Nanoscale On-Chip Decoupling Capacitors

Mikhail Popovich
Qualcomm, Inc.
5775 Morehouse Drive
San Diego, California 92121, USA
mikhailp@qualcomm.com

Eby G. Friedman
Dept. of Electrical and Computer Engineering
University of Rochester
Rochester, New York 14627, USA
friedman@ece.rochester.edu

***Abstract*— A distributed on-chip decoupling capacitor network is proposed in this paper to replace one large capacitor. A system of distributed on-chip decoupling capacitors is shown to provide an efficient solution for providing the required on-chip decoupling capacitance under existing technology constraints. In a system of distributed on-chip decoupling capacitors, each capacitor is sized based on the parasitic impedance of the power distribution grid. To be effective, on-chip decoupling capacitors should be charged before the next switching cycle. A design space for estimating the maximum parasitic resistance of the power distribution grid during discharge and the charging phase has been determined. Related simulation results for typical values of the on-chip parasitic resistance are also presented. An analytic solution is shown to provide accurate parameters of the distributed system. The worst case error is 0.002% during discharge and 0.08% during the charging phase as compared to SPICE.**

I. Introduction

Decoupling capacitors are widely used to manage the power supply noise [1]. Decoupling capacitors are an effective way to reduce the impedance of power delivery systems operating at high frequencies [2]. A decoupling capacitor acts as a local reservoir of charge, which is released when the power supply voltage at a particular current load drops below some tolerable level. Since the inductance scales slowly [3], the location of the decoupling capacitors significantly affects the design of the power/ground (P/G) networks in high performance integrated circuits (ICs) such as microprocessors. At higher frequencies, a distributed system of decoupling capacitors are placed on-chip to effectively manage the power supply noise [4].

As described in [5], to be effective, an on-chip decoupling capacitor should be placed such that both the power supply and the current load are located inside the appropriate effective radius. The efficient placement of on-chip decoupling capacitors in nanoscale ICs is the subject of this paper. Unlike the methodology for placing a single on-chip decoupling capacitor presented in [5], a system of *distributed* on-chip decoupling capacitors is proposed to replace a single decoupling capacitor.

Decoupling capacitors have traditionally been allocated into the white space (those areas not occupied by the circuit elements) available on a die [6]. In this way, decoupling capacitors are often placed at a significant distance from the current load. The efficacy of the distant on-chip decoupling capacitors is greatly reduced. As a result, larger on-chip decoupling capacitors are required. The conventional allocation strategy, therefore, results in increased power noise, compromising the signal integrity of an entire system.

To be effective, a decoupling capacitor should be placed physically close to the current load. This requirement is naturally satisfied in board and package applications, since large capacitors are much smaller than the dimensions of the circuit board (or package) [7]. In this case, a lumped model of the decoupling capacitor provides sufficient accuracy [8].

The size of an on-chip decoupling capacitor, however, is directly proportional to the area occupied by the capacitor and can require a significant portion of the on-chip area. The minimum impedance between an on-chip capacitor and the current load is fundamentally affected by the magnitude (and therefore the area) of the capacitor. Systematically partitioning the decoupling capacitor solves this issue. A system of distributed on-chip decoupling capacitors is illustrated in Fig. 1.

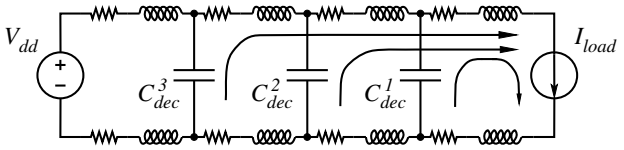


Fig. 1. A network of distributed on-chip decoupling capacitors. The magnitude of the decoupling capacitors is based on the impedance of the interconnect segment connecting a specific capacitor to a current load.

In a system of distributed on-chip decoupling capacitors, each decoupling capacitor is sized based on the impedance of the interconnect segment connecting the capacitor to the current load. A particular capacitor only provides charge to a current load during a short period. The rationale behind the proposed scheme can be explained as follows. The capacitor closest to the current load is engaged immediately after the switching cycle is initiated. Once the first capacitor is depleted, the next capacitor is activated, providing a large portion of the total current drawn by the load. This procedure is repeated until the last capacitor becomes active. Similar to the hierarchical placement of decoupling capacitors presented in [4], the proposed technique provides an efficient solution for providing the required on-chip decoupling capacitance based on specified capacitance density constraints. A system of distributed on-chip decoupling capacitors should therefore be utilized to provide a low impedance, cost effective power delivery network in nanoscale ICs.

