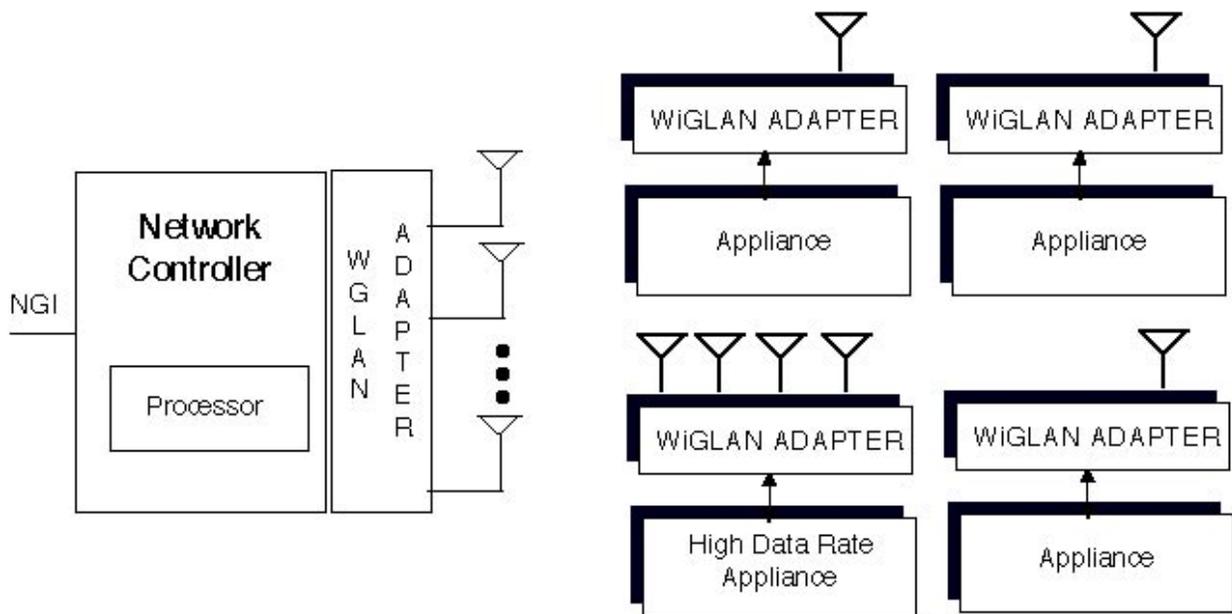

Integrated Circuits and Systems



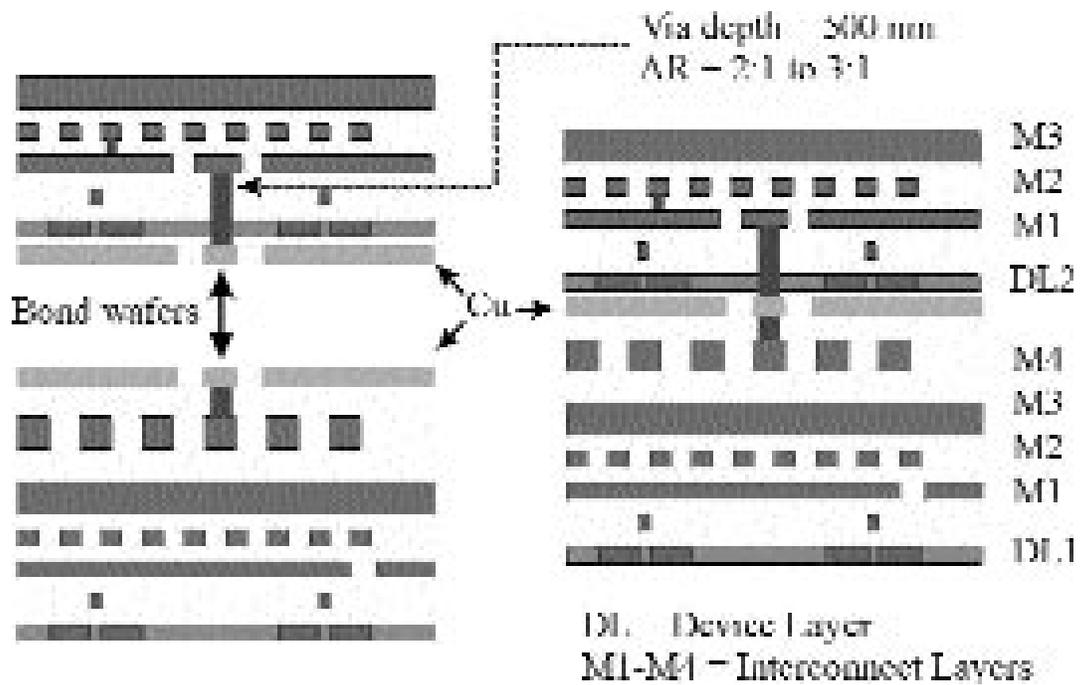
Integrated Circuits and Systems

- *The MIT μ AMPS Project*
- *Energy-Efficient Communication Protocols for Wireless Microsensor Networks*
- *Energy-Scalable Source Tracking for Wireless Networked Sensors*
- *Optimal Collaborative Strategies for Distributed Sensor Networks*
- *An Energy-Efficient Link Layer Protocol for a Wireless Sensor Network*
- *Dynamic Power Management in Wireless Sensor Networks*
- *Hardware Architecture for a Power Aware Microsensor Node*
- *Reconfigurable Architectures for Energy-Efficient, Algorithm-Agile Cryptography*
- *Ultra Low Power Radio for Wireless Devices*
- *A Low Power Transmitter for Wireless Sensor Networks*
- *Energy Minimization for Wireless Microsensor Systems*
- *Analog-to-Digital Conversion of High-Frequency Signals for RF Communication*
- *Wireless Gigabit Local Area Network*
- *Circuit Design and Technological Limitations of Silicon RFICs*
- *A Linear 5.8GHz Power Amplifier with Variable Output Control*
- *5.8 GHz Wideband Receiver for Wireless Gigabit LAN*
- *Analog Base-Band Processor for Gigabit Wireless LAN*
- *Design of a Power-Scalable Digital Least-Mean-Square Adaptive Filter*
- *Circuits for Optical Clock Distribution*
- *Active GHz Clock Network using Distributed PLLs*
- *Reduction of Interconnect Power and Delay Through Coding*
- *Low-Voltage Field Programmable Gate Arrays (FPGAs)*
- *Optimum Supply and Threshold Voltage Scaling for Low Power Digital Circuits*

Integrated Circuits and Systems

- *Energy-Efficient Hardware Reconfigurable Digital Signal Processing Architecture for Wireless Sensor Arrays*
- *Oversampled Pipeline A/D Converters with Mismatch Shaping*
- *Superconducting Bandpass Delta-Sigma A/D Converter*
- *Energy-Efficient Hardware Reconfigurable Digital Signal Processing Architecture for Wireless Sensor Arrays*
- *Oversampled Pipeline A/D Converters with Mismatch Shaping*
- *Superconducting Bandpass Delta-Sigma A/D Converter*
- *Low Power Reconfigurable Analog-to-Digital Converter*
- *Substrate Noise Shaping in Mixed-Signal Systems*
- *Analog Circuit Design with Scaled CMOS Devices*
- *Intelligent Transportation Systems*
- *Recognition of Three-dimensional Compressed Images without Decompression*
- *Priority Control for Networks*
- *Fusion of Machine Vision and Radar Sensors*
- *A Programmable, Wide Dynamic Range CMOS Imager with On-Chip Automatic Exposure Control*
- *A Differential Passive Pixel Imager*
- *Characterization Methodology of CMOS Processes for Image Sensor Applications*
- *Three-Dimensional Integration: Analysis, Modeling and Technology Development*
- *Software Tools for Process-Sensitive Reliability Assessments of IC Designs*

Integrated Circuits and Systems



Cross sectional view of a proposed three-dimensional integrated circuit formed by low-temperature wafer bonding.

Courtesy: A. Rahman, A. Fan, S. Das, K-N. Chen and R. Tadealli (R. Reif). Research sponsored by MARCO Focused Center on Interconnect (MARCO/DARPA)

The MIT μ AMPS Project

Personnel

A. Chandrakasan

Sponsorship

DARPA, ARL Advanced Sensors Federated Lab, Texas Instruments

Recent advances in MEMS-based sensor technology, low-power analog and digital electronics, and low-power RF design will enable the development and deployment of relatively inexpensive, low-power wireless microsensors. Microsensors are not as reliable or as accurate as their expensive macrosensor counterparts, but their size and cost enable applications to network hundreds or thousands of nodes in order to achieve high quality and fault-tolerant sensing. Reliable environment monitoring is important in a variety of commercial and military applications such as machine monitoring and target detection.

Microsensor networks pose several new design challenges, which are being addressed by the μ AMPS project. First, the limited available energy makes direct communication to the remote basestation infeasible. Collaboration among nodes must be integrated into the protocol stack as a means of reducing system energy. This involves adaptive system partitioning, considering computation vs. communication trade-offs as well as efficient data aggregation and routing strategies. Second, the availability of resources (e.g., battery capacity) can be time-varying. Hence, software, algorithms and architectures are structured so that they are amenable to energy / quality scalability. Third, microsensor systems utilize low data rates and must be designed for low duty cycles. Minimizing idle mode and transient power is critical. Finally, limited battery volume coupled with long lifetime requirements suggest the use of alternate energy sources; it will become critical to scavenge energy from the environment and convert ambient energy to electric energy. A guiding principle in our design is one of power awareness, a concept that goes beyond the integration of low-power components but focuses in great depth on domain-specific energy reduction techniques, as discussed above. Power-aware methodologies leverage the diversity in the environment and operating conditions of the node to deliver

energy savings and hence, maximum system lifetime.

Our initial architecture based on commercial, off-the-shelf components is used to develop a framework for power aware microsensor system design. We have developed prototypes and demonstrations of innovative energy-optimized solutions at all levels of the system hierarchy including transceiver design, packetization and encapsulation, multi-user communication with emphasis on energy scalability, data routing and aggregation schemes, and a power aware micro-OS. We are optimizing vertically across the protocol stack and energy-scalable programmable solutions. Future directions include a sensor API and CAD support for exposing power awareness to the networks' users and designers, and integration of power aware methodologies and node architecture into ever-smaller form factors, culminating in a system-on-a-chip.

Energy-Efficient Communication Protocols for Wireless Microsensor Networks

Researchers have been studying wireless networks for a number of years and have developed fairly sophisticated

Personnel

W. Heinzelman (H. Balakrishnan, A. P. Chandrakasan)

Sponsorship

DARPA, ARL Advanced Sensors Federated Lab, Kodak Fellowship

protocols using clustering, multi-hop routing, and direct communication approaches. We are developing communication protocols for sensor networks, which differ from traditional wireless networks in both the function that these networks serve and the fact that each node has a very limited energy supply, requiring energy-efficient DSP algorithms, and network protocols to maximize system lifetime. In a sensor network, it is not the individual nodes' data that is important, but the combined knowledge that reliably describes the environment the nodes are sensing. Therefore, the end-user only needs a function of the data, rather than all the individual data. By exploiting this sensor-specific function, we can ensure energy-efficiency in our algorithms and protocols.

We have developed a framework for minimizing the energy dissipation of wireless protocols using energy models for computation and communication, optimization algorithms, application-specific communication optimizations, and joint optimization across different layers of the communication protocol stack. Using this framework, we created the LEACH (Low-Energy Adaptive Clustering Hierarchy) communication protocol, an adaptive clustering approach that includes localized coordination and control, rotating cluster-heads and associated clusters, and local data fusion and classification.

The cluster-heads in LEACH function as local control centers to coordinate communication within the cluster and perform local data processing. By rotating the cluster-head position, the energy load is equally distributed among all the nodes in the network; this reduces individual node failure by ensuring that there are no overly utilized nodes. LEACH also uses local data processing within each cluster to greatly reduce the amount of data that must be transmitted long distances to the base station, saving considerable amounts of energy. LEACH

achieves cross-stack optimization by combining application-level knowledge of the goal of the sensor network into the routing protocol, which allows for local processing to compress data in an intelligent manner. Simulations show that LEACH can achieve an order of magnitude reduction in energy dissipation compared with conventional approaches. LEACH is being incorporated in the μ AMPS system.

Energy-Scalable Source Tracking for Wireless Networked Sensors

Personnel

A. Wang (A. P. Chandrakasan)

Sponsorship

DARPA, ARL Advanced Sensors Federated Lab, Lucent Fellowship

Sensor collaboration in a network of sensors can be highly energy-efficient because redundant communication costs can be reduced. One application demonstrating sensor collaboration is source tracking using acoustic sensors. Within a sensor cluster, individual sensors detect an event, and transmit their data to the local cluster head. At the cluster head, the data is beamformed before transmitting the Line of Bearing result to the distant basestation or end-user. In this scenario, transmission energy is conserved since only the cluster head sensor is required to communicate to the distant basestation. However, this is done at the expense of computation energy done locally at the sensor cluster. Thus, it is important to develop energy-efficient signal processing algorithms for source tracking to be run at the sensor nodes.

Through our collaboration with the Army Research Lab (ARL), we have access to acoustic data of moving vehicles. We have developed a system-level technique to optimize energy by parallelizing computation through the network and by exploiting underlying hooks for power management. By parallelizing computation, the voltage supply level and clock frequency of the nodes can be lowered, which reduces energy dissipation. A 60% energy reduction for a sensor application of source localization is demonstrated. The results are generalized for finding optimal voltage and frequency operating points that lead to minimum system energy dissipation.

