Clock Tree Optimization in Synchronous CMOS Digital Circuits for Substrate Noise Reduction Using Folding of Supply Current Transients

Mustafa Badaroglu[1,4], Kris Tiri[1,5], Stéphane Donnay[1], Piet Wambacq[1],
Ingrid Verbauwhede[3], Georges Gielen[2], Hugo De Man[1,2]

[1]IMEC, Kapeldreef 75, B-3001 Leuven, Belgium
[2]ESAT - KU Leuven, Kasteelpark Arenberg 10, B-3001 Heverlee, Belgium
[3]EE Dept. - UCLA, 7440B Boelter Hall, P.O. Box 951594, Los Angeles, CA 90095-1594, USA
[4]Also Ph.D. Student at ESAT - KU Leuven, Leuven, Belgium
[5]Also Ph.D. Student at EE Dept. - UCLA, Los Angeles, USA
badar@imec.be

ABSTRACT

In a synchronous clock distribution network with zero latencies, digital circuits switch simultaneously on the clock edge, therefore they generate substrate noise due to the sharp peaks on the supply current. We present a novel methodology optimizing the clock tree for less substrate generation by using statistical single cycle supply current profiles computed for every clock region taking the timing constraints into account. Our methodology is novel as it uses an error-driven compressed data set during the optimization over a number of clock regions specified for a significant reduction in substrate noise. It also produces a quality analysis of the computed latencies as a function of the clock skew. The experimental results show >x2 reduction of substrate noise generation from the circuits having four clock regions of which the latencies are optimized.

Categories and Subject Descriptors

B.5.1 [**Register-Transfer-Level Implementation**]: Design – *data-path design.* B.6.1 [**Logic Design**]: Design Styles – s*equential circuits.* B.6.3 [**Logic Design**]: Design Aids – *optimization, simulation.* B.7.1 [**Integrated Circuits**]: Types and Design Styles-*VLSI.* B.8.2 [**Performance and Reliability**]: Performance Analysis and Design Aids*.*

General Terms

Algorithms, Design, Performance, Reliability.

Keywords

Substrate noise, di/dt noise, low-noise digital design, clock distribution networks, supply current shaping and optimization.

1. INTRODUCTION

There is a trend towards single-chip integration of more complex mixed-signal systems, higher speeds and lower supply voltages

where the signal-integrity analysis becomes a challenging task. In mixed-signal ICs, substrate noise degrades the performance of the analog circuits due to noisy digital circuits in the same substrate.

A few publications concentrate on reducing the noise generation at the source. There are publications on low-noise logic cell design, such as low-voltage logic [1], current-mode logic [2], and CMOS gates with guard wiring and decoupling [3]. Speed degradation and lower noise margins are the drawbacks for low-voltage logic. Static power consumption increase is a major drawback of current-mode logic, not tolerable in large digital systems. The gates with decoupling and guard wiring have a drawback of increase in area and additional supply rails. Up till now, no good methodologies exist to reduce the substrate noise at its source without drawbacks such as increase in area and power increase, speed degradation.

Decreasing the peak and the slope of the supply current will reduce the substrate noise since a large part of the noise is generated due to the ringing of the damped LC-tank, which is formed by the on-chip capacitance and the package inductance with series resistance on the supply. Flattening the supply current profile requires an estimation of current waveforms, which is a complex task since:

- At a given time, one or more gates can switch simultaneously, depending on the inputs and the state of the circuit.

- For a gate, the supply current differs due to the gate, the load, the supply, the input transition time and the input state.

- The waveform width depends on the penetration depth of the switching into the combinatorial logic.

Designing regular blocks, where each block has the same delay and the same supply waveform pattern, can solve the problem. One must introduce a lot of logic redundancies to achieve this goal, which is not acceptable. Another possibility is to introduce different latencies in the different regions in a clock tree to make the supply current flatter. To optimize the latencies, the use of the total transient data of the supply current is not acceptable due to the complexity of an exhaustive search for optimum clock latencies. It is therefore necessary to find a representative current waveform of all clock cycles to reduce the number of points. There exist a number of techniques to find a representative supply current waveform. A pattern-independent algorithm, iMAX [4], generates a maximum instantaneous peak current for one clock

cycle. The drawback of lacking logical constraints in iMAX has been solved by adding signal correlations [5]. Simulation based approaches have also been proposed to estimate either the maximum peak or RMS value of the supply current [6]. However, those techniques require not only an extensive simulation time due to the large number of input patterns but also the produced vectors are still large to be used for an exhaustive optimization. Besides this, none of the techniques described above gives an error interval on the estimation. The algorithm described in [7] compresses the supply currents of M current sources into C compression sets, with a higher compression than a single-cycle for a single compression set, assuring also a user-specified error bound. However, this algorithm can result in a high number of compression sets, which is not acceptable for our optimization, as the number of compression sets is highly dependent on the temporal locality and periodicity of the currents across different clock cycles.