The paper is organized as follows. The design of an on-chip distributed decoupling capacitor network is presented in Section II. Charging the system of distributed on-chip decoupling capacitors is discussed in Section III. Related simulation results for typical values of the on-chip parasitic resistance

978-1-4244-2596-9/08/$25.00 ©2008 IEEE

are discussed in Section IV. Some specific conclusions are summarized in Section V.

II. Design of a System of Distributed On-Chip Decoupling Capacitors

Similar to a chain of tapered buffers, a network of on-chip decoupling capacitors with progressively increasing magnitude is proposed to replace a large on-chip decoupling capacitor, as illustrated in Fig. 2. The magnitude of each decoupling capacitor in the distributed network is determined based on the parasitic resistance of the power/ground lines connecting the decoupling capacitors and the current load. Note that as the capacitor is placed farther from the current load, the magnitude of the on-chip decoupling capacitor is increased due to the larger available area.

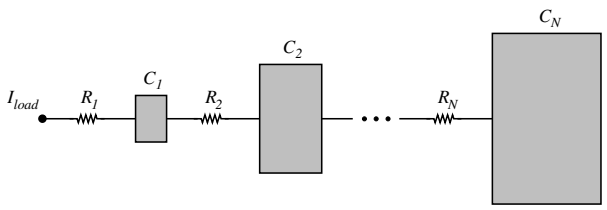


Fig. 2. A physical model of the proposed system of distributed on-chip decoupling capacitors.

A circuit model of the proposed system of distributed on-chip decoupling capacitors is shown in Fig. 3. For simplicity, two decoupling capacitors are assumed to provide the required charge drawn by the current load. The impedance of the metal lines connecting the capacitors to the current load is modeled as resistors, R_1 and R_2. Note that the on-chip inductance of the power/ground lines is neglected due to the significantly increased complexity of the characterizing system of equations. A system of distributed on-chip decoupling capacitors which includes the parasitic inductance of the metal lines is analytically non-tractable. A triangular current source is assumed to model the current load. The magnitude of the current source increases linearly, reaching the maximum current I_{max} at rise time t_r, *i.e.*, $I_{load}(t) = I_{max}\frac{t}{t_r}$. The maximum tolerable ripple at the current load is 10% of the power supply voltage.

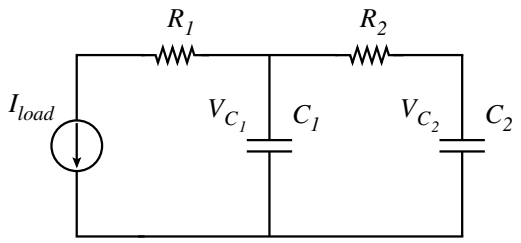


Fig. 3. Circuit model of an on-chip distributed decoupling capacitor network. The impedance of the metal lines is modeled as R_1 and R_2, respectively.

The voltage across the decoupling capacitors at the end of the switching cycle ($t = t_r$) can be determined from Kirchhoff's laws [9]. Writing KVL and KCL for each of the loops (see Fig. 3), the system of differential equations for the voltage across C_1 and C_2 at t_r is

$$\frac{dV_{C_1}}{dt} = \frac{V_{C_2} - V_{C_1}}{R_2 C_1} - \frac{I_{load}}{C_1}, \tag{1}$$

$$\frac{dV_{C_2}}{dt} = \frac{V_{C_1} - V_{C_2}}{R_2 C_2}. \tag{2}$$

Simultaneously solving (1) and (2) and applying the initial conditions, the voltage across C_1 and C_2 at the end of the switching activity can be determined as described in [10].

In general, to determine the parameters of the system of distributed on-chip decoupling capacitors, the following assumptions are made. The parasitic resistance of the metal line(s) connecting capacitor C_1 to the current load is known. R_1 is determined by technology constraints (the sheet resistance) and by design constraints (the maximum available metal resources). Note that the voltage across C_2 after discharge is also a design parameter. Since the capacitor C_2 is directly connected to the power supply (a shared power rail), the voltage drop across C_2 appears on the global power line, compromising the signal integrity of the overall system.

