Voltage and Frequency Island Optimizations for Many-Core/Networks-on-Chip Designs

Wooyoung Jang, Duo Ding, David Z. Pan
Dept. of Electrical and Computer Engineering
The University of Texas at Austin
Austin, Texas 78712
{wyjang, ding}@cerc.utexas.edu, dpan@ece.utexas.edu

***Abstract*— Many-core chips interconnected by networks-on-chip (NoC) are increasingly challenged by the tight power consumption constraints. The concept of voltage and frequency island (VFI) which has been recently introduced for achieving fine-grain core-level power management fits well with an NoC design style. This paper will discuss some recent advancement of VFI optimizations for many-core/NoC designs. We will also discuss other research challenges for low-power many-core/NoC designs from an electronic system level (ESL) perspective.**

I. INTRODUCTION

The continuing advancements in the semiconductor technology enable to integrate many cores on a single die. As hundreds to thousands of cores are integrated for parallel computing, efficient on-chip communication among cores and economic power consumption strategies become key factors in the success of system-on-chip (SoC) design. Recently, networks-on-chip (NoC) based on a modular packet-switching mechanism has been proposed as a promising solution for on-chip communication [1, 2]. In NoC, a core is interconnected to a network router and communicates with different other cores through the network router. Indeed, as a better SoC platform for scalable system integration, NoC provides more competitive features than the previous dedicated point-to-point and shared bus-based architectures.

On the other hand, NoC in itself does not provide low power consumption for modern and future SoC designs. It may consume more communication energy since packets detour or pass more routers. Moreover, it is inefficient for all cores on a network to operate at a single supply/threshold voltage level and a single clock speed because all cores are not busy. Fortunately, NoC fits well with the concept of voltage-frequency island (VFI) design [3]. VFI enables fine-grained core-level power optimization by utilizing a unique supply/threshold voltage and operating clock for each island. Consequently, significant improvements in energy-efficiency are enabled by fine-grained dynamic voltage/frequency scaling (DVFS) whereas the performance cost of using the VFI organization is negligible. This is possible because static complementary metal–oxide–semiconductor (CMOS) logic used in the vast majority of current processors has a voltage-dependent maximum operating frequency. Thus, when used at a reduced frequency, the processors can operate at a lower supply/higher threshold voltage. The power is supplied by an off- or on-chip source and can be controlled independently for each VFI. This may be achieved by using either on-chip voltage regulators or multiple power grids. Therefore, the use of VFIs in the NoC design provides better performance-power tradeoffs than its single voltage and clock. A lot of modern processors such as Intel's XScale [4], AMD's Athlon [5], IBM's CU-08, -45HP and -65HP [6] and Transmeta's Crusoe [7] are employed with the VFI and DVFS functionality.

VFI based NoC is good to use globally asynchronous locally synchronous (GALS) design paradigm since each VFI is locally synchronous whereas it is globally asynchronous with other VFIs [8, 9]. The communication across different VFIs is achieved through a mixed clock first input, first output (MCFIFO) buffer and a voltage level converter (VLC) [10, 11]. They provide the flexibility to scale the frequency and voltage of various VFIs in order to minimize power consumption. Therefore, the GALS design makes the clock distribution and timing closure problems more manageable in VFI based NoC.

In this paper, we survey VFI optimizations for many-core/NoC designs. In the many-core/NoC designs, many cores/NoC tiles cannot be controlled independently and all of the links cannot be equipped with an MCFIFO buffer and VLC. Therefore, it is essential to partition many cores/tiles to several VFIs and then to control each VFI with fewer MCFIFO buffer and VLC. In Section II, we present various DVFS schemes. VFI optimizations for many-core/NoC designs are introduced in Section III. In Section IV, we survey temperature/process variation-aware VFI/DVFS optimization. In Section V, on-chip nanophotonic interconnect is briefly discussed as a promising technology to extend current VFI based NoC design paradigm. Finally, we draw the conclusion and point out some research direction in Section VI.