In this paper we present a novel methodology for substrate noise reduction, which is based on an error-driven optimization of the clock tree latencies using supply current profiles taking timing constraints and the clock skew into account. There are techniques using clock latencies in order to reduce the peak current [8] and the ground bounce [9]. However, they suffer from a large number of the constraints, given as the total number of the flip-flops. Importantly they do not give a value for the number of the clock regions, which is set by the relation between the major resonance frequency of the circuit and the rise/fall time of the supply current.

First we describe the substrate noise simulation methodology that will be used to evaluate the results of the latency optimization. Next, the steps of the latency optimization are described together with their computational complexity. Then we define figure-of-merits for the clock skew sensitivity of the computed latencies. Finally we present our experimental results and draw conclusions.

2. OVERVIEW OF SUBSTRATE NOISE SIMULATION METHODOLOGY

For large designs, it is not feasible to use a transistor-level simulation of the substrate noise generation with detailed substrate models. We have proposed a methodology called SWAN [10] to simulate the substrate noise generation from large digital circuits. The accuracy of SWAN has been verified with measurements [11].
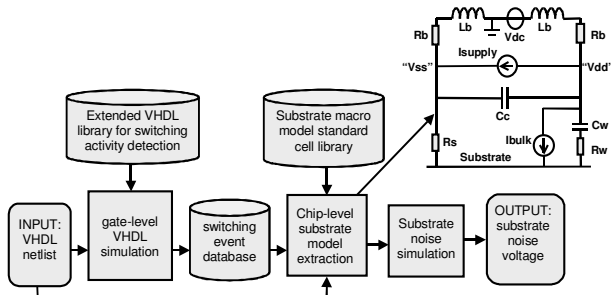

Figure 1. SWAN flow and chip-level substrate model [10]

In this paper, we will use SWAN to simulate the supply current and substrate noise. An overview of the methodology is shown in Figure 1. For every gate a substrate macro model is characterized once using transistor level model of each gate. It macro model contains two substrate noise injection mechanisms, the bulk and the supply current, together with the coupling impedances between VDD, VSS and the substrate [10]. A chip-level substrate model is

extracted (see Figure 1) by using the macro models and the switching data generated by an event-driven simulator. Finally, we simulate the substrate noise using the chip-level substrate model.

3. CLOCK TREE LATENCY OPTIMIZATION METHODOLOGY

The coupling to the substrate from the supply current is due to the ringing of the supply voltage caused by the Ldi/dt noise and the resistive coupling via the ground contact resistance. Therefore, decreasing both the amplitude and the time derivative of the supply current will reduce the substrate noise generation. The RMS value of substrate noise is proportional to the integral of its spectrum, resulting from the multiplication of the supply current spectrum and the supply current transfer function to the substrate. Since most of the noise power is due to the major resonance frequency in the transfer function, reducing the spectrum power under this resonance will also reduce the substrate noise (see Figure 2).
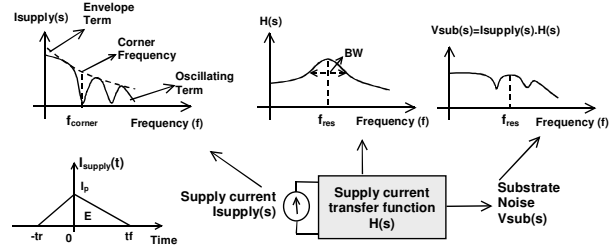

Figure 2. Effect of supply current on the substrate noise

For synchronous CMOS circuits, the total supply current in the time domain can be approximated by a triangular waveform as shown in Figure 2 where I_p, t_r, t_f and E are the peak current, rise time, fall time and total charge respectively. Fourier transformation of the supply current in Figure 2 gives: (1)

$$I_{supply}(f) = \frac{E}{2\pi^2.(t_r + t_f).f^2}\left[\frac{(1 - \exp(j.2\pi f.t_r))}{t_r} + \frac{(1 - \exp(-j.2\pi f.t_f))}{t_f}\right]$$

As the fastest oscillating harmonic in equation (1) determines the first local minimum on the oscillating term, we can define the corner frequency (f_{corner}) in the supply current spectrum as the minimum of $1/t_r$ and $1/t_f$. One can adjust the corner frequency, by modifying I_p, t_r, and t_f, in order to eliminate the major resonance frequency (f_{res}), which is set by the chip-level substrate model (see Figure 1). An optimum value for the rise/fall time is computed by:

$$\left(f_{corner} = \min\left(\frac{1}{t_r}, \frac{1}{t_f}\right)\right) < \left(f_{res} = \frac{1}{2\pi\sqrt{2L_b(C_w + C_c)}}\right) \quad (2)$$

The reduction becomes more by choosing the corner frequency as the notch points in (1) as:

$$\left(f_{corner} = \frac{K_r}{t_r} = \frac{K_f}{t_f}\right) < f_{res}, \quad K_r, K_f \in N \quad (3)$$

Without looking at the timing implications, the required minimum number of the clock regions, M, is found by the ratio of the actual rise/fall time over the optimum rise/fall time. The actual rise/fall time is computed after the triangular approximation of the total supply current of the circuit. Normally, the timing implications and the multiple peaks on the supply current will avoid having an optimum rise/fall time by using M or more clock regions. This requires an optimization of the latencies on these clock regions.

