

Distributed On-Chip Power Delivery

Selçuk Köse, *Member, IEEE*, and Eby G. Friedman, *Fellow, IEEE*

Abstract—The performance of an integrated circuit depends strongly upon the power delivery system. With the introduction of ultra-small on-chip voltage regulators, novel design methodologies are needed to simultaneously determine the location of the on-chip power supplies and decoupling capacitors. In this paper, a unified design methodology is proposed to determine the optimal location of the power supplies and decoupling capacitors in high performance integrated circuits. Optimization algorithms widely used for facility location problems are applied in the proposed methodology. The effect of the number and location of the power supplies and decoupling capacitors on the power noise and response time is discussed.

Index Terms—Distributed power delivery, heterogeneous integrated circuit (IC), on-chip decoupling capacitor, point-of-load power supply.

I. INTRODUCTION

POWER consumption has become one of the primary design constraints with the proliferation of mobile devices as well as server farms where the performance per watt is the fundamental benchmark [1], [2]. The quality of the voltage delivered to the many circuit blocks has a direct effect on the performance of an integrated circuit (IC). The voltage down-converted and regulated by the off-chip and on-chip voltage regulators is distributed throughout a power distribution system to the billions of load circuits. Due to the finite parasitic impedance of a power distribution network, voltage drops and bounces can occur in the supply voltage. The frequency and amplitude of these voltage fluctuations depend upon several factors, including the characteristics of the load current, parasitic impedance of the power distribution network, output impedance of the power supplies, and effective series resistance and inductance of the decoupling capacitors. To reduce the amplitude of these voltage fluctuations, the power supplies are supported by locally distributed decoupling capacitors, which serve as a nearby reservoir of charge, providing current to the load circuits [3].

The complexity of high performance power delivery systems has increased significantly with the integration of diverse

Manuscript received June 29, 2012; revised August 16, 2012; accepted October 06, 2012. Date of current version December 10, 2012. This work was supported in part by the National Science Foundation under Grant CCF-0811317 and Grant CCF-0829915, in part by grants from the New York State Office of Science, Technology and Academic Research to the Center for Advanced Technology in Electronic Imaging Systems, and in part by Grants from Intel Corporation and Qualcomm Corporation. This paper was recommended by Guest Editor S. Mitra.

S. Köse is with the Department of Electrical Engineering, University of South Florida, Tampa, FL 33620 USA (e-mail: kose@usf.edu).

E. G. Friedman is with the Department of Electrical and Computer Engineering, University of Rochester, Rochester, New York 14627 USA (e-mail: friedman@ece.rochester.edu).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/JETCAS.2012.2226378

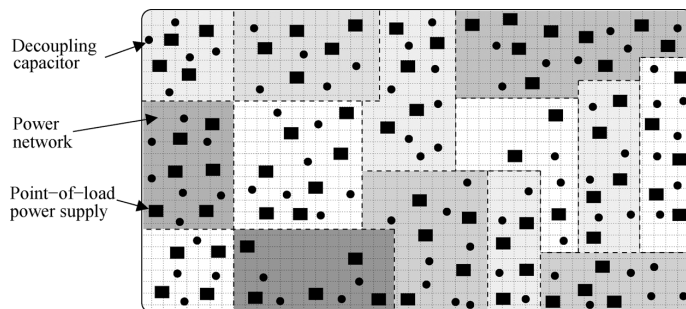


Fig. 1. Next generation power delivery network with local point-of-load power supplies supported by decoupling capacitors, providing current to billions of load circuits within different voltage islands.

technologies on a single die, forming a heterogeneous system. The required voltage levels and noise constraints vary significantly for different technologies. Novel voltage regulator topologies have recently been proposed [4]–[10], enabling not only the integration of on-chip power supplies but also multiple distributed on-chip point-of-load power supplies [10], [11]. These on-chip point-of-load power supplies provide the necessary voltage close to the load circuits, greatly reducing the parasitic impedance between the load circuits and power supplies, and enhancing the efficiency of the overall power delivery system [12]. There are three primary advantages of a distributed point-of-load power delivery system. First, the voltage is generated close to the load circuits, reducing the noise caused by the parasitic impedances within the power network. Second, discrete voltages can be generated for various circuit blocks built in different technologies in heterogeneous circuits. Third, a high granularity power management system are realized where the voltage levels can be individually controlled for different circuit blocks. This paper primarily exploits the first and second advantages.

Next generation power delivery networks for high performance circuits will contain tens to hundreds of on-chip power supplies supported by many on-chip decoupling capacitors to satisfy the current demand of billions of load circuits within different voltage islands, as illustrated in Fig. 1. The design of these complex systems would be greatly enhanced if the available resources, such as the physical area, number of metal layers, and power budget, were not severely limited. The continuous demand over the past decade for greater functionality within a small form factor has imposed tight resource constraints while achieving aggressive performance and noise targets [13].

Several techniques have been proposed for efficient power delivery systems, typically focusing on optimizing the power network [13], [14] and placement of the decoupling capacitors [15]–[17]. Recently, Zeng *et al.* [18] proposed an optimization technique for designing power networks with multiple on-chip voltage regulators. The design of these on-chip voltage regulators and the effect of these regulators on high frequency

voltage fluctuations and mid-frequency resonance have been investigated. The interactions between the power supplies and decoupling capacitors, which can significantly affect the performance of an IC, have, however, not been considered [18]. These interactions are quite critical in producing a robust power distribution network [10]. Decoupling capacitors and on-chip power supplies exhibit several distinct characteristics, such as the response time, area requirements, and parasitic output impedances. Circuit models of these components should accurately capture these characteristics, while being sufficiently simple to not computationally constrain the optimization process.

In this paper, the optimum location of the on-chip power supplies and decoupling capacitors for different constraints is determined using facility location optimization algorithms [19]–[21]. The constraints of this power network co-design problem depend upon the application and performance objectives. The optimization goal can be to minimize the maximum voltage drop, total area, response time for particular circuit blocks, or total power consumption. Multiple optimization goals can also be applied for smaller or midsize ICs.

The rest of the paper is organized as follows. A recently developed point-of-load voltage regulator is briefly described in Section II. The facility location problem is introduced with some exemplary applications in Section III. A proposed methodology to determine the optimum location of the power supplies and decoupling capacitors is examined in Section IV. The optimum location of the power supplies and decoupling capacitors, exemplified on several benchmark circuits, is presented in Section V. A brief discussion of the proposed optimization technique and possible enhancements are offered in Section VI. The paper is concluded in Section VII.

II. POINT-OF-LOAD VOLTAGE REGULATORS

Placing multiple point-of-load power supplies is challenging since the area occupied by a single power supply should be small and the efficiency sufficiently high. Guo *et al.* proposed an output capacitorless low-dropout regulator which occupies an on-chip area of 0.019 mm^2 [6]. The authors of this paper recently proposed a hybrid point-of-load voltage regulator, occupying an on-chip area of 0.015 mm^2 [8], [22]. A microphotograph of this hybrid point-of-load regulator is shown in Fig. 2. These area efficient voltage regulators provide a means for distributing multiple local power supplies across an IC, while maintaining high current efficiency and small area. With point-of-load voltage delivery, on-chip signal and power integrity are significantly enhanced while providing the capability for distributing multiple power supplies. Design methodologies are therefore required to determine the location, size, and number of these distributed on-chip power supplies and decoupling capacitors.

