SoC Issues for RF Smart Dust

Wireless sensor nodes, each a self-powered system performing sensing, communication, and computation, form reliable mesh networks coordinating efforts to add intelligence to the environment.

By Ben W. Cook, Student Member IEEE, Steven Lanzisera, Student Member IEEE, and Kristofer S. J. Pister

ABSTRACT | Wireless sensor nodes are autonomous devices incorporating sensing, power, computation, and communication into one system. Applications for large scale networks of these nodes are presented in the context of their impact on the hardware design. The demand for low unit cost and multyear lifetimes, combined with progress in CMOS and MEMS processing, are driving development of SoC solutions for sensor nodes at the cubic centimeter scale with a minimum number of off-chip components. Here, the feasibility of a complete, cubic millimeter scale, single-chip sensor node is explored by examining practical limits on process integration and energetic cost of short-range RF communication. Autonomous cubic millimeter nodes appear within reach, but process complexity and substantial sacrifices in performance involved with a true single-chip solution establish a tradeoff between integration and assembly.

KEYWORDS | Low-power circuits; low-power RF; Smart Dust; wireless mesh networks; wireless sensor networks; wireless sensors

I. INTRODUCTION AND HISTORY

The term “Smart Dust” has come to be used to describe a wide range of wireless sensor network hardware at a small scale down to a handful of cubic millimeters [1]. Each wireless sensor node, or “mote,” contains one or more sensors, hardware for computation and communication, and a power supply (Fig. 1). Motes are assumed to be autonomous, programmable, and able to participate in multihop mesh communication.

The genesis of Smart Dust was a workshop at RAND in 1992 in which a group of academics, military personnel, and futurists were chartered to explore how technology revolutions would change the battlefield of 2025 [2]. By this time it was clear that MEMS technology was going to revolutionize low-cost, low-power sensing. Moore’s law was accurately predicting CMOS digital circuit performance improvements with no end in sight, and the wireless communication revolution, already firmly established in two-way pagers, was beginning to make its way into handheld cellphones. The confluence of these three technological revolutions in sensing, computation, and wireless communication placed the major sensor mote functions on asymptotic curves down to zero size, power, and cost over time. Furthermore, the potential for cointegration of CMOS and MEMS made single-chip sensors with integrated signal conditioning possible at low cost [3]–[11].

In 1996, the term “Smart Dust” was coined to describe the ultimate impact of scaling and process integration on the size of an autonomous wireless sensor [12]. Several DARPA-sponsored workshops in the mid-1990s fleshed out some of the implementation and application details of the 1992 vision, and key research proposals were written and funded at the University of California, Los Angeles (UCLA); the University of California, Berkeley; and the University of Michigan, Ann Arbor. It was clear to the community at that time that low-cost ubiquitous wireless sensor networks would have a revolutionary impact on military conflict. What was not as clearly anticipated was the potential impact on commercial and industrial applications.

The first wireless sensor motes, called COTS (commercial-off-the-shelf) Dust, were built early in the Smart Dust project using printed circuit boards and off-the-shelf components. It was shown that these inch-scale devices could perform many of the functions predicted in the 1992 workshop, including multihop message passing and mote localization [13]. COTS dust and other macro-scale motes were developed to explore sensor network software and individual mote architecture as well as deploy small scale networks [14]–[16].
While great strides were made in miniaturization and power reduction of the hardware, perhaps the most important event during this early period was the observation that networks of autonomous sensor motes represented a ubiquitous, embedded computing platform [17]–[20], and they needed a new operating system to match. Proposed in the Endeavour project [21], the TinyOS operating system [22] was developed under DARPA funding and put into the public domain, along with all of the COTS Dust hardware designs, and a thriving open-source sensor networking community was born.

Meanwhile, in 1999 the IEEE formed the 802 working group 15, with a charter to develop standards for wireless personal area networking (WPAN), from which the low-rate WPAN 802.15.4 standard emerged. The 802.15.4 standard was designed from the beginning to be a low-power, low-complexity solution for sensor networking in industrial, automotive, and agricultural applications [23]. As a result of the HomeRF, focused on home automation applications, created the Zigbee standard in 2004. Zigbee 1.0 [24] is based on the 802.15.4 standard radio [25]. With the blessings of the IEEE on a radio standard, a consortium of large companies defining applications, and the help of the venture capital community, a new industry was born.

