

Design of a 10GHz Clock Distribution Network Using Coupled Standing-Wave Oscillators

Frank O'Mahony
Stanford University

CIS 017
Stanford, CA 94305
(650) 725-8566

fomahony@stanford.edu

C. Patrick Yue
Aeluros, Inc.

201 San Antonio Circle, #172
Mountain View, CA 94040
(650) 917-2007

yuechik@alumni.stanford.edu

Mark A. Horowitz
Stanford University

Gates 306
Stanford, CA 94305
(650) 725-3707

horowitz@ee.stanford.edu

S. Simon Wong
Stanford University

CIS 202
Stanford, CA 94305
(650) 725-3706

wong@ee.stanford.edu

ABSTRACT

In this paper, a global clock network that incorporates standing waves and coupled oscillators to distribute a high-frequency clock signal with low skew and low jitter is described. The key design issues involved in generating standing waves on a chip are discussed, including minimizing wire loss within an available technology. A standing-wave oscillator, a distributed oscillator that sustains ideal standing waves on lossy wires, is introduced. A clock grid architecture comprised of coupled, standing-wave oscillators and differential, low-swing clock buffers is presented. The measured results for a prototyped standing-wave clock grid operating at 10GHz and fabricated in a 0.18 μ m 6M CMOS logic process are presented. A technique is proposed for on-chip skew measurements with sub-picosecond precision.

Categories and Subject Descriptors

B.7.1 [Integrated Circuits]: Types and Design Styles – microprocessors and microcomputers, VLSI.

General Terms

Measurement, Performance, Design, Reliability, Theory, Experimentation.

Keywords

Clock distribution, resonant clocking, salphasic, standing wave, coupled oscillators, distributed oscillators, on-chip phase measurement.

1. INTRODUCTION

The design of global clock distributions for multi-GHz microprocessors has become an increasingly difficult and time-consuming task. As the frequency of the global clock continues to increase, the timing uncertainty introduced by the clock network – the skew and jitter – must reduce proportionally with the clock

period. However, the clock skew and jitter for conventional, buffered H-trees are proportional to latency, which has increased for recent generations of microprocessors [1].

A global clock network that uses standing waves and coupled oscillators has the potential to significantly reduce both skew and jitter. Standing waves have the unique property that the phase is the same at all points, meaning that there is ideally no skew. They have previously been used for board-level clock distribution [2], on coaxial cables [3], and on superconducting wires [4] but have never been implemented on-chip due to the large losses of on-chip interconnects. Networks of coupled oscillators have a phase-averaging effect that reduces both skew and jitter. However, none of the previous implementations of coupled oscillator clock networks use standing waves and some require considerable circuitry to couple the oscillators [5]-[9].

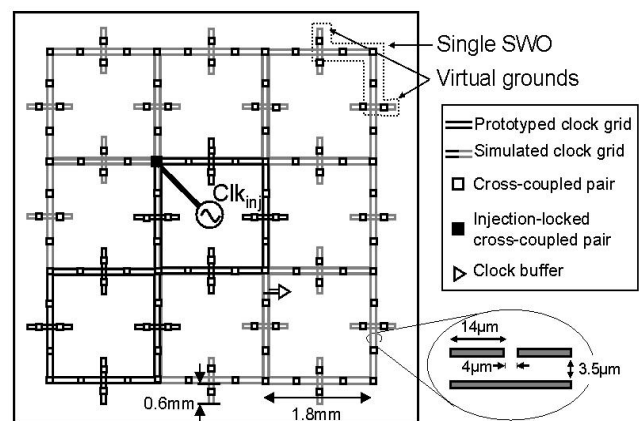


Figure 1. 10GHz standing-wave clock distribution.

