Proximity Communication

Robert J. Drost, Robert David Hopkins, and Ivan E. Sutherland
Sun Microsystems, Inc.
Mountain View, California, U.S.A.

Abstract

This paper reports results from wireless chip to chip communication experiments. Sixteen bit words pass from one chip to another in parallel without detectable error at 1.35 billion data items per second for a total data rate of 21.6 Gigabits per second. The experiment transmits pseudo random patterns between chips built in 350nm CMOS technology. Chips touch face-to-face to communicate. The same pseudo random data pattern is loaded onto both chips so that the receiving chip can check the accuracy of every bit communicated. Each communication channel consumes a static power of 3.6 milliWatts, and a dynamic power of 3.9 picoJoules per bit communicated. The channels lie on 50 micron centers. Because the capacitive communication works through covering oxide, ESD protection is unnecessary. Vernier position measuring circuits built into the chips indicate the relative position of transmitting and receiving arrays to assist mechanical alignment. The test chip includes a Vernier circuit that provides inter-chip position measurements with a resolution of 1.4 microns.

Background

On-chip performance has been increasing much more rapidly than off-chip communication bandwidth because both on-chip transistor density and clock frequency are increasing faster than off-chip input/output (I/O) density and frequency [2]. This difference occurs because off-chip bonding and wiring are about two orders of magnitude larger than on-chip wiring. On-chip wiring pitch is on the order of 1 micron, while off-chip wiring and ball-bond pitches are on the order of 100 microns. The performance gap between on-chip and off-chip bandwidth makes off-chip bandwidth a performance bottleneck.

Fig. 1 shows a comparison of the density of I/O pads for proximity communication versus area ball bonding [2]. Because proximity communication structures use the lithographic pitch of on-chip wires, proximity communication can achieve a density about 60 times greater than area ball bonds.

Proximity Communication

Proximity communication is based on the observation that faster, lower-cost communication is possible over shorter distances. We name this communication method "Proximity" because chips are placed in a face-to-face arrangement with the transmitter and receiver circuits aligned with only microns of distance between them.

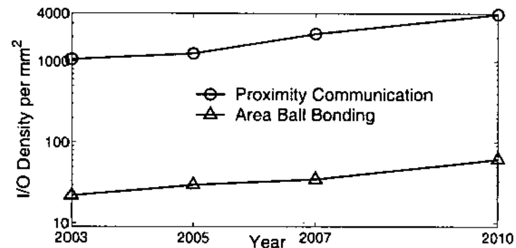

Figure 1. I/O Density Comparison

Fig. 2 illustrates two chips positioned for proximity communication. Because the transmitter and receiver are close to each other and communicate by capacitive coupling rather than through large off-chip wires, the transmitter and receiver circuits are small. Moreover, the transmitter and receiver pads are protected from exposure to electrostatic discharge (ESD) events by the top layer dielectric and passivation. Thus the transmitter and receiver omit ESD protection, further reducing the parasitic capacitance and power consumption.

We chose to couple the transmitter pads directly to the receiver pads [1][5], in contrast to [3][4] in which the transmitter couples to a multi-chip module (MCM) substrate trace which couples in turn to a receiver pad. Coupling the transmitter pad directly to the receiver pad provides a stronger signal, improving signal integrity and reducing power.

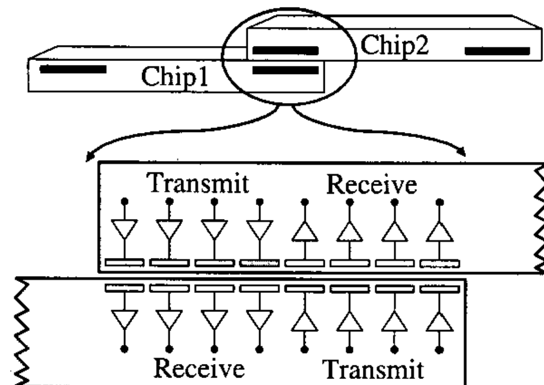

Figure 2. Cross-section of chips

19-1-1

Alignment

Alignment is a critical issue. Misalignment reduces received signal strength by reducing coupling area and increasing distance. Cross-talk also increases with misalignment because the voltage swings on adjacent transmitter and receiver pads couple into the receiver pad. In order to couple the transmitter and receiver pads best, the pads must be placed in correct alignment. Fig. 3 shows the six alignment dimensions, three translational and three rotational.

We use an on-chip Vernier measurement system to measure the relative positions of two chips in the x, y, and θ dimensions. Verniers are common in mechanical systems, and a micro electro-mechanical system (MEMS) Vernier strain gauge has demonstrated a resolution of 10 nm [6].