II. DEVELOPMENTS IN SENSOR MOTE HARDWARE

The Mica mote (Fig. 2), the most popular mote used in research, was developed to mimic the expected architecture of a highly integrated mote while using off-the-shelf parts mounted on a common PC board to reduce development time. This mote includes a microcontroller, RF transceiver, and the ability to interface to a variety of sensors. The mote is powered by a pair of AA batteries, and these take up the majority of the unit’s volume [14]. Similar in-chip motes utilizing primarily off-the-shelf components are now commercially available from numerous sources [26]–[30].

Development of highly integrated sensor mote components started in the mid-1990s and resulted in multichip systems that could be assembled to create a mote. At UCLA, MEMS devices were combined with commercial CMOS chips that provided sensor control and readout as well as communication [31]. At the University of Michigan, Ann Arbor, a 10 000-mm³ device containing sensors, computation, and RF communication using multiple chips in a single package was developed and demonstrated [32]. Other wireless sensor multichip units or components have been demonstrated for a variety of industrial, commercial, and defense applications [33]–[42].
To minimize energy, passive optical communication was explored for early Smart Dust motes. The smallest optical mote to date (Fig. 3.) displaced only 4 mm$^3$ and contained an 8-bit ADC, an optical receiver, a corner cube reflector passive optical transmitter, a light sensor, an accelerometer, a multivoltage solar cell power source, and limited computation [43]. A newer generation sensor mote, called the Spec mote, contained a microprocessor, SRAM, an RF transmitter, and an 8-bit ADC integrated onto a single CMOS die [41]. More recently, highly integrated chips with a complete RF transceiver, microprocessor, ADC, and sensor interface have been reported [44], and even commercialized [45]. Even these highly integrated chips still require an off-chip battery, some passive components, a crystal timing element, and an RF antenna, resulting in a complete package at the centimeter to inch scale.

III. WIRELESS SENSOR NETWORK APPLICATIONS

Today’s sensor networks rely on a wired infrastructure to provide power and transfer data. The high cost of running wire for power and communication often dramatically exceeds the cost of the sensors themselves, slowing the adoption of sensor networks for all but the most critical applications. By drastically reducing installation costs, reliable low-cost wireless mesh networking places sensor networks on the same technology curves as the rest of the IT revolution.

Wireless connectivity for sensors has been an attractive option for years, but, due to problems with reliability, adoption has been limited to applications where occasional loss of connectivity and data is acceptable. The current revolution in wireless sensing is being driven by the dramatic improvement in reliability and lifetime possible with wireless mesh networking. This is an echo of the Internet revolution, where point-to-point wired communications were replaced by multihop wired communication. The insensitivity of the Internet mesh to the loss of a path or a node is a key part of what makes the Internet reliable. The same concept applied to wireless sensor networks improves reliability.

In commercial and consumer applications, motes can be used to eliminate the wiring cost for light switches, thermostats, and fire alarms. Fig. 4 illustrates the wireless routing mesh blueprint from an actual sensor network deployment. In this application, motes were installed throughout a health clinic in just 2 h to implement a low-cost air temperature and energy consumption monitoring system with a simple Web browser based control interface [46].

In applications such as inventory monitoring, motes will not be fixed in space. A primary concern of the network will be determining the location of motes on boxes or pallets on demand and this requires location discovery capability to be built into the network [47].

In some entertainment applications inertial sensing motes may be worn by humans to detect and interpret movements as communication gestures or control signals [48]–[52]. Similarly, wearable motes have been used to interpret human motion as musical gestures, allowing users to create music interactively in real time [53]. In these systems the latency requirements are more stringent than in typical monitoring scenarios and, since humans will be wearing the sensor mote, a small form factor is important.

Defense applications drove much of the initial research in sensor networks. The Igloo White system was a wired sensor network employed from 1966 to 1972 along the Ho Chi Minh trail during the Vietnam War. In a more modern military application, wireless sensors were distributed throughout a mock urban battlefield to pinpoint a sniper’s location by acoustically detecting the arrival time of the muzzle blast at several different points in the field [54]. Sensor networks have also been proposed for position tracking and identification of people and fast-moving vehicles in both civilian and military scenarios [55].

