Ultra-low power circuit techniques
for a new class of sub-mm^3 sensor nodes

Yoonmyung Lee, Gregory Chen, Scott Hanson, Dennis Sylvester, David Blaauw

University of Michigan, Ann Arbor, MI 48109

Abstract — Bell's Law predicts continual reductions in the size of computing systems. We investigate the status of the next paradigm shift that will usher in ubiquitous computing – sub-mm^3 sensor nodes. However, this form factor remains beyond the capabilities of modern integrated circuit design techniques due to battery size. This paper describes new ultra-low power circuit techniques applied to digital processors, memory, power management, and a special focus on standby mode operation, that will bring mm^3 sensor nodes to reality.

I. Introduction

Since the invention of the transistor, continuous technology scaling has led to the integration of growing computational capabilities in an increasingly small volume. This has been leveraged to create both very small and yet highly capable systems, as well as new multi-core and/or networking technologies that push the upper limits of modern computing performance. This, in turn, has produced in a diversification of computing platforms, ranging from portable handheld devices to building-scale data centers. According to Bell's Law, a new class of smaller computers is developed approximately every decade by using fewer components or fractional parts of state-of-the-art chips [1]. Along with the emergence of personal computers in 1980s and portable handheld devices in 1990s and 2000s, wireless sensor networks have been recently developed. Looking ahead, the next a new class of the miniature computing system is poised to be unveiled – sub-mm^3 sensor nodes (Figure 1).

Wireless sensors can vary greatly in application and distribution but universally benefit from longer device lifetimes, smaller size, and reduced cost. Low power operation is vital for sensors used to monitor flow rates in oil pipelines [2] or in heating, ventilation, and air conditioning (HVAC) systems [3] since the sensors are inaccessible and battery replacement often requires disassembling the infrastructure. It is equally important for implanted medical sensors, where battery replacement requires costly, invasive surgery and high power densities can cause tissue heating and damage [4]. Tiny sensor nodes are needed for many applications to collect and communicate environmental data without interfering with the subject under study. For example, a pebble-sized sensor can be attached to a bumblebee to track a colony's territory without impeding the insect's movement [5]. Similarly, tiny sensors can be mixed into concrete to measure a building's structural integrity during an earthquake without compromising the concrete's strength [6]. Lower sensor cost is vital for large wireless sensor networks with many sensors. Cheaper sensors make it more economical to monitor conditions as individual items are transported
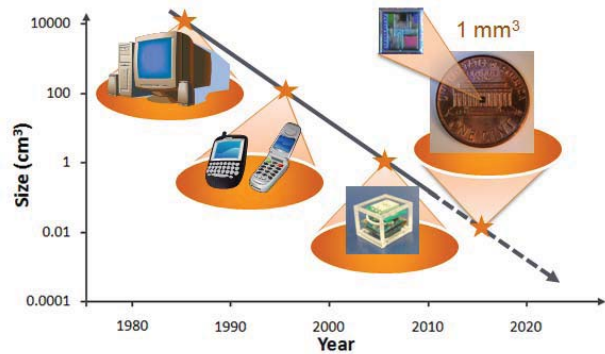

Figure 1. Bell's Law predicts continuous scaling of minimal-sized computing systems

through a supply chain or to use wireless sensors to track inventory in a store.

However, today's wireless sensors are composed of multiple components on a printed circuit board (PCB). Bulky batteries are included in the system to power the circuit components with adequate lifetimes. The result is a milliwatt-powered system that is centimeters or tens of centimeters on a side. New ultra-low power circuit design advances are creating exciting opportunities to dramatically reduce the size and cost of future wireless sensors without affecting device lifetime. Continued scaling of transistor, sensor, and packaging technologies will enable unprecedented integration, decreasing size, interconnect power, and total cost. Robust low-power circuit design will enable the use of smaller, less-expensive power sources while still increasing device lifetime to reduce maintenance costs for battery replacement or recharging. The result will be a sub-mm^3 wireless sensor suitable for a multitude of applications not feasible today.
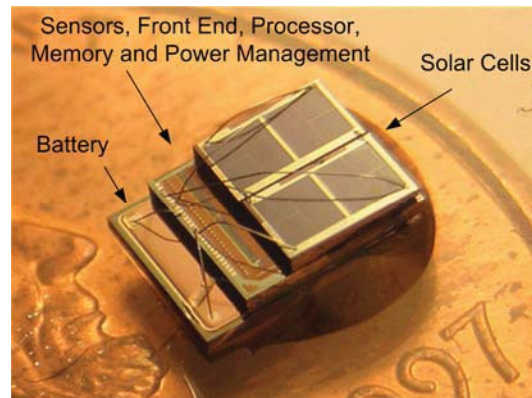

Figure 2. A highly-integrated millimeter scale low-power sensor [7].

Consider a hypothetical intraocular pressure (IOP) monitor as an example to better understand the possibilities created by, and constraints placed on, future sub-mm^3 wireless sensor nodes. Intraocular sensors provide the opportunity for continuous pressure monitoring to detect and track the progression of glaucoma and other diseases. Current eye pressure measurement techniques are invasive and must be infrequently performed at a doctor's office. With an implanted sensor, intraocular pressure can be recorded nearly continuously (i.e., every few minutes) using a capacitive MEMS sensor. This provides doctors with a much more realistic picture of the eye pressure during normal daily activities. It also allows them to customize medication levels for different times of the day to adjust to the circadian rhythms in eye pressure of the patient and allows them to determine patient compliance with the prescribed medication regime.