The system of equations to determine the parameters of an on-chip distributed decoupling capacitor network as depicted in Fig. 3 is

$$V_{load} = V_{C_1} - I_{max} R_1, \tag{3}$$

$$V_{C_1} = f(C_1, C_2, R_2), \tag{4}$$

$$V_{C_2} = f(C_1, C_2, R_2), \tag{5}$$

$$\frac{I_{max} t_r}{2} = C_1 \left(V_{dd} - V_{C_1}\right) + C_2 \left(V_{dd} - V_{C_2}\right), \tag{6}$$

where V_{C_1} and V_{C_2} are the voltage across C_1 and C_2, respectively. Equation (6) states that the total charge drawn by the current load is provided by C_1 and C_2. Note that in the general case with the current load determined *a priori*, the total charge is the integral of $I_{load}(t)$ from zero to t_r. Solving (3) for V_{C_1} and substituting into (4), C_1, C_2, and R_2 are determined from (4) to (6).

III. Charging Distributed On-Chip Decoupling Network

Once discharged, the decoupling capacitors must be charged to support the current demands during the next switching event. If the charge on the capacitors is not fully restored during the relaxation time between two consecutive switching events (the charge time), the system of distributed decoupling capacitors will gradually be depleted, becoming ineffective after several clock cycles. A maximum resistance between the power supply and the distributed decoupling capacitor network therefore exists, which is determined during the charging phase for a target charge time. Similar to discharge, the system of distributed on-chip decoupling capacitors should be placed close to the power supply (within the maximum effective radii [5]) to be effective.

A circuit charging a system of distributed on-chip decoupling capacitors is shown in Fig. 4. The impedance of the metal

lines connecting the capacitors to the power supply is modeled as resistors R_2 and R_3. Note that the initial voltage $V_{C_1}^0$ and $V_{C_2}^0$ are determined by the voltage drop during discharge. Also note that for a specific charge time t_{ch}, R_3 is the only design parameter (R_1, R_2, C_1, and C_2 are determined from the discharge phase).

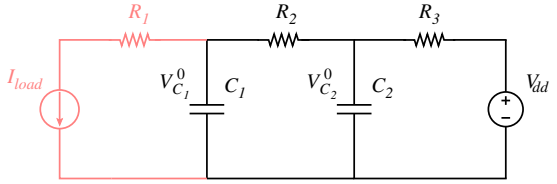


Fig. 4. A circuit charging a system of distributed on-chip decoupling capacitors. The impedance of the metal lines is modeled as R_2 and R_3, respectively.

The voltage across the decoupling capacitors during the charging phase can be determined from KVL and KCL for each loop (see Fig. 4). Similar to discharge, the system of differential equations for the voltage across C_1 and C_2 at t_{ch} is

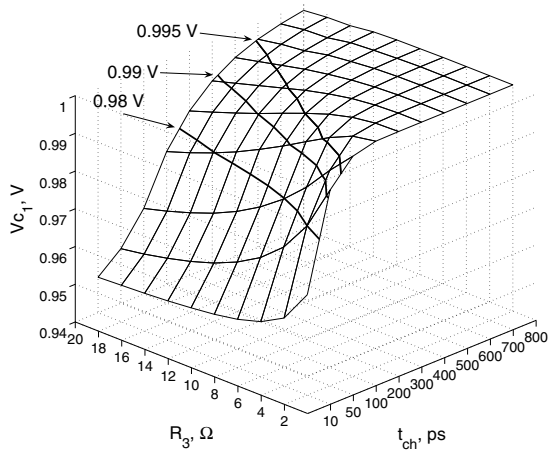
$$\frac{dV_{C_1}}{dt} = \frac{V_{C_2} - V_{C_1}}{R_2 C_1}, \tag{7}$$

$$\frac{dV_{C_2}}{dt} = \frac{V_{dd} - V_{C_2}}{R_3 C_2} - \frac{V_{C_2} - V_{C_1}}{R_2 C_2}. \tag{8}$$

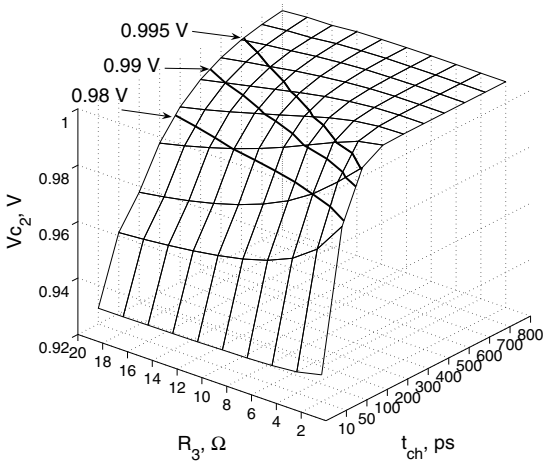
Simultaneously solving (7) and (8) and applying the initial conditions, the voltage across C_1 and C_2 at the end of the charging phase can be determined. The maximum resistance R_3 between C_2 and the power supply is determined such that both C_1 and C_2 are charged to the power supply voltage.