IV. APPLICATION REQUIREMENTS AND HARDWARE IMPLICATIONS

Applications for wireless sensor networks can be broken down into two categories: wire replacement and wirelessly enabled. In the former case, the cost of hardware for a wireless solution is generally dramatically lower than the comparable cost of running wiring. Once secure, reliable, low-power solutions are demonstrated in this domain, adoption is limited by caution, rather than cost. Wirelessly
enabled applications, on the other hand, may require novel technologies such as dynamic mote localization.

A. Reliable Data Delivery

Reliability in a multihop RF mesh sensor network can be defined in terms of end-to-end delivery of time-stamped sensor data with a specified worst case latency. Time stamping requires some form of network synchronization, but the primary hardware impact of the reliability requirement is on the choice of radio and the use of spectrum. The majority of motes will operate in regulated but unlicensed bands, such as 902–928 MHz in North America, and 2.4–2.485 GHz throughout most of the world. Because these bands are open to transmitters putting out as much as 1 W, and motes are likely to have an output on the order of 1 mW to extend battery life, it is critical that motes be able to avoid high-power interferers to maintain adequate reliability. For example, even the spreading gain of the direct sequence spread spectrum 802.15.4 radio will not prevent an 802.11 transmitter from jamming several channels over distances of tens of meters [56]. Multipath propagation effects indoors cause similar problems for reliability, with time-varying

---

**Fig. 3.** Conceptual drawing and SEM of the optical Smart Dust node presented in [43]. This multichip node displaced only 4 mm² and featured a solar cell power source, temperature, light and acceleration sensors, an 8-bit ADC, and bidirectional optical communication.
narrowband fading of many tens of dB commonly observed [57].

High-powered interferers and unpredictable fading preclude the use of fixed-frequency radios in high-reliability applications. Reliable solutions will have the ability to avoid or work around those parts of the spectrum which are jammed or deeply faded. For relatively narrowband radios like 802.15.4, this implies some form of channel hopping at the medium access layer in addition to the direct sequence spreading defined in the physical layer of the radio.

B. Low-Power Consumption

From a system deployment perspective, mote lifetimes measured in years are required for most applications in building and industrial automation. Operation from batteries and/or scavenged power is required. To avoid high battery replacement costs, this dictates a battery lifetime of between one and ten years. An AA-sized battery contains roughly 250 μA-years of charge or about 12 000 J. For some lithium chemistries, the internal leakage is low enough that supplying this charge as a current of 25 μA for a decade is possible while common alkaline chemistries have shorter lifetimes. The average power consumption of an inch-scale mote, then, must be in the range of tens to hundreds of microwatts or just a few joules per day.

Achieving a total current consumption of tens of microamps requires deep duty cycling, on the order of 1% or less with off-the-shelf hardware [58]–[60]. Deep duty cycling implies that the hardware should be able to quickly transition between the powered state and the unpowered (and low leakage) state. At very low duty cycles, leakage power in the digital circuits, predominantly the SRAM, can dominate the system energy budget. Though the power required for active digital circuits is scaling down with minimum feature size of standard CMOS, leakage power is growing. Leakage power sets a lower bound on average power consumption of sensor motes and is a major obstacle to the scaling of digital circuits. In 130-nm bulk CMOS, for example, leakage is on the order of 1 μW per kilobyte with a standard 6T SRAM [61]. Silicon-on-insulator (SOI) is a CMOS device technology offering substantial leakage reduction that has yet to be adopted into mainstream commercial use [62].

In addition to leakage issues, multihop mesh networking with radio communication in a deeply duty cycled environment presents major challenges to algorithm and software developers. Turning the radio off 99% of the time is easy, but knowing exactly when to turn it on again is not. Hardware support for some combination of mote-to-mote time synchronization, fast radio polling, or low-power detection of RF energy is desirable.