The data logged by the sensor is stored into memory by an on-sensor microprocessor and is periodically transmitted wirelessly to the doctor's or patient's personal computer. Additionally, the microprocessor can perform signal processing on the pressure data to check for abnormally high or low pressures, as well as sharp changes in pressure. In case an abnormally high ocular pressure is detected, the sensor transmits a warning signal to a personal computer, which relays the message to the patient and physician so that the proper medical actions are taken with lower response time. The intraocular sensor is powered by a thin-film lithium battery which could be stacked with the integrated circuits used in the system as well as energy harvesting elements or antenna, as recently demonstrated and shown in Figure 2[7].

Since the sensor is implanted in the eye, its volume is heavily constrained and a cubic-millimeter or smaller sensor node is required. This limits the area of its thin-film battery to 1mm2. Even if the patient fully recharges this battery daily and the battery has an energy density of 1.5μAh/mm2, the power budget of the sensor is 240nW. Unfortunately, even simple circuit blocks such as a band-gap reference generator or an oscillator far exceed this power budget. Moreover, the sensor must perform a multitude of functions, including collecting, processing, storing and transmitting intraocular pressure data, all using its limited energy supply. To meet the stringent power requirements of miniature wireless sensors, significant strides have been made in each wireless sensor component to achieve robust ultra low-power operation. In this paper, we present recent research findings of a new class of ultra-low-power circuits and techniques aimed at providing these unprecedented low power levels. Using these techniques, an intra-ocular pressure sensor with a 1mm^3 volume and lifetimes of months to years can be possible.

Due to space constraints this paper concentrates on digital processor, memory, power management, and standby mode operation. For discussions on wireless communication and sensors, readers can refer to following papers [8][9][10]. In Section II we will examine microprocessors that compress
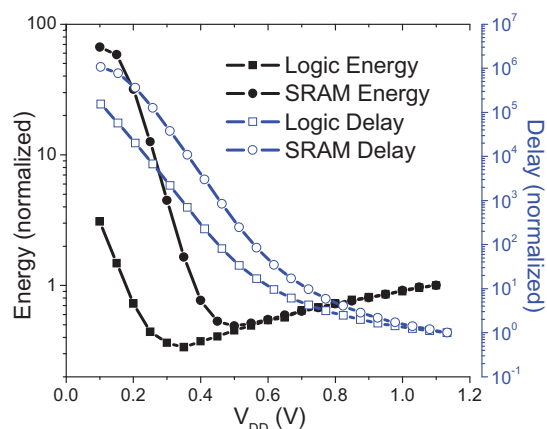

Figure 3. Dynamic energy and leakage power decrease as V_{DD} is scaled down. Latency and leakage energy per cycle increase. The minimum total energy per cycle is achieved at an intermediate voltage (V_{MIN})

and analyze the raw sensor data. Section III discusses low-power memory to store data on the sensor and Section IV describes the power electronics needed to supply low-power sensor nodes and Section V details standby mode operation. We then examine open challenges in this area in Section VI and conclude in Section VII.

II. ULTRA-LOW POWER PROCESSOR DESIGN

Sensors can gather large amounts of data about their surrounding environment. However, only a small portion of this data is typically useful. For example, motion detecting sensors for surveillance only need to report the measured data when any motion is detected. Therefore, most of the measurement data indicating there is no change can be ignored. Moreover, storing and transmitting raw measurement data is power-inefficient as it requires larger memories and more wireless communication using power-hungry radios. To avoid these effects, sensor nodes should perform on-chip digital signal processing to extract and compress useful sensor data with a microprocessor before storing in memory or transmission.

To operate with stringent energy budget, minimizing energy consumption is the critical concern for microprocessors for sub-mm^3 sensor node. By reducing supply voltage (V_{DD}), dramatic energy reduction can be achieved as dynamic switching energy scales quadratically with V_{DD}. Several designs have reduced power consumption through aggressively scaling supply voltage [11][12][13]. However, voltage scaling also increases latency, especially when V_{DD} is scaled below the device threshold voltage (V_{TH}) to the subthreshold region. Although leakage power reduces with V_{DD}, leakage energy per cycle increases because of this increased latency. In typical operation (V_{DD} ~1V today), active energy is orders of magnitude higher than leakage energy, however as shown in Figure 3 the competing trends in dynamic and leakage energy results in a intermediate voltage (V_{MIN}) where total energy per instruction (E_{MIN}) is

minimized [14][15][16].

Aggressive voltage scaling not only enables much higher energy efficiency but also reduces robustness of circuits due to the reduced noise margin and increased susceptibility to process variations. Noise margins degrade with supply voltage because of a reduced on-current to off-current ratio. Also low voltage devices are more sensitive to process variations, further reducing noise margins. V_{TH} variations due to random dopant fluctuations (RDF) dominate other sources of variation at such low voltages. While removing high fan-in gates, series transmission gates, and dynamic logic is sufficient to maintain functional robustness, process variations create up to a 300% delay variation in a subthreshold logic gate, leading to high margins to meet timing yield. Since RDF variation is uncorrelated, its effects on critical path delay can be decreased by increasing the logic depth between pipeline stages [17].

III. ULTRA-LOW POWER MEMORY DESIGN

Memory in sensor nodes is used for temporary storage of measurement data until the measurement log is collected and can also be used as a scratchpad for complex data processing. Hence, larger memory allows for more infrequent data collection and more complex data processing. For a fixed volume, dense memory is desired but operating power must meet stringent limitations. For this reason non-volatile memories such as flash are undesirable since they require charge pumps with very large write power that can quickly dominate total sensor power consumption. This section discusses two types of CMOS-compatible volatile memories: static random access memory (SRAM) and embedded dynamic random access memory (eDRAM).