Optimal Collaborative Strategies for Distributed Sensor Networks

Personnel

M. Bhardwaj (A. P. Chandrakasan)

Sponsorship

DARPA, IBM Fellowship

Increasingly integrated and lower power electronics are enabling a new sensing paradigm - wireless networks composed of tens of thousands of highly integrated sensor nodes allow sensing that is far superior, in terms of quality, robustness, economics and autonomous operation, to that offered by using a few, ultra high precision macro-sensors. Such sensing networks are expected to find widespread use in a variety of applications including remote monitoring (of climate, equipment etc.) and seismic, acoustic, medical and intelligence data-gathering. Since these integrated sensor nodes have highly compact form factors and are wireless, they are highly energy constrained. Furthermore, replenishing energy via replacing batteries on up to tens of thousands of nodes (in possibly harsh terrain) is infeasible. Hence, it is well accepted that the key challenge in unlocking the potential of such networks is conserving energy with a view to maximizing their post-deployment active lifetime - the central theme of our work.

To begin with, we ask a fundamental question concerning the limits of energy efficiency of sensor networks - what is the upper bound on the lifetime of a sensor network that collects data from a specified region using a certain number of energy-constrained nodes? The answer to this question is valuable for two main reasons. First, it allows calibration of real world data-gathering protocols and an understanding of factors that prevent these protocols from approaching fundamental limits. Secondly, the dependence of lifetime on factors like the region of observation, the source behavior within that region, basestation location, number of nodes, radio path loss characteristics, efficiency of node electronics and the energy available on a node, is exposed. This allows architects of sensor networks to focus on factors that have the greatest potential impact on network lifetime. By employing a combination of theory and extensive simulations of constructed networks, we show that in all data gathering scenarios presented,

Continued

An Energy-Efficient Link Layer Protocol for a Wireless Sensor Network

Personnel

E. Shih (A. P. Chandrakasan)

Sponsorship

DARPA

there exist networks which achieve lifetimes equal to or >95% of the derived bounds. Hence, depending on the scenario, our bounds are either tight or near-tight.

We have also developed a framework to determine the optimal manner of assigning roles (sensing, relaying, aggregating, compressing, sleeping etc.) to nodes and changing this assignment dynamically such that lifetime is maximized. The proposed formalisms and techniques allow data-gathering lifetimes that are potentially an order of magnitude greater than those achieved by previously proposed schemes.

The purpose of the link layer is to provide reliability and framing for data that needs to be sent over the physical network. Because different networks have different physical characteristics, different levels of reliability maybe needed. Over a wireless channel, errors can occur during transmission of data if unprotected. In the link-layer, reliability is often improved through the use of error correction coding (ECC) or retransmissions. If reliability must be ensured and energy and latency are not constraints, retransmission is the best approach to use since duplicate messages can be sent until the message is received correctly. However, the amount of communication energy used may be enormous, and the latency may be unacceptable. ECC, on the other hand, may require more computational energy, but may reduce the overall transmission energy required to successfully send the message.

We will attempt to show that by choosing between various ECC schemes, power can be partitioned between the power amplifier and the ECC computation. By choosing the correct parameters during a communication session, it is possible to minimize the system energy used for computation and transmission for a certain BER. In addition to the type of ECC used, we will try to demonstrate that energy can be minimized through software manipulation of a number of different parameters, such as packet length. Currently, theoretical analysis and simulations are being developed to test our hypotheses. At the same time, a testbed that will include a real 2.4 GHz radio is being designed and built.

Dynamic Power Management in Wireless Sensor Networks

Personnel

A. Sinha (A. P. Chandrakasan)

Sponsorship

DARPA

The first step towards a dynamic power managed sensor network is to have a power aware sensor node description. A power aware sensor node sleep model essentially describes the power consumption in different levels of node sleep state. Every component in the node can have different power modes, e.g. the processor can be in active, idle or sleep mode; the radio can be in transmit, receive, standby or off mode. Each node sleep state corresponds to a particular combination of component power modes. Every component power mode is associated with a latency overhead for transitioning to that mode. Therefore each node sleep mode is characterized by an overall power consumption and latency overhead. However, from a practical point of view not all the sleep states are useful. Each of these node sleep modes correspond to an increasingly deeper sleep state and are therefore characterized by an increasing latency and decreasing power consumption. These sleep states are chosen based on actual working conditions of the sensor node e.g. it does not make sense to have the memory in the active state and everything else completely off. The design problem is to formulate a policy of transitioning between states based on observed events so as to maximize energy efficiency.

A set of sleep time thresholds corresponding to the states can be computed such that transitioning to a sleep state from the active state will result in a net energy loss if the idle time is less than a threshold. This assumes that no productive work can be done in the transition period, which is invariably true, e.g. when a processor wakes up the transition time is the time required for the PLLs to lock, the clock to stabilize and the processor context to be restored.

We are working on embedded operating systems in the sensor nodes that can use these energy gain sleep time thresholds along with observed event arrival statistics to formulate an optimal power aware sleep state transition strategy. Figure 1 shows the results of our simulations for a 1000 node network scattered over a 100x100m area. The left graph shows the spatial distribution of events while the right graph shows the spatial energy consumption in the network. It can be seen that the node energy consumption tracks the event probability. In the non-power managed scenario we would have a uniform energy consumption in all the nodes.

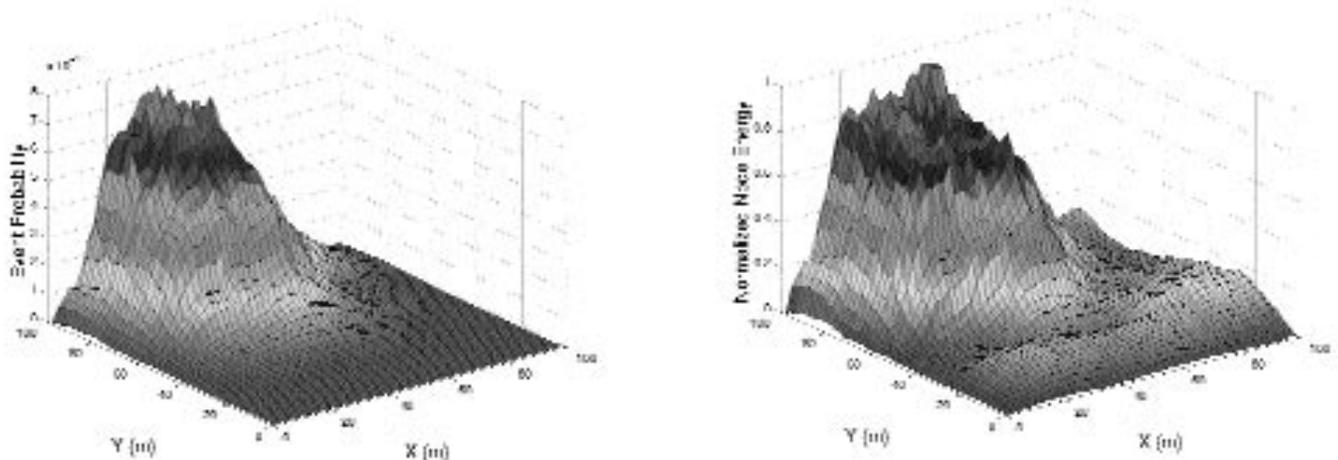


Fig. 1 : Event probability and spatial node energy.

Continued

Hardware Architecture For a Power-Aware Microsensor Node

Personnel

R. Min (A. P. Chandrakasan)

Sponsorship

DARPA, NDSEG Fellowship

The first prototype of a power-aware microsensor node has been developed for the MIT μ AMPS project. This prototype is constructed with commercial, off-the-shelf components for rapid implementation and provides a hardware fabric for the immediate demonstration of power aware design methodologies at all levels of the system hierarchy.

Power. Switching and linear regulators generate 5V, 3.3V, and adjustable 0.9-1.5V supplies from a 3.6V battery. The 5V supply powers the analog sensor circuitry and A/D converter. The 3.3V supply powers all digital components on the sensor node with the exception of the processor core. The core is powered by a digitally adjustable switching regulator that can provide 0.9V to 1.6V in twenty discrete increments. This variable-voltage regulator allows the SA-1100 to control its own core voltage, enabling dynamic voltage scaling.

Sensors. The node includes seismic and acoustic sensors. The seismic sensor is a MEMS accelerometer capable of resolving 2 mg. The acoustic sensor is an electret microphone with low-noise bias and amplification. The analog signals from these sensors are conditioned with 8th-order analog filters and are sampled by a 12-bit A/D. The high-order filters eliminate the need for oversampling and additional digital filtering in the SA-1100.

Processor and memory. A StrongARM SA-1100 microprocessor is selected for its low power consumption, sufficient performance for signal processing algorithms, and static CMOS design. The memory map mimics the SA-1100 "Brutus" evaluation platform, including both RAM and ROM. The lightweight, multithreaded " μ OS" running on the SA-1100 is an adaptation of the eCOS microkernel that has been customized to support demonstrations of power-aware methodologies. The μ OS, data aggregation algorithms, and networking firmware are embedded into ROM.

Radio. The radio module is based on a commercial sin-

gle-chip transceiver optimized for ISM 2.45 GHz wireless systems. The PLL, transmitter chain, and receiver chain are capable of being shut-off under software or hardware control for energy savings. The radio module is capable of transmitting up to 1 Mbps at a range of up to 15 meters.

The digital processing sections of the sensor node have been prototyped with commercial, off-the-shelf components as illustrated in Figure 2.

Dynamic voltage scaling (DVS) exploits variabilities in processor workload and latency constraints and realizes this energy-quality tradeoff at the circuit level. The switching energy of any particular computation on CMOS logic is $E_{\text{switch}} = C_{\text{tot}} V_{\text{dd}}^2$, a quantity that is independent of time. Reducing V_{dd} offers a quadratic savings in switching energy at the expense of additional propagation delay through static logic. Hence, we can reduce V_{dd} and the processor clock frequency together to trade off performance for energy savings.

Figure 3 illustrates the regulation scheme on our sensor node for DVS support. The μ OS running on the SA-1100 selects one of the above eleven frequency-voltage pairs in response to the present and predicted workload. The regulator controller receives a 5-bit voltage request and typically drives the new voltage on the buck regulator in 100 ms or less. At the same time, the SA-1100's PLL is commanded to lock to a new clock frequency through a control register write. Relocking the PLL requires 150 ms.

Our implementation of the above system demonstrates energy-quality tradeoffs with DVS. For example, the quality of a real-time FIR digital filter is varied by adjusting the number of filter taps. Reducing the number of taps reduces the result quality and processor workload simultaneously. When DVS is applied to

Continued

exploit this workload variability, the SA-1100 consumes up to 60% less energy than a fixed-voltage counterpart.

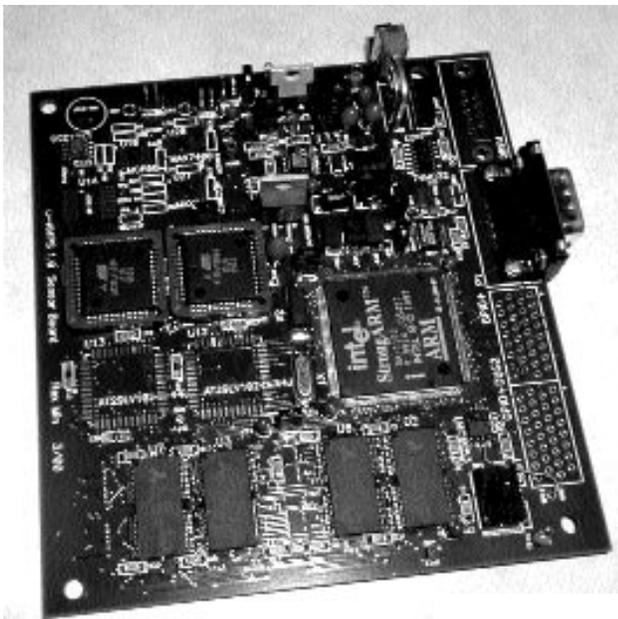


Fig. 2: Prototype of the power-aware digital processing sections of the first μ AMPS sensor node.

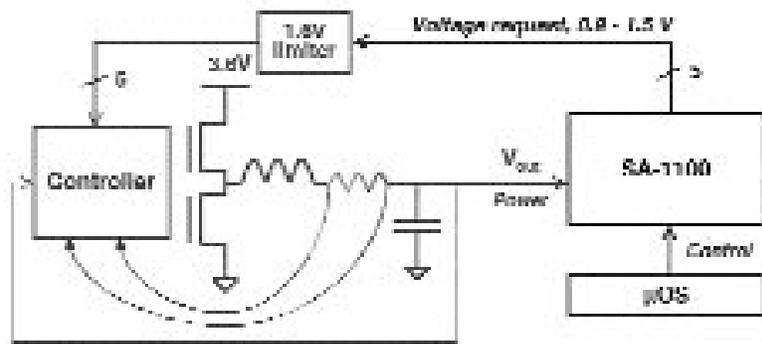


Fig. 3 Feedback for dynamic voltage scaling.