III. FACILITY LOCATION PROBLEM

Every complex system is an ensemble of small components, typically with simple structures. The interactions and aggregation of these components form a highly sophisticated system. The efficiency of this system strongly depends not only upon the physical properties of the individual components but also on the spatial location of these components since the placement

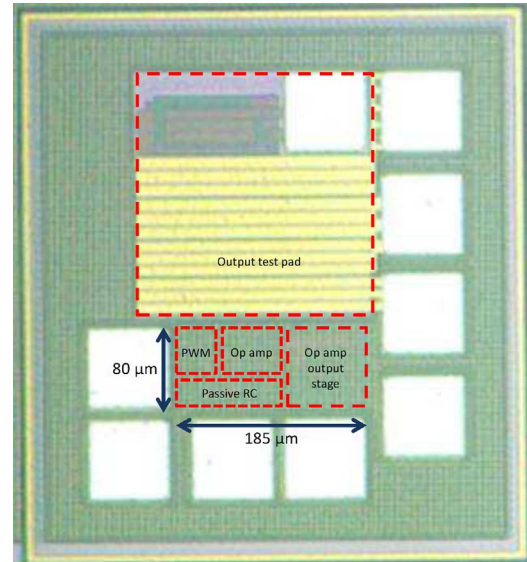


Fig. 2. Microphotograph of the hybrid voltage regulator [8].

of these components significantly affects the multiple interactions within the system. In most systems, these components can be grouped into two categories: 1) facilities and 2) customers. Facility location problems, to determine the location, size, and number of facilities that minimize the cost of providing a high quality service to customers, have been well studied over the last several decades [19].

Mathematical models have been widely used to determine the optimal number, location, and size of the facilities as well as to allocate facility resources to those customers that minimize or maximize an objective function [19]–[21]. The problem can be categorized depending upon the interconnection network (discrete or continuous) and the input (static or dynamic). The objective is typically to minimize the average (or maximum) distance from the facilities to the customers, determine the minimum number of facilities that serve a particular number of customers at fixed locations, or maximize the minimum distance from a facility to the customers.

The design of an on-chip power delivery network for heterogeneous circuits exhibits significant similarities to the design of electrical distribution networks in larger scale systems, such as the electric power distribution grid of a city. The electricity generated at a power plant is downconverted and distributed to substation transformers, typically outside a city. The output voltage of these substation transformers is further downconverted and regulated by the local power supplies, as shown in Fig. 3. This voltage can be either delivered to industrial customers at a high voltage level or further downconverted and regulated at smaller substations and distributed to the local power grid within the city. Large capacitors are integrated within this electrical distribution system to reduce voltage fluctuations. Alternatively, in an IC, the board level voltage regulators downconvert the output voltage of the board level power supply unit. This voltage is delivered to the on-chip voltage regulators or directly to the on-chip power grid, which provides current to the load circuits. The required voltage levels and noise constraints are technology and design dependent. The on-chip power delivery system is designed to deliver different voltage levels within specified noise constraints. Decoupling capacitors are distributed throughout

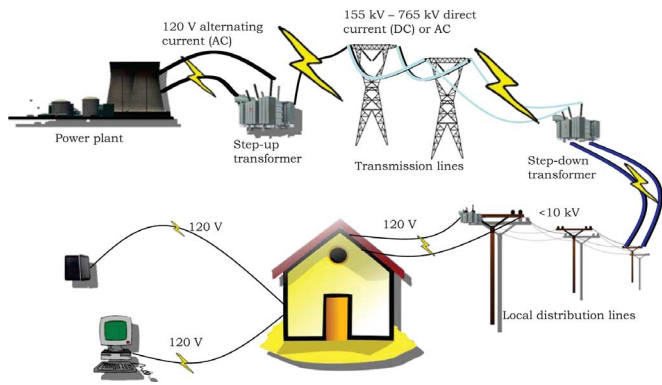


Fig. 3. Large scale electric power distribution system.

the on-chip power delivery network to support the power distribution system by providing local charge to the load circuits. A parallel can be drawn between the transformers and off-chip voltage regulators, the small substations and on-chip voltage regulators, and the large capacitors and on-chip decoupling capacitors. Additionally, the voltage requirements within an IC vary in a similar manner as the voltage requirements of different industrial and residential regions within a city.

Several optimization algorithms have been proposed to provide an optimal solution to this problem. Due to the similarity between the electrical distribution network of a city and the power distribution network of a heterogeneous circuit, analogous algorithms can be applied to the design of these systems. Since facility location algorithms are widely used to design electrical distribution networks [23], [24], these city planning algorithms are leveraged here in designing on-chip power networks within high performance ICs.

IV. PROPOSED OPTIMIZATION METHODOLOGY

The primary objective of the proposed optimization methodology is to determine the optimal location of the on-chip power supplies and decoupling capacitors that minimize the maximum power noise and response time to certain blocks while maintaining the area constant. Other design objectives can, however, be incorporated into the objective function while minimizing the maximum voltage drop and response time for certain blocks, such as minimizing the 1) power consumed by the power distribution networks, or 2) on-chip area.

Placing the power supplies and decoupling capacitors close to the load circuits reduces the power noise. Minimizing the sum of the weighted distances between a source point and a set of destination points is known as a Weber problem [25]–[27]. In the proposed optimization, the weighted sum of distances between the power sources and the load circuits is minimized similar to a Weber problem while also considering the individual characteristics of the power supplies, decoupling capacitors, and load circuits.

The cost of delivering power from a power supply or a decoupling capacitor to a load circuit depends upon 1) the parasitic impedance of the power distribution network, 2) the amount of current delivered to the load circuit, and 3) the parasitic impedance to the power supplies and decoupling capacitors. A Euclidean or Manhattan distance is widely used in

the cost function of facility location problems. Alternatively, a closed-form impedance model [12], which accurately captures the power grid characteristics and physical distance, is utilized to determine the effective impedance within the power grid from the power supplies and decoupling capacitors to the load circuits.

Multiple power supplies and decoupling capacitors can provide current to a single load circuit. The contribution depends upon both the effective impedance among these components and the requirements of the load circuit. For example, when the current profile exhibits a fast transition time, a decoupling capacitor is a better choice due to the faster response of these structures. The contribution of current from multiple power supplies and decoupling capacitors to each load circuit is estimated considering the effective impedance among the load circuit and power sources, as in (5) and (6) where G_{ij} is the equivalent conductance between the i th voltage supply (or decoupling capacitor) and the j th current load.

An objective function $F(n, m, k)$ is proposed to determine the optimum location of the power supplies and decoupling capacitors. $F(n, m, k)$ is comprised of three terms. Minimizing the first and second terms optimizes the location of the power sources for minimum noise whereas minimizing the third term optimizes the location to provide a fast response to specific blocks. The solution of the weighted sum of these three terms provides the optimum location of the power sources for both minimum power noise and fastest response. Alternatively, the solution guarantees a minimum power noise level or fastest response time based upon the weight of the terms in the optimization function.