Note that the expressions for determining R_3 are transcendental. A closed-form solution, therefore, cannot be determined. A design space can be graphically obtained for determining the maximum tolerable resistance R_3 (or physical distance), permitting the capacitors to be successfully charged by the target charge time. The voltage across C_1 and C_2 during the charging phase as a function of R_3 and t_{ch} is depicted in Fig. 5.

Observe from Fig. 5 that the voltage across C_1 and C_2 increases exponentially with the charge time, approaching the power supply voltage for large t_{ch}. Note that as R_3 increases, a larger t_{ch} is required to restore the charge on the decoupling capacitors. Alternatively, the maximum effective frequency of the distributed on-chip decoupling network is reduced. Ideally, the decoupling capacitors should be charge to V_{dd}, substantially increasing the required charge time. In practical applications, however, a lower voltage across the decoupling capacitors can be tolerated, significantly increasing the maximum parasitic resistance for a specific t_{ch} (see the black equipotential lines plotted in Fig. 5).



(a)



(b)

Fig. 5. Voltage across C_1 (a) and C_2 (b) during the charging phase as a function of R_3 and t_{ch}: $C_1 = 2\,\text{pF}$, $C_2 = 10\,\text{pF}$, $R_1 = 5\,\Omega$, $V_{C_1}^0 = 0.91\,\text{V}$, $V_{C_2}^0 = 0.95\,\text{V}$, and $V_{dd} = 1\,\text{V}$.

IV. Case Study

The dependence of a system of distributed on-chip decoupling capacitors on the current load and the parasitic impedance of the power delivery system is described in this section to quantitatively illustrate the previously presented concepts. Resistive power and ground lines are assumed and are modeled as resistors. The load is modeled as a ramp current source with a 100 ps rise time. The minimum tolerable voltage across the load terminals is 90% of the power supply voltage. The magnitude of the on-chip decoupling capacitors for several parasitic resistances of the metal lines connecting the capacitors to the current load during discharge is listed in Table I. The voltage across C_1 and C_2 as a function of R_3 during the charging phase is listed in Table II. Note that the values of R_1, R_2, and R_3 are typical parasitic resistances of an on-chip power distribution grid for a 90 nm CMOS technology.

The parameters of the system of distributed on-chip decou-

TABLE I

THE MAGNITUDE OF ON-CHIP DECOUPLING CAPACITORS AS A FUNCTION OF THE PARASITIC RESISTANCE OF THE POWER/GROUND LINES CONNECTING THE CAPACITORS TO THE CURRENT LOAD

R_1	R_2	I_{max}	C_1	C_2	V_{load} (mV)		Error	V_{C_2} (mV)		Error
(Ω)	(Ω)	(A)	(pF)	(pF)	V_{load}^{min}	SPICE	(%)	$V_{C_2}^{min}$	SPICE	(%)
0.5	4.5	0.01	0	9.99999	900	899.999	0.0001	950	949.999	0.0001
0.5	6	0.01	1.59747	6.96215	900	899.986	0.002	950	949.983	0.002
0.5	8	0.01	2.64645	4.97091	900	899.995	0.0006	950	949.993	0.0004
0.5	10	0.01	3.22455	3.87297	900	899.997	0.0003	950	949.996	0.0004
0.5	12	0.01	3.59188	3.17521	900	899.998	0.0002	950	949.997	0.0003
0.5	14	0.01	3.84641	2.69168	900	899.998	0.0002	950	949.997	0.0003
0.5	16	0.01	4.03337	2.33650	900	899.999	0.0001	950	949.998	0.0002
0.5	18	0.01	4.17658	2.06440	900	899.998	0.0002	950	949.998	0.0002
0.5	20	0.01	4.28984	1.84922	900	899.999	0.0001	950	949.998	0.0002

$V_{dd} = 1\,\text{V}$ and $t_r = 100\,\text{ps}$

TABLE II

THE VOLTAGE ACROSS C_1 AND C_2 AS A FUNCTION OF R_3 DURING THE CHARGING PHASE

R_3	V_{C_1} (mV)		Error	V_{C_2} (mV)		Error
(Ω)	Analytic	SPICE	(%)	Analytic	SPICE	(%)
2	999.973	999.978	-0.0005	999.957	999.965	-0.0008
4	998.988	999.073	-0.009	998.735	998.842	-0.001
6	996.272	996.424	-0.02	995.688	995.864	-0.02
8	992.735	992.857	-0.01	991.908	992.043	-0.01
10	989.111	989.127	-0.002	988.137	988.154	-0.002
12	985.714	985.583	0.01	984.661	984.521	0.01
14	982.640	982.347	0.03	981.554	981.242	0.03
16	979.900	979.441	0.05	978.806	978.322	0.05
18	977.467	976.847	0.06	976.382	975.733	0.07
20	975.306	974.533	0.08	974.240	973.434	0.08