Custom-designed circuits leveraging the relaxed performance specifications unique to sensor network

Fig. 4. Deployment of a wireless sensor network in a health clinic. The network monitors air temperature and energy consumption and has a convenient central control interface [46].
applications have been developed to drastically extend sensor mote lifetimes and/or reduce cost and size by minimizing energy consumption [63]–[66]. Fig. 5 provides a comparison of the energy consumption per operation of published custom ICs and off-the-shelf parts. In both commercial and custom solutions, the energetic cost of RF communication dwarfs that of other sensor node operations, making RF a bottleneck for size, cost, and lifetime improvements. In Section VI, the energy requirements of RF communication are explored and a system-level optimization of energy per transferred bit of a generic transceiver is performed.

Unfortunately, the custom ICs presented in Fig. 5 operate optimally at different supply voltages and were not developed in the same CMOS process. Integration of these devices would require redesign in one process and efficient dc level conversion from a battery [67]. Ideally, the hardware would operate efficiently with the lithium cell potential and deep duty cycling. Since lithium chemistries generally provide over 3-V cell potential, this presents a challenge for integration in deep submicrometer CMOS, where normal supply voltages are half of the lithium potential or less. The computational requirements of a mote are generally consistent with MHz rather than GHz operation, allowing digital circuits to run as low as 0.5 V or less. Efficient dc to dc conversion from a constant 3 V supply to a duty cycled 0.5–1.8 V output will allow future systems-on-chip to achieve battery-referenced energy efficiencies similar to those shown in Fig. 5.

C. Security

Security in sensor networks shares many of the same problems as IT security in general, with the beneficial exception that fewer humans are involved. As sensor networks come to be used in commercial, industrial, and defense applications, their security requirements will likely be just as stringent as those required of the information systems they feed [68]–[71]. Security requirements include access control, data encryption, message authentication, key exchange, and certification of trust.

Link-level encryption and message authentication can be performed in software, but these low-level, time critical, and computationally intensive operations are a natural target for silicon support. The hardware for the Advanced Encryption Standard (AES) [72] is already incorporated in chips which support the 802.15.4 standard [60], [73].

For key exchange and certification, software implementations of the public key algorithms RSA and ECC have been demonstrated on 8- and 16-bit processors common in sensor network applications [74], [75]. Execution times are on the order of seconds to tens of seconds, and the memory requirements are substantial for a mote. The addition of integer multiplication units with large operand size will speed execution and reduce memory requirements roughly as the square of the operand size.

D. Location Discovery

As the cost of motes falls and the number of wireless sensors increases, the cost of locating installed sensors will drive the development of automatic location discovery. This capability is critical for asset tracking applications and for many of the “sprinkle deployment” military and environmental monitoring applications envisioned for the technology. Furthermore, many applications require mobile motes with the ability to dynamically update position information [76]. Asset management and other tracking applications may require an accuracy of 1 m to find a person, laptop, or record file in an office building or hospital, several meters to find a crate in a warehouse, or many tens of meters to find a cargo container in a shipping yard.

Acoustic localization systems with good performance have been implemented [77], but the physics of acoustic
by increasing frequency and making the antenna resonate, leading many to investigate integrated resonant antennas at and above 10 GHz [85]–[87]. Even at appropriately high frequencies, integrated antennas have thus far only demonstrated low efficiencies. Furthermore, propagation losses are generally worse at higher frequency. As a result, an integrated antenna will incur a substantial power penalty with current technology.

If a modest size increase and some assembly are acceptable, commercially available miniaturized antennas may provide the best combination of cost-effectiveness, size, and efficiency. Efficient dielectric chip antennas displacing only about 10 mm$^3$ are commercially available for use at low-GHz frequencies from a variety of sources [88], [89]. However, as designed, these chip antennas require sizable ground planes for good performance.

The timing reference is another element of RF transceivers not amenable to integration. RF transceivers typically use a resonant quartz crystal to synthesize high frequency signals needed for transmission and reception. The geometry of crystal references is precisely controlled to create a mechanical resonance that is stable across a wide temperature band. There are no conventional circuit elements that can offer precision comparable to a crystal. However, MEMS resonators are currently being explored in industry and academia as a quartz crystal replacement technology because of their potential for integration and cost reduction [90]–[93]. Currently, the temperature stability of MEMS resonators is not as good as quartz crystals, but temperature compensation may be employed to mitigate this problem [94]. As this technology matures, MEMS components may supplant not only the crystal timing element, but filters, mixers, and RF oscillators as well [95].