Reconfigurable Architectures for Energy-Efficient, Algorithm-Agile Cryptography

Personnel

J. Goodman (A. P. Chandrakasan)

Sponsorship

DARPA, National Semiconductor Fellowship

In the past there have been several standards for implementing various asymmetric techniques such as the ISO, ANSI (X9.*), and PKCS standards. The variety of standards has resulted in a multitude of incompatible systems that are based upon different underlying mathematical problems. The IEEE 1363 Standard for Public Key Cryptography recognizes three distinct families of problems upon which to implement asymmetric techniques: integer factorization, discrete logarithms, and elliptic curves. Each family has its advantages and disadvantages (e.g., IF and DL have been around for many years, allowing them to be thoroughly scrutinized for flaws, whereas EC appears to be much more resilient to cryptanalytic attacks but is still relatively new so users should be less willing to trust it).

As a result, system developers have had to either utilize software-based techniques in order to achieve the algorithm agility required to maintain compatibility, or use special purpose hardware and restrict themselves to only providing secure communications with compatible systems. However, software-based approaches lead to computationally intensive implementations that are very energy-inefficient. Hardware-based implementations on the other hand are very inflexible and capable of supporting only a single algorithm.

We attempt to compromise between these two extremes by taking advantage of the fact that the range of required operations is small enough that we can develop domain specific reconfigurable hardware that is capable of implementing the various algorithms. Furthermore, we do this in an energy-efficient manner that enables us to operate in the portable, energy-constrained environments where this algorithm agility is required most of all.

The resulting implementation is known as the Domain

Specific Reconfigurable Cryptographic Processor (DSRCP). The DSRCP differs from conventional reconfigurable implementations (e.g., Field Programmable Gate Arrays (FPGA)) in that its reconfigurability is limited to the domain of asymmetric cryptography. This domain requires only a small set of configurations. As a result the reconfiguration overhead is small in terms of performance, energy efficiency, and reconfiguration time. The resulting design has been fabricated in a $0.25\ \mu\text{m}$ CMOS technology with 5 levels of metallization. Figure 4 depicts a microphotograph of the processor whose core contains 880,000 devices and measures $2.9 \times 2.9\ \text{mm}^2$. At 50MHz the processor operates at a supply voltage of 2V and consumes at most 75 mW of power (the power consumption of the processor is a function of both the instruction being executed and the operand sizes being processed). In ultra low power mode (3MHz @ 0.7V) the processor consumes at most 525 μW .

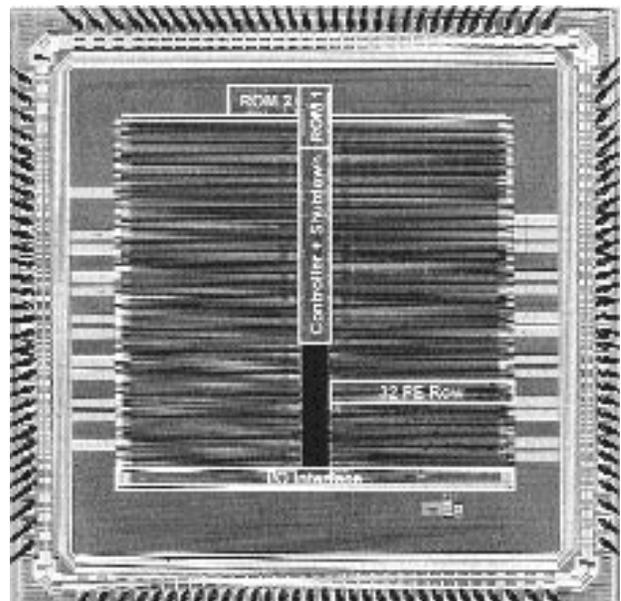


Fig. 4: DSRCP die photograph.

Ultra Low Power Radio for Wireless Devices

Personnel

A. Chandrakasan, H. Lee, C. Sodini

Sponsorship

ABB

There is an explosive growth in the use of battery operated wireless devices such as cellular phones, PDA's, etc. While power efficient radio solutions exist for high rate applications, very little work has been done on low rate applications such as sensors. The goal of this project is develop a complete communication module (basic building blocks as well as the protocol stack implementation) for emerging sensor applications.

There are three major system constraints. First, we assume that a significant number (100's) of energy-constrained sensors in a small area (5m x 5m) may communicate to a high powered base station. As a result of the short communication distance, the energy for the transmitter electronics becomes a significant issue in addition to the transmitted RF power. Second, reliability is very critical and the system must provide a very low Telegram Error Rate (e.g., 10^{-12}). Finally, the packet sizes are very small (10's of bits) compared to multimedia traffic. To meet the specifications, we are developing physical layer electronics for both the energy-constrained sensor as well as the base station (Figure 5). Optimizing the multiple-access protocols and modulation schemes will enable us to reduce sensor complexity at the cost of basestation complexity.

In current radio systems, there is significant overhead in

turning on and off the transmitter circuitry, resulting in a large response time. The energy dissipation for transmitting short packet sizes is dominated by the turn on/off energy. Therefore, for the transmitter, the focus is on circuit techniques to achieve a swift response. Various approaches have been explored including a fast turn on time frequency synthesizer combined with direct modulation of the VCO (open loop approach).

The base station electronics focuses on two components, the design of a high sensitivity receiver and a high resolution IF frequency data converter. The impact of placing more computation at the base station to relax the requirements of the transmitter is also being investigated. Various receiver architectures will be explored along with the technology choice of CMOS vs. BICMOS to obtain the highest sensitivity. The A/D work for the base station will focus on high speed and high resolution conversion. The base station electronics will be designed for the 2.4GHz carrier frequency and the goal is to demonstrate a complete integrated radio system.

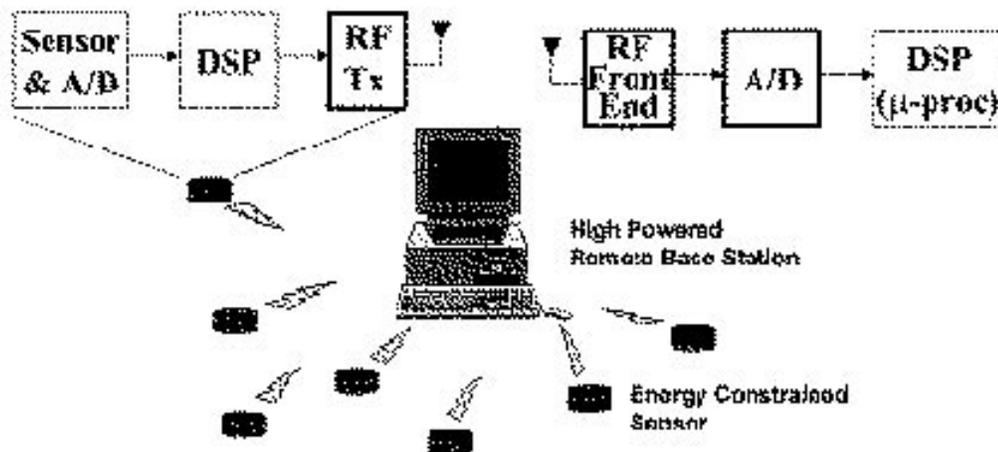


Fig. 5: Block diagram of the wireless sensor system highlighting the radio sub-system.

A Low Power Transmitter for Wireless Sensor Networks

Personnel

S. Cho (A.P. Chandrakasan)

Sponsorship

ABB

The communication module of a wireless sensor must be designed for low duty cycle activity. For short range transmission at GHz carrier frequencies, the power is dominated by the radio electronics (frequency synthesizer, mixers, etc.) and not the actual transmit power. In order to save power in the radio module, the electronics must be turned off during idle periods. Unfortunately, frequency synthesizers require a significant overhead in terms of time and energy dissipation to go from the sleep state to the active state. For short packet sizes, the transient energy during the start-up can be significantly higher than the energy required by the electronics during the actual transmission. Hence we are developing techniques to reduce the startup transient and power consumption of the frequency synthesizer. Specifically, we are exploring a variable loop bandwidth sigma-delta synthesizer architecture where we exploit trade-offs between divider, VCO and sigma-delta complexity. A 5.8GHz frequency synthesizer with integrated VCO will be implemented in 0.25 μm CMOS process, which has fast start-up transient ($\sim 10\mu\text{s}$) and low power consumption ($\sim 10\text{mW}$).

In addition to designing the physical layer of the radio, we have been exploring power efficient MAC protocols and modulation schemes. By using a detailed communication model of the radio which includes non-ideal behavior such as start-up time, various MAC layer protocols and modulation approaches are evaluated from an energy perspective. The overall goal is to develop sensor network that provides more than an order magnitude power reduction compared to conventional approaches by optimizing at all levels of abstraction.

Energy Minimization for Wireless Microsensor Systems

Personnel

A. Wang (C. G. Sodini)

Sponsorship

ABB and NSF Fellowship

The design of wireless microsensor systems have gained increasing importance for a variety of civil and military applications. With the objective of providing short-range connectivity with significant fault tolerances, these systems find usage in such diverse areas as environmental monitoring, industrial process automation, and field surveillance.

The main design objective is maximizing the battery life of the sensor nodes while ensuring reliable operations. For many applications, the sensors need to “live” for 1-5 years without battery replacement. To achieve this goal, the microsensor system has to be designed in a highly integrated fashion and optimized across all levels of system abstraction. This also means that all the characteristics particular to the microsensor system must be exploited. One such characteristic is that the RF output power is small due to the short transmission distance, which makes the transmitter electronics the dominant source of energy dissipation.

This research develops a unique design methodology that brings the system and circuit issues together to analyze the transmitter energy consumption as a whole. The effect of various modulation schemes on the total energy dissipation is analyzed, and the energy is minimized on the global level. It is found that significant energy savings can be achieved by trading off RF output power, which depends on the modulation technique, with transmitter electronics complexity. This approach is general enough that it can be applied to all short-range wireless systems. Current research focuses on building key RF front-end components to verify the above analytical results.

Analog-to-Digital Conversion of High-Frequency Signals for RF Communication

Personnel

Susan Dacy (H-S. Lee)

Sponsorship

Lucent Fellowship and ABB

In wireless receiver design, digitizing the received signal at the intermediate frequency (IF) or higher permits the channel selection filters and demodulation to be performed in the digital domain. This results in a highly programmable receiver, often coined 'software radio.' In pursuit of this completely digitally programmable 'software radio,' there has been a push to convert at higher IFs in a superheterodyne receiver. This has created a demand for high performance, high speed analog-to-digital converters (ADCs). Bandpass sigma-delta modulators with high center frequencies have become popular for such narrowband high-resolution analog-to-digital conversion. Of particular interest for radio applications are converters whose sampling frequency is four times the carrier frequency, called $f_s/4$ bandpass-sigma-delta modulators. This ratio of IF to sampling frequency makes mixing down to baseband particularly simple in the digital domain. As the sampling frequency increases, continuous-time sigma-delta modulators have been preferred because they do not have the op-amp settling-time constraints or the fast, high-precision sample and hold (S/H) requirements that limit the maximum clock rate in discrete-time sigma-delta modulators and upfront sampled ADCs. Furthermore, with high IF inputs, sampling jitter degrades the maximum achievable SNR in not only continuous-time sigma-delta modulators, but in any ADC. As the input frequency increases, clock jitter becomes the dominant limit on SNR performance of ADCs. In this research, we focus on techniques that reduce the sensitivity to random clock jitter.

Previous work has analyzed the fundamental SNR limit in discrete-time and continuous-time sigma-delta modulators. Single bit continuous-time sigma-delta modulators have been shown to have lower SNR than discrete-time sigma-delta modulators or upfront sampled converters for the same amount of clock jitter. A few recent high speed continuous-time sigma-delta modulator designs have indeed been limited by jitter.

In practice they have only been able to achieve noise shaping up to an oversampling ratio of approximately 15, above which they are limited by in-band white jitter noise rather than shaped quantization noise.

Jitter in the DAC feedback pulse limits performance in continuous-time sigma-delta modulators. This jitter modulates the width of the feedback current pulse which changes the amount of charge fed back. This research uses a non-rectangular discrete-time feedback signal as the DAC feedback pulse in a continuous-time sigma-delta modulator. Theoretical analysis has shown that the new continuous-time sigma-delta modulator shows significantly decreased clock jitter sensitivity over upfront sampled ADCs as well as conventional continuous-time sigma-delta modulators for high input frequencies. Thus the 'fundamental' jitter limit is broken, offering the ability to build high speed, higher resolution converters.

Continued

Wireless Gigabit Local Area Network

Personnel

A. Chandrakasan, H. -S. Lee, C. G. Sodini, and G. W. Wornell

Sponsorship

Center for Integrated Circuits and Systems

The exploding number of electronic devices or “appliances” requiring high bandwidth communication will continue to drive the need for higher speed (Gigabit-per-second, Gb/s) networking. We assume that the Next Generation Internet (NGI) will carry high-speed data to and from the home or office. However, a local area network (LAN) within these structures is necessary to continue high-speed data transmission to and from end-use devices, such as cameras, displays, printers, high resolution video, mobile communicators, and novel devices. The enabling technology for this rich set of applications is a wireless Gb/s LAN, (WiGLAN), connected to the NGI.

The WiGLAN offers several research challenges. First, there is a wide range of data rates, quality of service, and need for real time transmission to and from the appliances. For example, voice transmission over the network will not require high data rates but may require low power dissipation for portability. Interactive video transmission requires real time transmission and very high data rates especially as high resolution video and 3D graphics become available.