A. SRAM

Along with the voltage scaling of processors, both dynamic and leakage power of memory can be reduce with voltage scaling (Figure 3). However, low-voltage SRAM is prone to functional failures because process variations lead to destructive read operations and insufficient write margin [18]. In addition, since many sensor systems require large amounts of SRAM in which the vast majority of bitcells must function for chip yield, SRAM bitcell yield must be extremely high for acceptable overall yield. A write operation in the common differential 6-transistor (6T) bitcell (Figure 4a) is performed by raising the wordline (WL) voltage and asserting a differential value on the bitlines (BL). A read operation is performed by precharging bitlines to V_{DD}, letting them float, and then asserting the WL so the bitcell can drive its value on the BLs. Write margin is improved by increasing the strength of the pass gates (A3,A4) relative to the pull-up devices (A1,A2). Read stability is improved by increasing pull-down strength (A5,A6) relative to the pass gates. Designing the bitcell for higher write margin generally decreases read stability and vice versa, creating a fundamental robustness limit. Device sizing and SRAM assist circuits such as dual-VDD wordline circuits can improve low-voltage SRAM
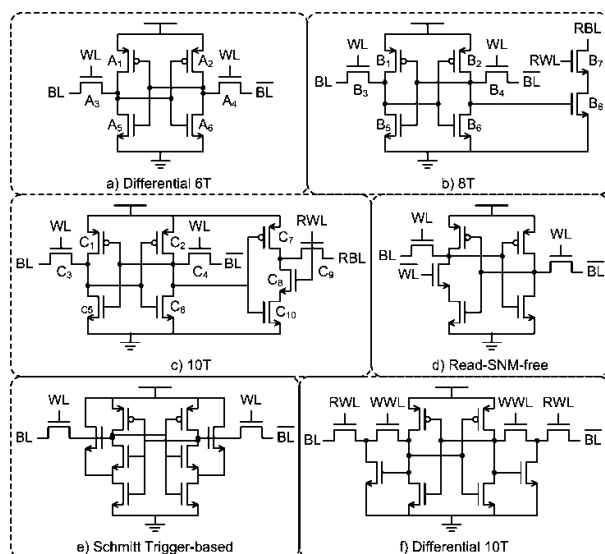

Figure 4. SRAM designs for low-voltage low-power operation

robustness, but are not sufficient to enable robust near-threshold or subthreshold SRAM [19].

The 8-transistor (8T) bitcell (Figure 4b) achieves higher low-voltage robustness at the expense of lower density by using a separate buffer for read accesses (B7,B8) [18][20][21]. This allows the cross-coupled inverters and pass-gates to be sized optimally for writes, while making the bitcell virtually immune to destructive reads. A 4kB commercial 8T sub-array was demonstrated with 295MHz operation at 0.41V [21]. Although the 8T bitcell is tolerant to destructive reads, read operations can still fail if the bitcell is unable to pull the read bitline (RBL) down quickly enough to meet timing constraints. Another read-failure mechanism is undesired RBL discharge because of leakage from unaccessed bitcells. This failure mode is aggravated when V_{DD} is scaled because of reduced on-to-off current ratio. To reduce contention from unaccessed bitcells, the 10T bitcell (Figure 4c) lowers read buffer leakage current by placing two off NFET devices (C_8,C_9) in series between the read bitline and ground and also pulls the RBL up during the read of a one to prevent undesired discharge [22]. Bitline leakage can be mitigated further through compensation with leaking column headers or footers [23].

While 8T and 10T bitcells improve read and write margins, write margins can still be an issue, hold margins are not addressed, and bit interleaving is more cumbersome. The read-SNM-free bitcell (Figure 4d) improves read margin by cutting the feedback loop during accesses [24]. Hold errors occur when SRAM state is lost between accesses and hold margins can be improved by incorporating Schmitt triggers into the cross-coupled inverters (Figure 4e) [25]. Pseudo-write errors occur when unaccessed bitcells on the same wordline as accessed bitcells experience destructive read errors. The 8T and 10T bitcells are not tolerant to pseudo-write errors since unaccessed bitcells experience read-

conditions similar to the differential 6T bitcell. In these bitcells the wordline must be segmented if column multiplexing employed, making bit interleaving difficult. Bit interleaving is desirable to prevent single-event upsets caused by neutrons from corrupting multiple bits of one word. Having at most one error per word allows these failures to be fixed with error correction codes (ECCs). A subthreshold differential 10T SRAM (Figure 4f) that tolerates pseudo-write accesses can be implemented with series pass gates driven with WLs in orthogonal directions [26]. Both pass gates are turned on during write and only one pass gate is turned on and driven through separate devices during read.

B. Embedded DRAM

While 8T and 10T SRAM have higher robustness than differential 6T SRAM at low voltages, they are considerably larger in area because of increased device count. The increased bitcell area can limit memory sizes in sensor nodes that require a small form factor. Embedded DRAM (eDRAM) (Figure 5) is fully compatible with CMOS logic and only requires 2 or 3 devices per bit, substantially increasing memory densities. While traditional DRAMs connect the data storage capacitor to the read bitline (RBL) with a pass gate, gain cells use the transconductance of a gain device (M_{GAIN}) to increase read speeds [28]. Read speed is further improved in boosted gain cells since the voltage on the data storage node is boosted during a read operation. During a read operation, RBL transitions from V_{DD} to V_{SS}. This voltage change couples charge into the storage node (V_{NODE}) through the gate-source capacitor of M_{GAIN}. Boosting the storage node increases the overdrive of the gain device and decreases read latency by 41% [27]. In addition to increased density, eDRAM can achieve higher read and write margins than differential 6T SRAM and has lower leakage power due to reduced device counts.