Multiple parameters such as the parasitic impedance of the power network, output impedance of the power supply, effective series resistance of a decoupling capacitor, and load current characteristics significantly affect the power noise and/or response time. These parameters are therefore considered in the first and second terms of the optimization function to minimize the power noise throughout a circuit. Alternatively, the response time of the power delivery network to transient changes in the current within certain blocks is minimized by placing the decoupling capacitors physically close to these blocks. The third term is therefore included within the objective function to place the decoupling capacitors close to those circuit blocks demanding a fast transient current. The contribution of the decoupling capacitor to the circuit blocks, the normalized transition time within the circuit blocks, and the sum of the equivalent impedance of the power network and effective series resistance of the decoupling capacitors are considered in the third term. Intuitively, since the transition time of the current within the blocks with a fast switching activity is smaller, reducing the effective impedance between the decoupling capacitors and these blocks decreases the cost function. Moving the decoupling capacitors close to those circuit blocks requiring a faster transition time minimizes the objective function. The proposed objective function is

$$\begin{aligned} & \text{Minimize} \\ & F(n, m, k) \\ & = K_1 \sum_{j=1}^m \sum_{i=1}^n C_{P_{ij}} (R_{\text{out}}(P_i) + R_{\text{eff}}(P_i, L_j)) \end{aligned}$$

$$\begin{aligned}
& + K_2 \sum_{j=1}^m \sum_{i=1}^k C_{D_{ij}} (R_{\text{esr}}(D_i) + R_{\text{eff}}(D_i, L_j)) \\
& + K_3 \sum_{j=1}^m \sum_{i=1}^k \text{cap}_{D_i} N_{trL_j} (R_{\text{esr}}(D_i) + R_{\text{eff}}(D_i, L_j))
\end{aligned} \tag{1}$$

Subject to

$$\begin{aligned}
R_{\text{eff}} & \frac{(\text{node}_\alpha, \text{node}_\beta)}{r} \\
& = \frac{\sqrt{1}}{2\pi} [\ln((x_1 - x_2)^2 \\
& \quad + (y_1 - y_2)^2) + 3.44388] - 0.033425
\end{aligned} \tag{2}$$

$$1 < x_{\alpha,\beta} < (\text{Grid size})_X \tag{3}$$

$$1 < y_{\alpha,\beta} < (\text{Grid size})_Y \tag{4}$$

$$C_{P_{ij}} = \frac{G_{ij}}{\sum_{i=1}^n G_{ij}} \tag{5}$$

$$C_{D_{ij}} = \frac{G_{ij}}{\sum_{i=1}^k G_{ij}} \tag{6}$$

$$\sum_{j=1}^m C_{P_{ij}} \leq \text{cap}_{P_i} \tag{7}$$

$$\sum_{j=1}^m C_{D_{ij}} \leq \text{cap}_{D_i} \tag{8}$$

$$\sum_{i=1}^n C_{P_{ij}} + \sum_{i=1}^k C_{D_{ij}} = \sum_{i=1}^m I_i \tag{9}$$

$$\sum_{i=1}^n \text{cap}_{P_i} + \sum_{i=1}^k \text{cap}_{D_i} \geq \sum_{i=1}^m I_i \tag{10}$$

where the definition of the aforementioned parameters are listed in Table I.

The maximum voltage drop and/or response time is minimized using the objective function $F(n, m, k)$, where the effective resistance is defined in (2) [12]. By applying constraints (3) and (4), the optimum location of the power supplies and decoupling capacitors is maintained within the dimensions of the power grid. Constraints (7) and (8) ensure that the total contribution of current from a power supply or a decoupling capacitor cannot exceed, respectively, the capacity of that particular power supply or decoupling capacitor. Furthermore, the total current demand from all of the load circuits is equal to the total contribution from the power supplies and decoupling capacitors, as guaranteed by (9). Additionally, by applying constraint (10), the total capacity of the power supplies and decoupling capacitors is maintained greater than or equal to the total current demand of the circuit.

In the proposed optimization function, K_i (see Table I) provides the flexibility to optimize the power distribution system for different objectives, such as minimizing the maximum voltage drop or response time. When K_3 (or K_1 and K_2) is equal to zero, the location of the power supplies and decoupling capacitors is chosen to minimize the maximum voltage drop (or response time). When the total capacity of the available

TABLE I
DEFINITION OF THE PARAMETERS IN (1)–(10)

Parameter	Definition
P_i	i^{th} power supply
D_i	i^{th} decoupling capacitor
L_i	i^{th} circuit block
$R_{\text{eff}}(\text{node}_1, \text{node}_2)$	Effective resistance between node_1 and node_2
(x_1, y_1)	Coordinates of node_1
(x_2, y_2)	Coordinates of node_2
r	Unit resistance within the power grid
n	Number of power supplies
k	Number of decoupling capacitors
m	Number of load circuits
$R_{\text{out}}(P_i)$	Output resistance of i^{th} power supply
$R_{\text{esr}}(D_i)$	Effective series resistance of i^{th} decap
G_{ij}	Equivalent conductance
K_i	Weighting parameter
$C_{P_{ij}}$	Contribution of i^{th} power supply to j^{th} load
$C_{D_{ij}}$	Contribution of i^{th} decap to j^{th} load
cap_{P_i}	Capacity of i^{th} power supply
cap_{D_i}	Capacity of i^{th} decap
N_{trL_j}	Normalized transition time of the j^{th} load circuit
I_i	Current demand of i^{th} load
$(\text{Grid size})_X$	Power grid size in horizontal direction
$(\text{Grid size})_Y$	Power grid size in vertical direction

power supplies and decoupling capacitors is greater than the total current demand of the IC, the current can be supplied either from the decoupling capacitors or power supplies, which satisfies (10). For example, when the power consumption is the primary bottleneck rather than the physical area occupied by the power supplies and decoupling capacitors, adding more decoupling capacitors instead of on-chip power supplies is a better option if the noise constraints are satisfied. In this case, K_1 should be greater than K_2 to ensure that the weight of the first term in (1) (i.e., the cost function of the power supplies) is greater than the weight of the second term in (1) (i.e., the cost function of the decoupling capacitors). K_i can therefore be treated as weighting parameters to balance the optimization process for different design constraints.

Please note that the voltage delivered to the on-chip point-of-load power supplies from the dedicated power pads is assumed here to be perfect. From a noise perspective, this assumption is reasonable since the on-chip power supplies have both line and load regulation. Since the output voltage of the on-chip power supplies is regulated, the input power to these supplies does not significantly affect the power noise. From a power consumption perspective, however, the distribution of the power to the on-die power supplies has a significant effect. Since the location of the power supplies and decoupling capacitors that minimize the noise and response time is determined, the assumption of perfect power delivery to the power supplies does not significantly affect the optimum locations.

