# Dynamic Power Management with Power Network-on-Chip

Inna Vaisband and Eby G. Friedman
Department of Electrical and Computer Engineering
University of Rochester
Rochester, New York 14627
{vaisband, friedman}@ece.rochester.edu

***Abstract*—A critical challenge in multifunctional *heterogeneous* systems-on-chip is delivering high quality dynamically controlled power to support power efficient and portable systems. Efficiently providing multiple point-of-load on-chip voltages requires fundamental changes to the power delivery and management process. The concept of a power network on-chip (PNoC) is introduced here as a general scalable platform for delivering and managing power. This network provides enhanced power control and real-time management of resource sharing for systematic and scalable management of the power delivery process in heterogeneous integrated circuits. To evaluate the performance of the proposed power delivery and management system, a PNoC with four power routers is investigated which supplies power to four power domains. The behavior of the individual power domains is modeled with IBM power grid benchmarks. The proposed architecture and circuits for sensing, routing, and dynamic control of the on-chip power are described. The built-in modularity of the PNoC is exploited to apply dynamic voltage scaling, illustrating the scalability of the PNoC platform while exhibiting power savings of up to 32%.**

## I. Introduction

The delivery of high quality power to the on-chip circuitry with minimum energy loss is a fundamental requirement of modern integrated circuits (ICs). The system-wide quality of the power supply can be efficiently addressed with distributed point-of-load (POL) power delivery [1], and requires the on-chip integration of multiple power supplies. Distributed power delivery requires the co-design of hundreds of power converters with many thousands of decoupling capacitors and billions of nonlinear current loads within multiple power domains, significantly increasing the design complexity of existing power delivery systems. The number of transistors purposely being under utilized in future multicore ICs is expected to rapidly increase. 21% at the 22 nm technology node, and more than 50% at the 8 nm technology node have recently been reported [2], requiring methodologies to manage power on-chip. Per core dynamic voltage scaling (DVS) and dynamic voltage and frequency scaling (DVFS) are primary design objectives with hundreds of power domains and thousands of cores to efficiently manage the power budget, requiring the on-chip integration of compact controllers, further increasing the design complexity of the power delivery system. While in-package and on-chip power integration has recently become a primary objective [3], research remains focused on developing more compact and efficient power supplies, lacking a methodology to effectively integrate and manage the in-package and on-chip power delivery system. The scalability of both static and dynamic power delivery systems, and the granularity of power management in DVS/DVFS multicore systems are limited in existing *ad hoc* approaches.

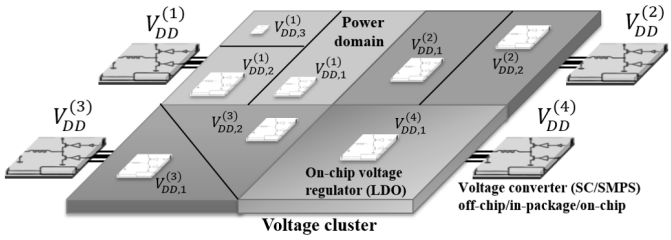
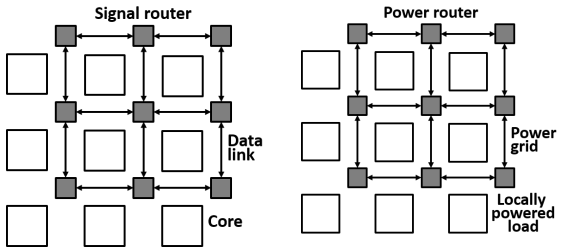



Fig. 1. Heterogeneous power delivery with two switching mode power supplies (SMPS), two switched capacitor (SC) voltage converters, seven low dropout regulators (LDO), and six power domains grouped into three voltage clusters.