However, eDRAM stores data as a floating charge that must be periodically refreshed, requiring dynamic power. Therefore, data retention time is a critical metric for determining refresh rate and overall memory power. To increase retention time, the subthreshold leakage through the write device (M_{WR}) should be minimized. Subthreshold current is especially large in cells containing a ZERO during

Figure 5. Embedded DRAM is denser than SRAM [27].

the write of a ONE to an adjacent cell sharing the same write bitline (WBL). A boosted voltage on the write wordline (WWL) can be used to super-cutoff the write device, but this requires a costly charge pump or external power supply. Super-cutoff NMOS devices have negative gate-source voltages (positive for PMOS), resulting in lower-than-off drain currents [29]. Alternately a lower WBL voltage can be used to write a ONE. One convenient voltage to use is the steady-state voltage of V_{NODE} when storing a ONE. Writing a ONE with this voltage keeps unaccessed M_{WR} devices super-cutoff, but does not increase the off-current during the read of a ONE [27].

IV. POWER MANAGEMENT FOR ULP SENSOR NODE

In a sensor node, each circuitry requires different supply voltage for its energy optimal operation. For example, microprocessors and other digital circuits can be run at energy optimal point with subthreshold or near-threshold supply voltage which ranges from 300mV to 600mV. Analog components require higher supply voltages to ensure proper headroom and noise margins which can range from 1.2V to 2.5V. Meanwhile common power sources incorporated in sensor nodes, such as batteries and fuel cells, are limited in their output voltages by their chemistries and their voltages degrade with use. Lithium (3.3-4.2V) and alkaline and zinc-air (1.5V) battery chemistries are common. Thin-film batteries are created by depositing layers of electrolytic materials using semiconductor manufacturing techniques [30]. These batteries can be very small in size with reasonably high energy densities. While many thin-film batteries are planar, the energy density can be further increased by using three-dimensional processing to increase the surface area between the electrolytic materials [31]. Fuel cells draw their energy through chemical reactions with an external fuel source such as ethanol and can be integrated with CMOS processes [32]. Since battery voltages do not usually match the desired circuit supply voltages, DC-to-DC converting power electronics are necessary. Most power electronics are designed for high output power levels and do not efficiently convert the low levels of power as low as sub µA needed by sensor nodes. For effective sensor nodes, power electronics must be specifically designed for low power applications.
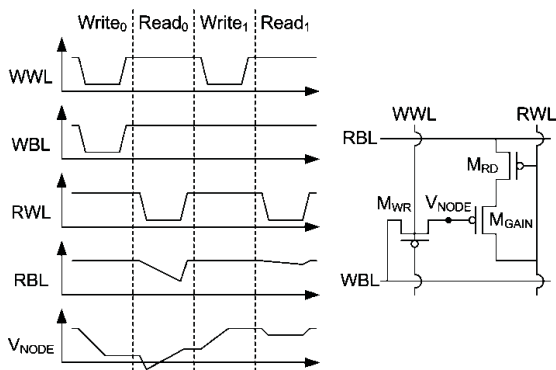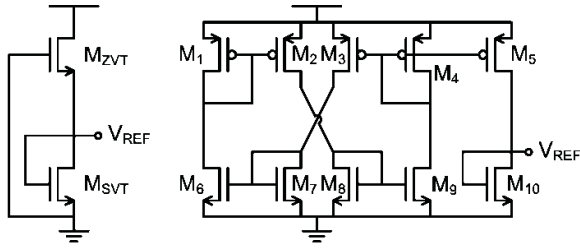
A. Linear Regulation

Linear regulator is a type of regulator which generates output voltage by amplifying reference voltage and the amplifier is powered by higher input supply voltage. Therefore ideal efficiency of a linear regulator is determined by the ratio of the output to input supply voltage. However, there is additional power overhead for quiescent current of amplifier. For high regulating efficiency, it is important to minimize this overhead, but in turn, it will reduce the bandwidth of the amplifier resulting in regulator's degraded transient response to load current fluctuations. This can increase the noise on the output power rails which can affect the robustness of SRAM and increase the latency of the load

a) Constant Reference   b) CTAT Reference

Figure 6. References for voltage outputs for linear regulators [36][37]

Figure 7. Fibonacci switched capacitor network for DC-DC conversion [41].

circuits [33][34]. To minimize regulator power while maintaining proper amplifier bandwidth, bias current can be dynamically increased whenever load power surge is detected [35]. In this regulator design, power surges are detected by dropping the supply voltage across a diode. This diode voltage controls both switches that limit the supply surges and the tail current devices used in the linear regulator.

The output voltage of a linear regulator is determined based on its input voltage, usually supplied from a reference voltage generator. Band gap and constant-g_m references are commonly used in power management systems, but tend to have microamp-level quiescent currents that can singlehandedly exceed the power budgets of low-power sensor nodes. A pico-amp voltage reference (Figure 6a) with low V_{DD} and temperature sensitivity can be generated based on the subtraction of threshold voltages [36]. Although device threshold voltage changes with temperature, the threshold voltage of devices with different V_{TH}'s scale together so the difference is constant with temperature. One method of obtaining this voltage is to connect a diode-connected nominal-V_{TH} NFET device (M_{SVT}) in series with a super-cutoff zero-V_{TH} NFET device (M_{ZVT}). When the devices are properly sized, the output voltage will settle to one half of the difference of the V_{TH}s because of the equal current condition. Since both devices are in the subthreshold region, this 19.4ppm/°C temperature reference can be realized with only 2.2pW power consumption. Other voltage references with low temperature coefficients can be created by combining a complementary to absolute temperature (CTAT) device threshold voltage with a multiple of the proportional to absolute temperature (PTAT) thermal voltage [38]. While these voltage references were realized with low temperature coefficient, a CTAT power supply can be used in subthreshold circuits to keep the frequency of CMOS logic constant with temperature [37]. This CTAT reference voltage (Figure 6b) supplies a constant current through a diode connected device (M_{10}). The temperature independent current is generated by equating the currents of a high V_{TH} subthreshold device (M_6, M_8) with a low V_{TH} saturated devices (M_7, M_9). When temperature increases, the device threshold of M_{10} and the reference voltage decrease. The CTAT power supply balances the effects of temperature on subthreshold logic, which is faster at high temperatures because threshold voltage decreases, resulting in less than 5% frequency variation with temperature.
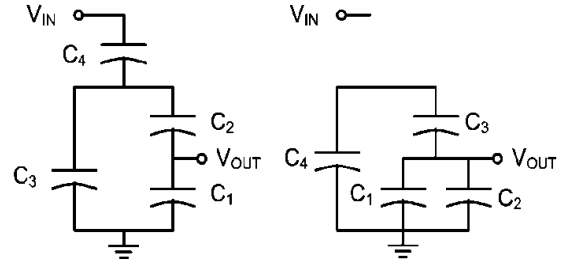
B. Switched Capacitor Networks

