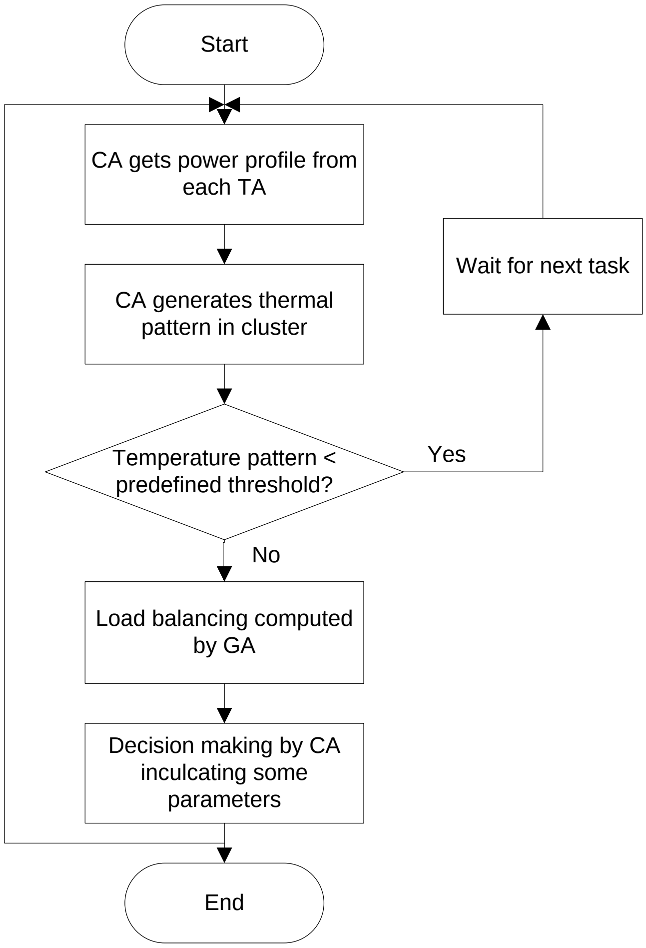


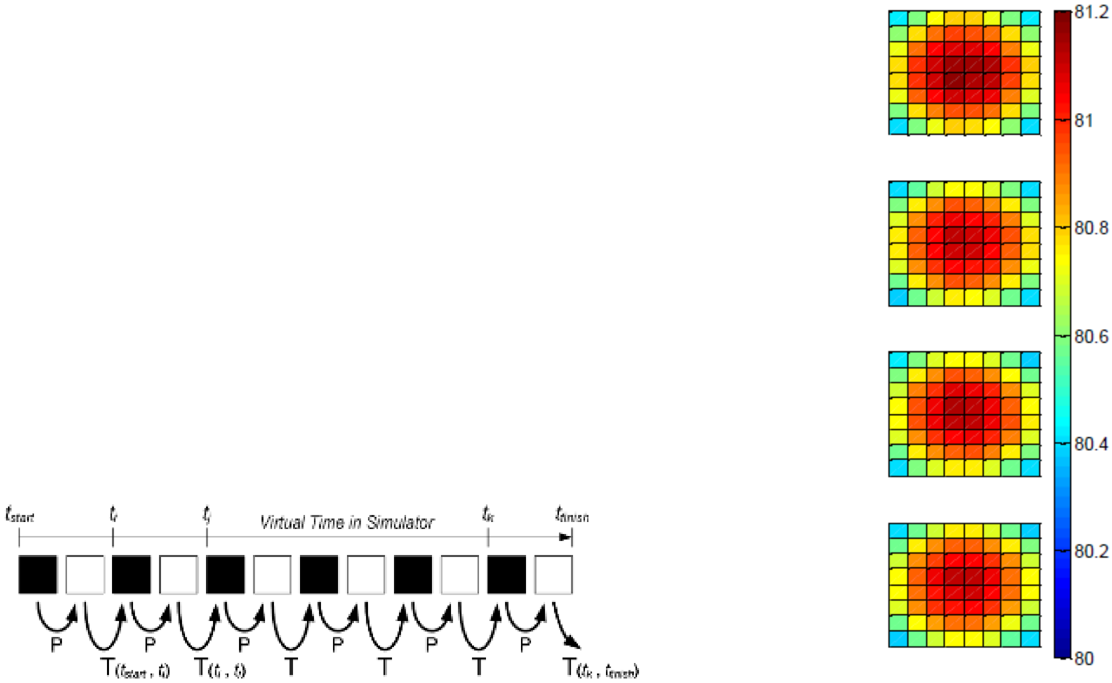
Figure 3. Flowchart of proposed architecture

Noxim determines temperature change in discrete time of a specified predefined interval to determine thermal energy consumption (Figure 4(a)). Using chip floorplan, power trace at specific interval and its discrete time, thermal energy, T profile is generated for thermal management strategy [20]. From the figure, the black block represents network traffic simulation whereas white block represents chip thermal simulation in the simulator. Figure 4(b) shows the router temperature profile of 4x4 tile NoC at 4 different simulation intervals by the simulator. The figure shows the growing imbalance from initially lower temperature variations among cells to a higher temperature imbalance core cells.