This paper describes the operation and design of a global, standing-wave clock distribution network that is comprised of coupled oscillators and intended for multi-GHz clock frequencies (Figure 1). First, we discuss why generating ideal standing waves on lossy interconnects is difficult and show how distributed amplification can compensate for these losses. A new type of distributed oscillator that sustains ideal standing waves on lossy interconnects is introduced and a design example is given. Next, we show how these oscillators can be coupled together in a simple way to create a grid of standing waves and propose a clock buffer that converts the low-swing clock signal to digital levels. Finally, we present the design and results of a 10GHz standing-wave clock

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

DAC 2003, June 2-6 2003, Anaheim, California, USA.
Copyright 2003 ACM 1-58113-688-9/03/0006...\$5.00.

grid that is fabricated in a 0.18 μm 6-metal CMOS logic process. The technique used to measure the skew of the prototyped clock grid with sub-picosecond resolution is also described.

2. STANDING-WAVE OSCILLATOR

2.1 Standing waves

A standing wave is formed when two waves travelling in opposite directions with identical amplitude and frequency interact. The general case of two waves travelling in opposite directions with arbitrary phase and with amplitudes $A_1 \geq A_2$ is described by

$$A_1 \cos(\omega t - \beta z) + A_2 \cos(\omega t + \beta z + \phi) = \quad (1)$$

$$\underbrace{2A_2 \cos(\omega t + \frac{\phi}{2}) \cos(\beta z + \frac{\phi}{2})}_{\text{Standing wave}} + \underbrace{(A_1 - A_2) \cos(\omega t - \beta z)}_{\text{Travelling wave}}$$

The travelling-wave term in (1) reduces to zero when the amplitudes of the two waves $-A_1$ and A_2 are identical. Unlike a travelling-wave, which has phase that varies linearly with position, a standing wave has the same phase regardless of position, but amplitude that varies sinusoidally with position.

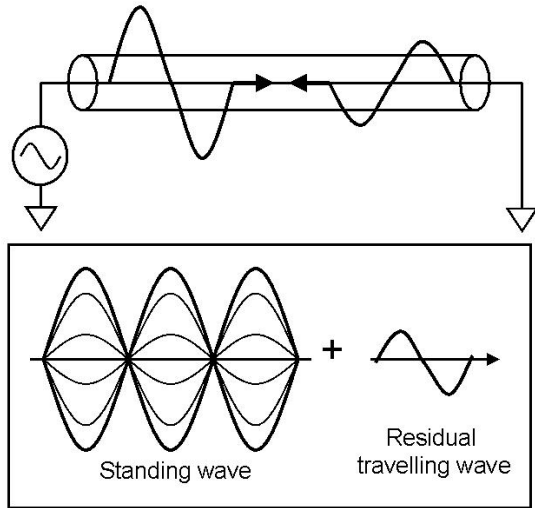


Figure 2. Generating a voltage standing wave.

A simple way to generate a voltage standing wave is to send an incident wave down a transmission line and reflect it back with a lossless termination such as a short circuit (Figure 2). However, wire losses cause amplitude mismatch between the incident and reflected waves, resulting in a residual travelling wave. The amplitude of the travelling wave – and hence the skew – is directly related to the loss and length of the wire. Previous standing-wave implementations achieve low skew by using very low-loss transmission lines or distances that are small relative to a wavelength. However, standing-wave clock networks need to span multiple wavelengths on lossy, on-chip interconnects.

2.2 Compensating for wire losses

The effect of wire loss on signal amplitude can be offset by placing lumped transconductors along the wire and spacing them sufficiently close ($\leq \lambda/10$) to present a distributed transconductance. The equivalent, lumped model for a

transmission line with distributed transconductance is shown in Figure 3, where R , L , and C are the distributed transmission-line parameters and g'_d is the effective transconductance per unit length. The loss can be calculated by

$$\alpha = \text{Re} \sqrt{(R + j\omega L)(-g'_d + j\omega C)} \quad (2)$$

$$\approx \frac{R}{2Z_o} - \frac{g'_d Z_o}{2}$$

where the approximate expression is valid when Rg'_d is small relative to the other terms in the quadratic. For a correct choice of transconductance in (2), the wire becomes effectively lossless.

