Active On-Die Suppression of Power Supply Noise

Gökçe Keskin, Xin Li and Larry Pileggi

Carnegie Mellon University

Dept. of ECE, 5000 Forbes Ave. Pittsburgh, PA 15213 USA

Email: {gkeskin, xinli, pileggi}@andrew.cmu.edu

Abstract- An active on-chip circuit is demonstrated in 130nm CMOS for the suppression of on-chip power supply noise due to power distribution resonance. Testchip measurement results indicate up to 40% reduction in power supply noise during clock/power gating at a 2% power and 6% area overhead cost. Oscillation time is reduced by 50%. Simulation results show that comparable overshoot/undershoot and ringing control via on-chip decoupling would require significantly more area and power due to leakage, particularly at 90nm and below.

Keywords: power supply noise, L.di/dt noise, damping and decoupling capacitor

INTRODUCTION

Even with reduced power supply voltage for emerging process nodes, the maximum power consumption of high-performance circuits has been projected to continue to increase for the foreseeable future (Fig. 1, [1]). Therefore, such circuits generally include clock and power gating schemes to reduce power consumption [2]. However, when idle regions are switched back on as required under normal operation, the current demand increases rapidly in a short period of time (at most a few nanoseconds). Ultimately, this extra current has to be supplied by the board to the chip through the inductive bonding connections between the chip, package, and the board. This current step creates noise on the on-chip power rails, commonly called as L.di/dt noise, also known as simultaneous switching noise.

The traditional solution for reducing power supply noise is to use on-chip decoupling capacitors, along with on-package and on-board capacitors to supply instantaneous current demand [3]. However, addition of these capacitors causes undesired resonances in the frequency domain, which translate to oscillations in the transient response. The most dominant of these resonances is due to the package inductance and on-chip decoupling capacitance that is generally observed around 100-150MHz. Fig. 2 shows transient and frequency domain simulation results for a simplified model of a microprocessor under clock gating. The positive and negative power rail peaks in the transient simulation can result in timing and reliability problems, as well as loss of stored data.

As power supply voltages scale down and noise margins become tighter in new generation process nodes, even more on-chip decoupling (generally in the form of MOS capacitors) is required for high-performance circuits to supply enough charge for noise suppression. Unfortunately, there are diminishing returns for adding more decoupling capacitors on the chip. Fig.3 shows the simulated total power supply noise overshoots in a 130nm test circuit for different values of on-chip decoupling.

This diminishing return is partially attributable to the supply rail oscillations due to the underdamped nature of the power grid distribution network. Note that this low frequency oscillation (Fig. 2) is a function of the chip and package, but excited by the step responses due to the clock/power gating. Additional damping may be provided by introducing dissipative elements, rather than more capacitance. In 2003, Ji described a passive resistor in series with the decoupling capacitors, but this approach reduces the efficiency of the on-chip capacitors for controlling high frequency, localized power rail noise [4]. Gabara proposed the use of active devices in series with the inductive bonding to introduce resistance; however this is not applicable in high performance circuits where the IR drop would be significant [5]. Larsson discussed adding a passive resistance in parallel with the decoupling capacitors, but deemed it infeasible due to excessive DC power dissipation [6].

In this paper we describe the results for an *active* resistor in parallel with the on-chip decoupling capacitors. This active resistor provides good damping in the AC domain at a significantly reduced DC power dissipation penalty.
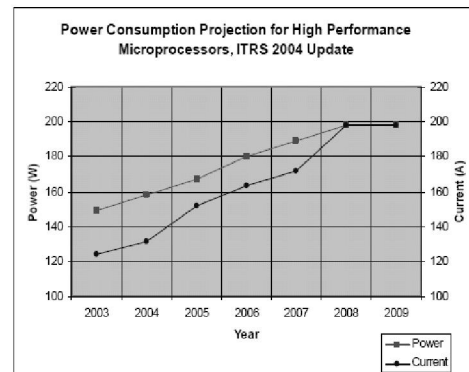