V. CASE STUDY AND BENCHMARK CIRCUITS

In this section, the optimum location of the power supplies and decoupling capacitors for different circuits is determined utilizing the proposed optimization methodology. The voltage drop maps of the related circuits with the decoupling capacitors and power supplies located at the predetermined locations are

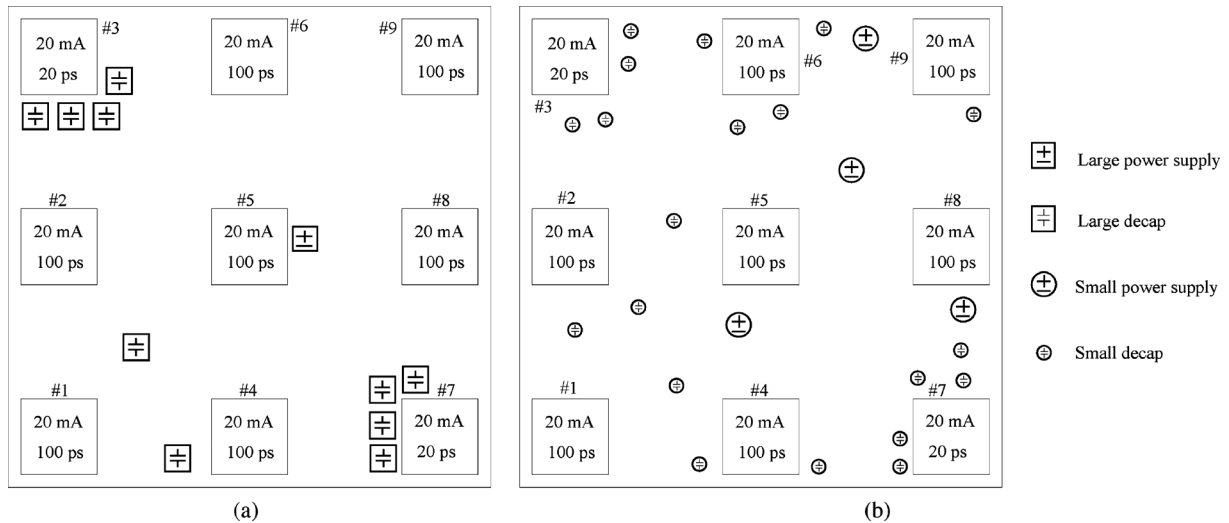


Fig. 4. Floorplan of the example circuit with two different power delivery networks: (a) one large power supply with ten decoupling capacitors, and (b) four relatively smaller distributed power supplies with 20 small decoupling capacitors.

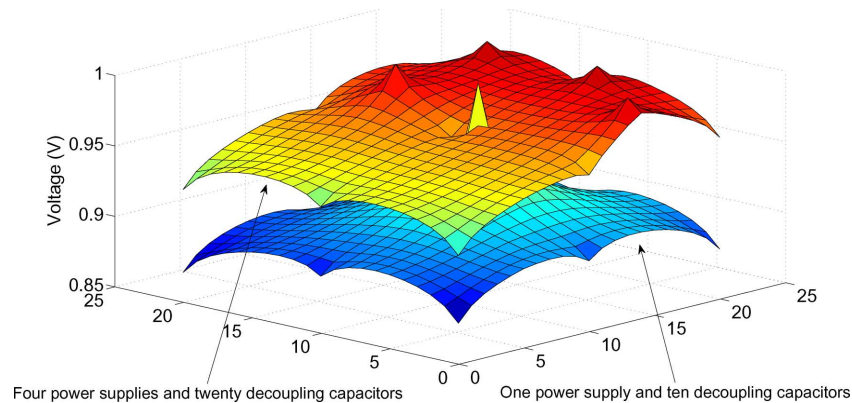


Fig. 5. Map of voltage drops within the sample circuit for two different cases, one large power supply with ten decoupling capacitors, and two relatively smaller distributed power supplies with 20 small decoupling capacitors. The maximum voltage drop is reduced when the number of power supplies and decoupling capacitors is increased due to the distributed nature of the power delivery network.

obtained using SPICE. The node voltages, which are determined by the SPICE simulations, are produced from MATLAB.

The optimum number and location of the power supplies and decoupling capacitors that minimize the voltage drop and response time within certain blocks are determined for a small sample circuit, as shown in Fig. 4, to provide an intuitive understanding of the proposed methodology. The sample circuit is composed of nine circuit blocks with different current profiles. The third and seventh blocks have current profiles with a faster transition time (i.e., 20 ps) than the rest of the circuits which have a relatively slower transition time (i.e., 100 ps). Since the decoupling capacitors provide immediate charge, intuitively, the decoupling capacitors should be placed close to those blocks with a fast transition time to provide a fast response to transient changes in the current. The optimum location of the power supplies and decoupling capacitors that minimizes both the maximum voltage drop and response time for certain blocks (the third and seventh blocks) is used, where K_1 , K_2 , and K_3 are set to one. The optimum location of one large on-chip power supply and ten decoupling capacitors (case a) is shown in Fig. 4(a). The power supply is located at a central location to reduce the maximum physical distance to each of the circuit blocks. The

decoupling capacitors, however, are placed physically close to the third and seventh blocks. Most of the current demand of these blocks is provided by the surrounding decoupling capacitors. The optimum location of the four relatively low current power supplies and 20 small decoupling capacitors (case b) is also determined, as shown in Fig. 4(b). In this case, the third and seventh circuit blocks are surrounded by local decoupling capacitors whereas the power supplies are distributed to ensure that the maximum distance from the power supplies to the remaining blocks is minimized. The voltage drop map for these two cases is shown in Fig. 5, where increasing the number of power supplies and decoupling capacitors significantly reduces the voltage drop. The maximum voltage drop is 133 mV and 77 mV, respectively, for cases a and b. More than a 40% reduction in the maximum voltage drop is achieved by increasing the number and distributing the location of the power supplies and decoupling capacitors.

The area of an on-chip power supply is typically dominated by the output pass transistors [22], where the size of these pass transistors changes linearly with the maximum output current demand. The size of an on-chip power supply therefore changes linearly with the maximum output current capacity. Addition-

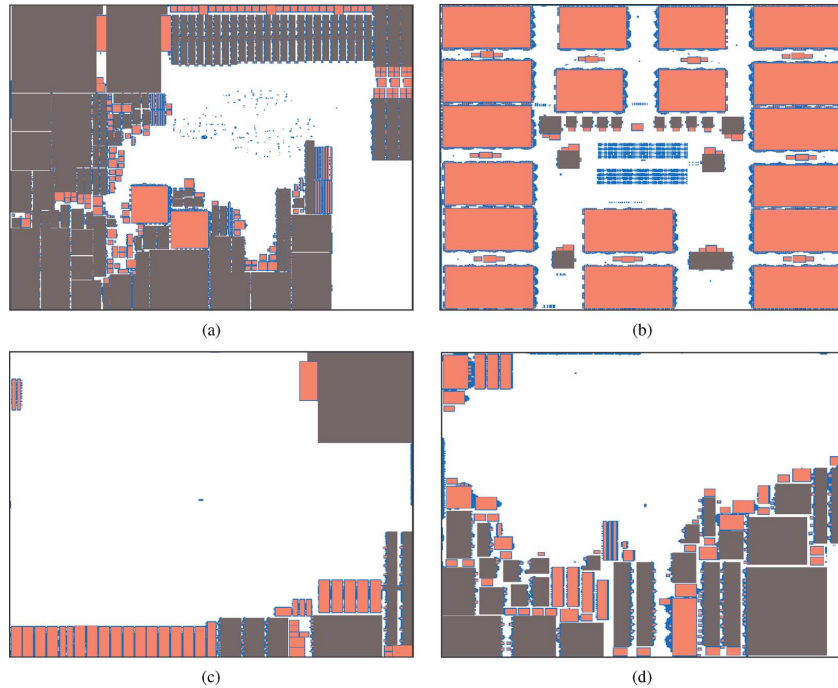


Fig. 6. Floorplan of ISPD'11 circuits [28] (a) superblue5, (b) superblue10, (c) superblue12, and (d) superblue18.