Fig. 4 shows the cross-section of our Vernier measurement system. The Vernier measurement system consists of Vernier bar structures and circuits that connect to the bars. The Vernier circuits apply positive and negative edges to alternate bars, shaded white and gray respectively, of Chip1 and sense the polarity of the capacitively coupled signals received by Chip2. The polarity of the received signal reveals whether a receiver bar is closer to the positive or negative transmitter bar.

In our chips we have 9 bars in the receiver under 10 bars in the transmitter spanning a distance of 126 microns. Although the receiver bars are on a 14 micron pitch, a relative movement between the chips of 1.4 micron flips the polarity sensed by one Vernier receiver. This measurement resolution is sufficient for our purposes. Finer resolutions may be obtained by using finer metal bar patterns.

The receiver circuit uses a simple inverter that has a large value resistor connecting its output to input. The transmit and receiver strobe signals, Tstrobe and Rstrobe, are inactive for a long time between strobe events. During this inactive period, the feedback resistor causes the input inverter to self-bias its input signal to its threshold voltage, cancelling its input offset voltage. When the Tstrobe event occurs, the transmit pads couple their transitions onto the receiver pads. Each receiver pad will see either a net positive or negative coupled transition depending on whether the receiver is closer to a positive or a negative polarity transmitter pad. The feedback resistor value is chosen so that the coupled transition decays with a time constant of about 1 us. The two inverter stages in the receiver amplify the coupled transition to a digital level and a Rstrobe event captures the digital levels in latches. We read out the latch values using a JTAG-like scan chain. Simulations of our Vernier receiver circuit showed that it could sense a net difference of about 0.1fF between the coupled capacitance to the positive and negative transmit bars. Our simulations included transistor mismatch [7].

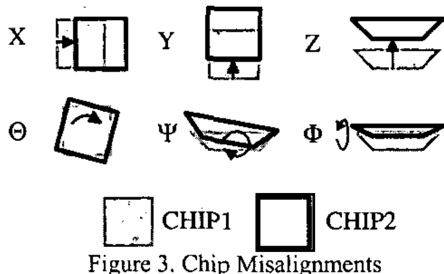

Figure 3. Chip Misalignments

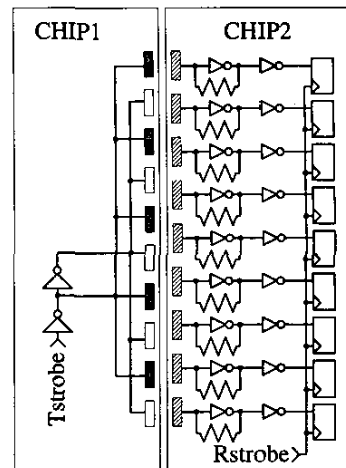

Figure 4. Vernier circuit diagram

The Vernier structure as described measures alignment along one axis. Our chip includes four Verniers to enable multi-axis measurements. Two Verniers, one rotated 90 degrees from the other, measure x and y alignment. A second set of two Verniers separated from the first set indicates θ alignment. The Verniers cannot measure separation in z nor angular misalignment in ψ and φ dimensions.

Transmitter and Receiver Circuits

Fig. 5 shows the transmitter and receiver circuits along with the signal capacitor, Cs, and the parasitic capacitors on the transmitter and receiver pads, Cpt and Cpr, respectively.

The transmitter is simply an inverter driving the transmitter pad. The inverter is sized to achieve an equivalent fanout of two. The parasitic capacitor on the transmitter pad, Cpt, loads the driver but otherwise does not matter. In contrast, the parasitic capacitor on the receiver pad, Cpr, forms a capacitor divider with the coupling signal capacitor Cs, and reduces the received signal swing on the receiver pad.

The receiver circuit includes an input inverter and a feedback inverter. These inverters generate positive feedback to hold the receiver state until the next transition occurs on the

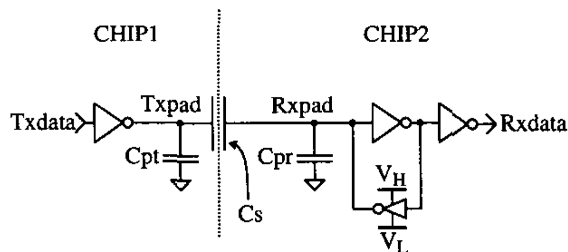

Figure 5. Transmitter and Receiver Circuit Diagram

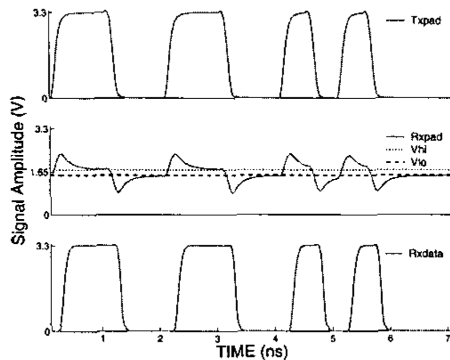

Figure 6. Spice Simulation Waveforms of Transmitter and Receiver Signals

transmitter pad. Thus a DC signal may be communicated. A third inverter in the receiver provides additional amplification.