The clock latency optimization (see Figure 3) consists of 3 main parts: (1) assignment of every instance into M clock regions, (2) folding of supply current transients, (3) optimization of latencies. The next three sections will describe these steps in more detail.
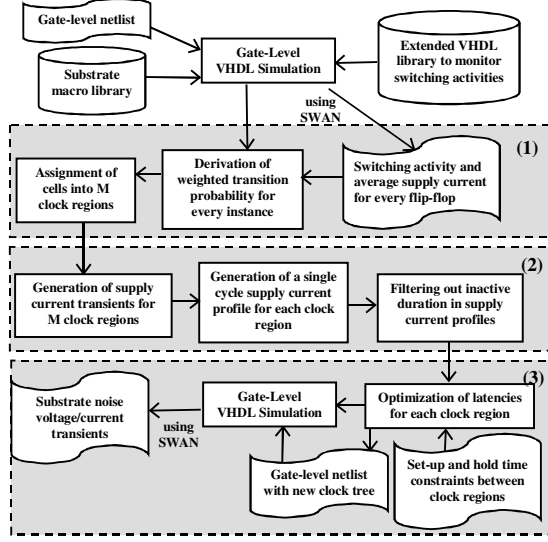

Figure 3. Clock latency optimization methodology

3.1 Clock region assignment

It is important to balance the instances over the clock regions for a significant reduction in the substrate noise generation. First, we divide the instances in sets FF_i, the set of all instances (u_{ij}) that have a data dependency on the driving flip-flop (ff_i) in FF_i. The set u_{ij} also contains the driving flip-flop (ff_i). Note that in general the intersection of two sets may not be empty. Instance assignment is based on transitive fan-out analysis from each flip-flop (ff_i). We define the weighted transition probability, a figure indicating switching statistics for all instances assigned to clock region-m (CR_m) weighted with individual contribution of the instances during a single switch, as follows:

$$P_{m,transition} = \frac{\sum_{i=1}^{N}\sum_{j=1}^{F(i)} a_{ij}.n_{ij,switch}.A_{ij,RMS}}{\sum_{i=1}^{N}\sum_{j=1}^{F(i)} n_{ij,switch}.A_{ij,RMS}} \qquad a_{ij} = 1 \text{ if } u_{ij} \in CR_m \qquad (4)$$

where $n_{i,switch}$ is the number of switching activities of the instance (u_{ij}), $A_{ij,RMS}$ is the average RMS value of the supply current for u_{ij} over all switching occurrences of u_{ij}. N is the total number of the flip-flops. F(i) is the total number of the instances in set FF_i. $A_{ij,RMS}$ factor reflects the average contribution of u_{ij} on the RMS value of the total supply current waveform during a single switching of u_{ij}. An instance u_{ij} is strictly an element of CR_m, if all sets (FF_i) where u_{ij} belongs are in CR_m. To derive $A_{ij,RMS}$ and $n_{ij,switch}$ values for each instance in the network, a gate-level simulation of the initial netlist using SWAN is performed. The sets of FF_i are joined into appropriate clock regions such that the weighted transition probabilities are balanced by 1/M. In a condition where $u_{ij} \in \{FF_{i1}, FF_{i2}\}$, FF_{i1} and FF_{i2} are assigned to different clock regions, e.g. $FF_{i1} \rightarrow CR_{m1}$ and $FF_{i2} \rightarrow CR_{m2}$, then u_{ij} is assigned to CR_{m1} if weighted transitional probability of u_{ij} is caused mostly by CR_{m1} or vice versa. It is vital to reduce the shared set of the cells as much as possible to reduce the possible glitches, which cause an increase in power, integrity problems and

the error term in the supply current compression described in Section 3.2.