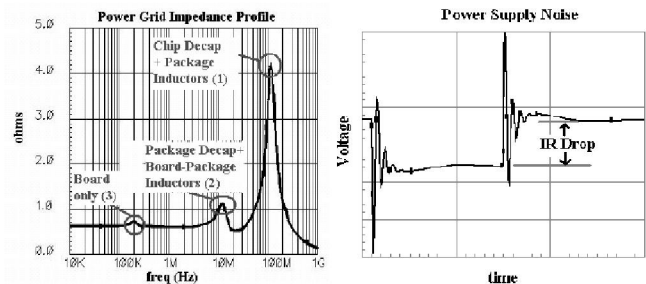Fig.1. Power consumption trend in microprocessors [1]

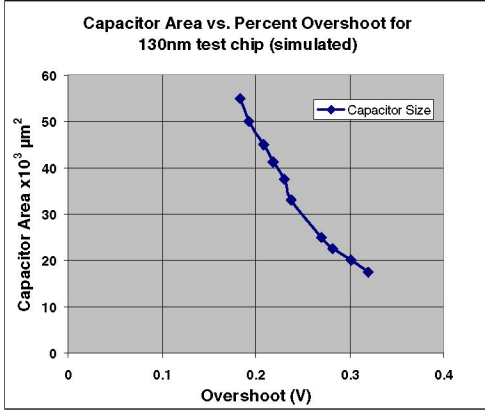Fig.2. Power supply noise due to resonance

Fig.3. Power Supply Noise vs. on-chip decoupling

Fig.4. Simplified small-signal power grid network model

ACTIVE DAMPING

A simplistic small-signal model of an IC power grid distribution is a parallel RLC circuit (Fig. 4). L represents the inductive bonding connections, C is the on-chip decoupling capacitance, and R is the added damping resistance. I is the step current disturbance caused by clock-gating.

The transfer function from input I to output V_{out} is:

$$Z(s) = \frac{s/C}{s^2 + 2\zeta w_0 s + w_0^2} \qquad (1)$$

where:

$$w_0 = \frac{1}{\sqrt{LC}}, \qquad \zeta = \frac{1}{2R}\sqrt{\frac{L}{C}} \qquad (2)$$

The step response of this system is:

$$V_{out}(t) = \frac{w_0 L}{\sqrt{1-\zeta^2}} e^{-\zeta w_0 t} \sin\left(w_0\left(\sqrt{1-\zeta^2}\right)t\right) \qquad (3)$$

One can determine that as the damping ratio ζ increases, the overshoot in the step response decreases. For the parallel RLC circuit, this can be achieved by a low R value. Adding extra damping reduces the impedance of the power grid distribution in the frequency domain and that translates to smaller peaks with shorter duration in the transient response. The peaks of the impedance profile (Fig. 2) are reduced, resulting in a flatter response.

The choice of the amount of resistance to be added depends on the L and C values of the distribution. In an over-damped system, where there are no oscillations present, ζ should be greater than 1. From (2), we get the upper bound of R as:

$$\zeta = \frac{1}{2R}\sqrt{\frac{L}{C}} > 1 \Rightarrow R < \frac{1}{2}\sqrt{\frac{L}{C}} \qquad (4)$$

The lower the resistance, the lower the overshoots; however that comes with an increased power consumption trade-off. Adding a conventional resistor in parallel to the power grid network would increase the power consumption significantly. In reality, damping is only required in the frequency domain around the resonance frequency, rather than at all frequencies as provided by a conventional resistor. We can exploit this damping requirement by using active devices.