If the feedback inverter was capable of driving the Rxpad to Vdd or Gnd, then the coupled signal amplitude would be too small to cross the input inverter threshold voltage. To increase the receiver sensitivity, the feedback inverter drives Rxpad to reduced bias levels, VH and VL. These reduced bias levels are generated on-chip and are set to be about 100 mV above and below the switching threshold.

The transmitter and receiver circuits create a bandpass filter. The output resistance of the feedback inverter and the capacitance on the receiver pad node set the low frequency cutoff. Although the transmitted signal can contain frequencies down to zero, the signal transitions on the transmit pad must be fast relative to this cutoff frequency, lest they suffer too much attenuation to be recognized. The low frequency cutoff provides a welcome benefit: noise on power supplies is attenuated if its frequency content is lower than the cutoff. On-chip power supply networks severely attenuate the high frequency content of noise signals. As a result, capacitively coupled communication circuits may tolerate significant noise from on-chip digital logic circuits.

Fig. 6 shows the results of spice simulation of the transmit and receive circuits shown in Fig. 5. The top graph plots the signal on the Txpad node transitioning between Vdd and Gnd. The middle graph plots the signal on the Rxpad node and the VH and VL bias voltages. During a transition on Txpad, the Rxpad node shows an attenuated transition. After the transition, the Rxpad node is driven back to the VH or VL level by the feedback inverter. The bottom graph shows the output signal, Rxdata, which is amplified to full digital levels.

Test Chip

In order to test the Vernier measurement system and the transmitter and receiver circuits, we fabricated a test chip via Mosis in the TSMC 350 nm CMOS technology. Fig. 7 shows the test chip and floorplan. In addition to transmitter and receiver arrays, the chip contains control logic and a set of probe pads The control logic block includes two pattern

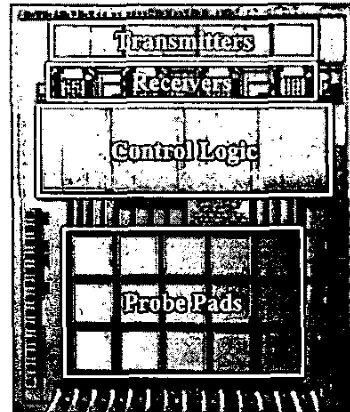

Figure 7. Chip Photograph

generators, a comparator, and a bit error counter. Transmit and receive timing signals, Tx Clock and Rx Clock, are generated off-chip and dictate the frequency and phase of the pattern generators. The pattern generators provide data patterns to the transmit arrays and the comparator. The comparator checks for differences between the received pattern and the on-chip generated pattern. A bit error counter tracks the number of differences seen.

The transmitter and receiver contain 4-by-4 arrays of circuits and pads. Each of the 16 transmitters outputs a unique bit pattern. We implemented different transmitter pad areas to test the effect of varying amounts of capacitive coupling. The receiver pads are all the same size.

Experimental Setup

The test setup consists of micro manipulators and mounting blocks that hold two printed circuit boards (PCBs) containing test chips. The micro manipulators allow translation with micron resolution in all three spacial dimensions, as well as rotation about each of the three axes. The two identical printed circuit boards each hold a test chip glued and wirebonded directly to the board. Fig. 8 shows a close-up photograph of the chips after they have been aligned; the wirebonds and printed circuit boards can be seen. In the photograph, Chip1 is facing away from and Chip2 is facing towards the camera. In the picture the chips are separated by about a millimeter, but are moved together for testing.