System resources will need to be adaptive in order to support this wide range of appliances. Second, since many of the appliances will require portability, low power design techniques at the circuit, chip architecture and overall system level will be required. Third, this research requires synergy between a variety of disciplines including, communication system design at the physical layer, low power circuit and system design, digital signal processing algorithm and IC design, mixed signal IC design, and RFIC design. It also lends itself to a number of demonstration projects using some of the technology which results from this research. Besides the educational component of the PhD researchers directly involved, this program will generate a number of IC’s and algorithms which can be demonstrated by Masters student design projects.

A block diagram of the Wireless Gigabit Local Area Network, WiGLAN, is shown in Fig. 5a. We envision a network server being the gateway between the NGI and the local area network. Each appliance is attached to the network through a WiGLAN adapter, which is capable of providing a wireless connection to the network. This adapter should be physically small, implying a high degree of integration of the electronic functions required to interface digital data from the appliance to and from the network. The quality of service, QoS, which is a function of data rate and bit error rate, should be scaleable with power dissipation to permit battery operation of many appliances.

The network requirements of high bandwidth efficiency and real time transfer led to our choice of a multi-carrier modulation, such as orthogonal frequency division multiplexing, (OFDM) using M-quadrature amplitude modulation, (MQAM) signal constellations. We plan to digitize the entire signal bandwidth (150 MHz) available at the 5.8 GHz ISM band and adapt the bit rate (change M) within sub-bands according to the available signal-to-noise ratio (SNR) and interference in the sub-band. A programmable digital signal processor will perform this adaptive modulation.

The adaptive bit rate processor located in the network

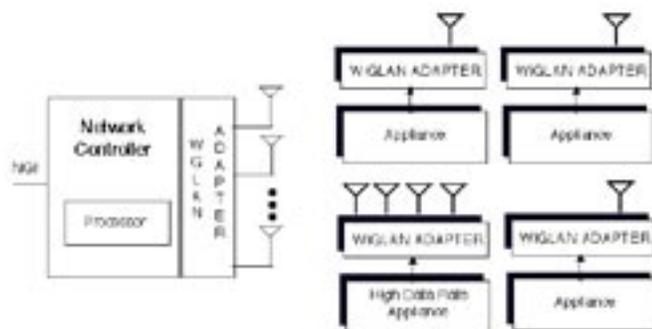


Fig. 5a

Continued

server will estimate the channel capacity by measuring the SNR and interference within sub-bands across the entire 150 MHz signal band. The channel estimation algorithm is a subject of this research. Depending on the SNR and interference, data modulation will range from simple phase shift keying (PSK) up to 256 level QAM with intermediate levels of QAM, (i.e. 4-QAM, 16-QAM, etc.) allowing for transmission of approximately 1b/Hz for PSK up to 8b/Hz for 256-QAM.

In order to provide the capacity enhancements required to support the target data rates, the system to be developed will make extensive use of multiple-element antenna arrays for both transmission and reception. A key component of the proposed research will therefore be the development of computationally and power efficient space-time coding and space-time processing algorithms that exploit the substantial diversity benefit inherent in the use of such antenna arrays. At the implementation level, multiple-element antenna arrays require a separate receive and transmit channel for each antenna element. To efficiently meet this requirement we propose to build a system of parallel radios divided into three distinct Integrated circuits, namely RF, Mixed signal, and DSP.

The WiGLAN network adapter consists of three functions, digital signal processing for multi-carrier adaptive bit rate QAM, a baseband analog processor performing data conversion and filtering, and an RF transceiver function which interfaces the modulated baseband data to a 5.8 GHz carrier. We will design and characterize integrated circuits to perform these functions.

Circuit Design and Technological Limitations of Silicon RFICs

Personnel

D. A. Hitko (C. G. Sodini)

Sponsorship

SRC, Hughes Research Laboratories, and LLC

Wireless products and communications systems have thrived on the increased utility enabled by semiconductor technologies, the demand for which is necessitating ever higher communication channel frequencies to obtain wider bandwidth and to alleviate interference. This places still greater demands on the technologies used to implement the wireless systems; however, for a given application, the market determines the acceptable end product price based on convenience, functionality, and a comparison with substitutes, effectively setting a bound on the technologies that can be used. The limitations of these technologies in part determine the achievable performance, which then in turn may confine the very convenience and functionality being sought through the wireless system.

In this interplay of circuits and systems with technology where both price and performance are crucial, we are exploring two aspects that can yield significant improvements in the design of wireless systems. The first of these is the optimization of a broad range of RF circuits at the device level. By working through the exercise of designing the components needed to realize a 5.8GHz receiver, generalities are being sought which link technological parameters with system level performance. Using the receiver application to frame circuit constraints, device level issues are being studied to determine the physical origins of circuit limitations. Design techniques are then being investigated to mitigate these limitations to the extent possible, a procedure that aims to both optimize the circuits and underscore the degree to which the limitations are fundamental.

Next, the knowledge of those factors that are limiting circuit performance are then used to devise methods of implementing circuits in which the performance, measured in terms of system level parameters, can be dynamically tuned to match real-time conditions in

a power efficient manner. This facility provides the mechanism by which energy can be conserved in RF circuits by sensing and using information about the environment, communications channel, and data to be transmitted. Incorporation of adaptability at the circuit and system levels is paramount in expanding capabilities and increasing utilization of wireless communication links, and yet remains a largely untapped resource in this field.

This project entails the application of these concepts to the development of the component-level circuits required in an integrated RF front end for a 1Gbit/s wireless network operating at 5.8GHz. The typical components in a receiver, namely, low noise amplifiers (LNAs), voltage-controlled oscillators (VCOs), and mixers are considered in this work. A set of VCOs and LNAs have been designed into a 0.5 μm SiGe BiCMOS technology as data points to illustrate device and circuit optimization trade-offs. Direct comparisons of CMOS versus bipolar, the impacts of transistor f_T/f_{MAX} , and the implications of design methodologies based upon linear time variant models are some of the issues being explored. The approach that has been developed along with information gleaned from the experimental circuits can provide a basis for shaping future integrated circuits, technologies, and system designs for wireless applications.

A Linear 5.8 GHz Power Amplifier with Variable Output Control

Personnel

A. Pham (C. G. Sodini)

Sponsorship

MIT Presidential Fellowship, Center for Integrated Circuits and Systems

Power amplifier design can often limit the overall performance of a wireless system. In the Wireless Gigabit LAN project we require variable output power from 0-20 dBm of a 150MHz signal centered at the 5.8 GHz band. This feature allows us to optimize the output power based on the required data rate and bit error rate. The power amplifier also must be ultra-linear across this bandwidth in order to allow for adaptive Quadrature Amplitude Modulation.

In designing power amplifiers, there is always a trade-off of between output level, efficiency and linearity. Since linearity is our biggest concern in this application, we are approaching the design using the following methodology:

- Chose a topology with the best linear characteristic.
- Chose a fabrication technology that fits our frequency band.
- Design matching and biasing circuitries that give us the best efficiency for the chosen topology.
- Use design techniques to improve the linearity performance such as filters and feedback.

Our initial design will be fabricated in a SiGe technology.

5.8 GHz Wideband Receiver for Wireless Gigabit LAN

Personnel

L. Khuon (C. G. Sodini)

Sponsorship

Center for Integrated Circuits and Systems(CICS) and SRC

The receiver for the Wireless Gigabit LAN (WiGLAN) performs amplification, filtering and downconversion of the 150 MHz signal centered at 5.8 GHz. The receiver downconverts the radio frequency (RF) signal to a baseband frequency that is fed to the analog baseband processor where it is equalized and digitized. The receiver uses a dual conversion architecture that allows for optimized frequency planning to reduce the effect of image frequencies. The design approach for the receiver is based upon block level analyses that consider the gain, noise, and linearity tradeoffs necessary for the WiGLAN's adaptive modulation scheme.

The focus of this research is the design and optimization of the components of the receiver. In particular, the emphasis is on the design of the critical front-end low noise amplifier (LNA) and mixer circuits. The LNA provides the necessary signal amplification before conversion and dominates the noise performance of the receiver. The mixer performs the actual downconversion to the IF frequency but can produce unwanted spurious signals. At present, the design of the LNA and mixer shall be fabricated with Silicon-Germanium BICMOS technology. Optimization of these circuits shall consider how linearity and noise, vary with power consumption, and size of the devices.

Analog Base-Band Processor for Gigabit Wireless LAN

Personnel

M. Spaeth (H-S. Lee)

Sponsorship

SRC

The baseband analog processor performs necessary signal processing on the 150 MHz baseband signal in the transmit and the receive signal paths for a wideband wireless local area networks. In the receive (Rx) section, the wideband amplifier amplifies the received signal from the RF transceiver network. The channel equalization filter then follows. The channel characteristic depends on the RF signal fade and interference. Broadcasting to more than one appliance requires channel equalization at the receiver. The amplitude of the signal after the channel equalization filter can vary greatly depending on the channel condition. Therefore, a programmable gain amplifier adjusts the signal amplitude to better match the dynamic range of the subsequent analog-to-digital conversion. The demodulation of the multi-carrier QAM signal is carried out in the digital domain by the DSP.

There are tremendous technical challenges in the development of the baseband analog processor. The analog circuits in both the receive and the transmit sections of the processor must handle the signals up to 150MHz with a high signal-to-noise ratio. These analog circuits include the wideband amplifier (WBA), the programmable gain amplifier (PGA), the anti-alias filter, the channel equalization filter, the A/D converter in the Rx section, and the D/A converter and the reconstruction filter in the Tx section. In order to digitize 150MHz-wide signal band, the A/D converter must have an effective sampling rate over 300 MHz, preferably around 600 MHz to ease the anti-alias and digital filtering requirements. Our preliminary estimate of the resolution of the A/D converter is at least 12 bits to handle the wide dynamic range of the received signal. At present, such high performance is beyond the capability of monolithic silicon IC's. Also, any harmonic and intermodulation distortion in the signal path produces spurious signals in other sub-bands. Therefore, the WBA, the PGA, the anti-alias filter, the channel equalization filter, and the A/D converter not only must have signal bandwidth of at least 150 MHz, but also must exhibit

very high spurious-free dynamic range (SFDR). In the Tx section, the D/A converter and the reconstruction filter must possess similar performance levels. In order to address these technical challenges we propose to investigate innovative techniques for the baseband analog processor.

Instead of employing separate blocks for the PGA, the channel equalization filter and the A/D converter, in the Rx section, we will explore combining the functions of all three blocks. Since the received signal will eventually be decomposed into many sub-bands, it is possible to employ an array of bandpass A/D converters operating in parallel. The bandwidth of each converter may cover one or multiple sub-bands. There are a number of advantages of this approach. Since the baseband is eventually separated into a number of sub-bands, the array of bandpass converters fit very well with the proposed network. Also, the array of a bandpass converters greatly facilitates the channel equalization. A simple adjustment of the gain of each bandpass A/D converter shapes the frequency response of the receive path which otherwise would have required a high-order programmable filter. Such a filter with 150 MHz bandwidth and the requisite linearity would have been unfeasible in silicon technologies at present. In the proposed approach, the array of bandpass A/D converters replace the PGA, channel equalization filter, and the A/D converter. Since the non-linearity and noise of the anti-alias filter must be very low, we will explore RC or RLC passive filters. This is possible because the signals are effectively oversampled due to a large ratio between the bandwidth of each converter in the array and the sampling rate, thus low order anti-alias filters will suffice.

Design of a Power-Scalable Digital Least-Mean-Square Adaptive Filter

Personnel

C. Ng (A. Chandrakasan)

Sponsorship

NSF

DSP based modem applications such as gigabit Ethernet transceivers require channel equalization. Because of the high rate and computation complexity involved, adaptive equalization filters consume a lot of power. Currently, equalization is typically hardwired instead of using a digital signal processor because of the large complexity. One of the approaches to building low-power systems is to develop hardware that scales its power consumption to different operating scenarios. We have used adaptive tap length and precision techniques to design a power aware digital adaptive equalizer.

The design is based on synthesis and place-and-route tools. Part of the work was to develop a synthesis based design flow. The tools in this flow included Synopsys for high-level simulation and logic mapping and Cadence Silicon Ensemble, for place-and-route of standard cells. Power estimation was performed using the PowerMill transistor level simulator. We demonstrated a trade-off between power dissipation and the quality of ISI-cancellation in an LMS adaptive filter. The adaptive filter accepts a 10-bit data stream and a 10-bit error stream, and produces a 10-bit filter output stream. The filter adapts its coefficients so that the mean-square of the error stream is minimized using a Least-Mean-Square algorithm. Clocked at 33MHz, the circuit demonstrated a power scalability from 6.4 to 20.4 mW with the corresponding output quality measured by standard deviation of error ranging typically from 0.28 to 0.1. Our simulations show that shorter filter length and smaller tap precision lead to larger standard deviation of error, but also consumes less power.

Circuits for Optical Clock Distribution

Personnel

T. Simpkins (A. Chandrakasan)

Sponsorship

NDSEG Fellowship

Clock distribution has become a major problem in integrated circuits for two main reasons. As clock cycle times have decreased, the portion of the timing budget allocated to uncertainty in the arrival time of the clock has remained constant. Therefore, the percentage of the budget devoted to this uncertainty has become significant. Furthermore, as the number of transistors on a chip has increased, so has the load presented to the clock buffers. Since the energy dissipated by a circuit is directly proportional to the size of the load being driven and the frequency of operation, the amount of energy dissipated by the clock network now accounts for nearly half of the energy consumed by a modern microprocessor.