TABLE II
PROPERTIES OF ISPD BENCHMARK CIRCUITS

circuit	# of blocks	Reduced # of blocks	Coverage of reduced floorplan	Power grid size	# of nodes in the power grid
superblue5	95,041	89	82.5 %	774 X 713	551,862
superblue10	214,223	49	89.5 %	638 X 968	617,584
superblue12	15,349	70	98.4 %	444 X 518	229,992
superblue18	41,047	83	94.4 %	381 X 404	153,924

ally, when the on-chip power supplies are sufficiently small, the ultra-small power supplies are combined to form a larger power supply with a higher output current. In this paper, the size of a power supply is assumed to change linearly with the maximum output current capacity.

The optimal location of the power supplies and decoupling capacitors for several ISPD'11 placement benchmark suite circuits is evaluated with the proposed distributed power delivery methodology for a different number of power supply and decoupling capacitor configurations [28]. The floorplan of these circuits is illustrated in Fig. 6. More than 15 000 individual circuit blocks exist in these circuits. As shown in Fig. 6, a significant portion of the floorplan is occupied by several large circuit blocks. To reduce the complexity of the proposed optimization problem, only the large circuit blocks are considered in the proposed co-design methodology. The actual and reduced number of circuit blocks are listed in Table II. Although the reduced number of blocks corresponds to less than 0.5% of the actual number of blocks, these fewer number of blocks occupies more than 82% of the total active circuit area.

The size of the power distribution networks and total number of nodes in these benchmark circuits are listed in Table II. Each circuit block is modeled as a single current load where the maximum current demand is proportional to the size of the circuit

TABLE III
FIVE DIFFERENT POWER SUPPLY AND DECOUPLING CAPACITOR ARRANGEMENTS

	# of power supplies	# of decoupling capacitors
Case1	1	2
Case2	1	10
Case3	3	10
Case4	3	20
Case5	20	32

block. Each current load, representing a circuit block, is connected to the power grid from the node physically closest to the center of that particular circuit block.

The general algebraic modeling system (GAMS) is used as the optimization tool [29]. The proposed optimization methodology is modeled as a mixed integer nonlinear programming problem. The location of the power supplies and decoupling capacitors that minimizes the maximum voltage drop is determined for a different number of power supplies and decoupling capacitors for four different ISPD'11 benchmark circuits. These results are listed in Tables IV–VII, respectively, for superblue5, superblue10, superblue12, and superblue18. The total area of the power supplies and decoupling capacitors is maintained the same for all of the test cases to provide a fair comparison.

TABLE IV
OPTIMUM LOCATION OF POWER SUPPLIES AND DECOUPLING CAPACITORS THAT MINIMIZE THE AVERAGE VOLTAGE DROP FOR SUPERBLUE5

# of power supplies	# of decoupling capacitors	Power supply location (x,y)	Decoupling capacitor location (x,y)
1	2	(267,246)	(396,86), (90,608)
1	10	(141,360)	(748,626), (90,608), (21,277), (324,59), (89,97), (90,608), (40,462), (90,608), (69,630),(422,47)
3	10	(761,586), (3,331), (254,142)	(30,98), (90,610), (619,389), (89,98), (114,90), (499,46), (113,114), (736,694), (736,694), (88,98)
3	20	(87,454), (346,465), (373,447)	(576,311), (761,623), (404,131), (761,589), (725,604), (581,71), (499,46), (42,462), (532,187), (422,47), (23,278), (30,98), (422,47), (83,299), (42,372), (500,47), (31,97), (713,305), (250,41), (23,277)

TABLE V
OPTIMUM LOCATION OF POWER SUPPLIES AND DECOUPLING CAPACITORS THAT MINIMIZE THE AVERAGE VOLTAGE DROP FOR SUPERBLUE10

# of power supplies	# of decoupling capacitors	Power supply location (x,y)	Decoupling capacitor location (x,y)
1	2	(297,253)	(564,73), (564,894)
1	10	(258,211)	(320,860), (74,725), (533,442), (564,393), (398,894), (564,73), (563,895), (564,251), (331,582), (111,210)
3	10	(77,73), (238,795), (396,73)	(77,72), (79,71), (378,647), (469,547), (493,487), (597,884), (563,716), (401,791), (209,597), (564,894)
3	20	(564,894), (564,715), (398,694)	(397,695), (447,570), (76,399), (79,257), (78,71), (327,590), (240,894), (399,895), (564,395), (3,651), (399,895), (417,796), (202,479), (394,699), (76,401), (239,796), (75,400), (432,467), (237,694), (564,717)

TABLE VI
OPTIMUM LOCATION OF POWER SUPPLIES AND DECOUPLING CAPACITORS THAT MINIMIZE THE AVERAGE VOLTAGE DROP FOR SUPERBLUE12

# of power supplies	# of decoupling capacitors	Power supply location (x,y)	Decoupling capacitor location (x,y)
1	2	(434,117)	(385,439), (369,36)
1	10	(380,408)	(241,34), (370,37), (385,441), (295,34), (353,90), (415,104), (371,36), (183,28), (386,440), (369,37)
3	10	(297,15), (386,440), (381,101)	(385,439), (296,13), (267,33), (329,466), (431,120), (369,36), (421,30), (326,16), (418,116), (384,439)
3	20	(385,439), (421,23), (304,281)	(8,448), (307,18), (384,440), (210,26), (435,116), (329,466), (461,449), (319,18), (124,16), (420,100), (385,439), (12,20), (269,35), (430,57), (384,441), (267,34), (385,439), (345,78), (385,439), (329,466)

TABLE VII
OPTIMUM LOCATION OF POWER SUPPLIES AND DECOUPLING CAPACITORS THAT MINIMIZE THE AVERAGE VOLTAGE DROP FOR SUPERBLUE18

# of power supplies	# of decoupling capacitors	Power supply location (x,y)	Decoupling capacitor location (x,y)
1	2	(323,61)	(123,93), (325,61)
1	10	(50,169)	(132,23), (257,3), (265,165), (87,28), (66,172), (27,183), (48,75), (334,229), (188,3), (375,231)
3	10	(266,13), (50,169), (318,202)	(85,28), (273,150), (30,29), (14,376), (30,27), (13,361), (30,29), (17,162), (3,383), (31,167)
3	20	(66,61), (323,61), (179,4)	(48,75), (82,103), (37,378), (85,28), (291,180), (254,72), (52,39), (29,29), (37,378), (130,23), (29,28), (85,28), (30,27), (324,60), (325,61), (38,172), (30,29), (17,162), (24,391), (24,392)