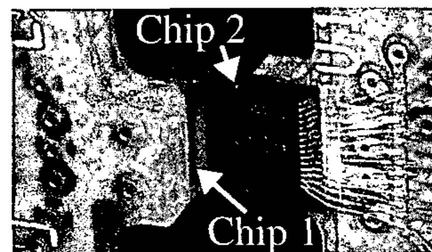

Figure 8. Detail of aligned chips

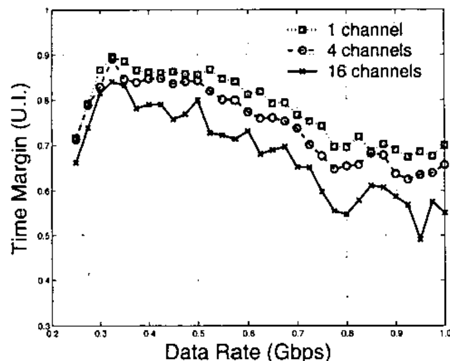

Figure 9. Effect of # of channels on time margin

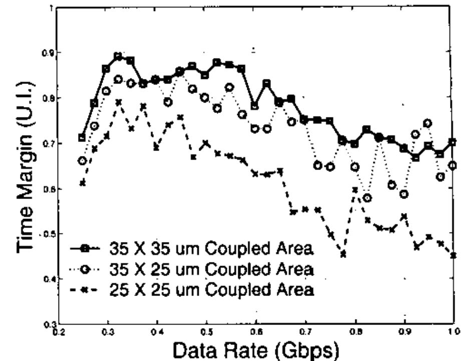

Figure 10. Effect of coupled area on time margin

We first align the chips visually using a 10x magnification stereoscopic microscope that views the chips along their top edges. During this step the rotation micrometers are used to establish approximate coplanarity between the chip surfaces. A small amount of non-coplanarity can be tolerated because the boards and mounting hardware flex slightly when the chips are pressed together. Once the chips are visually aligned and pressed together, the Vernier measurement system is used to align the chips in x, y, and θ to within a few microns. A small amount of oil between chips increases the capacitive coupling. The oil has a dielectric constant of about 2.4. We tested the chips in a laboratory room that had no special dust control. We found that dust on the chip surface was squeezed sufficiently thin to permit communication.

Test Results

The prototype test chip has demonstrated reliable communication over 16 simultaneous data channels with an aggregate bandwidth of 21.6 Gigabits per second (Gbps). We measured a bit-error rate of less than $10e^{-10}$ with each of the 16 channels communicating simultaneously at a data rate of 1.35 Gbps. Each channel communicated a different pseudo random pattern. The pattern repeats every 16 cycles, but because patterns are initialized from off-chip, we have tested a wide variety of patterns. Each channel consumed a static power of about 3.6 mW due to receiver bias currents and a dynamic power of 3.9 pJ per bit due to switching activity.

Fig. 9 shows a plot of the time margin as a function of data rate. The time margin is shown in unit intervals (U.I.), where one U.I. is equal to the bit period. The graph shows the time margins for the cases of 1, 4, or 16 channels operating simultaneously. Simultaneous transmission through multiple channels reduces the time margin. A number of factors may cause the time margin reduction. The comparator monitors all channels in parallel, forcing the overall time margin to be the minimum intersection for all enabled channels. Capacitive crosstalk among channels and local power supply droop may also degrade the time margin. Our test equipment for the time margin control limited the measurement to a data rate of 1 Gbps. However, our bit error rate measurements were taken up to 1.35 Gbps

We conducted a series of tests to determine the effect on the timing margin of coupled area, and hence coupling capacitance, between the transmit and receive pad. Fig. 10 shows the timing margin as a function of data rate for three coupled areas. For coupled area less than 25 X 25 microns, reliable transmission failed.

Conclusions

We have investigated capacitive coupling to communicate directly between two chips placed face-to-face. We have designed transmitter and receiver circuits that reliably communicate capacitively coupled signals. We have also designed a Vernier measurement system that achieves a resolution of 1.4 microns. Our 350nm CMOS test chip demonstrates 16 channels operating simultaneously, each communicating pseudo random patterns at a rate of 1.35 Gbps, for an aggregate bandwidth of 21.6 Gbps.

Acknowledgements

This work was supported in part by Defense Advanced Research Projects Agency (Contract No. NBCHC020055). We thank Robert Bosnyak, Ann Coulthard, and Jon Lexau for help in the layout and circuit design of the test chip, and Bill Coates for a thorough review of a draft of this paper.

References

(1) I. Sutherland, "Face to Face Chips," US Patent 6,500,696, Filed October 2, 2001, Granted December 31, 2002.

(2) 2001 International Technology Roadmap for Semiconductors.

(3) S. Mick, J. Wilson, P. Franzon, "4 Gbps high-density AC coupled interconnection," *IEEE Custom Integrated Circuits Conference*, pp. 133-140, 2002.

(4) D. Salzman, T. Knight, "Capacitively Coupled Mutlichip Modules," *Multi-Chip Module Conference Proceedings*, pp. 487-494, April 1994.

(5) K. Kanda, D. Antono, K Ishida, H. Kawaguchi, T. Kuroda, T. Sakurai, "1.27Gb/s/pin 3mW/pin Wireless Superconnect (WSC) Interface Scheme," ISSCC Digest of Technical Papers, pp. 186-7, Feb 2003.

(6) S. Jacobsen, M. Mladejovsky, C. Davis, J. Wood,R. Wyatt, "Advanced intelligent mechanical sensors (AIMS)," *International Conference on Solid-State Sensors and Actuators*, pp. 969-973, June 1991.

(7) M. Pelgrom, H. P. Tuinhout, and M. Vertregt, "Transistor matching in analog CMOS applications," *International Electron Devices Meeting*, pp. 915-8, 1998.