3.2 Folding of supply current transients

After the assignment of the instances to each clock region, the individual supply transient for each clock region is generated from the previously explained transient simulation of the supply current used during clock region assignment. This results in $M.(T_{clock}/\Delta t).n_{cycle}$ data points, where M is the number of clock regions, T_{clock} is the clock period (integer multiple of Δt), Δt is the unit time step of the simulation and n_{cycle} is the total number of clock cycles. We discretize a single clock cycle into K time intervals, where each time interval can be chosen as Δt resulting in $K = T_{clock}/\Delta t$. We define I(k,m,n) as the actual value of the supply current at time interval-k, clock cycle-n and clock region-m. For each clock region, the union of I(k,m,n) points is compressed into a set of supply current profiles of each having a single clock cycle representation. Every clock region contains at least one current profile defined over every time interval. The compression can create more than one profile in a clock region depending on the user error bound. The set of supply current profiles is defined as:

$$IP_m = \left\{ I_p(k,m,p) \right\} \quad \text{for } k = 1..K \quad \text{for } p = 1..P(m) \qquad (5)$$
$$IP = IP_1 \cup IP_2 \cup ... \cup IP_M$$

where P(m) is the number of elements in the set of supply current profiles, IP_m, defined for clock region-m. With tight assumption on temporal locality and periodicity of currents across different clock cycles, for the supply current profile one can choose a single supply current cycle with maximum peak-to-peak value, that is $I_p(k,m,p=1) = I(k,m,n_o)$ where n_o is the clock cycle number with maximum peak-to-peak value on the supply current. This choice of n_o leads to wrong results when the temporal locality assumption is violated. Better representation is to compute statistical properties such as mean, standard deviation and probability density function at each profile point $I_p(k,m,p)$ using all the points in the actual waveform. $I_p(k,m,p)$ contains a set of statistical functions given as:

$$I_p(k,m,p) = \{\mu(k,m,p), \sigma(k,m,p), h(k,m,p)\} \qquad (6)$$

where　　$\mu(k,m,p)$: mean of I(k,m,n),
　　　　$\sigma(k,m,p)$: standard deviation of I(k,m,n),
　　　　h(k,m,p): histogram of I(k,m,n) distribution,
　　　　over all clock cycles ($n = 1..n_{cycle}$).

Each bin in the histogram, h(k,m,p), is created by the quantization of I(k,m,n) values. To compute the probability of achieving a supply current value in a certain interval, one can normalize the histogram, h(k,m,p), by dividing by the total number of actual transient points used for the construction of $I_p(k,m,p)$. The supply current profile is represented by $\mu(k,m,p)$ values where $\sigma(k,m,p)$ is a measure, how the actual values deviate from the supply current profile-p at time interval-k and clock region-m. For example, having $\sigma(k,m,p)$ zero means: "The supply current profile-p perfectly represents the actual supply current at time interval-k, clock region-m". Figure 4 shows the folding procedure.

The error interval on the supply current folding is derived for every time interval by using the derived probabilities from the normalized histogram h(k,m,p). For a confidence percentage, P_c (such as 98%), one has to find a minimum value of ε from:

$$P((\mu(k,m,p) - \varepsilon) < I_p < (\mu(k,m,p) + \varepsilon)) > P_c \qquad (7)$$

As $n_{cycle} \rightarrow \infty$, a normal distribution is likely to occur when the circuit contains a large number of independent processes and has an input vector having normal distribution. In this case, the error interval on $I_p(k,m,p)$ is given as $\mu(k,m,p) \pm 2.32\sigma(k,m,p)$ with 98% confidence. Using error interval analysis on $I_p(k,m,p)$, the set of supply current profiles, IP, are extracted within a given error interval (ε_{user}). For each $I(k,m,n)$, the confidence percentage, P_c, on ε_{user} is tested using (7) by substituting $\varepsilon = \varepsilon_{user}$. If the inequality in (7) is not satisfied for all elements of IP_m set, an additional current profile, $I_p(k,m,p)$, is generated in IP_m using (6). $P(m)$ is incremented by 1 else $I_p(k,m,p)$ in (7) is computed for the current profile-p in IP_m where P_c is maximum. The trade-off in this error driven profile extraction is the decrease of the compression ratio.
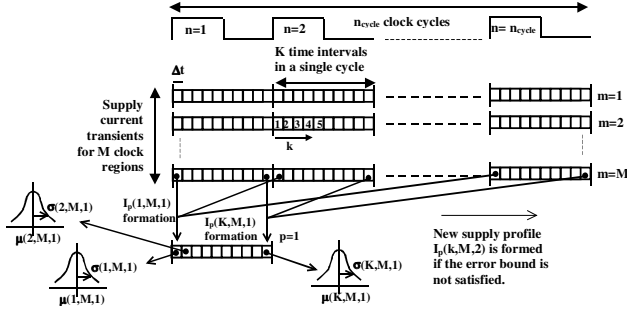

Figure 4. Folding algorithm for forming the current profiles

T_{clock} has to be chosen always larger than the maximum width of the supply current to satisfy timing constraints. For this reason, a filtering is performed on the supply current transients to increase the computational efficiency of the folding algorithm.