Linear regulation provides a stable level converted output voltage, but is limited in its power efficiency to the ratio of the load voltage to the input voltage. For systems where the conversion ratio is high, such as when down-converting a 3.6V Li-ion battery to subthreshold levels, high efficiency is not attainable. In these scenarios, switched capacitor networks (SCN) or buck down-converters can be used to attain higher power efficiencies [13][39]. SCNs connect the voltage input and output to capacitors in different configurations to convert DC voltage levels. The capacitors are connected in different configurations using switches, such as MOS devices, and can either up-convert or down-convert voltages. SCNs commonly alternate between two configurations, converting DC voltages at a fixed ratio. However, more than two configurations can be used to allow the circuit to convert DC values with several different ratios [40].

A hybrid SCN and linear regulator system was reported that can convert from a 3.6V Li-ion battery to subthreshold voltage levels for 5nA to 500nA loads with up to 55% efficiency, representing a 4.6× efficiency improvement over ideal linear regulation [41]. A Fibonacci SCN network, shown in Figure 7, is used to divide the battery voltage by 5. Normally, MOS switches in SCNs are large to reduce conductive losses within the network. However, since this SCN network was designed for very low power loads, nearly minimum sized devices are used for switches. This reduces the power overhead required to switch the gate capacitances of these switches and increases the overall efficiency of the system. Typical SCNs are clocked at MHz levels, but this network uses a 2kHz clock to reduce power overhead. The slow clock is efficiency generated using a slow timer circuit, discussed in Section V. The outputs of SCN networks are noisy unless the switching frequency of the network is very high. Since the power budget precludes such high frequency operation, a linear regulator is used to eliminate voltage ripples on the output supply.

V. STANDBY MODE OPERATION OF ULP SENSOR NODE

Many sensor measurements do not need to be taken continuously since environmental conditions can be periodically sampled. A sensor taking one image per second could adequately monitor automobile traffic, whereas one

measurement per hour would be adequate for monitoring water levels in reservoirs. Even when sensor nodes must forward data from other sensors in the same wireless network, it is likely that a sensor node will be idle for long periods of time. Turning off unnecessary circuitry during these idle periods can dramatically lower the total operating energy of the system. The sensor front end and wireless communication can be power gated, eliminating static currents used in amplifiers and reducing leakage in sensors and ADC circuitry. The microprocessor can also be power gated, although some SRAM and balloon (retentive) latches must remain powered to retain previously logged sensor data and system state. Power gating can be achieved by powering the front end, microcontroller, and wireless communication circuitry from virtual supply rails that are collapsed during standby mode using either high threshold voltage (HVT) header or footer power gates (Figure 8). For the microprocessor, the power gates must be properly sized to achieve both high active mode performance and low standby mode leakage. In analog circuits, the power gates must be turned on long enough before active system operation to allow the circuits to reach a stable DC state. While charging the power rails quickly has the benefit of faster system response time to external wakeup signals, it can cause resonance on the power rails due to the inductance of the supplies. Supply resonance can cause large voltage swings on the supply rails, resulting in the loss of data held in SRAMs and latches because the data-retentive voltage is violated. Supply resonance can be mitigated by gradually stepping the gate voltage of power gating devices [42] or by bypassing resonant current with supply monitors [43].

If the duty cycle of the sensor node is low, then the total system power can be dominated by standby mode power. Sensors have been reported with standby power as low as 30pW [12]. This power is limited by circuits that are not power gated during sleep and switching energy in the wakeup controller, which coordinates the sleep and wakeup procedure. To reduce overall standby leakage power, SRAM leakage and the amount of non-power-gated SRAM should be minimized. SRAM device V_{TH} selection and sizing is vital for leakage reduction. A 10T bitcell optimized for standby operation achieves a standby power of 3.3fW per bit while retaining state [7]. The bitcell shares the same configuration as the 10T bitcell discussed in Section III.A but uses HVT devices and gate-length biasing for the cross-coupled inverters and pass gates. The read buffer is power gated during sleep since it is not needed for data retention. A similar 14T bitcell uses transistor stacking in the cross-coupled inverters to reduce leakage [12].