II. DYNAMIC VOLTAGE/FREQUENCY SCALING

In most VFI generation algorithms [22-26], it is assumed that the optimum supply/threshold voltage level and clock speed of each core for executing any tasks are taken as inputs. DVFS is a popular technique to find the supply/threshold voltage level and clock speed. In addition, DVFS can be used

This work is sponsored in part by Texas ARP and Samsung Electronics.

to look for better supply/threshold voltage level and clock speed of the VFIs already composed [24]. DVFS-related works may be broadly divided into two categories.

The first category of techniques targets the chip-level DVFS [12-18] which is commonly used in commercial products [4-7]. In [12], DVFS is performed by the operating system (OS) at a task level to minimize power consumption while satisfying hard timing constraints of each task. It is assumed that task arrival times, workload and deadlines are known in advance. In [13, 14], either application or compiler support is required for performing DVFS. In [13], a software tool automatically converts a DVFS-unaware program into an equivalent low-energy program and the worst-case execution time is used to choose a supply voltage level and clock frequency. In [14], an intra-task DVFS technique using program checkpoints is proposed. Checkpoints are generated at compile time and indicate places in the code where the clock speed and voltage level should be recalculated. System-level DVFS techniques proposed in [15-18] do not make any assumption about applications or any support from compiler. The approaches mainly depend on runtime statistics made available by the platform or micro-architecture to model the task behavior. In [15-16], a microarchitecture-driven DVFS technique is proposed, where cache miss and instruction rate of a program execution drives the voltage scaling. However, the results in these two works are based on simulations. In [17], dynamic runtime statistics such as cache hit/miss ratio, instruction rate and memory access counts obtained from a performance monitoring unit are used to determine the appropriate voltage setting. In [18], an extremely lightweight DVFS technique targeted towards multi-tasking systems are proposed. The technique adapts runtime statistic and an online learning algorithm to estimate the best suited voltage level and clock speed setting at any given point in time.

The second category targets the per-chip DVFS [19-21]. Isci et al. [19] introduce a global power manager providing a hierarchical feedback-control mechanism to sense the per-core power at periodic intervals. Then, it sets the operating power level or mode of each core to enforce adherence to known chip-level power budgets. A fast static power management analysis tool is developed to evaluate the benefits and performance costs of various multi-core mode-management policies. In addition, they evaluate several different policies for global chip multiprocessor (CMP) power management under different objectives such as prioritization, fairness and optimized chip-level throughput. Juang et al. [20] illustrate how the use of local, per-tile DVFS techniques can result in tiles counteracting each others' power management policies, significantly hurting chip power-performance. Then, they propose a coordinated DVFS scheme for CMPs, which eliminates the oscillations and ensures efficient and resilient DVFS control. Specifically, their technique incorporates thread information collected at runtime across the chip. In addition, by extending a control-theoretic local DVFS control technique toward DVFS for CMPs, their technique prescribes DVFS settings formally at each tile, thus ensuring stable, distributed, coordinated DVFS control of a CMP. Herbert et al. [21] examine the tradeoffs involved in the choice of both DVFS control scheme and method by which the processor is partitioned into VFIs. They simulate real multithreaded commercial and scientific workloads, demonstrating the large real-world potential of DVFS for CMPs. They also show that the benefits of per-core DVFS are not necessarily large enough to overcome the complexity of having many independent DVFSs per chip.

III. VFI Optimizations for Many-Core/NoC Desings

VFI is a group of cores supplied by the same voltage level and clock speed, independently from the chip-level voltage level and clock speed. The use of VFIs permits operating different portions of the design at different voltage levels with different clock speeds to minimize the overall chip power consumption. For many-core/NoC designs, the VFI enables core-level power optimization by utilizing a voltage level and a clock speed that are unique from the rest of the design.

On the other hand, VFI causes the chip design process severely to be complicated with respect to static timing, power routing and clock tree. In particular, the design complexity grows significantly with the number of allowed VFIs as shown in Fig. 1. Since each VFI requires its own power grid, clock tree, MCFIFO buffer and VLC in order to communicate with other VFIs, those design overheads with respect to area, delay and energy are not avoidable. Therefore, many-core/NoC designs employing the concept of VFI design needs to cluster cores supplied by the same voltage level and clock speed and ensure that the created grouping does not violate other design constraints such as performance, timing and wiring congestion [31]. In [22-25], many researchers have focused on partitioning many cores/tiles to allowed VFIs and assigning supply/threshold voltage level and clock speed under performance constraints.