In communication systems, the concept of "separation of different concerns" is used to cope with design complexity. This fundamental and cost effective approach is used to increase performance and design productivity since the complexity of *ad hoc* solutions has become excessively high. Specifically, networks-on-chip (NoC) separate computation from communication, enhancing the performance, scalability, and control of the quality of service (QoS), while supporting heterogeneous integrated systems. Recently, a principle of separation of power conversion and regulation has been introduced [4] that addresses the issue of power efficiency in distributed power delivery systems. Consistent with this separation principle, power should be primarily converted with a few power efficient switching supplies, delivered to on-chip voltage clusters, and regulated with linear low dropout (LDO) regulators within the individual power domains, as illustrated in Fig. 1. The concept of a power network on-chip (PNoC) is introduced here based on this separation principle. The analogy between a NoC and PNoC is illustrated in Fig. 2 with simplified NoC and PNoC models. Similar to a network-on-chip, a PNoC decreases the design complexity of the power delivery system, while enhancing the control of the quality of power (QoP) and dynamic voltage scaling (DVS), and providing a scalable platform for efficient power management.

This research is supported in part by the Binational Science Foundation under Grant No. 2012139, the National Science Foundation under Grant No. CCF-1329374, and by grants from Qualcomm, Cisco Systems, and Samsung.

978-1-4799-4885-7/14/$31.00 ©2014 IEEE



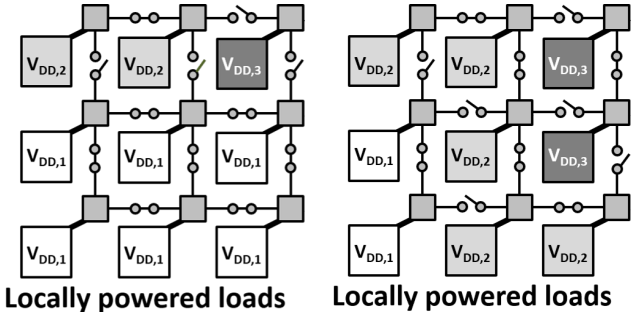

(a) (b)

Fig. 2. On-chip networks based on the approach of separation of functionality, (a) network-on-chip (NoC), and (b) power network-on-chip (PNoC).

The rest of the paper is organized as follows. The principles of the PNoC design methodology are described in Section II. In Section III, the performance of the PNoC architecture is compared with existing approaches based on the evaluation of several test cases. Design and performance issues of the PNoC architecture are also discussed. Some concluding remarks are offered in Section IV.

## II. Power Network-On-Chip Architecture

The key concept in systemizing the design of power delivery is to convert the power off-chip, in-package, and/or on-chip with multiple power efficient but large switching power supplies, deliver the power to on-chip voltage clusters, and regulate the power with hundreds of LDO regulators at the point-of-load [4]. A power network-on-chip is proposed here as a novel systematic solution to on-chip power delivery that leverages distributed point-of-load power delivery within a fine grained power management framework. The PNoC architecture is a mesh of power routers and locally powered loads, as depicted in Fig. 2. The power routers are connected through power switches, distributing current to those local loads with similar voltage requirements. A PNoC is illustrated in Fig. 3 for a single voltage cluster with nine locally powered loads and three different supply voltages, $V_{DD,1}$, $V_{DD,2}$, and $V_{DD,3}$. The power network configuration is shown at two different times, $t_1$ and $t_2$.




(a) (b)

Fig. 3. On-chip power network with multiple locally powered loads and three supply voltage levels (a) PNoC configuration at time $t_1$, and (b) PNoC configuration at time $t_2$.

The concept of a power network-on-chip is proposed here to virtually manage the power in SoCs through specialized power routers, switches, and programmable control logic, while supporting scalable power delivery in heterogeneous ICs. A PNoC is comprised of physical links and routers that provide both virtual and physical power routing. This system senses the voltages and currents, and manages the POL regulators through power switches. Based on the sensed voltages and currents, a programmable unit makes real-time decisions to apply a new set of configurations to the routers per time slot, dynamically managing the on-chip power delivery process. Novel algorithms are required to dynamically customize the power delivery policies through a specialized microcontroller that routes the power. These algorithms satisfy real-time power and performance requirements.