B. Sensors

For some applications, the sensors available in a standard integrated circuit process may be sufficient. Temperature, magnetic field, and capacitive fingerprint sensors have all been demonstrated in standard CMOS as well as megapixel cameras with on-chip image processing [96]–[100]. Integrated sensing of colored light can also be done in CMOS using metal grating patterns or variable depth PN junctions as a color filter [98], [101]. Imaging arrays are increasingly finding applications in noncamera applications, such as motion-flow sensing in computer mice. Imaging of legacy dials, knobs, and lights in industrial environments combined with local signal processing at the sensor to transmit only the dial position is a potentially low-power, low data rate application.

There are a host of miniaturized sensors possible with MEMS technology: linear and angular rate acceleration, pressure, chemical, fluid flow, audio microphones, and more [3], [4], [102]–[105]. While all of these sensors are also available off-the-shelf, custom designed MEMS sensors have the distinct advantages of low cost and size as well as the potential for integration with circuits.
Many methods of integrating MEMS devices with circuits have been demonstrated. One popular method is to etch away materials from commercially manufactured integrated circuit wafers to create mechanically free structures. The features of the resulting structures are defined using existing layers in the CMOS to selectively block etching. In [3], [4], [8], and [11], an electrochemical etching technique was applied to standard CMOS wafers after fabrication to create cantilevered beams and membranes for chemical, infrared, pressure, and other types of sensors. Dry etching techniques were applied to CMOS wafers in [7] and [106], to create inertial sensors and electrostatic actuators. An advantage of both of these integration techniques is low cost because no additional deposition or lithography is necessary after the circuits are fabricated (see Fig. 6). However, since the MEMS structures are defined by a stack of dielectrics and metals designed for CMOS, they may have undesirable mechanical properties.

The performance of resonant MEMS devices used for both sensing and RF applications is particularly sensitive to the mechanical properties of the constituent materials. Thus, many have investigated other integration methods that permit the use of mechanically advantageous materials. Adding thin films of polycrystalline materials to fabricated CMOS wafers, or surface-micromachining, is a powerful technique that combines the advantages of integration with CMOS and high performance resonant MEMS. Integrated high-Q MEMS resonators and resonant sensors made from both polycrystalline Silicon (poly-Si) and silicon–germanium (poly-SiGe) films have been demonstrated with surface-micromachining techniques [5], [9], [10]. Unfortunately, the elevated processing temperatures required for poly-Si (well above 400 °C) are too high for the aluminum metallization typical of standard CMOS. Thus, integrated poly-Si MEMS must either be machined into the wafers before metallization steps [9], [10] or added to fabricated CMOS without metallization [5]. However, the reduced processing temperatures of poly-SiGe are much more compatible with metallization, making postprocessed poly-SiGe a strong candidate for the future of CMOS-MEMS integration (see Fig. 6) [6], [104].

C. Scavenging and Storing Energy

Both electrostatic MEMS devices and PZT transducers have been used to harvest energy from ambient mechanical vibrations [107]–[110]. It should be possible to
integrate these devices, since vibration harvesting may be performed with simple electrostatic MEMS. The achievable power density with this method is strongly dependent on the environment and the design of the transducer. However, theory predicts a power density of 1.16 μW/mm² is available from a device mounted on the casing of a constantly operating microwave oven [110].

The most abundant and practical form of ambient power comes from the sun. In full sunlight, the available solar power per unit area is roughly 1 mW/mm² in the continental United States [111]. Simple silicon-based photovoltaic cells can convert this to electrical power with up to 25% efficiency [112]. The processing steps necessary to create silicon solar cells are quite compatible with standard IC manufacturing. In fact, the PN junctions inherent in the silicon of any integrated circuit are inadvertent solar cells. However, with standard CMOS, it is not straightforward to utilize these junctions as solar cells and simultaneously operate transistors on the same chip due to isolation issues. Integration of multijunction solar cells and CMOS circuitry has been demonstrated using silicon-on-insulator wafers with trench isolation [113]. Miniature, but not integrated, solar cells are currently available off-the-shelf from a variety of manufacturers. In particular, silicon-based, flexible, thin-film solar cells mounted on polymer substrates are now commercially available in custom sizes on the order of 1 mm² and up [114].