The sleep and wakeup controller must keep track of the time between sensor measurements. Crystal oscillators are widely used as process, voltage, and temperature insensitive frequency references, but they violate the power budgets of many sensor systems. Low-power timer circuits have been reported. One typical approach for slow timers is to create a
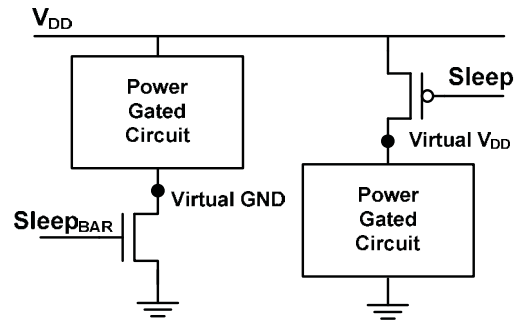

Figure 9. Power gating with NMOS footer and PMOS header [12].

ring oscillator with current starved devices (Figure 9a). These devices use NMOS footers, PMOS headers, or both that are driven with analog bias voltages to reduce device currents and oscillator speed. However, obtaining very slow speeds requires biasing the starving elements deeply into subthreshold mode with the consequences of reduced voltage swing and increased frequency variation due to process and temperature variation. Also, bias voltage generation consumes additional power. Full-swing slow timers with reduced sensitivity to process can be realized using differential monostable multi-vibrational delay elements (Figure 9b) [41]. These delay elements are reset into a high-impedance mode by the previous stage of the ring oscillator. Leakage slowly initiates a positive feedback loop that turns on devices in the delay element and causes the output to switch. However, these delay elements are still highly sensitive to temperature changes. Self-calibrating temperature compensation schemes can be employed for more accurate frequency references [44]. Another topology uses device gate leakage to generate sub-Hz oscillations using less than 1pW [45]. Gate leakage is less sensitive to temperature changes than subthreshold drain current. However, it is highly-sensitive to process variations in oxide thickness and is often poorly modeled.
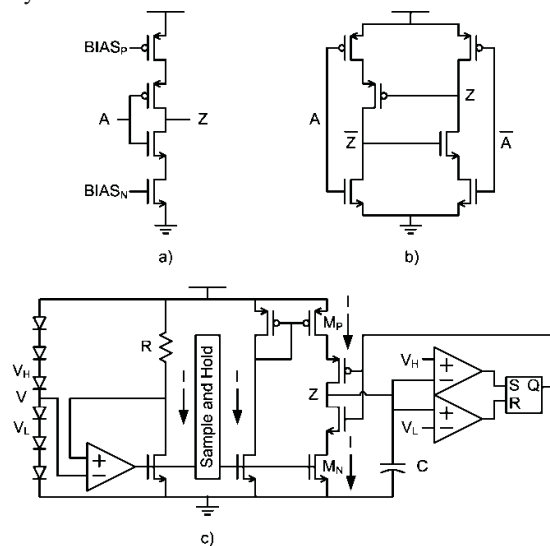

Figure 8. Timers for standby mode control [41] [44]

VI. Open Challenges In ULP Circuit Design

Important advances have been made to achieve millimeter-scale sensor nodes, but many design challenges remain. While these challenges are tackled the IC industry will inevitably evolve, providing new resources and obstacles in the field of millimeter-scale sensors. This section examines few open challenges in ultra-low power circuit design for sensor nodes.

A. CMOS Process Scaling and Optimization

Process technology has a large effect on the minimum energy point and delay variation for digital circuits. It also strongly influences the variability in analog circuit metrics due to process variation. While newer processes are denser, integrated circuits are often not the limiting factor for sensor size. In fact, trading area for lower power can reduce the total size of the system by allowing for smaller batteries. Newer processes have smaller device capacitances and lower dynamic energy, whereas older technologies exhibit lower subthreshold leakage and standby power. Thus sensor systems with high duty cycles and short idle periods achieve lower energy with newer processes. Conversely, sensors with low duty cycles and long idle periods (the more common case) are more energy efficient with older processes [46].

Apart from intelligent selection of process node, CMOS processes can be optimized specifically for subthreshold or near-threshold operation. A key factor in subthreshold CMOS devices is the subthreshold swing (S_S). The S_S measures the decrease in gate voltage needed to reduce drain current by $10\times$ in the subthreshold region. Smaller S_S leads to higher on-to-off current ratios, lower device leakage, and greater functional robustness. The ratio of device length (L_{EFF}) to oxide thickness (t_{OX}) and the device doping profile largely determine S_S and can be optimized for low-voltage operation, at the possible expense of higher-voltage performance. For standard superthreshold CMOS, L_{EFF} scales with each process generation to increase device speed and density. Recently, t_{OX} has not scaled down as rapidly because of gate tunneling leakage limitations. However, for subthreshold devices L_{EFF} has a weaker impact on performance, and S_S can be reduced as the ratio of L_{EFF} to t_{OX} increases. While increasing L_{EFF} improves S_S and leakage power, it increases gate capacitance and dynamic power. An L_{EFF} for a low voltage optimized technology node can be chosen where overall energy is minimized [47].

In conventional superthreshold CMOS processes, non-uniform doping profiles are used, including so-called halo implants at the edges of the channel, to reduce the effect of drain induced barrier lowering (DIBL). DIBL reduces the device V_{TH} for large drain to source voltages (V_{DS}) due to modulation of the source-to-channel energy barrier by the drain depletion region. For subthreshold operation, DIBL is not a large factor since V_{DS} is low. This allows for the removal of halo implants, which improves S_S and overall subthreshold device performance [48].

B. Beyond CMOS

MOSFET parameters can be optimized for lower S_S and better subthreshold performance, but the S_S of these devices is fundamentally limited to 60mV/decade at room temperature. New transistor types are being investigated that do not have this lower bound on S_S. Among several candidates is the heterojunction tunneling transistor (HETT) [49]. N-type HETT devices have a p+ source and n+ drain separated by a lightly doped channel. When the device is off, the energy bands of the source and drain do not overlap, creating a high energy barrier. However, when a potential is applied to the channel, electrons in the valence band of the source can tunnel into the conduction band of the channel, creating current flow. By exploiting band-to-band tunneling and the reduced bandgap of SiGe, both excellent turn-off characteristics (<30mV/decade) and reasonably high on-currents at low voltages can be achieved while maintaining CMOS compatibility. HETT devices exhibit unidirectional drain current because of their asymmetric nature. This does not greatly impact complementary logic, but makes differential 6T SRAM infeasible. A 7T SRAM was proposed using HETT devices that demonstrates higher subthreshold read and write margin due to the superior S_S [50].