One possible solution to the uncertainty and energy consumption problems of clock networks is to distribute the clock using pulses of light rather than electrical approaches. Although this idea has been previously proposed, it has not yet been successfully demonstrated and will therefore be the focus of this research. In addition to the circuits needed to implement the optical-to-electrical conversion, the research will also focus on methods for measuring clock uncertainty on-chip.

Active GHz Clock Network Using Distributed PLLs

Personnel

V. Gutnik (A. Chandrakasan)

Sponsorship

Interconnect Focus Center (MARCO-DARPA)

Clock skew and jitter continue to increase with scaling, and will consume an ever-larger fraction of the cycle time. We have developed a distributed clocking approach that allows faster clock speeds with lower random skew and jitter than possible with traditional clock tree distribution methods.

In our distributed clock approach, the clock signal is generated with phase locked loops at multiple points across a chip, and distribution happens only to small local tiles. Phase comparators at the boundaries of each tile produce an error signal that is summed by an amplifier in center of each tile and used to adjust the frequency of the node oscillator. Skew is introduced only by asymmetries in phase detectors instead of mismatches in physically separated buffers as with conventional tree based distribution. Also importantly, the clock is regenerated at each node, so jitter does not accumulate. To make this system functional, several issues had to be addressed including stability, power supply noise rejection, and locking. A proof of concept test chip (Figure 6) with 16 oscillators (4x4 array) was fabricated using 0.35 μm and showed stable operation at 1.4Ghz.

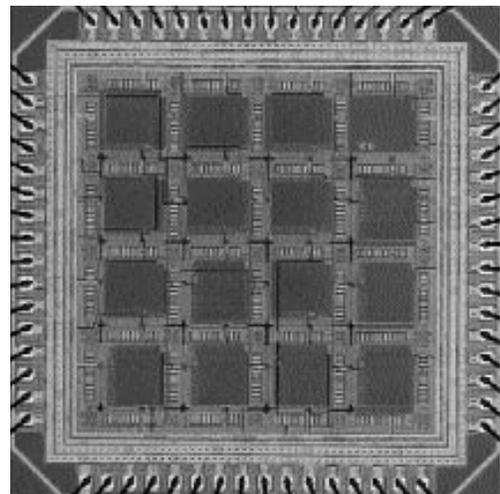


Fig. 6: Distributed clock test chip with 16 oscillators.

Reduction of Interconnect Power and Delay Through Coding

Personnel

P. Sotiriadis (A. P. Chandrakasan)

Sponsorship

Interconnect Focus Center (MARCO-DARPA)

Technology downscaling gives rise to important interconnect problems. Energy dissipation associated with driving long wires accounts for a significant fraction of the overall system energy. Signal delays and integrity become critical with higher clock rates.

Within this work new mathematical models have been developed to capture the deterministic and stochastic energy behavior and the signal transmission properties of data buses. The wires are modeled as coupled distributed RLC lines driven by independent sources. Closed formulas provide accurate estimates for delay and energy dissipation.

Using these models several new techniques based on coding theory have been proposed and optimized to actively reduce the delay and energy per bit that is needed to transmit data over long wires. We have shown a theoretical energy reduction of more than 40% and a delay reduction of 45%. A combination of coding and charge redistribution techniques have resulted in 60% energy reduction with a small delay increase.

Low-Voltage Field Programmable Gate Arrays (FPGAs)

Personnel

T. Konstantakopoulos (A. P. Chandrakasan)

Sponsorship

Interconnect Focus Center (MARCO-DARPA)

Field Programmable Gate Arrays are being used increasingly in embedded general purpose computing environments. The need for combining heterogeneous programmable architectures on a chip, makes FPGAs ideal for use in System-On-Chip environments. Although today's designs have been optimized for performance and density, energy efficiency has barely been taken into account.

A detailed breakdown of the power that is dissipated on existing designs is essential for the design of a low energy FPGA. After this essential first step, we will develop energy efficient reconfigurable architectures. The FPGA fabric will use aggressive sub-1V design to conserve power. Additionally, the basic logic cells will be designed in a power aware manner. Power reducing techniques such as charge recycling will also be used in order to reduce the component that is being dissipated in the interconnect. This portion is considerably larger when compared to conventional CMOS designs due to the elaborate routing and reconfiguration network.

Optimum Supply and Threshold Voltage Scaling for Low Power Digital Circuits

Personnel

J. Kao, M. Miyazaki (A. P. Chandrakasan)

Sponsorship

Hitachi, DARPA

In modern integrated circuits, the dominant source of power dissipation is the dynamic component, resulting from the charging and discharging of the circuit capacitances. Another increasingly important element of overall power dissipation is subthreshold leakage power, which increases exponentially as threshold voltages are scaled. For constant performance, it is possible to continue scaling supply voltages as well as threshold voltages until a minimum power point is achieved. At this extrema, the incremental decrease in dynamic power due to reduced supply voltage is equal to the incremental increase in leakage power due to reduced threshold voltages. Depending on the target frequency, this minimum power point will vary, i.e. for slow frequencies the power will be leakage dominated, while at high frequencies the power will be dynamic dominated.

If a triple well technology is available, it is possible to employ both forward and reverse body biasing adaptively to adjust both the chip's supply and threshold voltages to minimize overall power dissipation for different operating conditions. A framework for this optimal supply and threshold voltage biasing scheme is shown in Figure 7.

This scheme relies on a lookup table approach where supply voltages are tabulated as a function of varying workload and temperature conditions. This table can be generated through careful circuit simulations or can even be updated during a chip testing / calibration step during manufacturing. Although the supply voltage level is set in an open loop fashion, the target threshold voltage level, which is set by the body bias applied, is set in a closed loop fashion which can be used to compensate out chip to chip variations or time varying changes such as temperature or mobility degradation to ensure that the chip operates only as fast as necessary. The supply voltage choice can also be recalculated periodically (to reflect any large temperature changes) to ensure that the minimal power operating point will be met. Separating this control loop into a single dimensional feedback loop (closed loop around the body bias selection) is simpler than placing both the supply voltage and the body bias in a two dimensional feedback loop. Also, since threshold voltage variations will become increasingly large as one continues to scale supply and threshold voltages, it makes sense to control the threshold voltage level in a closed loop fashion.

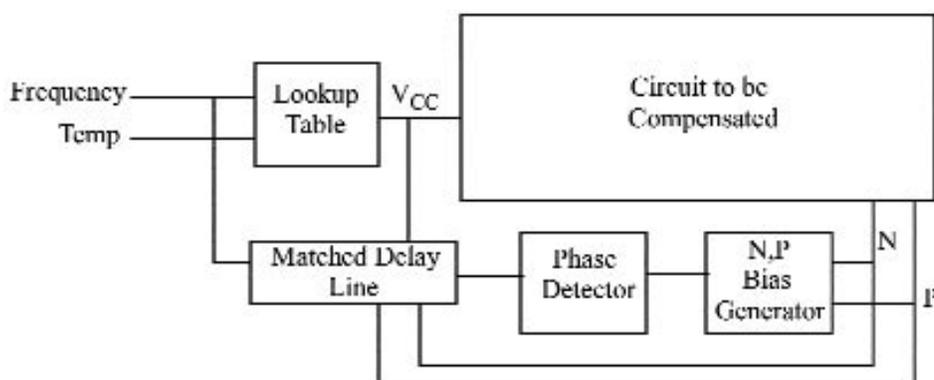


Fig. 7: Framework for optimal V_{CC}/V_t scaling for varying workload conditions.

Energy-Efficient Hardware Reconfigurable Digital Signal Processing for Wireless Sensor Arrays

Personnel

F. Honore (A.P. Chandrakasan)

Sponsorship

DARPA, Texas Instruments

To maximize the data-gathering lifetime of sensor nodes in a sensing network, energy-efficient hardware for data processing must be adaptable under current operating conditions. While a programmable DSP is well suited to these data processing tasks (filtering, FFT, etc), we believe finer grain control of the datapath is required to reduce the overhead present in register-based, fixed-width datapath designs of a programmable DSP. It is well known that custom hardware that is bit-width optimized can be orders of magnitudes more energy efficient than a programmable DSP. However, the custom approach must be designed to work in the worst case scenario. Our architecture strives to support limited hardware reconfigurability for a multiple functional unit design. With this approach we can take advantage of potential savings under better than worst case operating scenarios for improved energy efficiency over a hardwired custom datapath design while maintaining a modest amount of flexibility.

To achieve this goal, careful consideration is given to the overhead for reconfiguration of the functional units. Multiple coarse-grain functional units allow for greater optimization of each task common to a sensor node's data-processing requirements. Within each functional unit, key performance metrics are parameterized such as bit-width precision and rate. Processing is in data-flow fashion minimizing data storage between hardware blocks. A general purpose processor implements control functions and monitors operating conditions at a low duty cycle. We expect to show through simulation and eventual hardware implementation the benefits of this approach in this application domain. This project will employ high-level design methodology to allow algorithm development combined with micro-architectural and circuit-level optimizations.

Oversampled Pipeline A/D Converters with Mismatch Shaping

Personnel

A. Shabra (H-S. Lee)

Sponsorship

Center for Integrated Circuits & Systems (CICS)-

In recent years, delta-sigma modulators and pipeline converters have been considered as possible realizations of analog-to-digital converters for wide-band signals. In comparing these converters, we recognize a few important attributes. Due to the wide bandwidth of the input signal and limited circuit speed, delta-sigma converters afford only low oversampling ratios, which makes high-resolution conversion extremely difficult. The low oversampling ratio generally nullifies the primary advantage of delta-sigma converters; the tolerance to component mismatches. In this regard, remaining potential advantages of delta-sigma converters over pipeline converters now only include ease of anti-alias filtering and low quantization noise. It must be noted that the ease of anti-aliasing is not inherent to delta-sigma modulation. Rather, it is associated with oversampling. Therefore, pipeline converters can experience the same benefit of easy anti-aliasing by simply operating the converter at higher sampling rate than the Nyquist rate, i.e., oversampling. As for quantization noise in pipeline converters, the quantization noise can be made smaller by adding more stages at the end of the pipeline. Since the last stages of the pipeline do not contribute much thermal noise, they can be made extremely small and low power. Therefore, the quantization noise itself can be made arbitrarily small with negligible increase of area and power. Certainly, doing so will not improve the accuracy or thermal noise. However, it is no different in delta-sigma converters with low oversampling ratio.

Based on the above observation, we can conclude that delta-sigma converters do not possess any fundamental advantage over pipeline converters for wide-band applications that necessitate low oversampling ratios. At this low oversampling ratio many delta-sigma converters are incapable of providing good enough performance. While there are a few examples of delta-sigma converters with a low oversampling ratio, we believe that a more efficient approach would be to oversample

Continued

a standard pipeline converter, and shape the distortion due to mismatch out of the signal band, where it will be removed by a subsequent digital filter. Since no attempt is made to shape the quantization noise, there are none of the concerns associated with delta-sigma converters with a low oversampling ratio.

A test chip was fabricated in a 0.35 μm CMOS process to demonstrate a number of mismatch shaping concepts. A 77dB SFDR and 67dB SNDR is achieved at an oversampling ratio of 4 and a sampling rate of 60Msample/s. Mismatch shaping improves the converter SFDR by 12dB's and SFDR by 5dB's.

Superconducting Bandpass Delta-Sigma A/D Converter

Personnel

J. F. Bulzacchelli (H-S. Lee and M. Ketchen, IBM)

Sponsorship

Center for Integrated Circuits & Systems (CICS)

In this program, we present the design of a superconducting-bandpass delta-sigma converter for direct A/D conversion of-GHz RF signals. The schematic of the circuit is shown in Figure-8. The input signal is capacitively coupled to one-end of a superconducting microstrip transmission line, which-serves as a high quality resonator (loaded $Q > 5000$). The-current flowing out of the other end of the microstrip line is-quantized by a clocked comparator comprising two Josephson-junctions. If the current is above threshold, the lower-junction switches and produces a quantized voltage pulse known-as a single flux quantum (SFQ) pulse. If the current is below-threshold, the upper junction switches instead. The pattern-of voltage pulses generated across the lower Josephson-junction represents the digital output code of the delta-sigma-modulator. These voltage pulses also inject current back into-the microstrip line, providing the necessary "feedback" signal-to the resonator. At the quarter-wave resonance of the-microstrip line (about 2 GHz in our design), the resonator-shunts the lower junction with a very low impedance, the-"feedback" current to the resonator is maximized, and the-quantization noise is minimized. Because of the high speed of-Josephson junctions and the simplicity of the proposed-circuit, we expect sampling frequencies in excess of 20 GHz,-limited only by the digital circuitry needed to process the-output of the delta-sigma modulator.

Circuit performance at a 20 GHz sampling rate has been-evaluated with JSIM, a SPICE-like simulator for-superconducting circuits. A representative example of the A/D-converter's output spectrum is shown in Figure 9. In-this simulation, the A/D converter was driven by a large-(-0.8 dBFS) input near 2.13 GHz, just above the frequency band-of interest. The minimum in the quantization noise power-spectrum is located at 2.05 GHz. In-band noise is -53 dBFS-and -57 dBFS over bandwidths of 39 MHz and 19.5 MHz,-respectively. In addition to the minimum at 2.05 GHz, there-are minima

at other frequencies, which correspond to higher-order modes on the microstrip line. Intermodulation (IM) distortion was also studied with several long JSIM simulations. Over a 39 MHz bandwidth, in-band IM distortion is better than -69 dBFS. Other features of the circuit include unconditional stability and a full-scale input sensitivity of 20 mV (rms).