A PNoC composed of power routers connected to the global power grids and locally powered loads is illustrated in Fig. 4. The global power from the converters is managed by the power

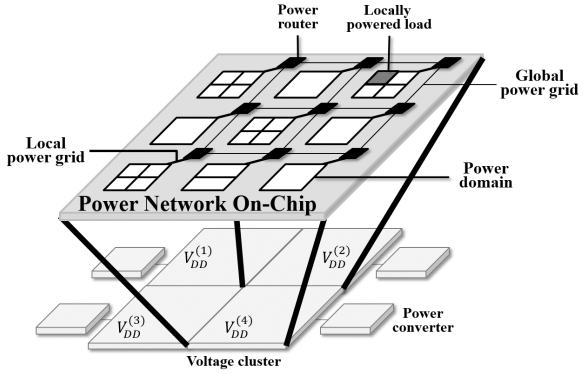



Fig. 4. On-chip power network with routers distributing the current over the power grid to the local loads.

routers, delivered to individual power domains, and regulated within the locally powered loads. These locally powered loads combine all of the current loads located within a specific on-chip power domain with the decoupling capacitors that supply the local current demand within that region. To support DVS, power switches in the proposed PNoC are dynamically controlled, (dis)connecting power routers within individual voltage clusters in real-time. Current loads powered at similar voltage levels therefore draw current from all of the connected power routers, lessening temporary current variations. Similar to a mesh based clock distribution network, the shared power supply distributes the effect of the on-chip parasitic impedances, enhancing the voltage regulation and QoP.

The power routers, local current loads, and power grids are described in the following subsections. Different PNoC topologies and specific design objectives are also considered.

### A. Power routers

The efficient management of the energy budget is dynamically maintained by the power routers. Each power domain is controlled by a single power router. A router topology ranges from a simple linear voltage regulator, shown in Fig. 5(a), to a complex power delivery system, as depicted in Fig. 5(b), with sensors, dynamically adaptable power supplies, switches,

and a microcontroller. The proposed structure features real-time voltage/frequency scaling, adaptable energy allocation, and precise control over the on-chip QoP. With the PNoC routers, the power is managed locally based on specific local current and voltage demands, decreasing the dependence on the remotely located loads and power regulators. The scalability of the power delivery process is therefore enhanced with the proposed PNoC approach.

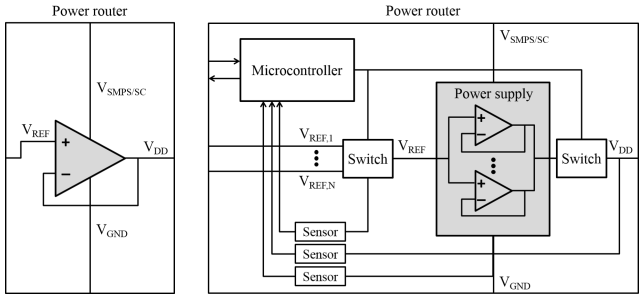



Fig. 5. Power routers for PNoC a) Simple topology with linear voltage regulator. b) Advanced topology with dynamically adaptable voltage regulator and microcontroller.

### B. Locally powered loads

Locally powered loads with different current demands and power budgets can be efficiently managed with this proposed PNoC. The local power grids provide a specific voltage to the nearby load circuits. The highly complex interactions among the multiple power supplies, decoupling capacitors, and load circuits need to be considered, where the interactions among the nearby components are typically more significant. The effective region for a point-of-load power supply is the overlap of the effective regions of the surrounding decoupling capacitors [1]. Loads within the same effective region are combined into a single equivalent locally powered load regulated by a dedicated LDO. All of the LDO regulators within a power domain are controlled by a single power router. A model and closed-form expressions of the interactions among the power supplies, decoupling capacitors, and current loads are required to efficiently partition an IC with billions of loads into power domains and locally powered loads.

### C. Power grid

A shared power grid with multiple LDO sources and dedicated power grids each driven by a separate LDO are considered. Dedicated power grids require fine grain distribution of the local power current. Alternatively, a shared power grid with multiple LDO sources is prone to stability issues due to current sharing, and process, voltage, and temperature variations. Specialized adaptive mechanisms are included within the power routers to stabilize a power delivery system that includes a multi-source shared power grid.