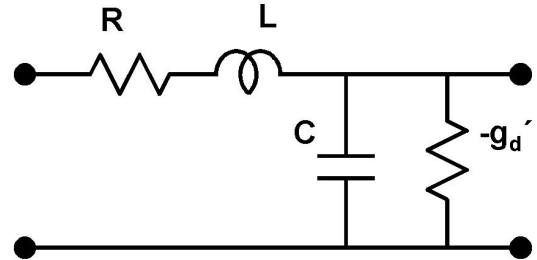


Figure 3. Lumped model of a transmission line with distributed transconductance.

2.3 Circuit Implementation

Although it is not practical to maintain a precise transconductance over an entire chip, it is possible to design transconductors that will reliably exceed the wire loss. For this reason, it is preferable to use distributed transconductors in an oscillator. The inherent amplitude saturation of the transconductors causes them to self-limit and exactly compensate for wire loss. A previously demonstrated rotary clock distribution also compensates for wire loss to create on-chip oscillators but generates travelling waves (not standing waves) [5].

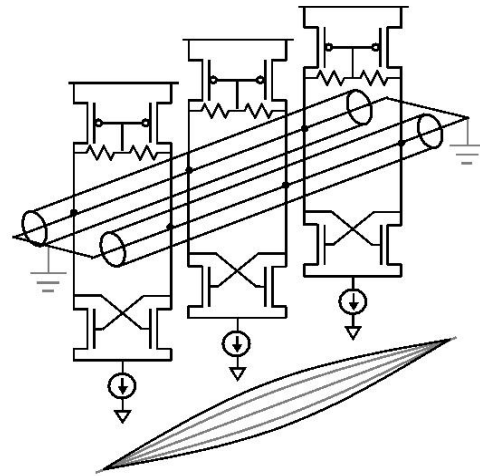


Figure 4. Standing-wave oscillator.

The circuit in Figure 4 is a distributed, standing-wave oscillator (SWO) that sustains ideal standing waves on lossy wires. NMOS cross-coupled pairs provide enough gain to compensate for wire losses, and PMOS diode-connected loads set the common-mode voltage. The differential transmission line forms a half-wave ($\lambda/2$)

resonator with virtual grounds at the ends. The parasitic capacitance of the transconductors, C_d , will load the SWO, thereby increasing the wire loss and decreasing the oscillation frequency. Therefore, the quality of the transconductors – quantified by g_d/C_d – should be maximized. Also, wire loss should be minimized in order to minimize power. The SWO will oscillate at the desired frequency, ω_{osc} , if it satisfies

$$\alpha(\omega_{osc}) < 0 \quad \text{and} \quad l = \pi / \beta(\omega_{osc}) \quad (3)$$

where: $\alpha = \text{Re}(\gamma)$ $\beta = \text{Im}(\gamma)$

$$\gamma = \sqrt{(R + j\omega L)(-ng_d / l + j\omega(C + nC_d / l))}$$

In the above equation, α , β , and γ are the loss, phase, and propagation constants for the interconnect, l is the length of the resonator, and n is the number of cross-coupled pairs that are distributed along the resonator.

2.4 Design Example

A SWO is straightforward to design from (3). Table 1 lists the parameters that will be used to design a SWO with five equally sized and equally spaced cross-coupled pairs. The transmission-line cross-section (Figure 1) is optimized for minimum loss given a total track width of $32\mu\text{m}$ and a distance of $3.5\mu\text{m}$ to the ground plane. A unit cross-coupled pair is defined to have NMOS and PMOS devices that are $18\mu\text{m}$ wide and $0.18\mu\text{m}$ long and is optimized for maximum g_d/C_d with 1.0mA of bias current. First, the required resonator length is obtained as a function of cross-coupled pair sizing by solving (3) for the given f_{osc} . Then, this length is used in (3) to calculate the effective wire loss (Figure 5). The design point is conservatively chosen so that the loss is zero for the 0.8mA bias current. The simulated waveforms from the end to center of the SWO at intervals of $l/10$ are shown in Figure 6. The free-running frequency of the oscillator is 9.6GHz , within 4% of the design goal. Note that the amplitude varies sinusoidally with position and the phase coherence is better than 1ps .