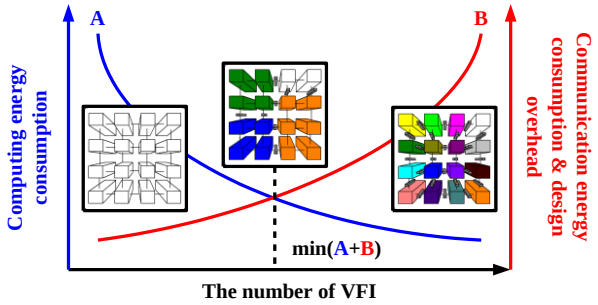


Figure 1. Computing and communication energy consumption and design overhead according to the number of VFIs [24]

There are many existing works that address the problem of VFIs generation for core-based SoC designs [22-23]. Hu et al. [22] partitions cores into several VFIs and floorplans both at a chip and island level. By using a graph-based representation, the partitioning and floorplanning steps are modeled in an integrated fashion. A simulated annealing-based algorithm is applied to give reasonable solutions with respect to power savings, area overhead and the number of voltage islands used in the design. Wu et al. [23] consider trade-off between power and design cost under timing requirement for a VFI generation problem. This problem is formulated as a voltage-partitioning problem which is NP-hard and then is solved by a two-step heuristic algorithm which combines dynamic programming with variable-size gridding.

As many cores have been recently interconnected by NoC, the concept of VFI design is employed in the NoC [24-26]. Ogras et al. [24] propose a design methodology for

partitioning a given NoC architecture into multiple VFIs and assigning the supply/threshold voltages and the corresponding clock speeds to each domain such that the total power consumption is minimized under given performance constraints. The basic idea employing a greedy algorithm is to start with an NoC configuration where each tile belongs to a different VFI characterized by a given supply/threshold voltage and local clock speed. Then, their approach finds two candidate VFIs to merge such that the decrease in the energy consumption is the largest among all possible merges, while performance constraints are still satisfied. This process is repeated until a single VFI implementation is obtained. Finally, among all VFI partitioning, the one providing the minimum energy consumption is selected as being the solution of the VFI partitioning problem. Leung et al. [25] also work on the design of NoC with voltage islands. The approach simultaneously solves three problems, i.e. tile mapping to determine the corresponding tile that each processor should be mapped to, routing path allocation to obtain the legal routing path for each communication transaction, physical voltage island generation and voltage assignment for each voltage island and communication link.

Jang et al. [26] propose a systematic VFI-aware energy optimization framework as shown in Fig. 2. In most previous works, VFI partitioning and voltage/frequency assignment algorithms are performed after mapping (or floorplanning) and allocating routing paths as shown in Fig. 3(a) [22-25]. As a result, all cores using the same voltage level and clock speed cannot be merged to a single VFI as shown in Fig. 3(b). Therefore, their solution spaces are severely limited since any cores runs at higher voltage level. On the other hand, in [26], VFI-aware partitioning is carried out with VFI assignment in advance. Then, mapping and routing path allocation are modeled and executed by a VFI-aware manner. The VFI-aware mapping is performed based on a region growing method where a core is not separated from the VFI using the same voltage and clock. In addition, the VFI-aware routing path allocation seeks to further reduce VFI overheads such as an MCFIFO buffer and a VLC by changing NoC topologies. Consequently, all cores using the same voltage level and clock speed are clustered to a single VFI as shown in Fig. 3 (c). Moreover, unique interconnection is built for key traffic paths between VFIs to minimize the overhead of VFI.