VII. Conclusion

To enable the new era of sub-mm^3 sensors, system architecture and all circuit components must be reconsidered for ultra-low power operation. Intra-Occular pressure sensing (IOP) is representative of emerging volume-constrained sensing applications with aggressive power budgets, asymmetric communication requirements, and low duty cycles. Myriad sensor applications have similar use models to the IOP device, collectively pointing to the need for millimeter-scale systems to replace today's bulky, expensive, and power-hungry wireless sensors. These miniaturized sensors offer end users the promise of a wealth of environmental data, and open the door to truly ubiquitous wireless sensing.

References

[1] Bell, G., "Bell's Law for the Birth and Death of Computer Classes", *Communications of the ACM*, January 2008, Vol 51, No. 1, pp 86–94.

[2] N. Mohamed, et al., "A Fault Tolerant Wired/Wireless Sensor Network Architecture for Monitoring Pipeline Infrastructures," *2nd International Conference on Sensor Technologies and Applications*, pp. 179-184, Aug. 2008

[3] Y. Tachwali, et al., "Minimizing HVAC Energy Consumption Using a Wireless Sensor Network," *33rd Annual Conference of the IEEE Industrial Electronics Society*, pp. 439-444, Nov. 2007

[4] D. Malan, et al., "CodeBlue: An ad hoc sensor network infrastructure for emergency medical care," *International Workshop on Wearable and Implantable Body Sensor Networks*, June 2004

[5] B.A. Warneke, et al., "An Autonomous 16 mm3 Solar-Powered Node for Distributed Wireless Sensor Networks," *Proceedings of IEEE Sensors*, pp. 1510-1515, June 2002

[6] N.G. Elvin, et al., "Feasibility of structural monitoring with vibration powered sensors," *Smart Materials and Structures*, vol. 15, pp. 977-986, June 2006

[7] G. Chen, et al., "A Millimeter-Scale Nearly-Perpetual Sensor System with Stacked Battery and Solar Cells," *IEEE International Solid-State Circuits Conference*, pp. 288-289, Feb. 2010.

[8] G. Chen, et al., "Circuit Design Advances for Wireless Sensing Applications," *Proceedings of the IEEE*, to appear.

[9] D.D. Wentzloff, et al., "A 47pJ/pulse 3.1-to-5GHz All-Digital UWB Transmitter in 90nm CMOS," *IEEE International Solid-State Circuits Conference*, pp. 118-119, Feb. 2007.

[10] S. Hanson, et al., "A 0.45–0.7V sub-microwatt CMOS image sensor for ultra-low power applications," *Symposium on VLSI Circuits, pp. 176-177*, June 2009.

[11] S.C Jocke, et al., "A 2.6-μW sub-threshold mixed-signal ECG SoC," *Symposium on VLSI Circuits*, pp. 60-61, June 2009.

[12] S. Hanson, et al., "A low-voltage processor for sensing applications with picowatt standby mode," *IEEE Journal of Solid-State Circuits*, vol. 44, no. 4, pp. 1145-1155, April 2009.

[13] J. Kwong, et al., "A 65 nm sub-Vt microcontroller with integrated SRAM and switched capacitor DC-DC converter," *IEEE Journal of Solid-State Circuits*, vol. 44, no. 1, pp. 115-126, Jan. 2009.

[14] B. Zhai, et al., "Theoretical and practical limits of dynamic voltage scaling," *Proceedings of the 41st Annual Design Automation Conference*, vol. 873, p. 868, June 2004.

[15] A. Wang, et al., "Optimal supply and threshold scaling for subthreshold CMOS circuits," *IEEE Computer Society Annual Symposium on VLSI*, pp. 5-9, April 2002.

[16] B.H. Calhoun, et al., "Characterizing and modeling minimum energy operation for subthreshold circuits," *International Symposium on Low Power Electronics and Design*, pp. 90-95, Aug. 2004.

[17] B. Zhai, et al., "Analysis and mitigation of variability in subthreshold design," *Proceedings of the International Symposium on Low Power Electronics*, pp. 20-25, Aug. 2005.

[18] L. Chang, et al., "An 8T-SRAM for Variability Tolerance and Low-Voltage Operation in High-Performance Caches," *IEEE Journal of Solid-State Circuits*, vol. 43, no. 4, pp. 956-963, April 2008.

[19] G. Chen, et al., "Yield-driven near-threshold SRAM," *IEEE Transactions on Very Large Scale Integration Systems*, in press.

[20] N. Verma, et al., "A 65nm 8T sub-Vt SRAM employing ense-amplifier redundancy," *IEEE International Solid-State Circuits Conference*, pp. 328-329, Feb. 2007.

[21] L. Chang, et al., "A 5.3GHz 8T-SRAM with Operation Down to 0.41V in 65nm CMOS," *Symposium on VLSI Circuits*, pp. 252-253, June 2007.

[22] B.H. Calhoun, et al., "A 256kb Sub-threshold SRAM in 65nm CMOS," *IEEE International Solid-State Circuits Conference*, pp. 259-260, Feb. 2006

[23] T.-H. Kim, et al., "A High-Density Subthreshold SRAM with Data-Independent Bitline Leakage and Virtual Ground Replica Scheme," *IEEE International Solid-State Circuits Conference*, pp. 330-331, Feb. 2007.

[24] K. Takeda, et al., "A read-static-noise-margin-free SRAM cell for low-VDD and high-speed applications," *IEEE Journal of Solid-State Circuits*, vol. 41, no. 1, pp. 113-121, Jan. 2006.