Table 1. Parameters for SWO design example.

Parameter	Value
R (Ω/mm)	4.65
L (pH/mm)	184
C (fF/mm)	250
α (units/mm @ f_{osc})	0.084 Np / 0.73 dB
g_d (mS/unit ccp)	3.60 ($I_{bias}=0.8\text{mA}$) 4.05 ($I_{bias}=1.0\text{mA}$)
C_d (fF/unit ccp)	76.3 ($I_{bias}=0.8\text{mA}$) 78.6 ($I_{bias}=1.0\text{mA}$)
n (ccp)	5
f_{osc} (GHz)	10.0

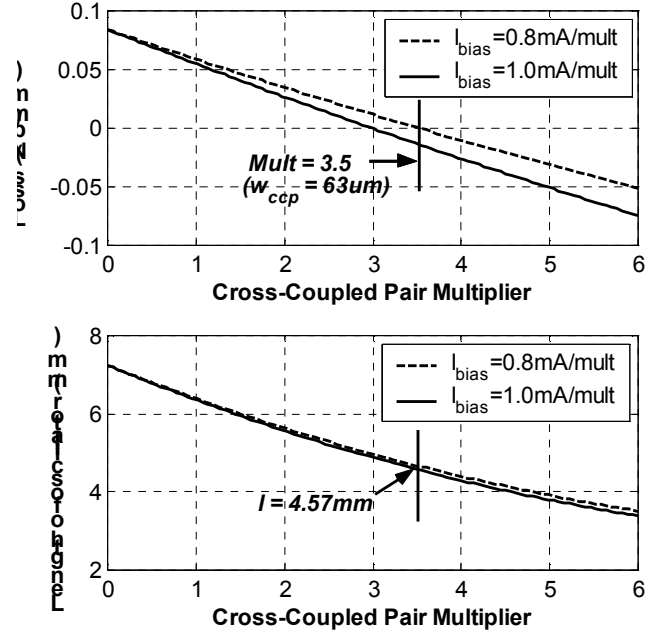


Figure 5. Loss as a function of cross-coupled pair sizing

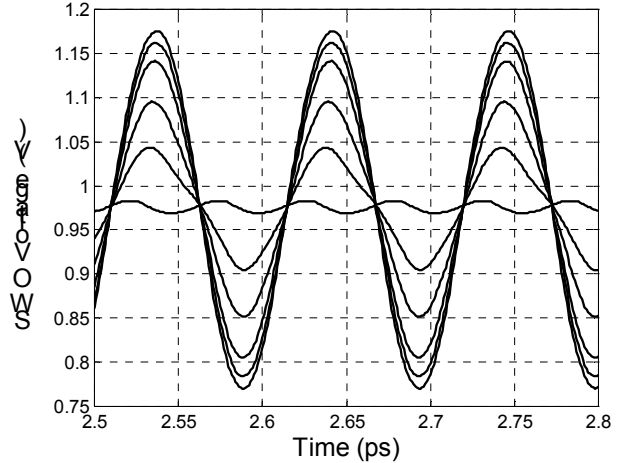


Figure 6. Simulated voltage for the SWO design example.