The voltage drop maps of the ISPD'11 circuits with the power supplies and decoupling capacitors distributed throughout these circuits, as listed in Tables IV–VII, are, respectively, shown in Figs. 7–10. The maximum and average voltage drop for five different cases [i.e., five different arrangements of power supplies and decoupling capacitors (see Table III)] is listed in Table VIII. The maximum voltage drop is greatest for each circuit when only one power supply and two decoupling capacitors are included within the power delivery network. Increasing the number of power supplies and/or decoupling capacitors significantly reduces the maximum and average voltage drops. When the number of decoupling capacitors

increases from two to ten with one power supply, the reduction in the maximum voltage drop is, respectively, 21.6%, 45.2%, 30%, and 23.7% for superblue5, superblue10, superblue12, and superblue18. Alternatively, the reduction in the maximum voltage drop is, respectively, 22%, 8.1%, 10%, and 35% for superblue5, superblue10, superblue12, and superblue18 when the number of decoupling capacitors increases from ten to 20 with three power supplies. The greatest noise reduction is achieved when the 20 power supplies and supported by 32 decoupling capacitors, as illustrated in Table VII.

The average voltage drop throughout the power distribution network for five different cases is also listed in Table VIII. When

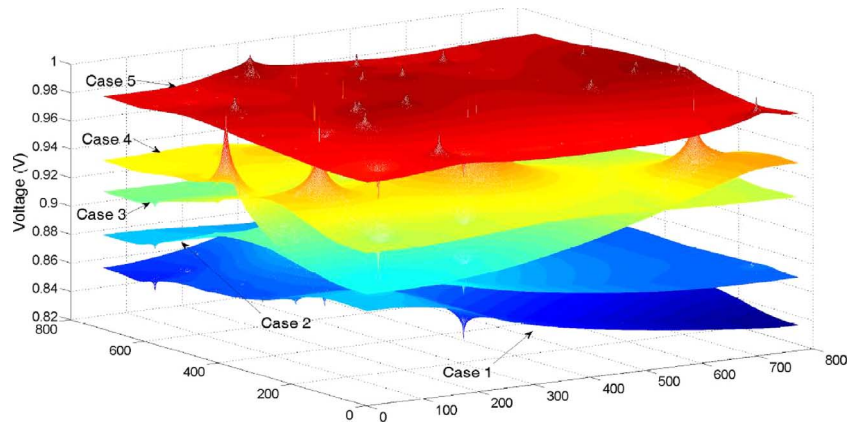


Fig. 7. Map of voltage drops within superblue5 for five different cases. The maximum and average voltage drop is reduced when the power supplies and decoupling capacitors are distributed.

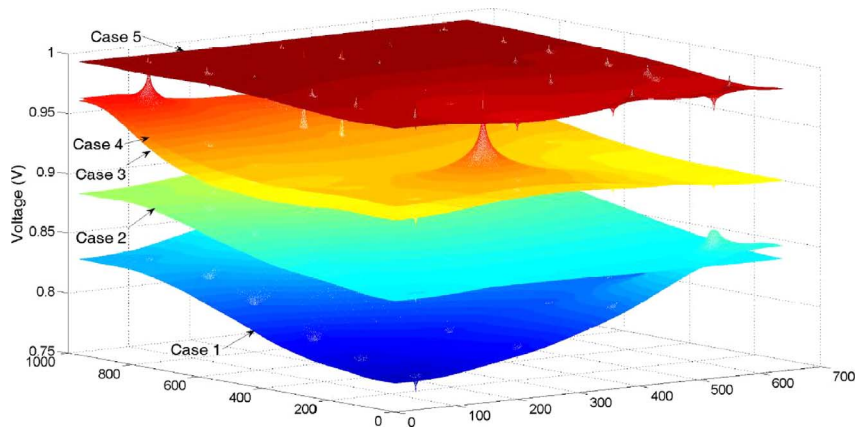


Fig. 8. Map of voltage drops within superblue10 for five different cases. The maximum and average voltage drop is reduced when the power supplies and decoupling capacitors are distributed.

the number of power supplies and decoupling capacitors increases, the power sources can be locally distributed throughout the large power distribution network, providing local current to the load circuits. Both the maximum and average power noise are therefore significantly reduced for different circuits with diverse floorplans.

VI. DISCUSSION

With the introduction of ultra-small on-chip voltage regulators [6], [22], the number of voltage regulators on a single die will increase significantly to maintain the increasingly stringent noise and power constraints in sub-20 nm ICs. Delivering a robust power supply voltage to circuits with varying noise and voltage constraints is crucial to maintaining the performance of next generation ICs. Local supply voltages are generated and regulated by point-of-load voltage regulators within a distributed power delivery system. The physical distance among the power sources and load circuits is less with a distributed power delivery system. Since the power source is placed physically closer to the load circuits, the inductive $L di/dt$ and resistive IR power noise is reduced.

In the proposed optimization methodology, minimizing the maximum voltage drop and response time for certain blocks is the primary optimization constraints. Other design constraints can also be incorporated within the proposed algorithm such

as minimizing the power consumption and on-chip area. The distinctive properties of the on-chip power supplies and decoupling capacitors should be further exploited to satisfy these constraints, while using limited system resources. Although the power supplies and decoupling capacitors both provide local charge to the load circuitry, a decoupling capacitor requires a power source to recharge after each clock cycle [3]. The decoupling capacitors provide a faster response with minimal power consumption (i.e., power is only consumed by the ESR of the decoupling capacitor). Alternatively, the power supplies dissipate significant power during voltage down conversion and regulation. A power supply, however, can provide continuous charge and does not need to be recharged after each clock cycle. Additionally, the proposed optimization techniques can handle a system with a maximum 20 power supplies and 32 decoupling capacitors, as listed in Table III, with the available computing resources and memory in feasible time (i.e., less than three hours). Heuristic algorithms will significantly increase the number of power sources that the proposed technique can handle.

VII. CONCLUSION

Distributed power delivery holds the promise of a significant paradigm shift, which will become necessary to achieve next generation power efficient systems. Circuit blocks with different

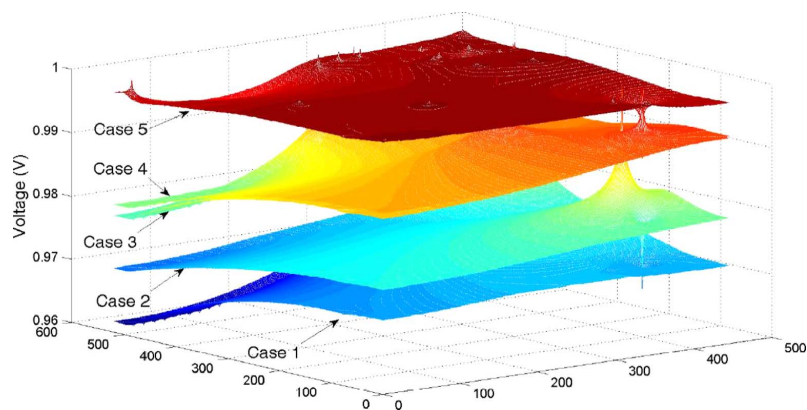


Fig. 9. Map of voltage drops within superblue12 for five different cases. The maximum and average voltage drop is reduced when the power supplies and decoupling capacitors are distributed.