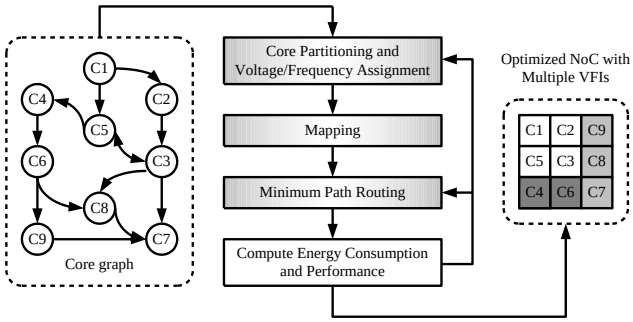


Figure 2. VFI-aware NoC methodology [26]

IV. Temperature/Variation-Aware VFI and DVFS

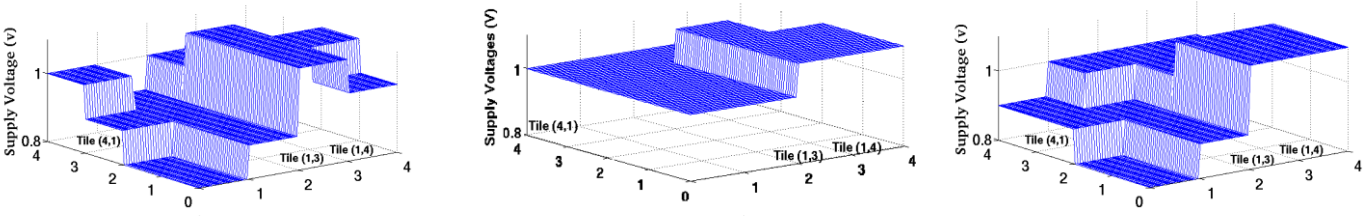
Temperature/process variations are an important parameter to be taken into consideration at voltage/clock selection. The variations can make VFI based many-core/NoC designs fail to satisfy clock speeds or power budgets. Therefore, temperature/process variations-aware VFI/DVFS can achieve more power/performance improvements.

Temperature has a critical impact on circuit delay and, implicitly, on frequency, mainly through its influence on carrier mobility and threshold voltage. Thus, a clock frequency relies on not only the voltage level but also the temperature. Hung et al. [27] present a hybrid optimization approach which targets peak temperature reduction and elimination of hot spots. They demonstrate that by carefully partitioning cores, assigning voltage levels, and allocating voltage islands, the thermal distribution of design can be improved. Bao et al. [28] propose an on-line temperature aware DVFS technique exploiting both static and dynamic slack and takes into consideration the frequency-temperature dependency. The approach consists of two parts: an off-line temperature aware optimization step and on-line voltage and frequency setting based on temperature sensor readings.

Exposing variability information to DVFS and VFI significantly improves energy-efficiency. Herbert et al. [29] propose a method of expressing the variations within a single voltage/frequency island by means of an effective process variation parameter was proposed. The benefits of exposing this variability information to existing DVFS control algorithms were demonstrated for two hardware DVFS controllers: a simple threshold-based controller and a complex greedy-search controller. Garg et al. [30] propose a theoretical framework to efficiently obtain the limits on the controllability and performance of DVFS controllers for multiple-VFI based NoC. They present results, using both real and synthetic benchmarks that explore the impact on temperature and reliability constraints, maximum inductive noise constraints and process variations.

V. Nanophotonic Interconnection for VFI

Electrical NoC provides great on-chip communication: however, it brings no true relief to power budget when the on-chip network scales in terms of complexity/size and signal bandwidth. Based on recent opto-electro material/device level break-throughs [32], on-chip nanophotonics offers compelling



(a) VFI map after mapping task (b) VFI partitioning and voltage and frequency assignment [24] (c) VFI-Aware NoC design [26]

Figure 3. VFI based NoC design on 4x4 NoC

high throughput/bandwidth communication and promising low power integration opportunities compared with traditional Cu/low-K interconnect, therefore is considered as a potential quantum leap towards the next generation on-chip interconnect [33].