3. STANDING-WAVE CLOCK GRID

3.1 Coupling and Injection-Locking SWOs

On-chip transmission-line resonators have an inherently modest Q that allows coupling and injection locking of the SWOs over a range of frequencies. For the SWO in the previous design example, Q is 2.7. SWOs can be coupled together by simply connecting their transmission lines. The coupling strength is largest when the oscillators are connected at the center and zero near the ends. Any detuning between coupled oscillators results in skew that is directly related to the coupling strength and Q [10]. Therefore low- Q resonators with strong coupling between should be used for clock distribution. These oscillators can also be injection locked to a reference signal. Injection locking allows the clock frequency to be dictated by an external clock source, such as a phase-locked loop (PLL), and stabilizes the otherwise noisy signal of this low- Q oscillator. In this work, we ac-couple the reference signal into the gate of the PMOS loads at the center of a

SWO. The locking range and the skew caused by driving an injection-locked oscillator off-resonance are also related to the coupling strength [10]. Again, strong coupling is preferable for low skew.

3.2 Grid Architecture

A resonant grid of coupled SWOs is shown in Figure 1. Choosing the coupling point is a tradeoff between the size of the grid and the coupling strength. Connecting the SWOs 15-20% from the short-circuits provides strong enough coupling to lock the segments together without causing excessive skew due to mismatches between SWOs. To make a grid pattern, the ends of the SWOs are folded at right angles to the grid. Due to the sinusoidal amplitude envelope of standing waves, the folded segments have the smallest voltage amplitude and hence are inappropriate for recovering the clock. The voltage standing-wave pattern for a portion of the grid is shown in Figure 7.

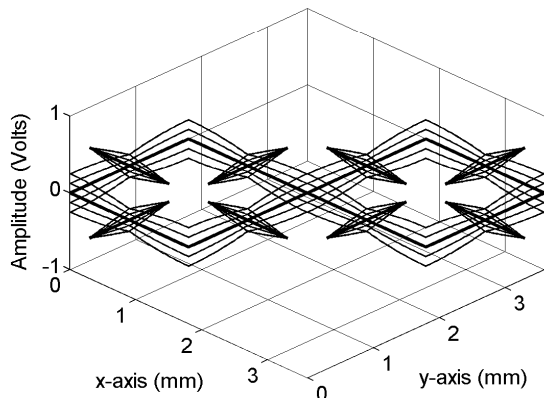


Figure 7. Voltage standing-wave pattern on prototyped grid.

A significant advantage of resonant clock distribution is that detuning of the grid – which is related to the skew and jitter – is primarily determined by the resonator and is not a strong function of the power supply. The impact of power supply variations on the tuning of SWOs is a second-order effect caused by variations in the parasitic capacitance of the cross-coupled pairs. In contrast, the delay – and hence the skew and jitter – of a buffered H-tree is a strong function of power supply variations due to the large sensitivity of inverter delay to power supply variations.

3.3 Phase Averaging in Grids

Within a grid of coupled oscillators, phase is averaged at each coupling point. Phase differences among the SWOs, either skew due to mismatch or jitter due to power supply variations, are reduced by this averaging process. In our approach, each SWO is coupled in up to three locations. The averaging effect is also directly related to the coupling strength. In order to test how well the coupled oscillators suppress jitter caused by localized supply noise, we simulated a single SWO, a grid of four SWOs, and the full grid in Figure 1, all injection locked at 10GHz. In each case, the power supply for one cross-coupled pair was reduced by 10% with a 100ps fall time. The resulting peak-to-peak jitter on the oscillator network was 0.41ps, 0.26ps, and 0.17ps, respectively, confirming that the phase averaging property of a coupled network of SWOs reduces jitter.

3.4 Clock Buffer

Standing-wave clock distribution is intended to interface with a conventional digital clock distribution at lower levels of the clock