To sustain reliable operation in the presence of fluctuating ambient solar or mechanical energy, a sensor mote must be able to store harvested energy. A promising technology for integrated energy storage is thin-film batteries. Work at Oak Ridge National Laboratory on lithium-based thin-film batteries [115] has led to commercial cells on the millimeter scale with high capacity and long cycle lives [116]–[120]. Packaging adds volume without increasing capacity, resulting in lower energy/volume ratios, but [115] reported 0.25 mA – hr/cm² at 4 V (or 36 mJ/mm²) in batteries as small as 5 mm² and only 15 μm thick without packaging.

Battery discharge rates as high as 40 mW/cm² are possible [115], and these cells can be laminated to a CMOS wafer, eliminating the need for packaging [117]. Cells as small as 50 μm × 50 μm have been demonstrated using standard lithographic techniques [121]. Other recent work in thin-film batteries has produced promising results with lower cell potentials, but the cell capacity and robustness is far behind solid-state lithium-based batteries [122], [123].

VI. ENERGY REQUIREMENTS OF WIRELESS COMMUNICATION

Based on a comparison of published solutions for RF transceivers and other sensor mote functions, the wireless communication circuits dominate the system energy budget.

This section explores the energy requirements of wireless communication and derives approximate energy targets.

As a first step, consider transmitting a single bit from one sensor mote to another over a distance r, using a carrier frequency f, and bitrate b. To determine the minimum required transmission power ($P_{TX,MIN}$), one must first determine how the signal power diminishes with distance, and then determine the minimum detectable signal power in the receiver ($P_{MDS}$). If a maximum communication range (r) is assumed, then $P_{TX,MIN}$ must be larger than $P_{MDS}$ by a factor equal to the transmission loss ($L_{PATH}$) due to propagation.

A. Transmission Loss Approximations: $L_{PATH}$

Electromagnetic theory states that the strength of a transmitted signal is attenuated with increasing distance according to the Friis equation [124]

$$L_{PATH} = \left(\frac{4\pi r}{\lambda}\right)^2.$$  \hspace{1cm} (1)

$L_{PATH}$ is the attenuation due to propagation and $\lambda$ is the wavelength at the frequency of interest ($\lambda = 30$ cm at $f = 1$ GHz). The Friis equation applies to free-space, line of sight propagation and, as such, underestimates path loss for nonideal conditions. In cluttered environments, path loss is much more complex. Several sources have utilized a modification to the Friis equation that roughly approximates propagation losses in less ideal environments, such as indoors [80]

$$L_{PATH} = \left(\frac{4\pi r_0}{\lambda}\right)^2 \left(\frac{r}{r_0}\right)^n.$$ \hspace{1cm} (2)

In this model, $r_0$ is a reference distance ($r_0 = 1$ m is often used) beyond which the inverse square characteristic of the Friis equation no longer governs propagation losses because of obstacles and multipath interference. The exponent n characterizes the attenuation beyond $r_0$ and has been measured for various propagation conditions. For short-range indoor propagation in the low GHz range, $n = 4$ is a common choice for the exponent [80], [125].

B. Minimum Detectable Signal Power: $P_{MDS}$

In any real receiver, there is a finite thermal noise power ($P_N$) inherent in the system that is proportional to both input bandwidth, BW, and the product $kT$; where $T$ is temperature in Kelvin and $k$ is Boltzmann’s constant

$$P_N = kT \cdot BW.$$  \hspace{1cm} (3)
The minimum detectable signal power in a receiver ($P_{MDS}$) is always greater than $P_N$. The product of two terms, noise factor (NF) and signal-to-noise ratio ($SNR_{MIN}$), quantifies the ratio by which $P_{MDS}$ must exceed $P_N$ for successful transmission. The noise performance of RF receivers is characterized by NF, defined as the ratio of the total equivalent noise power to the fundamental lower noise power limit $P_N$. In the best case, NF equals 1, but it is often in the range of 1.5–10 for real receivers. Intuitively, higher NF implies that the receiver’s internal noise generators dominate over the noise incident on the antenna.