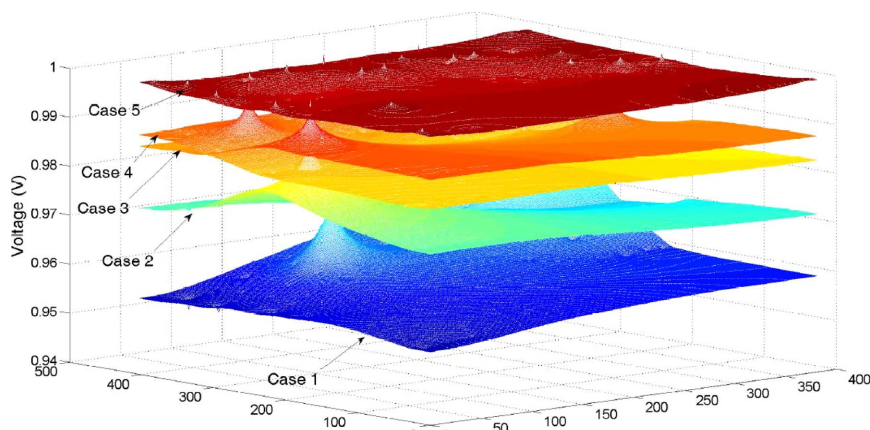


Fig. 10. Map of voltage drops within superblue18 for five different cases. The maximum and average voltage drop is reduced when the power supplies and decoupling capacitors are distributed.

TABLE VIII
MAXIMUM AND AVERAGE VOLTAGE DROP WITH 1 V POWER SUPPLY VOLTAGE WITHOUT ANY INCREASE IN AREA

	Case 1		Case 2		Case 3		Case 4		Case 5	
	Maximum voltage drop	Average voltage drop	Maximum voltage drop	Average voltage drop	Maximum voltage drop	Average voltage drop	Maximum voltage drop	Average voltage drop	Maximum voltage drop	Average voltage drop
superblue5	163 mV	130 mV	134 mV	115 mV	122 mV	73 mV	100 mV	69 mV	25 mV	9 mV
superblue10	241 mV	173 mV	166 mV	133 mV	106 mV	81 mV	98 mV	72 mV	22 mV	11 mV
superblue12	39 mV	28 mV	30 mV	24 mV	22 mV	12 mV	20 mV	13 mV	9 mV	3 mV
superblue18	47 mV	39 mV	38 mV	27 mV	27 mV	13 mV	20 mV	15 mV	10 mV	3 mV

voltage and noise constraints are commonly integrated onto a single die. With the introduction of ultra-small on-chip voltage regulators, a distributed on-chip power delivery system has become feasible. Novel techniques, however, are required to design and optimize this highly sophisticated and complex system. The similarity between the facility location problem and the design of heterogeneous ICs is exploited to determine the optimum number and location of the many distributed on-chip power supplies and decoupling capacitors in high performance ICs. An objective function based on the effective resistance among the power supplies, decoupling capacitors, and load circuits is proposed that minimizes the maximum voltage drop and response time throughout a high performance IC. This objective function considers the current contribution from the multiple power supplies and decoupling capacitors to each circuit block as well as the size of the individual circuit blocks. The optimal location

of the on-chip power supplies and decoupling capacitors is determined for four different ISPD'11 benchmark suite circuits. By exploiting the distributed nature of the local on-chip power supplies and decoupling capacitors, the local voltage fluctuations within a system with multiple power supplies and decoupling capacitors are minimized. The proposed methodology and techniques to determine the optimum location of the local power supplies and decoupling capacitors provide a means to realize more robust and efficient power delivery systems.

REFERENCES

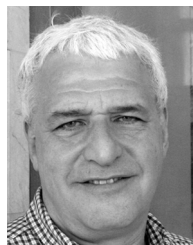
- [1] D. Meisner *et al.*, "Power management of online data-intensive services," in *Proc. ACM Int. Symp. Comp. Archit.*, Jun. 2011, pp. 319–330.
- [2] R. Jakushokas, M. Popovich, A. V. Mezhiba, S. Kose, and E. G. Friedman, *Power Distribution Networks with On-Chip Decoupling Capacitors*, 2nd ed. New York: Springer, 2011.