hierarchy. Therefore, a buffer is required to convert the low-swing, differential sinusoids to digital levels without adding significant amounts of timing error due to variations of the input amplitude. A two-stage clock buffer based on [11] is shown in Figure 8. The first stage differential pair has a small gate-overdrive that allows complete current switching even for the smallest expected input amplitude. It amplifies and limits the signal so the output amplitude is roughly independent of the input amplitude. A low-pass filter attenuates the harmonics added by the limiting amplifier that would otherwise cause amplitude-dependent skew. The second stage is a sine-to-square converter that uses cross-coupled inverters and a shunt resistor to achieve a well-controlled 50% duty cycle over process, temperature, frequency, and supply variation. Because the 0.18 μ m devices chosen for demonstration are not adequate to test the clock buffer at 10GHz, we simulated the buffer with a 2GHz sinusoidal clock. This clock period corresponds to an aggressive seven FO4 delays in this technology. The clock buffer exhibits 5.9ps skew (1.2% τ_{clk}) for the 30% voltage variation seen across the center 50% of a standing wave (Figure 9). Assuming similar performance using devices in a future process capable of 10GHz operation, the amplitude-dependent skew will be about 1ps.

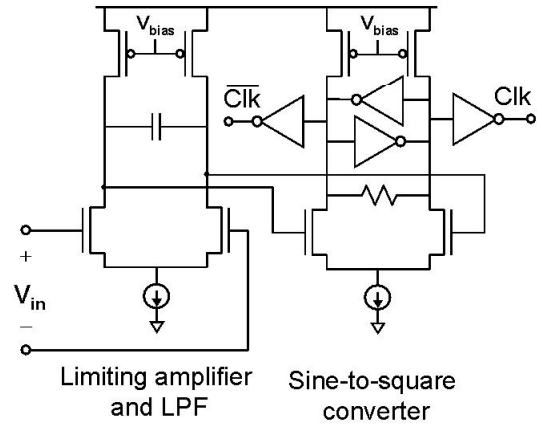


Figure 8. Clock buffer.

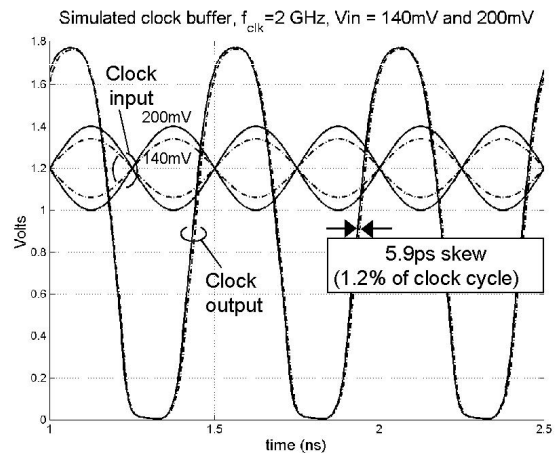


Figure 9 Simulated clock buffer performance.

4. EXPERIMENT

A 10GHz clock network comprised of eight coupled SWOs is prototyped in a 0.18 μ m, 1.8V CMOS process with six AlCu metal layers (Figure 1). Clock buffers are not integrated due to the speed

limitations of 0.18 μm devices at 10GHz but will be easily integrateable when devices are scaled for 10GHz operations. The differential $\lambda/2$ lines are 3mm long, 14 μm wide, and are 4 μm apart in metal six. Although the design parameters are identical to the ones used in the previous example, the additional loading of testing and tuning circuitry reduces the length required to oscillate at 10GHz and increases the necessary transconductance. Each SWO consists of five cross-coupled pairs with 90 μm -wide devices. The transconductance of each cross-coupled pair is variable from 0mS to 25mS by changing the bias current; 18mS is required to start oscillation. The grid is tunable from 9.8GHz to 10.5GHz (6.4% range) with accumulation-mode MOS varactors positioned 400 μm from the ends of the SWOs. Grid tuning extends the locking range and facilitates intentional skewing of specific grid segments for testing purposes. The SWOs consume 430mW, which is comparable to the CV^2f power if the grid were driven digitally at 10GHz. The measured sensitivity to power supply variations is $0.06\% \Delta f_{osc} / \% \Delta V_{supply}$.