The second term $SNR_{MIN}$, describes the minimum required ratio of signal power to noise power that must be maintained to properly detect signals with a certain probability. For example, to average less than one error for every 1000 bits (or $BER = 10^{-3}$), a theoretical minimum $SNR_{MIN}$ of 12 dB is required when noncoherent FSK is the modulation technique and no coding is done [126].

Under these assumptions, the minimum detectable signal power in the receiver is given by product

$$P_{MDS} = P_N \cdot (NF \cdot SNR_{MIN})$$

$$= kT \cdot BW \cdot NF \cdot SNR_{MIN}.$$  \hspace{1cm} (4)

### C. Minimum Transmission Energy per Bit: $E_{BIT,TX}$

Link margin ($LM$) quantifies the maximum path loss between transmitter and receiver that can be tolerated while maintaining a reliable link. $LM$ is given by the ratio of $P_{OUT}$ to $P_{MDS}$. At the maximum communication range ($r_{MAX}$), $LM$ is equal to $L_{PATH}$. Therefore, given $r_{MAX}$, the lower bound on transmitted power ($P_{TX,MIN}$) is simply the product of $L_{PATH}$ and $P_{MDS}$.

$$P_{OUT,MIN} = L_{PATH} \cdot P_{MDS}$$

$$= \left(\frac{4\pi \cdot r_0}{\lambda}\right)^2 \cdot \left(\frac{r_{MAX}}{r_0}\right)^n \cdot kT \cdot BW \cdot NF \cdot SNR.$$  \hspace{1cm} (5)

To convert $P_{OUT,MIN}$ to energy per bit ($E_{BIT,TX}$), we must assume a relationship between the bitrate and the receiver input bandwidth BW. Bitrate is generally proportional to input bandwidth and, depending on the modulation technique, may be higher or lower than BW. For simplicity, we assume the bitrate is equal to BW. Then, $E_{BIT,TX}$ is given by

$$E_{BIT,TX} = \frac{P_{OUT,MIN}}{bitrate}$$

$$\approx \left(\frac{4\pi r_0}{\lambda}\right)^2 \cdot \left(\frac{r_{MAX}}{r_0}\right)^n \cdot kT \cdot NF \cdot SNR_{MIN}.$$  Let bitrate $= BW$.  \hspace{1cm} (6)

To calculate the minimum $E_{BIT,TX}$, assume the base station is an ideal, noncoherent FSK receiver (i.e., let $NF=1$ and $SNR_{MIN}=12$ dB) located $r_{MAX}$ meters away and apply (6). Assuming $n = 4$, $r_0 = 1$ m, $r_{MAX} = 20$ m and a 1-GHz carrier signal, the minimum energy per transmitted bit is only 20 pJ—a factor of at least $10^2$ lower than any of the reported values from Section IV.

In this scenario, if a bitrate of 1 Mb/s is used, only 20 jJ must be transmitted to maintain a 20-m link. On the other hand, if a 2.4-GHz carrier is chosen, the minimum energy

![Fig. 7. Simplified block diagram of a low-IF or direct conversion RF transceiver.](image)
per bit increases to 114 pJ, because path loss is worse at higher frequencies. This calculation represents the minimum transmitted energy to reach a perfect receiver (i.e., $NF = 1$) 20 m away. The total consumed energy by the transmitter must be substantially higher due to overhead circuit power ($P_{OH,TX}$) and nonideal efficiency in the output amplifier ($e_{PA}$).

**D. Design Considerations and Practical Targets for $E_{BIT}$**

When calculating network energy cost per bit, the power consumption of both the transmitting and receiving motes should be included. Second, the models for transmitter and receiver should take overhead power and nonideal SNR, $NF$, and PA efficiency ($e_{PA}$) into account. A block diagram of a conventional direct-conversion or low-IF transceiver, labeled with sources of overhead power, is shown in Fig. 7.