In spite of its superior signaling speed and low power potentials, the on-chip nanophotonics faces major roadblocks for interconnecting on-chip computation resources. Major challenges in this field include but are not limited to: photonic network architecture design, low power high performance integration, device characterization and thermal reliability modeling/optimization. In [34], a low power design for a photonic network-on-chip architecture is proposed and analyzed. In [35], a CAD flow for on-chip optical waveguide planning is proposed in the presence of an optical interconnect library, which considers physical level device constraints such as loss figure, power and foot-print. Similarly, in [36] a physical-layer device library of key nanophotonic devices is build, which is then implemented into several key network components and applied to large scale on-chip network performance evaluations.

Under the considerations of the promising low power on-chip optical links, VFI based NoC optimization flows can be extended for further power efficiency. Other important future works along this direction may include thermal reliability optimizations for on-chip nanophotonic links and potential applications utilizing wavelength division multiplexing techniques.

VI. CONCLUSION AND FUTURE DIRECTION

VFI are essential to minimizing power consumption for many-core/NoC designs. In this paper, we show several DVFS control schemes and VFI-aware designs. Furthermore, we present temperature/process variation-aware VFI optimizations since the variation has an important impact on VFI performance and power budgets. We also extend VFI to nanophotonic NoC which need to be more researched for lower power consumption and higher performance.

Until now, VFI based many-core/NoC designs for general-purpose computing have not been comprehensively studied. One of the reasons is that optimizing the VFI overheads for all different applications is extremely challenging. There are still many research opportunities and open problems in the VFI based many-core/NoC designs for general-purpose computing.

REFERENCES

[1] Luca Benini and Giovanni De Micheli, "Network on chips: a new SoC paradigm," Computer, 2002, 35, pp. 70-78.

[2] William J. Dally and Brian Towles, "Route packets, not wires: on-chip interconnection networks," in Proc. Design Automation Confence, 2001.

[3] D. E. Lackey et al., "Managing power and performance for system-on-chip designs using voltage islands," in Proc. International Conference on Computer-Aided Design, 2002

[4] Intel, "Intel XScale core," http://www.intel.com.

[5] AMD, "Mobile AMD Athon 4 processor model 6," http://www. amd.com.

[6] IBM, "IBM ASIC Solutions," http://www.ibm.com.

[7] M. Fleischmann, "Longrun power management – dynamic power management for Crusoe processors," http://www.transmeta.com.

[8] D. M. Chapiro, "Globally asynchronous locally synchronous systems," Ph.D. dissertation, Dept. of Computer Science, Stanford University, Stanford, CA, 1984.

[9] Jens Muttersbach, Thomas Villager and Wolfgang Fichtner, "Practical design of globally asynchronous locally synchronous systems," in Proc. International Symposium on Advanced Research in Asynchronous Circuits and Systems, 2000.

[10] Tiberiu Chelcea and Steven M. Nowick, "A low latency FIFO for mixed-clock system," in Proc. IEEE Computer Society Annual Workshop on VLSI, 2000.

[11] Puru Choudhary and Diana Marculescu, "Hardware based frequency/voltage control of voltage frequency island systems," in Proc. International Conference on Hardware/Software Codesign and System Synthesis, 2006.

[12] G. Quan and X. S. Hu, "Minimum energy fixed-priority scheduling for variable voltage processors," in Proc. Design, Automation and Test in Europe, 2002.

[13] D. Shin, J. Kim and S. Lee, "Low-energy intra-task voltage scheduling using static timing analysis," in Proc. Desing Automation Conference, 2001.

[14] Ana Azevedo, Ilya Issenin, Radu Cornea, Rajesh Gupta, Nikil Dutt, Alex Veidenbaum, Alex Nicolau, "Profile-based dynamic voltage schedling using program checkpoints," in Proc. Design, Automation and Test in Europe, 2002.

[15] Diana Marculescu, On the sue of microarchitecture-driven dynamic voltage scaling," in Workshop on Complexity-Effecitve Design, 2000.

[16] Soray Ghiasi, Jason Casmira and Dirk Grunwald, "Uisng IPC variation in workloads with externally specified rates to reduce power consumption," in Workshop on Complexity-Effecitve Design, 2000.