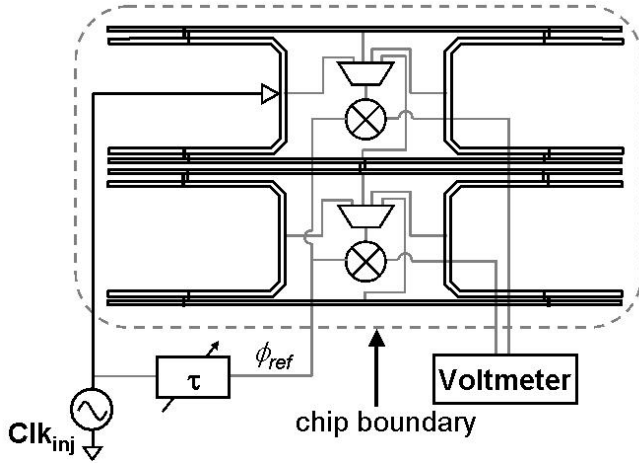


Figure 9. Prototype grid layout and timing circuit.

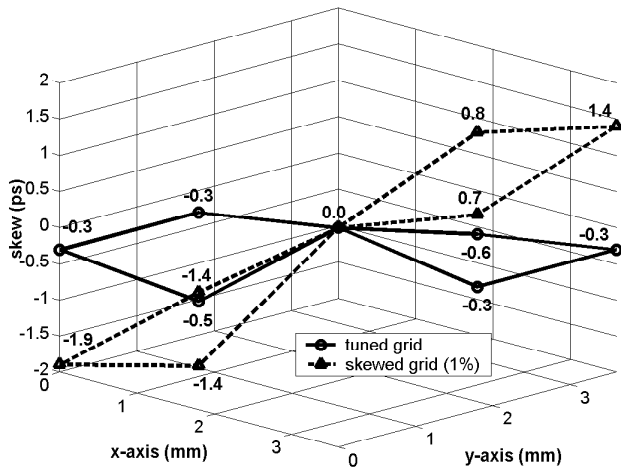


Figure 10. Measured skew on grid.

On-chip skew is measured with a homodyne technique that converts phase into dc voltage with a sensitivity of 60fs/mV (Figure 9). The clock signal is tapped at eight points around the grid and routed through length-matched wires and multiplexers to a pair of mixers. The mixers compare the phase of each clock signal to a reference phase, θ_{ref} , that is set by an off-chip phase

shifter. The grid is folded to minimize the distance from tapping points to the mixers. The measured skew is 0.6ps (0.6% τ_{clk}) when the grid is tuned to 10.0GHz with a single control voltage for all the varactors and 3.3ps (3.3% τ_{clk}) when half of the grid is detuned by 1% (Figure 10). The ability to smooth out skew is a key advantage of this type of clock distribution. The worst-case skew between any two adjacent points is 1.4ps (1.4% τ_{clk}) for the detuned grid. A die micrograph of the prototyped global clock grid is shown in Figure 11.

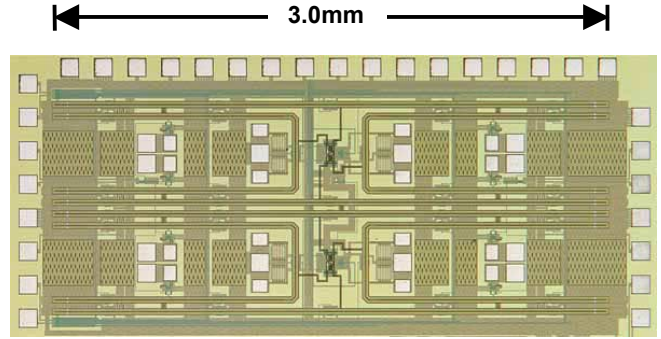


Figure 11. Die micrograph.