The proposed active resistor topology is given in Fig. 5. Transistors M2-M4 amplify the noise on the V_{dd} rail and apply this voltage to the gate of M1, which behaves as a $1/g_m$ resistance. The total small signal resistance of M1-M4 block is $1/(K.g_{m1})$, where K ($=g_{m4}/g_{m2}$) is the amplification factor of M2-M4. M3 is added to increase g_{m4} while keeping g_{m2} smaller; hence increasing K and lowering R. M1 is biased on the edge of conduction so that it only responds to positive peaks on the V_{dd} with minimal DC power dissipation. Transistors M5-M8 are similar to M1-M4, but they respond to negative peaks using a higher supply voltage (e.g. the I/O supply), V_{dd2}. If V_{dd2} is not available, M1-M4 and with M1 biased above its edge of conduction will also respond to negative peaks, but at the cost of increased DC power dissipation since M1 is a relatively large transistor to provide sufficient g_{m1}. This allows the elimination of the upper resistor block. If the clock/power gating signal can be anticipated *a priori*, active resistors can be shut off when the transient noise dies out and then turned on again before switching; saving extra static power. $V_{bias1,2}$ are referenced to ground, $V_{bias3,4}$ are referenced to V_{dd}.
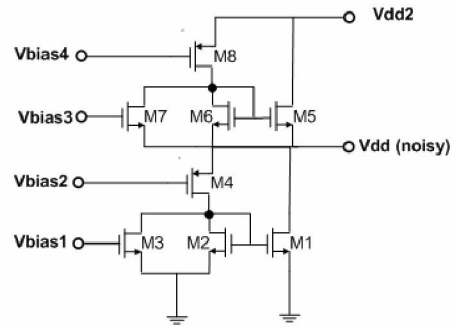

Fig. 5. Active resistor topology

TEST RESULTS

A test chip in 130nm CMOS has been designed and fabricated to verify the proposed method. The test chip

consists of high frequency (HF) ring oscillators (at 2.3GHz) that are gated in the chain with an AND gate that can be connected to either an on-chip low frequency (LF) ring oscillator (at 5MHz, Fig. 6) or an external gating signal. HF oscillators emulate the high speed switching circuits that are being gated in a modern processor, whereas the LF oscillator provides a clock gating signal at a low enough frequency to act as a step disturbance to observe the oscillations without switch-on and switch-off events affecting each other. The gating signal is distributed across the chip in an H-tree routing to provide the turn on of all HF oscillators at the same time. The switch-on event provides enough di/dt to observe the noise on the chip.
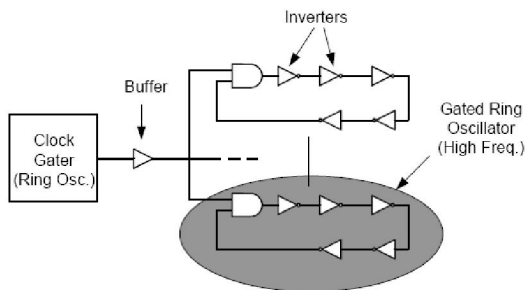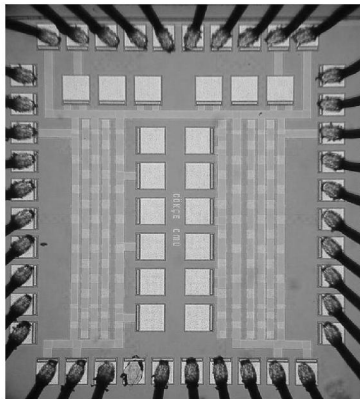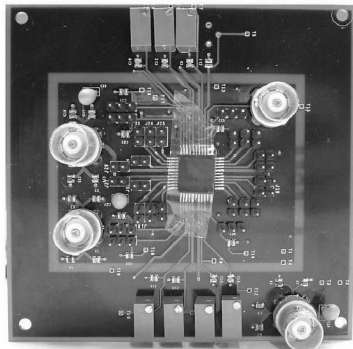

Fig. 6. Chip Diagram

Fig.7. Die photograph

Fig.8. PCB photograph

The die photograph is given in Fig. 7. The top three metal layers in the process are used for the power grid distribution and they are strapped at each layer using vias. There are redundant V_{dd}/Gnd pads on the chip to be connected to the package that can either be connected to the board power/ground or left alone (to control total inductance). Control signals carried to the die allow selective turning on a certain number of high frequency ring oscillators (to control di/dt) and the selection of either on-chip or external clock gating signals. Several internal V_{dd}/Gnd pads are provided for possible wafer probing.