We designed and laid out a bandpass modulator test chip, the functional block diagram of which is drawn in Figure 10. A 1:4 demultiplexer converts the 20 GHz 1-bit code of the bandpass modulator to a 4-bit parallel word at 5 GHz. This allows most of the test chip, including the programmable counter and the shift register memory banks, to run at an internal clock rate of 5 GHz instead of 20 GHz, where the timing margins for the digital circuitry would be much smaller. Because of the 1:4 demultiplexing, 128-bit memory banks A and B are organized as 4 parallel rows of 32-bit long shift registers. As just discussed, the number of clock cycles skipped between loading the A and B memory banks is set by a programmable counter, which is programmed by external control currents. Once the

shift registers have been loaded, a readout controller unloads the stored bits and transfers them to "high-voltage" drivers, which amplify the output signals up to about 2 mV, which is large enough to be detected by room-temperature electronics. The test chip employs over 4000 Josephson junctions and represents one of the most complex circuits ever designed in this technology.

Sixteen copies of the test chip were fabricated at HY-PRES, Inc. The chips were first tested at low frequency (2 kHz clock) in order to evaluate functionality. While yield was low, we have found one chip which passes all low-frequency testing. We have recently completed optimizing the parameters of the 34 computer-controlled current sources which provide power and control signals to the chip. We are now making final preparations for a high-speed test of the bandpass modulator with our optoelectronic clocking system. We expect to begin this high-speed test within a couple of weeks.

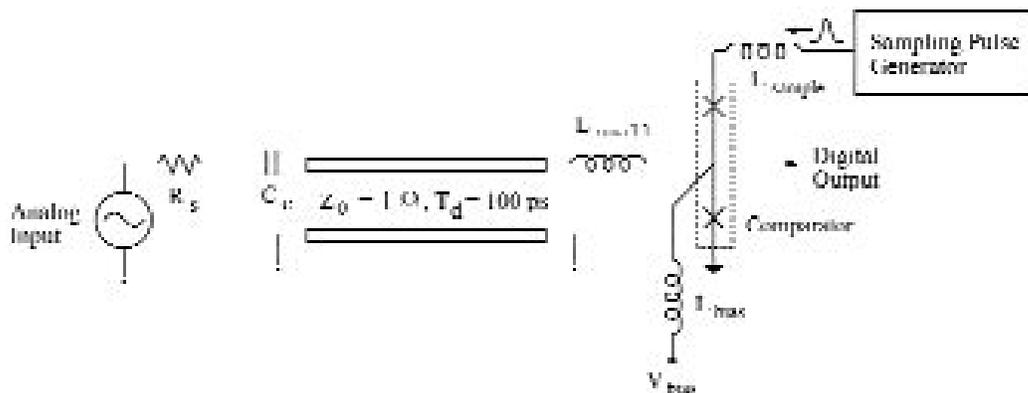


Fig. 8: Superconducting bandpass delta-sigma converter.

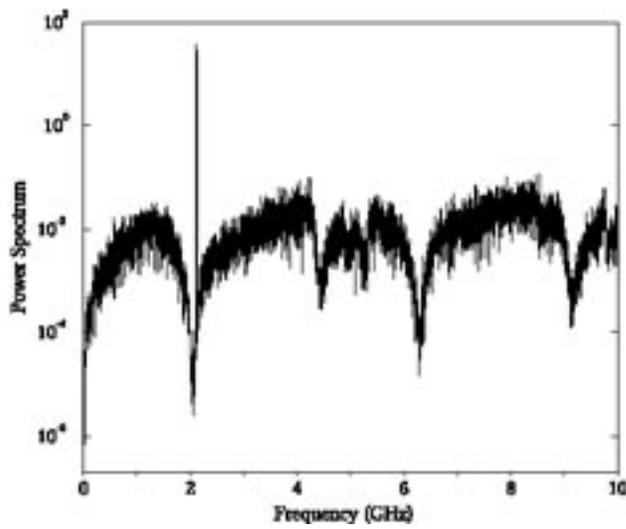


Fig. 9: Simulated output spectrum of proposed delta-sigma converter with large sinusoidal input.

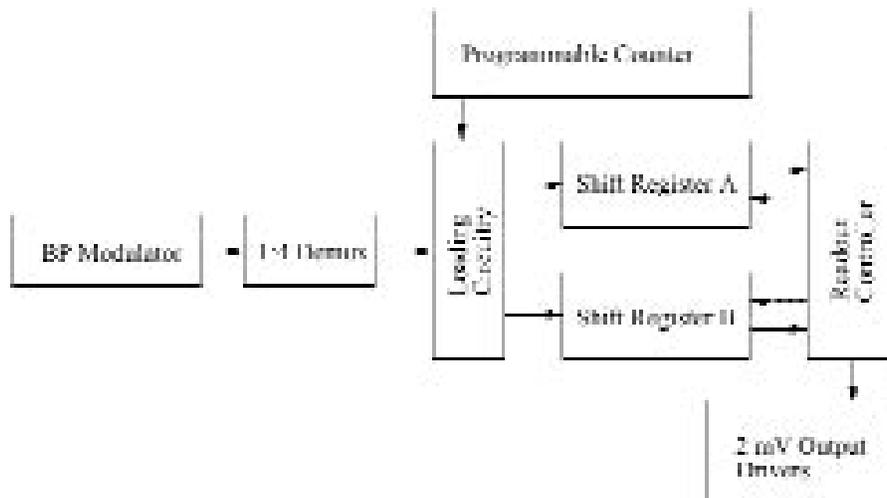


Fig. 10: Block diagram of superconducting bandpass modulator test chip.

Low Power Reconfigurable Analog-to-Digital Converter

Personnel

K. Gulati (H-S. Lee)

Sponsorship

Center for Integrated Circuits & Systems (CICS)

There are applications which require Analog to Digital Converters (ADC) that can digitize signals at a wide range of bandwidth at varying resolution with adaptive power consumption. Clearly, a conventional ADC with fixed topology and parameters cannot accomplish this task efficiently. An alternate approach is to employ an array of ADCs, each customized to work at narrow ranges of resolution and input bandwidth – such a system would occupy a prohibitively large area to achieve optimal power consumption at fine granularity over bandwidth and resolution. A single ADC with reconfigurable parameters and reconfigurable topology would be able to achieve the above goal. Prior reconfigurable ADCs, however, achieve very limited reconfigurability. The proposed ADC is designed to provide a significantly larger reconfigurability space. Its target resolution ranges from 6 to 16 bits and signal bandwidth from 0 to 10MHz.

The concept of this ADC stems from the observation that certain ADC architectures such as the pipeline, cyclic and sigma-delta ADC topologies are composed of the same basic components such as op amps, comparators, switches and capacitors. The sole difference between them, from a network perspective, is the interconnection between these devices. Thus, a converter composed of these basic building blocks in conjunction with a configurable switch matrix, can be made to construct these different topologies and work at different resolutions and bandwidths.

The reconfigurable ADC consists of several basic building blocks as shown in Figure 11. A user defined 'resolution word' that determines the resolution of the ADC is supplied to the main reconfiguration logic that then determines the global structure of the ADC and the state of each block. The PLL shown in the figure uses the frequency information in the clock and the resolu-

tion information from the main reconfiguration logic to determine the appropriate bias current of the opamps.

The ADC was fabricated in a TSMC 0.6mm DPTM CMOS process and occupies a total die area of 10.5x7.6 mm² (Figure 12). The reconfigurable ADC intrinsically requires an area only slightly larger than a 12-bit ADC, however, the prototype layout is optimized not for area but for testability. The resolution of the ADC can be varied from 6 - 15 bits while bias current can be varied over a range of about 3 orders of magnitude corresponding to a sampling rate range of 20KHz-20MHz. Table 1 provides a summary of representative measured results.

Table 1
Measured results at two performance points.

Process	0.6µm CMOS, DPTM
Die Area	10.5mm x 7.6mm
Power Supply	2.7V-4.0V
Parameter Reconfiguring Time	12 clock cycles
Sigma-Delta 15 bit Mode (3.3V)	
Resolution	15 bits
Fclock	10MHz
Fin	3.13KHz (1.5V p-p differential)
OSR	1024
Power	8.8mW
HD2	111.8dB
HD3	96.21dB
Pipeline 12 bit Mode (3.3V)	
Resolution	11 bits
Fclock	2.62MHz
Fin	1MHz (1V p-p differential)
Power	24.6mW
DNL	< +/- 0.5LSB
INL	< +/- 0.82LSB

Continued

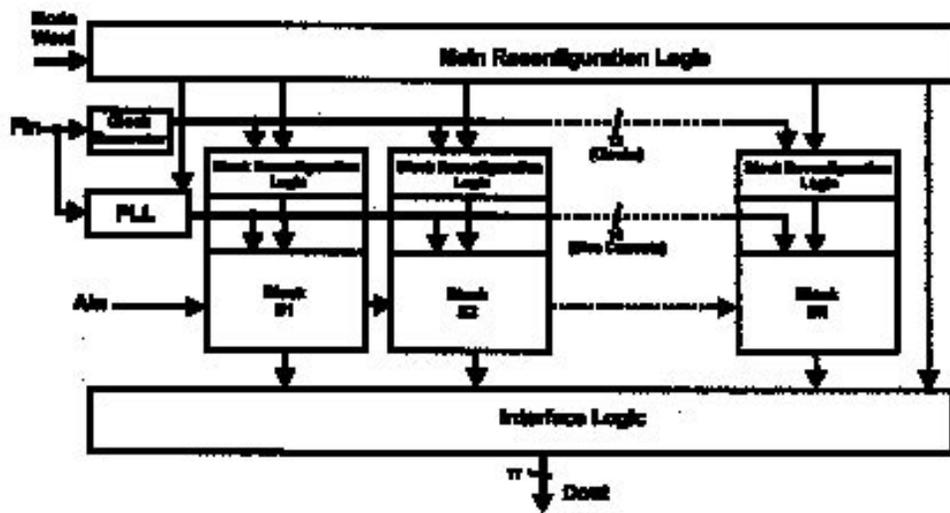


Fig. 11: ADC

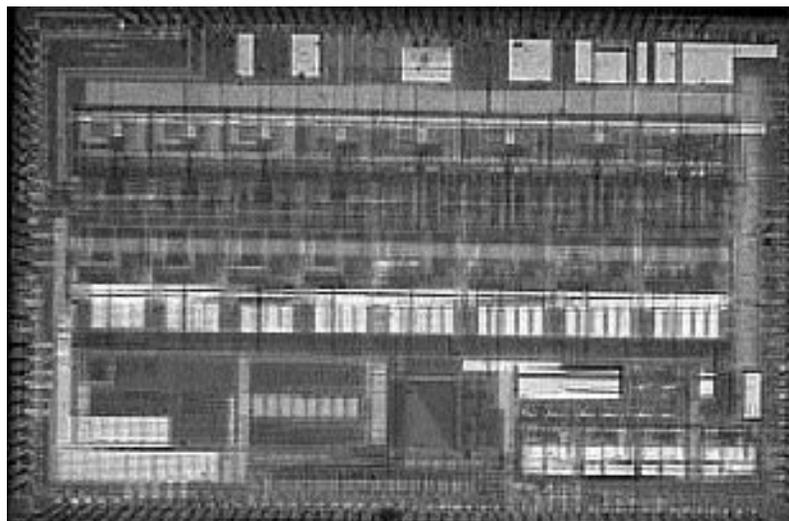


Fig. 12: ADC Microphotograph.

Substrate Noise Shaping in Mixed-Signal Systems

Personnel

M. Shane Peng (H-S. Lee) -

Sponsorship

Texas Instruments

The three main metrics of chip design — power, speed, and performance— all drive towards complete system integration on a single-substrate or what is called System on a Chip (SoC). This necessitates the integration of analog circuits with digital circuits. However, in this integration, the acute problem of substrate noise coupling arises. The noisy digital circuits tend to inject noise into the shared substrate which severely affects sensitive analog circuits. Uncontained, this noise will degrade performance severely, and in some cases destroy functionality.

Up to now, most efforts in addressing this problem have been to ensure that analog circuits are robust enough to withstand the digital noise. These techniques include physical separation, differential architectures, and simulation. Hardly any effort has been placed on reducing the magnitude of the noise. With this in mind, the focus of this research is to investigate a new way to cancel the noise. More specifically, this research proposes to use a feedback loop to shape the noise so that noise is reduced significantly in certain bands of interest. This is well suited for oversampling or bandpass type applications.

This type of system allows quick integration of analog circuits and digital circuits as they can be designed independently. Furthermore, the reduced noise comes at the expense of minimal area and minimal power which is highly desirable.

In order to demonstrate the concept of substrate noise shaping, a 16 bit delta-sigma A/D converter is integrated with a complex digital encryption circuit on the same substrate. The substrate noise shaping loop is based on a delta-sigma loop with the feedback D/A converter replaced by an array of noise-injecting inverters of varying strengths. The substrate noise shaping loop will remove noise out of the band of the 16 bit A/D converter. The prototype chip containing is currently being laid out.