- [3] M. Popovich, M. Sotman, A. Kolodny, and E. G. Friedman, "Effective radii of on-chip decoupling capacitors," *IEEE Trans. Very Large Scale (VLSI) Circuits*, vol. 16, no. 7, pp. 894–907, Jul. 2008.
- [4] K. N. Leung and P. K. T. Mok, "A capacitor-free CMOS low-dropout regulator with damping-factor-control frequency compensation," *IEEE J. Solid-State Circuits*, vol. 38, no. 10, pp. 1691–1702, Oct. 2003.
- [5] P. Hazucha *et al.*, "Area-efficient linear regulator with ultra-fast load regulation," *IEEE J. Solid-State Circuits*, vol. 40, no. 4, pp. 933–940, Apr. 2005.
- [6] J. Guo and K. N. Leung, "A 6- μ W chip-area-efficient output-capacitorless LDO in 90-nm CMOS technology," *IEEE J. Solid-State Circuits*, vol. 45, no. 9, pp. 1896–1905, Sep. 2010.
- [7] Y. Ramadass, A. Fayed, B. Haroun, and A. Chandrakasan, "A 0.16 mm² completely on-chip switched-capacitor DC-DC converter using digital capacitance modulation for LDO Replacement in 45 nm CMOS," in *Proc. IEEE Int. Solid-State Circuits Conf.*, Feb. 2010, pp. 208–209.
- [8] S. Kose and E. G. Friedman, "An area efficient fully monolithic hybrid voltage regulator," in *Proc. IEEE Int. Symp. Circuits Syst.*, May/Jun. 2010, pp. 2718–2721.
- [9] S. Kose and E. G. Friedman, "On-chip point-of-load voltage regulator for distributed power supplies," in *Proc. ACM Great Lakes Symp. VLSI*, May 2010, pp. 377–380.
- [10] S. Kose and E. G. Friedman, "Distributed power network co-design with on-chip power supplies and decoupling capacitors," in *Proc. Workshop Syst. Level Interconnect Predict.*, Jun. 2011, pp. 1–5.
- [11] S. Kose and E. G. Friedman, "Simultaneous co-design of distributed on-chip power supplies and decoupling capacitors," in *Proc. IEEE Int. System-on-Chip Conf.*, Sep. 2010, pp. 15–18.
- [12] S. Kose and E. G. Friedman, "Effective resistance of a two layer mesh," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 58, no. 11, pp. 739–743, Nov. 2011.
- [13] K. Wang and M. Marek-Sadowska, "On-chip power-supply network optimization using multigrid-based technique," *IEEE Trans. Computer-Aided Design Integr. Circuits Syst.*, vol. 24, no. 3, pp. 407–417, Mar. 2005.
- [14] X.-D. S. Tan and C.-J. R. Shi, "Fast power/ground network optimization based on equivalent circuit modeling," in *Proc. IEEE/ACM Design Autom. Conf.*, Jun. 2001, pp. 550–554.
- [15] M. D. Pant, P. Pant, and D. S. Wills, "On-chip decoupling capacitor optimization using architectural level prediction," *IEEE Trans. Very Large Scale (VLSI) Syst.*, vol. 10, no. 3, pp. 319–326, Jun. 2002.
- [16] M. Popovich, E. G. Friedman, R. M. Secareanu, and O. L. Hartin, "Efficient placement of distributed on-chip decoupling capacitors in nanoscale ICs," in *Proc. IEEE/ACM Int. Conf. Comput.-Aided Design*, Nov. 2007, pp. 811–816.
- [17] M. Popovich, E. G. Friedman, M. Sotman, and A. Kolodny, "Efficient distributed on-chip decoupling capacitors for nanoscale ICs," *IEEE Trans. Very Large Scale (VLSI) Syst.*, vol. 16, no. 7, pp. 1717–1721, Jul. 2008.
- [18] Z. Zeng, X. Ye, Z. Feng, and P. Li, "Tradeoff analysis and optimization of power delivery networks with on-chip voltage regulation," in *Proc. IEEE/ACM Design Autom. Conf.*, Jun. 2010, pp. 831–836.
- [19] M. S. Daskin, *Network and Discrete Location: Models, Algorithms, and Applications*. New York: Wiley, 1995.
- [20] Z. Drezner and H. Hamacher, *Facility Location: Applications and Theory*. New York: Springer, 2002.
- [21] R. Z. Farahani, M. S. Seifi, and N. Asgari, "Multiple criteria facility location problems: A survey," *Appl. Math. Model.*, vol. 34, no. 7, pp. 1689–1709, Oct. 2010.
- [22] S. Kose, S. Tam, S. Pinzon, B. McDermott, and E. G. Friedman, "Active filter based hybrid on-chip DC-DC converters for point-of-load voltage regulation," *IEEE Trans. Very Large Scale (VLSI) Syst.*, to be published.
- [23] D. S. Hochbaum, "Heuristics for the fixed cost median problem," *Math. Program.*, vol. 22, no. 1, pp. 148–162, Jan. 1982.
- [24] M. I. Brandeau and S. S. Chiu, "An overview of representative problems in location research," *Manage. Sci.*, vol. 35, no. 6, pp. 645–674, Jun. 1989.
- [25] W. Miehle, "Link-length minimization in networks," *Oper. Res.*, vol. 6, no. 2, pp. 232–243, Mar.-Apr. 1958.
- [26] L. Cooper, "Location-allocation problems," *Oper. Res.*, vol. 11, no. 3, pp. 331–343, May-Jun. 1963.
- [27] P. Hansen, D. Peeters, and J.-F. Thisse, "An algorithm for a constrained Weber problem," *Manage. Sci.*, vol. 28, no. 11, pp. 1285–1295, Nov. 1982.
- [28] N. Viswanathan *et al.*, "The ISPD-2011 routability-driven placement contest and benchmark suite," in *Proc. ACM Int. Symp. Phys. Design*, Mar. 2011, pp. 141–146.
- [29] A. Brooke, D. Kendrick, and A. Meeraus, *GAMS: A User's Guide*. San Francisco, CA: Scientific, 1992.



Selçuk Köse (S'10–M'12) received the B.S. degree in electrical and electronics engineering from Bilkent University, Ankara, Turkey, in 2006, and the M.S. and Ph.D. degrees in electrical engineering from the University of Rochester, Rochester, NY, in 2008 and 2012, respectively.

He is currently an Assistant Professor with the Department of Electrical Engineering, University of South Florida, Tampa. He was a part-time Engineer with the VLSI Design Center, Scientific and Technological Research Council (TUBITAK), Ankara, Turkey, where he worked on low-power ICs in 2006. During the summers of 2007 and 2008, he was with the Central Technology and Special Circuits Team in the enterprise microprocessor division of Intel Corporation, Santa Clara, CA, where he was responsible for the functional verification of a number of blocks in the clock network including the de-skew machine and optimization of the reference clock distribution network. In the summer of 2010, he interned in the RF, Analog, and Sensor Group, Freescale Semiconductor, Tempe, AZ, where he developed design techniques and methodologies to reduce electromagnetic emissions. His current research interests include the analysis and design of high performance integrated circuits, monolithic DC-DC converters, and interconnect related issues with specific emphasis on the design and analysis of power and clock distribution networks, 3-D integration, and emerging integrated circuit technologies. He is an Associate Editor of the *Journal of Circuits, Systems, and Computers*.



Eby G. Friedman (F'00) received the B.S. degree from Lafayette College, Easton, PA, in 1979, and the M.S. and Ph.D. degrees from the University of California, Irvine, in 1981 and 1989, respectively, all in electrical engineering.

From 1979 to 1991, he was with Hughes Aircraft Company, rising to the position of Manager of the Signal Processing Design and Test Department, responsible for the design and test of high performance digital and analog IC's. He has been with the Department of Electrical and Computer Engineering at the University of Rochester, Rochester, NY, since 1991, where he is a Distinguished Professor, and the Director of the High Performance VLSI/IC Design and Analysis Laboratory. He is also a Visiting Professor at the Technion-Israel Institute of Technology. His current research and teaching interests are in high performance synchronous digital and mixed-signal microelectronic design and analysis with application to high speed portable processors and low power wireless communications. He is the author of over 400 papers and book chapters, a dozen patents, and the author or editor of 15 books in the fields of high speed and low power CMOS design techniques, 3-D design methodologies, high speed interconnect, and the theory and application of synchronous clock and power distribution networks.

Dr. Friedman is the Regional Editor of the *Journal of Circuits, Systems and Computers*, a member of the editorial boards of the *Analog Integrated Circuits and Signal Processing*, *Microelectronics Journal*, *Journal of Low Power Electronics*, *Journal of Low Power Electronics and Applications*, and IEEE JOURNAL ON EMERGING AND SELECTED TOPICS IN CIRCUITS AND SYSTEMS, Chair of the IEEE TRANSACTIONS ON VERY LARGE SCALE INTEGRATION (VLSI) SYSTEMS steering committee, and a member of the technical program committee of a number of conferences. He previously was the Editor-in-Chief of the IEEE TRANSACTIONS ON VERY LARGE SCALE INTEGRATION (VLSI) SYSTEMS, a Member of the editorial board of the PROCEEDINGS OF THE IEEE, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II: ANALOG AND DIGITAL SIGNAL PROCESSING, and *Journal of Signal Processing Systems*, a Member of the Circuits and Systems (CAS) Society Board of Governors, Program and Technical Chair of several IEEE conferences, and a recipient of the University of Rochester Graduate Teaching Award and a College of Engineering Teaching Excellence Award. He is a Senior Fulbright Fellow.