3.3 Clock latency optimization

Clock latency optimization is based on an exhaustive search of all latencies for a minimum on the cost function, which will be described later. An exhaustive search is necessary as the problem is NP-complete.
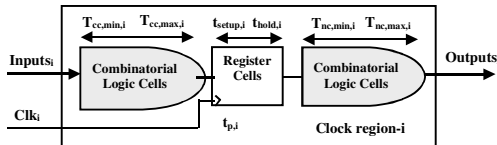

$T_{cc,min,i}$: Minimum delay of combinatorial cells before registers at clock region-i
$T_{cc,max,i}$: Maximum delay of combinatorial cells before registers at clock region-i
$t_{setup,i}$: Setup time of the register cells at clock region-i
$t_{hold,i}$: Hold time of the register cells at clock region-i
$t_{p,i}$: Propagation delay of the register cells at clock region-i
$T_{nc,min,i}$: Minimum delay of combinatorial cells after registers at clock region-i
$T_{nc,max,i}$: Maximum delay of combinatorial cells after registers at clock region-i

Figure 5. Timing values for clock region-i

The latencies have to be constrained with timing constraints defined in Figure 5. The clock region-i communicating with clock region-j has to satisfy the following constraints:

Setup time constraint: (8)

$$\Delta t_{clk,max}(i,j) < (T_{cycle} - T_{cc,max}(i) - t_{setup}(i) - t_p(j) - T_{nc,max}(j) - \delta)$$

Hold time constraint:

$$\Delta t_{clk,min}(i,j) > (t_{hold}(j) - T_{cc,min}(i) - t_p(j) - T_{nc,min}(j) + \delta)$$

where $\Delta t_{clk,max/min}(i,j)$ is the maximum/minimum allowed latency between Clk_i and Clk_j. δ is the clock uncertainty due to the

unexpected skew coming from the clock interconnect respectively. The constraint for each clock latency becomes:

$$\Delta t_{clk}(i,j) > (max(-\Delta t_{clk,max}(j,i), \Delta t_{clk,min}(i,j)) = t_{max}(i,j)) \quad (9)$$
$$\Delta t_{clk}(i,j) < (min(-\Delta t_{clk,min}(j,i), \Delta t_{clk,max}(i,j)) = t_{min}(i,j))$$

The optimization procedure is to find the best M latency bundle $(l_1, l_2, ..., l_m, ..., l_M)$ that gives minimum value on the cost function computation using the total supply current shifted with the latencies where l_m is defined as the latency value of clock region-m. One can freely set one of the latencies to zero such as $l_1 = 0$ such that one of the clock regions is aligned to the edge of the clock.

At each latency value, the cost function is evaluated as the product of the peak value and the slope of the total supply current. This comes as a direct result from equation (1), which states the direct proportionality of the supply current spectrum to the multiplication of the peak, $2E/(t_r+t_f)$, and the inverse of the rise/fall time, $1/t_r+1/t_f$, proportional to the slope of the supply current. The optimization tries to minimize this factor in order to reduce the spectral energy of the supply current, therefore the RMS value of the substrate noise. The optimization is performed on the constraint space formed by the latencies $(l_1, l_2, ..., l_M)$ as follows:

$$\min_l f_{cost}(l_1, l_2, ..., l_M) = max\left(\sum_{m=1}^{M} I_p([1,K] - l_m, m, 1)\right).$$
$$RMS\left(\sum_{m=1}^{M} (I_p([1,K] - l_i, m, 1) - I_p([1,K] - l_i - 1, m, 1))\right) \quad (10)$$
$$\text{s.t. } l_1 = 0, \ t_{max}(i,j) < l_i - l_j < t_{min}(i,j) \quad \forall i,j \in [1,M] \times [1,M]$$

4. COMPUTATIONAL COMPLEXITY

During the folding of the supply current transients of M clock regions, each having $K.n_{cycle}$, data points; the mean, the standard deviation and the probability histogram are computed recursively. Due to this recursion, the computational complexity is bound to $O(K.M.n_{cycle})$ whenever there is no error control. With error control, the complexity becomes $O(K.M.P_{MAX}.n_{cycle})$ where P_{MAX} is the maximum number of the supply profiles stored in IP_m. P_{MAX} increases by a decrease of the user error interval. The compression factor with and without error control becomes P_{MAX}/n_{cycle} and $1/n_{cycle}$ respectively. The complexity of the clock latency optimization is $O(K.(M+1).P_{MAX}.a^{M-1})$ where $a^{M-1} << K^{M-1}$ is the dimension of the search space of all M-1 latencies under set-up and hold time constraints while we set one of the latencies to zero. Overall complexity from the folding and latency optimization is $O(K.M.P_{MAX}.n_{cycle}) + O(K.(M+1).P_{MAX}.a^{M-1})$. If the number of compression sets in each clock region is bound to 1, computational complexity becomes $O(n) + O(a^{M-1}n/n_{cycle})$ where n is the number of the data points ($n = K.M.n_{cycle}$). The computational complexity approaches $O(n)$, first order dependency on the number of the data, with increasing number of clock cycles (n_{cycle}).