Analog Circuit Design with Scaled CMOS Devices

Personnel

J. K. Fiorenza and T. Sepke (H. -S. Lee and C. G. Sodini)

Sponsorship

Texas Instruments

As CMOS technology scales to sub-100nm features, new circuit techniques will be required to continue to achieve increasing performance. The control of transistor parameters is becoming extremely difficult leading to a wide variation of these parameters. In addition, the lowering of threshold voltages to accommodate the reduction of supply voltages results in devices with increased sub-threshold conduction. The continued scaling of gate oxide thickness has resulted in significant gate current, which threatens well-established circuit techniques such as dynamic logic and switched capacitor circuits. Analog circuit performance suffers due to restricted signal range coupled with increasing device noise. Small geometry transistors also exhibit far less voltage gain and greater threshold voltage mismatches.

A test chip is being constructed using a 0.13 micron CMOS process to characterize device parameters such as output resistance, transconductance, threshold voltage matching, subthreshold leakage, gate leakage and noise in order to determine their effect on analog circuits. The chip contains several carefully selected devices and circuits to allow the measurement of device parameters.

Differential amplifiers are used to measure the transistor matching properties, and a voltage reference circuit is implemented using a parasitic, bipolar transistor. A two-stage, operational amplifier is designed ignoring any non-ideal device behavior. It is not expected that the design will meet the performance goals, but it will provide insight into the relevant design issues. The information gathered from this test chip will help us make generalizations that provide an understanding of the analog behavior of devices and circuits in scaled CMOS technology.

Intelligent Transportation Systems

Personnel

B. K. P. Horn, H.-S. Lee, I. Masaki, T. B. Sheridan, C. G. Sodini, J. M. Sussman, and J. L. Wyatt

Sponsorship

Intelligent Transportation Research Center, MTL

US citizens are spending, on average, about \$1,000 per year for cars, trucks, and roads. The transportation is an important infrastructure for our society. The goal of this project is to develop a technical foundation for tomorrow's transportation systems. Currently we have a number of infrastructures which are independent from each other. Examples include infrastructures for transportation, communication, finance, health care, emergency care, and others. In the next generation, these independent infrastructures will be integrated more closely with advanced information technologies. For example, highway tolls can be charged to a driver, bank account automatically with electronic toll gates connected to bank computers. If a car accident occurs, as another example, the accident can be detected by an air-bag sensor and reported automatically through wireless network to ambulance stations. The ambulance and hospital will have teleconference on the way from the scene to the hospital for a quick care.

With this project, we are working on various research topics ranging from small-scale systems to large-scale systems as well as fundamental to application oriented subprojects. Examples of small-scale subprojects are an adaptive dynamic range image acquisition chip, and a time-to-collision chip. Medium-scale systems include a personal-computer-based real-time three-dimensional machine vision system, a fusion system of machine vision and radar sensors, and an image recognition system for compressed three-dimensional images without decompression. Examples of large-scale systems are a network for real-time image transfer, train control architecture, policies for inter modal systems which consists of cars, trucks, trains, airplanes, and other transportation means.

The research is being carried out at the Intelligent Transportation Research Center in MIT Microsystems Technology Laboratories. The Center is being sponsored by several member companies.

Recognition of Three-dimensional Compressed Images without Decompression

Personnel

N. S. Love (I. Masaki and B. K. P. Horn)

Sponsorship

Intelligent Transportation Research Center, MTL

Conventionally, image recognition and image compression were two different research areas which are independent from each other, and compressed images needed to be decompressed before recognition. With this project, we have developed a new image compression method so that the decompression process is not needed for recognition.

As another feature of this project, we are working on three-dimensional images, but not conventional two-dimensional images, which include distance information between each object and the camera system. We have developed a compact three-dimensional image data acquisition system which requires only a personal computer, two plug-in boards, and three TV cameras for real-time operation. The system focuses its computational power on only relevant regions in the acquired images to achieve its compactness without sacrificing its processing speed. For image compression, we proposed an edge-based method while the main stream of the conventional image compression methods use a spatial frequency scheme. The edges are points where the image intensity and/or color change significantly. The images are segmented into small regions by using edge information. The compression system calculates the attributes of each region including distance, color, and intensity.

Let's assume, as a recognition example, that we would like to find a red Ford Taurus from a large amount of three-dimensional image database. Our image database consists of the following four domains: edge, distance, color, and intensity. The recognition system first uses the color domain to find images which include red color regions. From the corresponding edge and distance domain information, the size of each red region is compared with the expected size. The images with high possibilities of including a red Ford Taurus are decompressed for detailed recognition.

We are now applying the system to a traf-

Continued

fic monitoring video network. The traffic monitoring network consists of a large number of wayside video cameras connected to hierarchical control centers. The centers monitor the numbers of vehicles per minute and average vehicle speeds, detect vehicle accidents, provide traffic condition information, and control wayside variable message boards.

Priority Control for Networks

Personnel

I. Mizunuma (I. Masaki)

Sponsorship

Intelligent Transportation Research Center, MTL

The goal of this project is to develop a new network architecture which is close to the conventional "guaranteed" network in reliability and also close to the conventional "best effort" network in efficiency.

As the first step, we have developed a network which transfer video images in near real-time by using NTCIP (National Transportation Communications for Intelligent Transportation Systems Protocol). The NTCIP is compatible with the protocol for conventional internet. A number of TV cameras are connected to local control centers, and the local centers are connected to broad-region centers through the NTCIP-based network. The broad-region centers send commands to control TV camera parameters such as camera's viewing direction, zoom/pan, image resolution, and others. Each local center consists of SNMP (Simple Network Management Protocol) agent, real-time middle-ware, MIB (Management Information Base), a camera controller, and an image compression board.

As the second step, we are now developing a priority control scheme for networks. Our scheme is based on automatic auctions at each network nodes. In case that the communication demand is larger than the communication capacity, the software agent associated with each message joins the automatic auctions. The bidding prices are decided by the software agents based on the natures of the messages. With this scheme, we can guarantee the delivery times for the urgent messages while offering low costs for the messages which are not sensitive to the delivery times. The first prototype has been developed for transferring video data, over the network, which have different time-sensitivities.

Fusion of Machine Vision and Radar Sensors

Personnel

Y. Fang and T. Kato (I. Masaki)

Sponsorship

Intelligent Transportation Research Center, MTL

Sensor fusion technology increases sensing capability by integrate multiple sensors. Machine vision and radar sensors have their own merits and demerits, respectively, and therefore it is expected that high performance is available by fusing them. We have developed two types of fusion systems: one with three-dimensional machine vision and radar sensors and the other with two-dimensional machine vision and radar sensors.

The three-dimensional vision system consists of two TV cameras and calculates the distance between each object and the camera system from the disparity between the right and left camera images. A time-consuming and difficult part of the distance measurement is to evaluate all possible correspondences between pixels in the right and left camera images. On the other hand, the radar sensor does not have such a correspondence problem but its lateral spatial resolution is much less compared to the machine vision sensor. We chose a collision warning system for an automobile as an application example. In case there are many cars in front at the different distances, both machine vision and radar sensors are confused but we could obtain highly reliable data by fusing both sensors.

The integration of two-dimensional machine vision and radar sensors provides higher reliability and efficiency compared to machine-vision-only or radar-only architecture. For example, if there are two cars at the same distance in front of your car, the radar system is used to detect that there are two objects at the same measured distance and the machine vision system is used to calculate whether the lateral distance between those two cars is wide enough for your car to drive through.

We are now extending this project to the fusion between binocular stereo and motion vision.

A Programmable, Wide Dynamic Range CMOS Imager with On-Chip Automatic Exposure Control

Personnel

P. M. Acosta Serafini (C. G. Sodini)

Sponsorship

Center for Integrated Circuits and Systems and National Semiconductor

Machine vision applications that use visual information typically need an image sensor able to capture natural scenes that may have an intensity dynamic range as high as four orders of magnitude. Reported wide dynamic range imager sensors may suffer from some or all of these problems: large silicon area, high cost, low spatial resolution, small dynamic range increase factor, poor pixel sensitivity, small intensity resolution, etc. The primary focus of this research is to develop a single-chip imager for machine vision applications which addresses these problems, but is still able to provide a wide dynamic range by implementing a novel pixel-by-pixel automatic exposure control. The secondary focus of the research is to make the imager programmable, so that its performance (light intensity dynamic range, spatial resolution, light intensity resolution, frame rate, etc.) can be tailored to suit a particular machine vision application.

The sensing array has pixels that can be independently read and reset. The proposed brightness adaptive algorithm then predictively scales the voltage in photodiodes that would saturate under normal circumstances based on information gathered in several readout checks. The total integration time is subdivided into several integration times (called integration slots) which are progressively shorter. If it is determined that the pixel will saturate at the end of the current integration slot, then the pixel is reset and it is allowed to once more integrate light, but for a shorter period of time. Each pixel has a small associated memory location needed to store an exponent that identifies the actual integration slot used. This information is used to appropriately scale the digitized pixel output.

A Differential Passive Pixel Imager

Personnel

I. L. Fujimori (C. G. Sodini)

Sponsorship

Lucent GRPW Fellow, Intelligent Transportation Research Center, MTL

Passive pixel sensors provide an alternative to the conventional active pixel sensor (APS) for high-density CMOS imaging arrays. Similar to the history of the single-transistor DRAM cell, this one-transistor pixel cell boasts one main advantage over the APS. It can achieve a high fill-factor in a smaller area, leading to a high-density array of pixels with high quantum efficiency. Experiments reveal a major weakness in passive pixels is a signal-dependent parasitic current that can contaminate charge signals in different parts of the array. In this project, we explain the origin of this parasitic current and demonstrate a correlated double sampling (CDS) circuit in a differential architecture that removes its effects.

The passive pixel consists of a high-efficiency n-well photodiode and one transistor for reset and row select. The charge difference between the output of a sensing pixel and a dummy cell kept in the dark is converted to a voltage with a sense amplifier at the bottom of every column. The differential architecture is advantageous in rejecting any common-mode signals such as ground bounce.

Passive pixels are plagued with a signal-dependent parasitic current caused by photogenerated electrons collected by the reverse-biased junction of the column line at each pixel. The combined effect of the charge leakage from 256 cells on the column line can be significant and will appear as a parasitic current at each column line. This parasitic current, which is also present in active pixel arrays, is catastrophic in passive pixels because charge-to-voltage conversion does not occur within the pixel. The parasitic charge of a bright pixel can thus contaminate the output of a dark pixel in other rows in the column line and cause smear in the image.

Two strategies were used to remove the effects of the parasitic current. The first was at the architectural level where a differential readout between a sensing and a dummy column rejects the parasitic current that is common between the two columns. The second part consisted of removing the difference in parasitic currents between adjacent columns. The latter was achieved with a CDS circuit that senses the signal with the parasitic current during the first sample phase and then senses the parasitic current only during the second sample phase. The difference between the output of these two sample phases then purely corresponds to the pixel signal and is no longer dependent on the parasitic current.

The improvements achieved with the differential CDS circuit were quantified in terms of column-to-column fixed-pattern noise (FPN). As expected, the dark FPN values are similar at 0.4% with and without CDS. The difference becomes more pronounced as the light intensity increases and the parasitic current mismatches result in a much higher FPN in the absence of the differential CDS circuit.

While this improved passive pixel imager addresses many of the problems that have plagued passive pixels in the past, it does not easily scale with increasing array sizes. The main limiting factor is the readout noise, which is proportional to the column line capacitance and inversely proportional to the pixel capacitance. The combined effect of these factors severely limits the intensity resolution that can be achieved for high-density arrays. The latest research efforts indicate that the noise may be suppressed by introducing a high load capacitance on the column amplifier. While requiring higher currents to achieve the original bandwidth, this method may allow passive pixels to remain in the imaging race for high-density arrays. CMOS image sensors have lower power consumption

Characterization Methodology of CMOS Processes for Image Sensor Applications

Personnel

C-C. Wang (C. G. Sodini)

Sponsorship

Center for Integrated Circuits and Systems and TSMC

and better circuit integration ability compared to CCD image sensors. However, standard CMOS processes are optimized for circuit applications rather than image sensing. Modifications from standard CMOS processes are often required for better image sensor performance. In order to select the most efficient photodiode structure and diagnose the effects of the modified process parameters, a two-stage characterization methodology is developed.

In the first stage, large photodiodes ($\sim 500\text{mm} \times 500\text{mm}$) with different junction structures, such as NW/Psub and N⁺/PW photodiodes, are implemented. This allows one to directly measure the fundamental junction properties of the diodes for image sensing. Two major parameters, quantum efficiency and leakage current are measured and compared. The large area of the diodes assures the measurement accuracy. Furthermore, the layouts of the diodes can be a bulk, strip or grid shape to study the area, edge, and corner components of the junction properties. In this first stage, the best junction type NW/pPsub is identified for further investigation at the pixel level.

In the second stage, a small test pixel array (e.g. a 64 x 64 array) is implemented to study the pixel perfor-

mance. Since this array is a miniature version of an imager, it contains all the characteristics of a large format imager. Its small size allows one to arrange several arrays with different designs on the same chip to remove wafer-to-wafer variation of the chips. Typical imager characteristics that can be measured with the test arrays include sensitivity, dark current, fixed pattern noise, random noise, and dynamic range. Process parameters, such as thermal treatment and sensor implant conditions, can also be fine-tuned with the test array.

This two-stage approach provides a methodology to select the best junction structure and process parameters for a CMOS image sensor process. This approach has been implemented in a 0.25 μm CMOS process and is currently under investigation.