$V_{dd} = 1\,\text{V}$, $t_{ch} = 200\,\text{ps}$, $R_2 = 5\,\Omega$, $C_1 = 2\,\text{pF}$, and $C_2 = 10\,\text{pF}$

pling capacitors are analytically determined from (3)–(6). The resulting power supply noise during discharge is estimated using SPICE and compared to the maximum tolerable level (the minimum voltage across the load terminals V_{load}^{min}). The maximum voltage drop across C_2 at the end of the switching activity is also estimated and compared to $V_{C_2}^{min}$. During the charging phase, the voltage across the distributed decoupling capacitors is analytically determined from (7) and (8) and compared to SPICE. Note that the analytic solution produces an accurate estimate of the on-chip decoupling capacitors for typical parasitic resistances of a power distribution grid. The maximum error in this case study is 0.002% for discharge and 0.08% for the charging phase, respectively.

V. CONCLUSIONS

A methodology for efficiently placing distributed on-chip decoupling capacitors to replace one large capacitor has been proposed in this paper. A distributed on-chip decoupling capacitor network is an efficient solution for providing the required on-chip decoupling capacitance under existing technology constraints in nanoscale ICs. In a system of distributed on-chip decoupling capacitors, each capacitor is sized based on the parasitic impedance of the power delivery system. Hierarchically allocating the on-chip decoupling capacitors greatly relaxes the technology constraints for physically distant capacitors, making the distant on-chip decoupling capacitors more effective.

Analytic expressions has been developed for determining the parameters of the system for both the discharge and charging phases. Related simulation results for typical on-chip parasitic resistances are also presented, demonstrating the high accuracy of the analytic solution. In the worst case, the maximum error is 0.002% during discharge and 0.08% during charging as compared to SPICE.

REFERENCES

[1] M. Popovich, A. V. Mezhiba, and E. G. Friedman, *Power Distribution Networks with On-Chip Decoupling Capacitors*, Springer Verlag, 2008.

[2] M. Popovich and E. G. Friedman, "Impedance Characteristics of Decoupling Capacitors in Multi-Power Distribution Systems," *Proceedings of the IEEE International Conference on Electronics, Circuits and Systems*, pp. 160–163, December 2004.

[3] A. V. Mezhiba and E. G. Friedman, "Scaling Trends of On-Chip Power Distribution Noise," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 12, No. 4, pp. 386–394, April 2004.

[4] M. Popovich and E. G. Friedman, "Decoupling Capacitors for Multi-Voltage Power Distribution Systems," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 14, No. 3, pp. 217–228, March 2006.

[5] M. Popovich, E. G. Friedman, M. Sotman, A. Kolodny, and R. M. Secareanu, "Maximum Effective Distance of On-Chip Decoupling Capacitors in Power Distribution Grids," *Proceedings of the ACM Great Lakes Symposium on VLSI*, pp. 173–179, April/May 2006.

[6] S. Zhao, K. Roy, and C.-K. Koh, "Decoupling Capacitance Allocation and Its Application to Power-Supply Noise-Aware Floorplanning," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 21, No. 1, pp. 81–92, January 2002.

[7] J. Kim, *et al.*, "Separated Role of On-Chip and On-PCB Decoupling Capacitors for Reduction of Radiated Emission on Printed Circuit Board," *Proceedings of the IEEE International Symposium on Electromagnetic Compatibility*, pp. 531–536, August 2001.

[8] T. Hubing, "Effective Strategies for Choosing and Locating Printed Circuit Board Decoupling Capacitors," *Proceedings of the IEEE International Symposium on Electromagnetic Compatibility*, pp. 632–637, August 2005.

[9] M. E. Van Valkenburg, *Network Analysis*, Prentice-Hall, 1974.

[10] M. Popovich, E. G. Friedman, R. M. Secareanu, and O. L. Hartin, "Efficient Placement of Distributed On-Chip Decoupling Capacitors in Nanoscale ICs," *Proceedings of the ACM/IEEE International Conference on Computer-Aided Design*, pp. 811–816, November 2007.