5. CLOCK SKEW SENSITIVITY

Due to clock routing, the load balancing or other random effects within different clock regions there will be an uncertainty or skew at each clock region. As the high slew rate of the supply current used during the optimization, the optimum point can have a high sensitivity to clock skew. To analyze the quality of the results to skew, we construct a skew radius around the optimum point. In addition, we exhaustively search the space around the optimum for a given radius (δ). Within this radius, we introduce the following skew figure on the quality of the results:

$$SB_{MAX,RMS}(\delta) = MAX, RMS\left(\sqrt{f_{cost}(l_{opt} \pm r)/f_{cost}(0)}\right), \quad \forall r \in [0,\delta] \quad (11)$$

where l_{opt} and r are the optimum latency bundle and the skew effect on the latencies respectively. $f_{cost}(0)$ is the value of the cost function (see equation (10)) before the optimization. The square root of the ratio is necessary to scale the cost function, which is given as $I_{peak}.(di/dt)_{RMS}$, appropriately in order to resemble the reduction factor of the substrate noise generation. $SB_{MAX,RMS}(\delta)$ is an indicator showing the maximum and RMS value of the reduction factor due to the clock skews, which is bound by δ.

6. EXPERIMENTAL RESULTS

The methodology is illustrated in a 4-bit Pseudo-Random-Noise-Generator (PRBS) implemented in a 0.35-μm CMOS process on an EPI-type substrate at 3.3V supply. Later, the results from ITC'99 benchmark circuits [12] and a test chip [13] are presented.
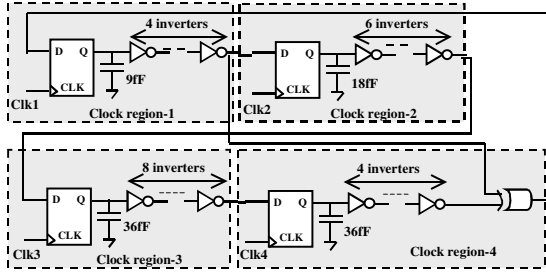

Figure 6. Pseudo-Random-Noise-Generator (PRBS) circuit

Figure 6 depicts the test-circuit and its division into four different clock regions. The supply current transfer function to the substrate has a resonance frequency of 2.3GHz. The 3dB bandwidth of the resonance stretches from 1.3GHz to 3.2GHz. The supply current has a corner frequency at 270MHz. So choosing 4 as the number of clock regions is appropriate as the initial corner frequency is already well below the resonance frequency. The design has a clock period of 4ns and a supply line parasitics of 5nH+0.5Ω. A single supply current profile has been constructed for each clock region using the actual supply current data from a total transient simulation of 105 clock cycles using SPICE. Choosing 105 clock cycles, considering the intrinsic periodicity of the 4-bit PRBS, results in an unbiased estimate.
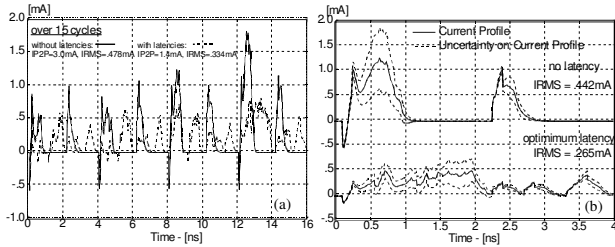

Figure 7. Simulated (SPICE) supply current (a) transients, (b) their profiles with/without latencies in the PRBS circuit

Figure 7 shows the transients (a) and the profiles (b) of the supply current with/without latencies. Figure 8 shows the simulated substrate noise transients (a) and the corresponding spectra (b) with/without latencies. The design with optimized latencies achieves factors of 2.10 and 1.75 reduction in the peak-to-peak and the RMS value of substrate noise respectively. The design with optimized latencies has a reduction of 8.9dB, 20dB, 18dB, 14dB, and 13dB at fundamental, 2nd, 3rd, 4th and 5th harmonics of the clock respectively. This is due to the reduction of the spectral power of the supply current as depicted in Figure 8. In this example, the spectral power of the supply current has been reduced by 6dB. Table 1 shows the clock skew sensitivity figures, explained in section 5, of the latency optimization results.