Three-Dimensional Integration: Analysis, Modeling and Technology Development

Personnel

A. Rahman, A. Fan, S. Das, K-N. Chen, R. Tadepalli (R. Reif)

Sponsorship

Interconnect Focus Center (MARCO-DARPA)

Three-dimensional (3-D) IC, devices are allowed to exist on more than one device layer, and they can be contacted from both top and bottom device layers. Flexibility to place devices along the third dimension allows higher device density and smaller chip area in 3-D IC. The critical interconnect paths that limit system performance can also be shortened by 3-D integration to achieve faster clock speed. By 3-D integration, active layers fabricated with different front-end processes can be stacked to form systems on a chip. A cross section of a proposed 3-D integrated circuit/system is shown in Figure 13.

System-Level Performance Modeling and Trade-off Analysis

Based on our simulation, 3-D integration results in narrower wire-length distribution, with more local (short) wires and less global (long) wires, than the conventional planar implementation. The average and total wire-lengths in 3-D integration are also shorter than 2-D integration. Wire-length distribution of 3-D IC with 21 million transistors/3.5 million logic gates, consistent with 0.18 mm technology generation, is shown in Figure 14.

Based on our analysis of scaled technologies, we find that the contribution of interconnect delay on local clock frequency in high-performance circuits such as microprocessors is going to be in the range of 30%-50%. Using 3-D IC with two device layers, ~ 15%-25% improvement in local clock frequency can be achieved. However, the contribution of interconnect delay on global/across-chip clock frequency can be in the range of 80%-90% and a much higher improvement in across-chip clock frequency can be achieved by 3-D integration.

Recent work has involved the exploration of circuit architectures suitable for 3-D integration. For example, our work has shown that field-programmable gate arrays (FPGAs) may benefit up to 50% in terms of interconnect delay and power dissipation and as much as 40% in logic density. Architectures that integrate dissimilar technolo-

gies may be enabled or improved by 3-D integration. Toward this end, we have begun research into aspects of the 3-D design flow, such as partitioning, placement, and layout.

Referring to Figure 13, the implementation of 3-D ICs involves vertical stacking of CMOS device layers using Cu-Cu wafer bonding at 400°C. All active layers are electrically interconnected using 2:1 or 3:1 aspect ratio vias. Metal (Cu) bumps on both wafers will serve as electrical contacts between the top wafer and Al interconnects on the bottom wafer. In addition, these metal bumps also function as the wafer bonding medium. Auxiliary Cu pads exclusive from inter-layer communication activities could possibly be used as ground planes or heat conduits for different Si active layers. Prior to bonding, the device wafers are assumed to contain multiple aluminum metal layers and inter-level dielectrics (ILD), thus requiring low-temperature bonding below 450°C to avoid Al degradation. A scanning electron micrograph of bonded wafers is shown in Figure 15.

Microstructures of the Cu-Cu interface can be examined using XTEM, as shown in Figure 16. Successful wafer bonding was achieved using Cu/Ta (300 / 50 nm) layers on Si at 400°C for 30 min and annealed at 400°C in N₂ for 30 min. The Cu film does not require special pre-bonding surface preparation, such as metal CMP or ultraviolet light exposures. The bonded pairs at 400°C exhibited good bonding strength when the razor blade test was applied.

Integrated circuits are currently designed using simple and conservative 'design rules' to ensure that the result-

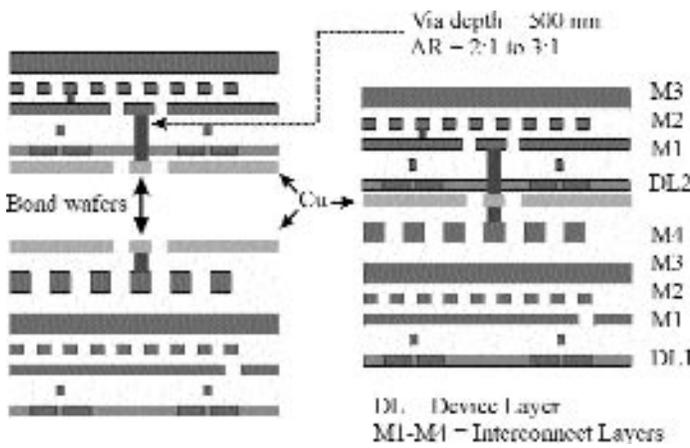


Fig. 13: Cross sectional view of a proposed three-dimensional integrated circuit formed by low-temperature wafer bonding.

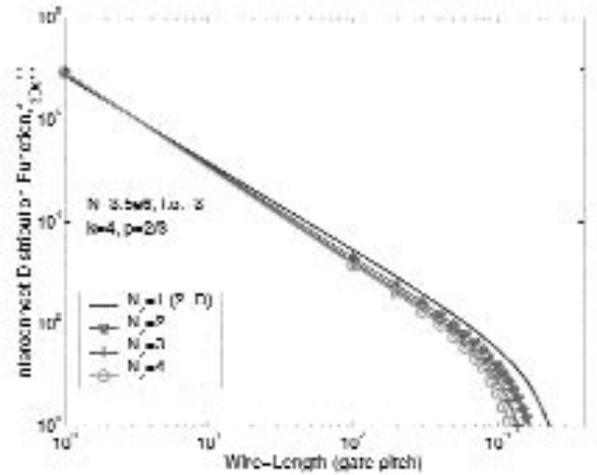


Fig. 14: Wire-length distribution of 2-D and 3-D IC of random logic networks. N_z is the number of device layers, N is the total number of logic gates, $f.o.$ is the average fan-in/out, and k and p are Rent's parameters. The gate pitch is a normalized unit, defined as the average separation between logic gates.

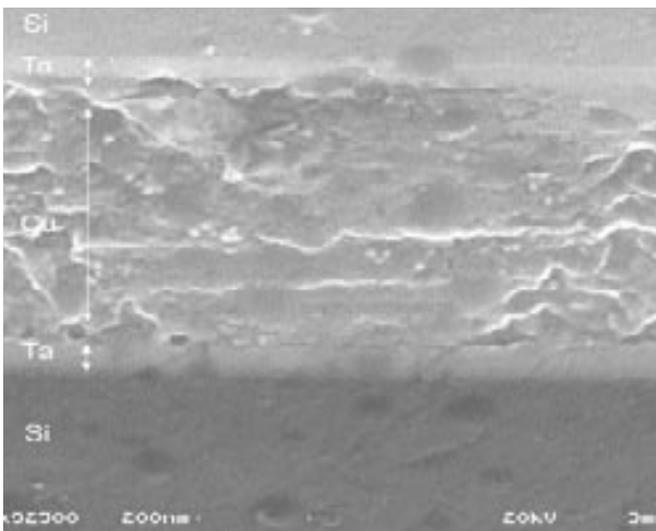


Fig. 15: Cross-sectional SEM of bonded Cu/Ta - Cu/Ta wafers. Cu/Ta = 300 / 50 nm, bonded at 400°C.

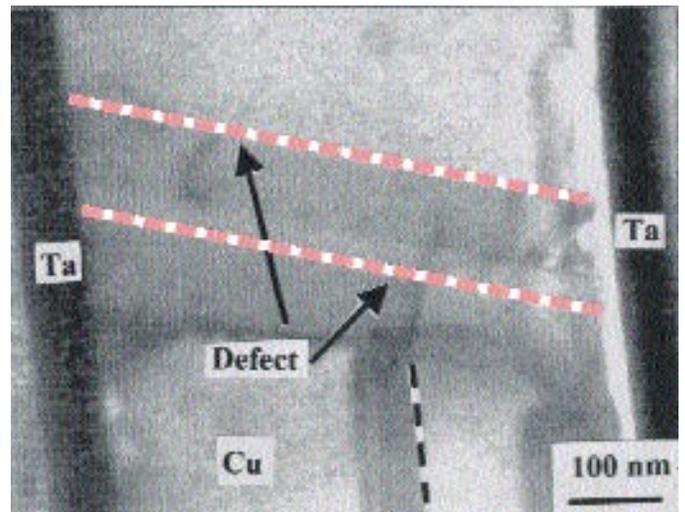


Fig. 16: TEM of bonded Cu/Ta - Cu/Ta wafers, exhibiting twins (dashed lines) perpendicular to the Cu-Cu bonding interface. Cu/Ta = 300 / 50 nm, bonded at 400 °C.

Software Tools for Process-Sensitive Reliability Assessments of IC Designs

Personnel

Y. Chery, S. Hau-Riege, S. Alam, D. Troxel, C.V. Thompson

Sponsorship

MARCO Focused Center on Interconnect (SRC/DARPA)

ing circuits will meet reliability goals. This simplicity and conservatism leads to reduced performance for a given circuit and metallization technology. We are developing a TCAD tool, ERNI, which will allow process-sensitive and layout-specific reliability estimates for fully laid out or partially laid out integrated circuits (see Figure 17).

Circuit-level reliability analyses require the assessment of the reliability of a large number of sometimes complexly connected interconnect trees. An interconnect tree is a continuously connected high conductivity metal, within one layer of metallization, bound by contacts or vias filled with diffusion barriers. We have shown through modeling and experiments that the resistance saturation observed in straight via-to-via lines, which can lead to immunity from electromigration-induced failure, also occurs in more complex interconnect trees. We have also shown that trees will be 'immortal' if their effective current-density line-length product, $(jL)_{eff}$ is below a critical value. This effective

jL product is defined as the maximum value of the sums of the jL products in individual lines taken over all the possible paths through a tree. The jL product that defines immortality can be determined from experimental characterization or simulation of the reliability of straight via-to-via lines.

Simple tests for tree immortality can be used in a hierarchical way to eliminate trees from further more computationally intensive reliability assessments. We have carried out a first-level analysis on microprocessor layouts available on the web, and found that at service conditions the majority of interconnect trees are immortal, even when the worst-case assumption is made that all the limbs of all the trees are at the maximum current density. Filtering of immortal trees significantly reduces the computations required for circuit-level reliability assessments.

After filtering of immortal trees, the reliability of mortal trees must be assessed. This can be done through simulations of the reliability of individual trees, but this

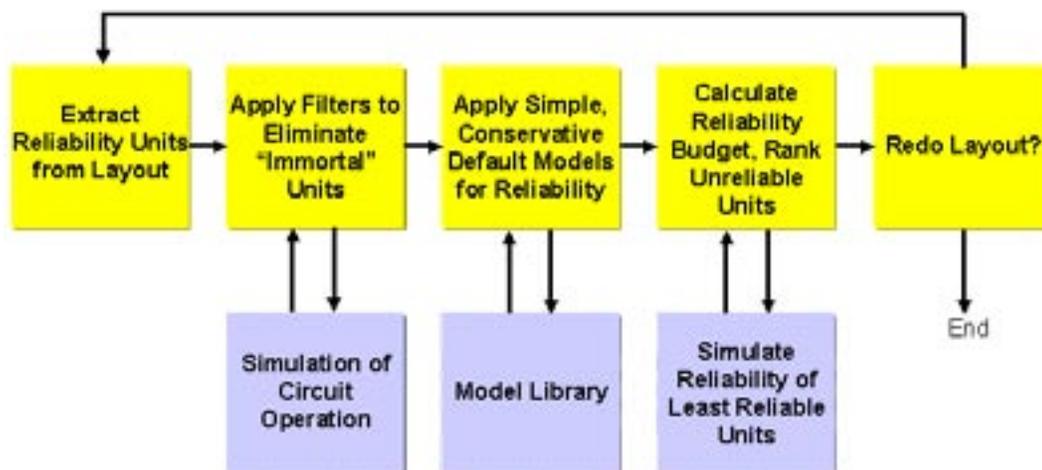


Fig. 17: A flow chart for a full hierarchical circuit-level reliability assessment, the basis for the prototype tool ERNI.

Continued

computationally intensive method should be reserved for the most problematic trees, those with the least reliability, and which are least convenient to 'fix' through layout modifications. We have suggested computationally simple and conservative 'default' models for assessment of tree reliabilities based on the Korhonen analysis and have tested models and simulations through experiments on simple interconnect trees. Our experimental results are consistent with both our analytic models and simulations. With the default models, a first prototype of ERNI has been developed.

Recent developments in semiconductor processing technology has enabled the fabrication of a single integrated circuit with multiple device-interconnect layers or wafers stacked on each other. This approach is commonly referred as 3D integration of ICs. Although there has been significant research on the impact of 3D integration on chip size, interconnect delay, and overall system performance, the reliability issues in the 3D interconnect arrays are largely unknown. We have extended ERNI to develop a novel Reliability Computer Aided Design (RCAD) tool ERNI-3D for reliability analysis of interconnects in a 3D IC. Using this tool, circuit designers can get interactive feedback on the reliability of their circuits associated with electromigration, 3D bonding, and joule heating. Based on a joint probability distribution, a full-chip reliability model combines all reliability figures from different components to give a useful number for the designers' reference.

The initial version of ERNI-3D treats 3D circuits with two wafers or device-interconnect layers in the stack. However, the data-structures and algorithms in the tool are generic enough to make it compatible with 3D circuits with more than two device-interconnect layers, and to allow incorporation of more sophisticated reliability models in the future. Since 3D integration technology is not yet widespread, and no CAD tool supports IC layouts for such a technology, a novel layout methodology has been implemented in 3DMagic by

extending MAGIC, a widely used layout editor in academia. Apart from the CAD tool work, this research has also led to development of, and interesting experiments with, some 3D circuits for testing of ERNI-3D. The test circuits investigated are a 3D 8-bit adder and an FPGA.