5. CONCLUSION

Preliminary simulation of thermal profile balancing has been established in this paper. Further research on achieving an efficient integration between agent based architecture and hierarchical monitoring agents will be conducted. The expected contributions for this research are; 1) novel contribution to balance thermal and dynamic power distribution in NoC by introducing hierarchical agent monitoring activity at runtime, thus reduces communication latency and hot spot formation at CMU. 2) achievement of network scalability by implementing self-organize and self-optimize technique for thermal distribution and power consumption at run- time thus allowing controlled autonomous decision at lower level agent to monitor its own activity (parametric monitoring activity).

The proposed design is expected to achieve significant improvement in thermal and power balancing design in NoC. It can also significantly self- optimize its thermal and power consumption via hierarchical agent-based technique as compared to centralized monitoring approach.



(a) (b)

Figure 4. **(a)** Thermal energy computation by AccessNoxim at discrete time [1] **(b)** Router thermal profile at four different intervals plotted using Matlab [1]

REFERENCES

[1] M. B. Taylor et al., The RAW microprocessor: A computational fabric for software circuits and general-purpose programs, *IEEE Micro*, vol. 22, no. 22, pp. 145-162, Feb. 2005.

[2] S. Pasricha, N. Dutt, On-chip communication architectures: System on chip interconnect, Morgan Kaufmann, April 2008.

[3] M. Daneshtalab, A. Sobhani, A. Afzali-Kusha, O. Fatemi, and Z. Navabi, NoC hot spot minimization using AntNet dynamic routing algorithm, in *IEEE International Conference on Application-specific Systems, Architectures and Processors ASAP '06*, 2006.

[4] S. Vangal et al., An 80-tile 1.28TFLOPS network-on-chip in 65nm CMOS, in *Proceeding of Solid-State Circuits Conference*, pp. 98-589, Feb. 2007.

[5] E. J. Kim et al., “Energy optimization techniques in cluster interconnects,” in *Proceeding of International Symposium of Low Power Electronic Design.*, pp. 459–464, August 2003.

[6] L. Shang, L.-S. Peh, and N. K. Jha, Power-efficient interconnection networks: Dynamic voltage scaling with links, in *IEEE Computer Architecture Letters*, vol. 1 no. 1, p. 6-6, January 2002.

[7] E. Beigne, F. Clermidy, S. Miermont, and P. Vivet, Dynamic voltage and frequency scaling architecture for units integration within a GALS NoC, in *Proceeding of International Symposium of Network-on-Chip*, pp. 129-138, 2008.

[8] U. Y. Ogras, R. Marculescu, D. Marculescu, and E. G. Jung, Design and management of voltage-frequency island partitioned networks-on-chip, *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 17, no. 3 pp.330–341, March 2009.

[9] S.-G. Yang, L. Li, Y.-A. Zhang, B. Zhang, and Y. Xu, A Power-aware adaptive routing scheme for network on a chip, *7th International Conference on ASIC, 2007. ASICON '07.*, pp. 1301–1304, Oct. 2007.

[10] Z. ZhuanSun, K. Li, and Y. Shen, An efficient adaptive routing algorithm for application-specific network-on-chip, *3rd International Symposium on Parallel Architectures, Algorithms and Programming (PAAP)*, pp. 333–338, Dec. 2010.

[11] C.C.N. Chu, and D.F. Wong, A matrix synthesis approach to thermal placement, in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 17 ,no. 11 pp. 1166–1174, Nov 1998.

[12] M. D. Osterman, and M. Pecht, Placement for reliability and routability of convectively cooled PWBs, in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 9, no. 7 pp. 734–744, Jul 1990.

[13] G. Chen, and S. Sapatnekar, Partition-driven standard cell thermal placement, *Proceeding of ISPD*, California, 2003.

[14] D. Sylvester, D. Blaauw, and E. Karl, ElastIC: An adaptive self-healing architecture for unpredictable silicon, in *IEEE Design & Test of Computers*, vol. 23, no. 6, pp. 484–490, June 2006.

[15] M.A. Al-Faruque, R. Krist, and J. Henkel, ADAM: Runtime agent-based distributed application mapping for on-chip communication, *45th ACM/IEEE Design Automation Conference*, pp. 760–765, 2008.

[16] M.A. Al-Faruque, J. Jahn, T. Ebi, and J. Henkel, Runtime thermal management using software agents for multi- and many-core architectures, in *IEEE Design & Test of Computers*, vol. 27, no. 6, pp. 58 – 68, Dec. 2010.

[17] T. Ebi, M. Faruque, and J. Henkel, TAPE: Thermal-aware agent-based power economy multi/many-core architectures, IEEE/ACM International Conference on Computer-Aided Design - Digest of Technical Papers, pp. 302–309, Nov. 2009.

[18] M. Nickray, M. Dehyadgari, and A. Afzali-kusha, Adaptive routing using context-aware agents for networks on chips, *4th IEEE International Design and Test Workshop (IDT)*, pp. 1 - 6, Nov. 2009.

[19] W. Liu, J. Xu, X. Wu, Y. Ye, X. Wang, W. Zhang, M. Nikdast, and Z. Wang, A noc traffic suite based on real applications, in *VLSI (ISVLSI), 2011 IEEE Computer Society Annual Symposium on,* July 2011, pp. 66 – 71.

[20] K.-Y. Jheng, C.-H. Chao, H.-Y. Wang, and A.-Y. Wu, Traffic thermal mutual coupling co-simulation platform for three-dimensional Network-on-Chip, in *Proceeding of International Symposium on VLSI Design Automation and Test (VLSI-DAT)*, pp.135-138, April 2010.