Table 1. Clock skew sensitivity of the latency results

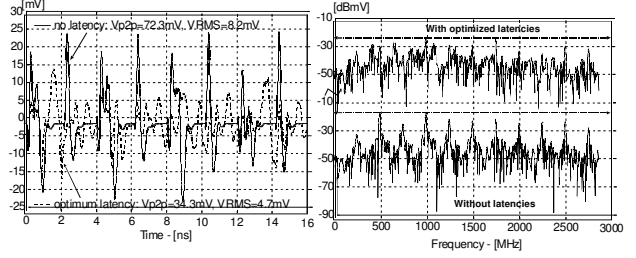| Skew radius	0ps	50ps	100ps	150ps	200ps
SB_{MAX}	0.43	0.54	0.63	0.66	0.67
SB_{RMS}	0.43	0.47	0.49	0.50	0.51

Figure 8. Simulated (SPICE) substrate noise (a) transients and (b) their spectra with/without latencies in the PRBS circuit
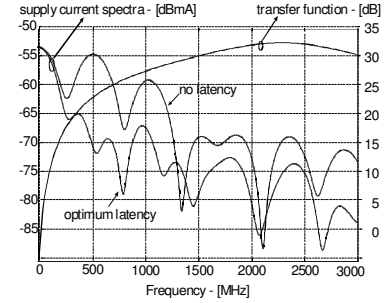

Figure 9. Supply current spectra with/without latencies and their transfer function to the substrate in the PRBS circuit

The reduction becomes more significant when the initial corner frequency is above the major resonance frequency and is shifted below the major resonance frequency. Figure 10 shows the transients and the corresponding spectra of a circuit, having a resonance at 950MHz, as a result of changing the amplitude and slope of its supply current. The supply current has been shaped from (50mA, 200ps) to (10mA, 1000ps), where we assume no timing constraints and equal rise and fall times. A significant reduction by a factor of 5.4 and 7.9 is achieved for the peak-to-peak and RMS value of the substrate noise respectively.
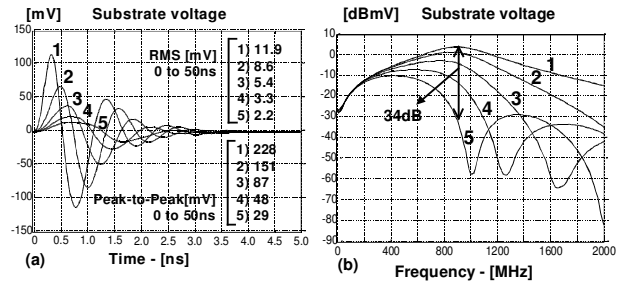

Figure 10. Simulated (SWAN) substrate noise (a) transients and (b) the corresponding spectra after supply current shaping in a circuit having a resonance frequency of 950MHz

We have also tested the introduction of the latencies on the substrate noise reduction figures of ITC99 benchmark circuits [12]. It is not anymore feasible to simulate those circuits using SPICE due to the increased complexity of the circuits relative to PRBS circuit illustrated above. Therefore, we use SWAN [10] to simulate supply current and substrate noise voltage transients using random test vectors as input. The circuits have been implemented in a 0.35-μm CMOS process on an EPI-type substrate using 1nH+0.1Ω package parasitics and 3.3V supply. The latencies of each design have been computed for the 4 clock regions. Table 2 shows the initial values of the substrate noise generation and the supply current together with the reduction percentages as a result of the introduced latencies. It is concluded that the reduction is around 37% in average for the RMS values of the substrate noise.

Table 2. ITC99 benchmark results

Circuit Name	#Gates/ FFs	Clk cycles/ Freq. [MHz]	#Sw. Act. Initial Final	Vs,rms [mV], % Vs,p2p [mV], %	Ipeak [mA], %
B04	606/66	199/50	131373 133242	22.30 [14%] 314.0 [15%]	49.70 [5%]
B05	977/34	199/50	56142 61704	6.875 [38%] 76.88 [30%]	17.57 [23%]
B06	61/9	199/50	9385 9663	4.204 [35%] 87.80 [43%]	4.96 [30%]
B07	422/49	199/50	32432 34206	7.186 [48%] 104.4 [2%]	27.80 [1%]
B08	168/21	199/50	10318 9813	4.942 [43%] 78.39 [16%]	8.89 [34%]
B09	160/28	199/50	22501 23746	7.525 [58%] 163.6 [43%]	12.44 [22%]
B10	190/17	199/50	7386 7381	0.420 [5%] 22.62 [32%]	1.98 [48%]
B11	484/31	199/50	77166 79405	7.095 [38%] 104.7 [5%]	19.97 [%13]
B12	1037/121	199/25	105424 108988	10.02 [58%] 123.2 [35%]	33.97 [18%]
B13	343/50	199/25	23962 22313	2.847 [30%] 112.8 [48%]	14.22 [39%]

We have also designed and measured a mixed-signal chip (see Figure 11), fabricated in a 0.35μm CMOS process on an EPI-type substrate, in order to compare several low-noise digital designs [13]. A comparison of two realizations, a reference circuit (REF) and a low-noise design with optimized clock latencies (LN1) in four clock regions, of a 5Kgate synchronous CMOS circuit, shows more than a factor of 2 reduction in the substrate noise generation.