## III. Case Study

To evaluate the performance of the power router, a PNoC with four power routers is considered, supplying power to four power domains. IBM power grid benchmarks [5] model the behavior of the individual power domains. To simulate a dynamic power supply in PNoC, the original IBM voltage profiles have been scaled to generate target power supplies between 0.5 volts and 0.8 volts. Target voltage profiles with four voltage levels (0.8 volts, 0.75 volts, 0.7 volts, and 0.65 volts) within a PNoC are illustrated in Fig. 6. The number

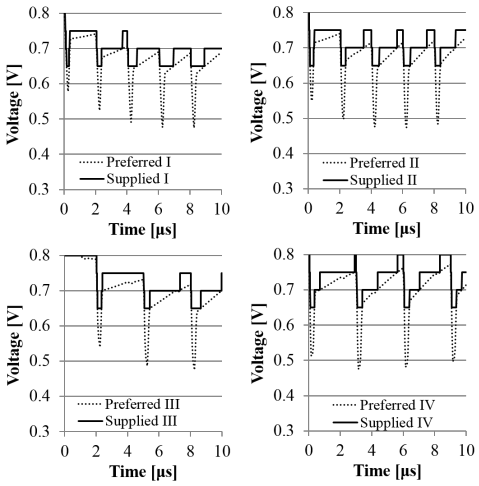



Fig. 6. Preferred and supplied voltage levels in PNoC with four power domains.

of power domains with each of the four voltages changes dynamically based on the transient power requirements of the power domains.

The PNoC is developed based on a 28 nm CMOS technology. A schematic of a PNoC with four power domains (I, II, III, and IV) and four power routers ($PR_I$, $PR_{II}$, $PR_{III}$, and $PR_{IV}$) is shown in Fig. 7. Each of the power routers is composed

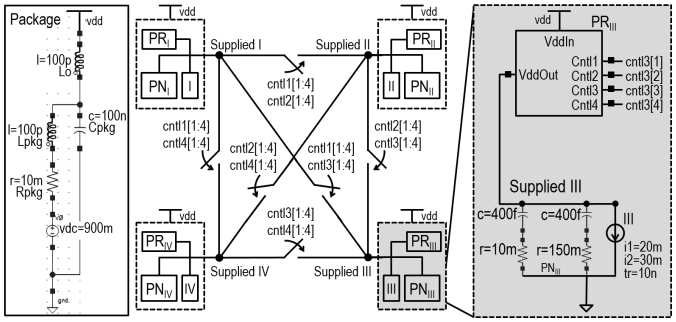



Fig. 7. Proposed PNoC with four power domains and four power routers connected with control switches.

of an LDO with four switch controlled reference voltages to support dynamic voltage scaling. In addition, the power routers feature adaptive RC compensation and current boost networks controlled by load sensors to provide quality of power control and optimization, as shown in Fig. 8. The adaptive RC

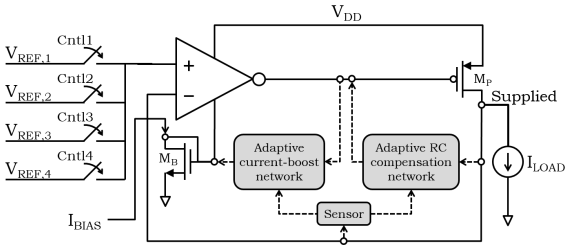



Fig. 8. Power router with voltage regulator, load sensor, and adaptive networks.

compensation network is comprised of a capacitive block connected in series with two resistive blocks, each digitally controlled. These RC impedances are digitally configured to stabilize the LDO within the power routers under a wide range of process variations. The proposed current boost circuit is composed of a sensor block that follows the output voltage at the drain of transistor $M_P$ (see Fig. 8), and a current boost block that controls the current through the differential pair within the LDO. When a high slew rate transition at the output of the LDO occurs, the boost mode is activated, raising the tail current of the differential pair within the LDO. Alternatively, during standard operation, no additional current flows into the differential pair, enhancing the power efficiency of the LDO.