[17] K. Choi, R. Soma and M. Pedram, "Dynamic voltage and frequency scaling based on workload decomposition," in Proc. International Symposium on Low Power Electrics and Design, 2004.

[18] Gaurav Dhiman and Tajana S. Rosing, "Dynamic voltage frequency scaling for multi-tasking systems using online learning," in Proc. International Symposium on Low Power Electrics and Design, 2007.

[19] P. Juang, Q. Wu, L.-S. Peh, M. Martonosi and D. W. Clark, "Corrdinated, distributed, formal energy management of chip multiprocessors," in Proc. International Symposium on Low Power Electrics and Design, 2005.

[20] C. Isci, A. Buyuktosunoglu, C.-Y. Cher, P. Bose and M. Martonosi, "An analysis of efficient multi-core global power management policies: maximizing performance for a given power budget," in Proc. International Symposium on Microarchitecture, 2006.

[21] Sebastian Herbert and Diana Marculescu, "Analysis of dynamic voltage/frequency scaling in chip-multiprocessors," in Proc. International Symposium on Low Power Electrics and Design, 2007.

[22] J. Hu, Y. Shin, N. Dhanwada and R. Marculescu, "Architecting voltage islands in core-based system-on-a-chip designs," in Proc. International Symposium on Low Power Electrics and Design, 2004.

[23] Huaizhi Wu, I-Min Liu, Martin D.F. Wong and Yusu Wang, "Post-placement voltage island generation under performance requirement," in Proc. International Conference on Computer-Aided Design, 2005

[24] Umit Y. Ogras, Radu Marcuescu, Puru Choudhary and Diana Marculescu, "Voltage-frequency island partitioning for GALS-based networks-on-chip," in Proc. Design Automation Conference, 2007.

[25] Lap-Fai Leung and Chi-Ying Tsui, "Energy-aware synthesis of networks-on-chip implemented with voltage lslands," in Proc. Design Automation Conf., 2007.

[26] Wooyoung Jang, Duo Ding and David Z. Pan, "A voltage-freqnecy island aware energy optimization framework for networks-on-chip," in Proc. International Conference on Computer-Aided Design, 2008.

[27] W.-L. Hung, G.M. Link, Y. Xie, N. Vijaykrishnan, N. Dhanwada and J. Conner, "Temperature-aware voltage island architecturing in system-on-chip design," in Proc. International Conference on Computer Design, 2005.

[28] Min Bao, Alexandru Andrei, Petru Eles and Zebo Peng, "On-line thermal aware dynamic voltage scaling for energy optimization with frequency/temperature dependency consideration," in Proc. Design Automation Conference, 2009.

[29] S. Herbert and D. Marculescu, "Variation-aware dynamic votlage/frequency scaling," in Proc. High Performance Computer Architecture, 2009.

[30] S. Garg, D. Marculescu, R. Marculescu and U. Ogras, "Technology-driven on DVFS controllability of multiple voltage-frequency island designs: a system-level perspective," in Proc. Design Automation Conference, 2009.

[31] R. Puri, L. Stok, J. Cohn, D. Dung, D. Pan, D. Sylvester, A. Srivastava, "Pushing ASIC performance in a power envelope," in Proc. Design Automation Conference, 2003.

[32] Y. Vlasov, "Silicon photonics for next generation computing systems," in European Conference on Optical Communications, 2008.

[33] David A.B. Miller, "Device requirement for optical interconnects to silicon chips," in Proc. IEEE Special Issue on Silicon Photonics, 2009.

[34] Assaf Shacham, Keren Bergman and Luca P. Carloni, "The case for low-power photonic networks on chip," in Proc. Design Automation Conference, 2007.

[35] Duo Ding, Yilin Zhang, Haiyu Huang, Ray T. Chen and David Z. Pan, "O-Router: an optical routing framework for low power on-chip silicon nano-photonic integration," In Proc. Design Automation Conf., 2009.

[36] J. Chan, G. Hendry, A. Biberman, K. Bergman and L.P. Carloni, "PhoenixSim: a simulator for physical-layer analysis of chip-scale photonic interconnection networks," in Prof. Design, Automation and Test in Europe, 2010.