7. Conclusions

Shaping the supply current is shown to be very effective as the coupling from the ringing of the supply into the substrate is dominant. In this paper, we have presented a methodology to optimize the clock tree latencies to reduce the substrate noise generation by using an error-driven compression of the supply current profiles. Before the optimization the number of the clock regions is computed based on the elimination the major resonance frequency set by on-chip circuit capacitance and the supply parasitics. Using compressed supply current transients, the computational complexity is reduced from $O((n/M)^M)$ to $O(n)$ where M, and n are the number of clock regions and the total number of transient data respectively. Experimental results show a factor of 2 reduction in the generated substrate noise on the designs with four clock regions. The efficiency of the methodology has been verified with measurements on a fabricated mixed-signal chip. The supply current shaping by the use of clock latencies is shown to be very effective if timing constraints allow shaping.
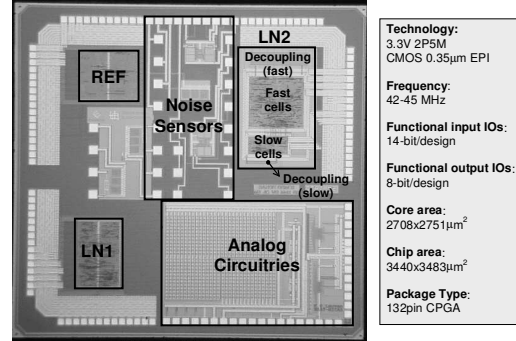

Figure 11. Microphotograph of the test chip

8. REFERENCES

[1] A. Matsuzawa, "Low-voltage and low-power circuit design for mixed analog/digital systems in portable equipment," *IEEE Journal of Solid State Circuits,* vol. 29, no. 4, pp.470-480, April 1994.

[2] D. J. Allstot, H. C. San, S. Kiaei, M. Shriyastawa, "Folded source-coupled logic vs. CMOS static logic for low-noise mixed-signal ICs," *IEEE Trans. on Circuits and Systems I: Fundamental Theory and Applications,* vol. 40, no.9, pp.553-563, Sep. 1993.

[3] M. Nagata, K. Hijikata, J. Nagai, T. Morie, A. Iwata, "Reduced substrate noise digital design for improving embedded analog performance," *ISSCC Dig. of Tech. Papers,* pp.224-225, Feb. 2000.

[4] H. Kriplani, F. Najm, I. Hajj, "Maximum current estimation in CMOS circuits," *Proc. of Design Automation Conf.,* pp.2-7, 1992.

[5] H. Kriplani, F. Najm, I. Hajj, "Pattern independent maximum current estimation in power and ground buses of CMOS VLSI circuits: Algorithms, signal correlations, and their resolution," *IEEE Tr. On Computer Aided Design of Integrated Circuits and Systems,* vol.14, no.8; Aug. 1995; pp. 998-1012.

[6] A. Krstic, T. C. Kwang, "Vector generation for maximum instantaneous current through supply lines for CMOS circuits," *Proc. of Design Automation Conference,* pp. 383-388, June 1997.

[7] R. Chaudhry, D. Blaauw, R. Panda, T. Edwards, "Current signature compression for IR-Drop Analysis," *Proc. of Design Automation Conference,* pp. 162-167, June 2000.

[8] P. Vuillod, L. Benini, A. Bogliolo and G. De Micheli, "Clock-skew optimization for peak current reduction," *IEEE Int. Symp. On Low Power Electronics and Design,* pp. 265-270, 1996.

[9] A. Vittal, H. Lia, F. Brewer, M. Marek-Sadowska, "Clock skew optimization for ground bounce control," *ICCAD Digest of Technical Papers,* pp. 395-399, 1996.

[10] M. van Heijningen, M. Badaroglu, S. Donnay, M. Engels, I. Bolsens, "High-level simulation of substrate noise generation including power supply noise coupling", *Proc. of Design Automation Conf.*, pp. 446-451, June 2000.

[11] M. van Heijningen, M. Badaroglu, S. Donnay, H. De Man, G. Gielen, M. Engels, I. Bolsens, "Substrate Noise Generation in Complex Digital Systems: Efficient Modeling and Simulation Methodology and Experimental Verification," *ISSCC Digest of Technical Papers,* pp.342-343, Feb. 2001.

[12] ITC99 benchmark circuits: http://www.cad.polito.it/tools/itc99.html

[13] M. Badaroglu, M. van Heijningen, V. Gravot, J. Compiet, S. Donnay, M. Engels, G. Gielen, H. De Man, "Methodology and Experimental Verification for Substrate Noise Reduction in CMOS Mixed-Signal ICs with Synchronous Digital Circuits," *ISSCC Digest of Technical Papers,* pp. 274-275, Feb. 2002.