The chip is packaged in a 44-pin LQFP package and soldered onto a 4-layer PCB where two intermediate layers are used for power/ground (Fig. 8). Bias voltages are generated by an external DC source, and measurements are taken using a Agilent 54855A oscilloscope with Agilent 1134A high impedance probe to prevent loading. Both on-chip decoupling in the form of nMOS capacitors (250pF) and on-board decoupling (c0805 capacitors with values ranging from 10pF to 10μF) are used. On-board capacitors are soldered as close as possible to the V_{dd}/Gnd pins of the package on the PCB to minimize the inductance of the PCB routing path. On-board ESD protection circuits are also implemented. Sense nodes on the board are connected to the chip rails, but not to the board rails, to provide access points for transient measurements. Positive peaks are reduced by 40%, whereas negative peaks are reduced by 15% (Fig. 9). Oscillation duration directly determines when the circuits are usable (when V_{dd} is stable), and since the resonance frequency is considerably lower than the clock frequency by a minimum of approximately ten times, longer oscillation times are highly undesirable. Longer oscillation durations result in the waste of many clock cycles. Positive peaks incur longer oscillations due to lower damping of the system, so the active resistors are very beneficial in this case providing around 50% reduction in oscillation time. The asymmetry in the droop and overshoot reduction is partially attributable to the inherent on-chip damping when all gates are actively switching.

For the measurements shown, V_{dd}=1.2V, V_{dd2}=2.2V, full current consumption of digital switching blocks is 49.53mA, and total current consumption of all active resistor circuits is 1.04mA. This translates to approximately 2% power overhead due to active resistors. The total area of the circuits and on-chip decoupling, including the active resistors, is $0.115mm^2$. Active resistors consume 6% of this area. It should be noted that in a production design, the same switching current would be realized by a larger die area since the activity factor of the ring oscillators is high, which would translate into a smaller area penalty than what is reported for this demonstration. From the simulation results of the test-chip we observe that 50% more decoupling capacitance would be required for 20% reduction in power rail overshoot. If the same design were implemented in 45nm CMOS, this would correspond to a 50% increase in gate area, hence gate-leakage current (which would be expected to be as high as 115mA for this design, [1]), whereas the active resistors would still require only 6% of extra gate area while providing 40% overshoot reduction.

CONCLUSIONS

An active resistor circuit is demonstrated for decreasing on-chip power supply noise. Peak noise amplitudes are reduced by 40% for overshoots and 15% for undershoots with power and area overheads which is substantially less than that required for comparable control with on-chip decoupling. Furthermore, the active control can be switched on and off in anticipation of clock/power gating for further power reduction.
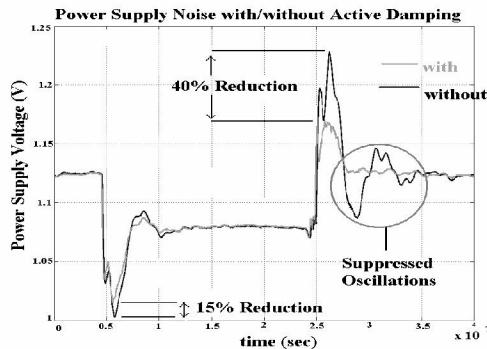

Fig. 9. Transient domain measurement results

REFERENCES

[1] International Technology Roadmap for Semiconductors, 2005 Ed., http://www.itrs.net/Common/2005ITRS/Home2005.htm, January 2006

[2] H. Jacobson et al., "Stretching the limits of clock-gating efficiency in server-class processors", High-Performance Computer Architecture, 11th International Symposium on, Feb. 2005 Pages:238 – 242.

[3] P. Gronowski, et al., "A 433-MHz 64-b quad-issue RISC microprocessor", IEEE Journal of Solid State Circuits, Vol. 31, No.11, Nov. 1996, Pages: 1687-1696.

[4] G. Ji, T. Arabi, and G. Taylor., "Design and validation of a power supply noise reduction technique", Electrical Performance of Electronic Packaging, Oct. 2003.

[5] T.J. Gabara, W.C. Fischer, J. Harrington, W.W. Troutman, "Forming damped LRC parasitic circuits in simultaneously switched CMOS output buffers", IEEE Journal of Solid State Circuits , Vol.32, No.3, March 1997, Pages: 407-418

[6] P. Larsson, "Resonance and damping in CMOS circuits with on-chip decoupling capacitance", IEEE Transactions on Circuits and Systems-I, Vol. 45, No.8, August 1998, Pages: 849-858.