5. SCALING EFFECTS

SWOs will benefit significantly from process scaling. As f_T increases g_d/C_d also improves, and the self-loading effect of the cross-coupled pairs – which can add significantly to the interconnect loss – decreases. As a result, less power and device area will be needed to compensate for the same amount of wire loss. Less self-loading will also allow larger grids since the effective wavelength will increase. Improved devices, high-conductivity metals and low- κ dielectrics will also allow designers to use less aggressive transmission line dimensions.

6. CONCLUSIONS

The first on-chip standing-wave clock distribution has been demonstrated. This approach benefits from the invariant phase property of standing waves and the phase averaging effect of coupled oscillators. A method for overcoming on-chip interconnect losses to generate ideal standing waves has been presented. The standing-wave oscillators can be coupled together to form a clock grid that injection-locks to an external clock source. All of the circuit blocks required for a 10GHz standing-wave clock network will be feasible in a future technology. A 10GHz clock grid was demonstrated that achieves low skew and jitter. Based on these results, we believe that standing-wave clock distribution will be an attractive and scalable alternative to H-trees for future microprocessors as clock frequency scales to 10GHz and beyond.

7. ACKNOWLEDGMENTS

The authors thank R. Chang, N. Talwalkar, B. Kleveland and T. Soorapanth of Stanford for helpful discussions, K. Soumyanath and M. Anders at Intel for support, TSMC for fabrication, and Intel and the MARCO Interconnect Focus Center for funding.

8. REFERENCES

- [1] P.J. Restle et. al., "A clock distribution network for microprocessors," *IEEE J. Solid-State Circuits*, vol. 36, no.5, pp. 792–799, May 2001.
- [2] V.L. Chi, "Salphasic distribution of clock signals for synchronous systems," *IEEE Trans. Comput.*, vol. 43, pp.597–602, May 1994.
- [3] M.E. Becker and T.F. Knight, Jr., "Transmission line clock driver," in *Proc. IEEE Int. Conf. Computer Design*, Oct. 1999, pp. 489–490.
- [4] M. Hosoya, W. Hioe, K. Takagi and E. Goto, "Operation of a 1-bit quantum flux parametron shift register (latch) by 4-phase 36-GHz clock," *IEEE Trans. Appl. Superconductivity*, vol. 5, no. 2, pp. 2831–2834, June 1995.
- [5] J. Wood, T.C. Edwards and S. Lipa, "Rotary traveling-wave oscillator arrays: a new clock technology," *IEEE J. Solid-State Circuits*, vol. 36, no.11, pp. 1654–1665, Nov. 2001.
- [6] I. Galton, D.A. Towne, J.J. Rosenberg and H.T. Jensen, "Clock distribution using coupled oscillators," in *Proc. IEEE Int. Symp. Circuits and Systems*, vol. 3, May, 1996, pp. 217–220.
- [7] L. Hall, M. Clements, W. Liu, and G. Bilbro, "Clock distribution using cooperative ring oscillators," in *Proc. 17th Conf. Advanced Research in VLSI*, Sept. 1997, pp. 15–16.
- [8] V. Gutnik and A.P. Chandrakasan, "Active GHz clock network using distributed PLLs," *IEEE J. Solid-State Circuits*, vol. 35, no.11, pp. 1553–1560, Nov. 2001.
- [9] M. Saint-Laurent, M. Swaminathoan and J.D. Meindl, "On the micro-architectural impact of clock distribution using multiple PLLs," in *Proc. IEEE Int. Conf. Computer Design*, Sept. 2001, pp. 214–220.
- [10] R.A. York, "Nonlinear analysis of phase relationships in quasi-optical oscillator arrays," *IEEE Trans. Microwave Theory and Tech.*, vol. 41, no. 10, Oct. 1993.
- [11] A. Maxim, B. Scott, E.M. Schneider, M.L. Hagge S. Chacko and D. Sturca, "A low jitter 125-1250MHz process independent and ripple-poleless 0.18- μ m CMOS PLL based on a sample-reset loop filter," *IEEE J. Solid-State Circuits*, vol. 36, no.11, pp. 1673–1683, Nov. 2001.