[25] J.P. Kulkarni, et al., ""A 160 mV robust schmitt trigger based subthreshold SRAM," *IEEE Journal of Solid-State Circuits*, vol. 42, no. 10, pp. 2303-2313, Oct. 2007.

[26] I.J. Chang, et al., "A 32kb 10T Subthreshold SRAM Array with Bit-Interleaving and Differential Read Scheme in 90nm CMOS," *IEEE International Solid-State Circuits Conference*, pp. 388-389, Feb. 2008.

[27] K.C. Chun, et al., "A sub-0.9V logic-compatible embedded DRAM with boosted 3T gain cell, regulated bit-line write scheme and PVT-tracking read reference bias," *Symposium on VLSI Circuits*, pp. 134-135, June 2009.

[28] D. Somasekhar, et al., "2 GHz 2 Mb 2T Gain Cell Memory Macro With 128 GBytes/sec Bandwidth in a 65 nm Logic Process Technology," *IEEE Journal of Solid-State Circuits*, vol. 44, no. 1, pp. 174-185, Jan. 2009.

[29] Y. Lee, et al., "Standby power reduction techniques for ultra-low power processors," *IEEE European Solid-State Circuits Conference*, pp. 186-189, Sept. 2008

[30] M. Armand, et al., "Building better batteries," *Nature,* vol. 451, no. 7, pp. 652-657, Feb. 2008.

[31] J.W. Long, et al., "Three-Dimensional Battery Architectures," *Chemical Reviews*, vol. 104, no. 10, pp. 4463-4492, Oct. 2004.

[32] M. Frank, et al., "An integrated power supply system for low-power 3.3V electronics using on-chip polymer electrolyte membrane (PEM) fuel cells," *IEEE International Solid-State Circuits Conference*, pp. 292-293, Feb 2009.

[33] R. Kanj, et al., "SRAM Yield Sensitivity to Supply Voltage Fluctuations and Its Implications on Vmin," *IEEE International Conference on Integrated Circuit Design and Technology*, pp. 1-4, May-June 2007.

[34] M. Agostinelli, et al., "Erratic fluctuations of sram cache vmin at the 90nm process technology node," *IEEE International Electron Devices Meeting*, pp. 655-658, Dec. 2005.

[35] G.K. Balachandran, et al., "A 110 nA Voltage Regulator System With Dynamic Bandwidth Boosting for RFID Systems," *IEEE Journal of Solid-State Circuits*, vol. 41, no. 9, pp. 2019-2028, Sept. 2006.

[36] M. Seok, et al., "A 0.5V 2.2pW 2-transistor voltage reference," *Custom Integrated Circuits Conference*, pp. 577-580, Sept. 2009.

[37] G. De Vita, et al., "A Sub-1-V, 10 ppm/°C, Nanopower Voltage Reference Generator," *IEEE Journal of Solid-State Circuits*, vol. 42, no. 7, pp. 1536-1542, July 2007.

[38] K. Ueno, et al., "A 300 nW, 15 ppm/°C, 20 ppm/V CMOS Voltage Reference Circuit Consisting of Subthreshold MOSFETs," *IEEE Journal of Solid-State Circuits*, vol. 44, no. 7, pp. 2047-2054, July 2009.

[39] G. Patounakis, et al., "A fully integrated on-chip DC-DC conversion and power management system," *IEEE Journal of Solid-State Circuits*, vol. 39, no. 3, pp. 443-451, March 2004.

[40] Y.K. Ramadass, et al., "Voltage Scalable Switched Capacitor DC-DC Converter for Ultra-Low-Power On-Chip Applications," *Power Electronics Specialists Conference*, pp. 2353-2359, June 2007.

[41] M. Wieckowski, et al., "A hybrid DC-DC converter for nanoampere sub-1V implantable applications," *IEEE Symposium on VLSI Circuits*, pp. 166-167, June 2009.

[42] S. Kim, S et al., "Minimizing inductive noise in system-on-a-chip with multiple power gating structures," *European Solid-State Circuits Conference*, pp. 635-638, Sept. 2003.

[43] K.-I. Kawasaki, et al., "A Sub-μs Wake-Up Time Power Gating Technique With Bypass Power Line for Rush Current Support," *IEEE Journal of Solid-State Circuits*, vol. 44, no. 4, pp. 1178-1183, April 2009.

[44] Y.-S. Lin, et al., "A 150pW program-and-hold timer for ultra-low-power sensor platforms," *IEEE International Solid-State Circuits Conference*, pp. 326-327, Feb. 2009.

[45] Y.-S. Lin, et al., "A sub-pW timer using gate leakage for ultra low-power sub-Hz monitoring systems," *Custom Integrated Circuits Conference*, pp. 397-400, Sept. 2007.

[46] M. Seok, et al., "Optimal technology selection for minimizing energy and variability in low voltage applications," *International Symposium on Low Power Electronics and Design*, pp. 9-14, Aug. 2008.

[47] S. Hanson, et al., "Nanometer device scaling in subthreshold circuits," *Proceedings of the 44th annual Design Automation Conference*, pp. 700-705, June 2007.

[48] B.C. Paul, et al., "Device optimization for digital subthreshold logic operation," IEEE *Transactions on Electron Devices*, vol. 52, no. 2, pp. 237-247, Feb. 2005.

[49] O.M. Nayfeh, et al., "Design of Tunneling Field-Effect Transistors Using Strained-Silicon/Strained-Germanium Type-II Staggered Heterojunctions," *IEEE Electron Device Letters*, pp. 1074-1077, Sept. 2008

[50] D. Kim, et al., "Low power circuit design based on heterojunction tunneling transistors (HETTs)," *International Symposium on Low Power Electronics and Design*, pp. 219-224, Aug. 2009.