The power routers are connected with control switches to mitigate load transitions in those domains with similar supply voltages. To model the RLC parasitic impedances of the package and power network, non-ideal LDO input and output impedances ($PN_I$, $PN_{II}$, $PN_{III}$, and $PN_{IV}$) are considered, as shown in Fig. 7. The load current characteristics are listed in Table I for each of the four domains. The proposed PNoC

TABLE I
OUTPUT LOAD IN A PNOC WITH FOUR POWER DOMAINS.

| Domain | I | II | III | IV |
|---|---|---|---|---|
| Minimum output current [mA] | 10 | 75 | 20 | 20 |
| Maximum output current [mA] | 50 | 10 | 30 | 20 |
| Output transition time [ns] | 50 | 50 | 10 | N/A |

is evaluated in SPICE. The simulation results are presented in Fig. 9, exhibiting a maximum error of 0.35%, 2.0%, and 2.7% for, respectively, the steady state dropout voltage, load regulation due to the output current switching, and load regulation due to dynamic PNoC reconfigurations. Good correlation with the required power supply in Fig. 6 is demonstrated. The power savings in each of the power domains range between 21.0% to 31.6% as compared to a system without dynamic voltage scaling. These average power savings show that the PNoC architecture can control the power supplies in real-time, dynamically optimizing the power efficiency of the overall power delivery system.

## IV. SUMMARY

Four primary components are necessary to achieve efficient real-time multi-voltage power delivery: (a) a distributed architecture for heterogeneous scalable on-chip power delivery,

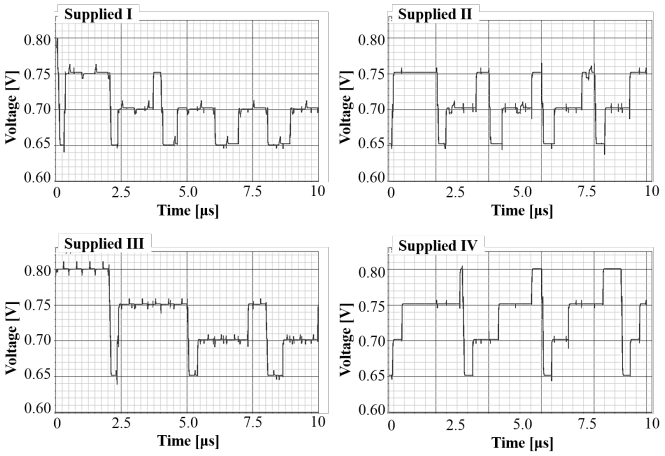



Fig. 9. Voltage levels in PNoC with four power domains.

(b) a systematic methodology that exploits the distributed nature of the proposed architecture to provide real-time adaptive and efficient control over the quality of power, (c) software/firmware based architectural solutions for on-chip power management, and (d) specialized circuits such as power routers, microcontrollers, and ultra-small voltage converters.

The primary objective of the proposed on-chip power network solution is to provide a fine grain power management framework comprising a variety of circuit, firmware, and architecture level techniques that controls power routing and switching, while optimally allocating power among a variety of different power domains at run time. A PNoC case study with four power domains and four power routers is evaluated based on a 28 nm CMOS technology and demonstrated on IBM power grid benchmarks, exhibiting power savings of up to 32%. The power routers feature adaptive RC compensation and current boost networks controlled by load sensors to enhance the stability of the distributed power delivery system, while controlling and optimizing the quality of power. PNoC systems enhance voltage regulation and quality of power while providing scalability.

## REFERENCES

[1] S. Kose and E. G. Friedman, "Distributed On-Chip Power Delivery," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, Vol. 2, No. 4, pp. 704–713, December 2012.

[2] H. Esmaeilzadeh, E. Blem, R. St. Amant, K. Sankaralingam, and D. Burger, "Dark Silicon and the End of Multicore Scaling," *Proceedings of the ACM International Symposium on Computer Architecture*, pp. 365–376, June 2011.

[3] F. Waldron, J. Slowey, A. Alderman, B. Narveson, and S. C. O'Mathuna, "Technology Roadmapping for Power Supply in Package (PSiP) and Power Supply on Chip (PwrSoC)," *Proceedings of the IEEE International Applied Power Electronics Conference and Exposition*, pp. 525–532, February 2010.

[4] I. Vaisband and E. G. Friedman, "Heterogeneous Methodology for Energy Efficient Distribution of On-Chip Power Supplies," *IEEE Transactions on Power Electronics*, Vol. 28, No. 9, pp. 4267–4280, September 2013.

[5] S. R. Nassif, "Power Grid Analysis Benchmarks," *Proceedings of the IEEE/ACM Asia and South Pacific Design Automation Conference*, pp. 376–381, January 2008.