The outlined portions of Fig. 7 represent sources of power overhead. Though these blocks are needed for functionality, they constitute overhead in the sense that increasing power spent in them does not directly increase link margin. In both transmitter and receiver, a large portion of the overhead power is dedicated to generating a stable RF signal with a voltage controlled oscillator (VCO). Other significant sources of overhead are RF mixers for modulation and channel selection, ADCs, DACs, and low-frequency filters. The power overhead of the VCO and RF mixers is relatively independent of BW. However, the overhead power in the DAC, ADC, and low-frequency filters for channel selection and baseband processing will depend on BW. Radios designed specifically for sensor networks in [63] and [127]–[130] reported numbers for power overhead between 0.17 and 0.9 mW in receive mode, 0.3–7 mW in transmit mode for bitrates of 300 kb/s and below.

In contrast, increased power in the PA and LNA does directly increase link margin. In general, increased power in the LNA makes the receiver more sensitive by decreasing $NF$, but the proportional noise benefit steadily diminishes at high power levels as $NF$ asymptotically approaches its minimum value of 1. On the other hand, the output of a PA can be roughly proportional to power consumed over a wide range. Efficient PA design over a broad range of power outputs is discussed in [131]. Power output of a PA can then be simply modeled by the product of efficiency ($e_{PA}$) and power consumed ($P_{PA}$). PA efficiencies ($e_{PA}$) of 40% or higher have been reported for various PAs with output power from 200 μW to 10 mW and beyond [127], [129], [132].

**E. Optimal Bandwidth to Minimize $E_{BIT}$**

Fig. 8 shows a first-order graphical representation of the power-performance tradeoffs in a simple RF transceiver.
transceiver. The figure is labeled with reported values of $P_{OH,TX}$, $P_{OH,RX}$, $P_{MDS}$ and $e_{PA}$ from [127]. The simplified model is useful for demonstrating tradeoffs and deriving approximate energy consumption targets. Measured data reported in [63] is shown in Fig. 9 for comparison.

The equations describing this model are given below. The term $\gamma$ is dependent on antenna impedance, supply voltage, and other circuit parameters (see [131]), but is equal to 2 mW for the transceiver in [127]

$$P_{OH} = P_{OH,RF} + P_{BB} \left( \frac{1 + BW}{BW_0} \right) \quad (7)$$

$$P_{MDS} = kT \cdot BW \cdot SNR_{MIN} \cdot \left( 1 + \frac{\gamma}{P_{LNA}} \right) \quad (8)$$

$$P_{OUT} = e_{PA} \cdot P_{PA} \quad (9)$$

The first question we wish to address is: Given a fixed power budget for a link, how should power be distributed between PA and LNA to maximize link margin ($L_M$)? Dividing (9) by (8), we get an equation for $L_M$ in terms of power consumption in PA and LNA. The goal is to maximize $L_M$ when the sum $P_{PA} + P_{LNA}$ is held constant, and the resulting equation, optimally relating LNA and PA power consumption, is shown below

$$\max_{P_{PA} + P_{LNA} = C} \{L_M\} \Rightarrow \frac{dL_M}{d(P_{LNA})} = 0$$

$$\Rightarrow P_{PA} = \frac{P_{LNA}^2}{\gamma} + P_{LNA}. \quad (10)$$

This ratio is independent of the path loss exponent assumed in (2). It is important to note that we have implicitly assumed a time synchronized network, where
receiver and transmitter duty cycles are approximately equal. By setting $L_M = L_{PATH}$ from (2), we can use (8) and (9) to relate transceiver power consumption to range ($r_{MAX}$) and bitrate (again, assume bitrate $= BW$)

$$r_{MAX} = r_0 \left( \frac{\lambda}{4\pi r_0} \right)^\frac{1}{4} \cdot \left( \frac{e_{PA} \cdot P_{PA}}{kT \cdot BW \cdot SNR_{MIN}} \cdot \frac{P_{LNA}}{P_{LNA} + \gamma} \right)^\frac{1}{4}$$

(11)

Now, using (10) to relate $P_{PA}$ and $P_{LNA}$

$$P_{LNA,OPT} = \left( \frac{r_{MAX}}{r_0} \right)^\frac{1}{4} \cdot \left( \frac{\lambda}{4\pi r_0} \right)^\frac{1}{4} \cdot \frac{(\lambda \cdot kT \cdot BW \cdot SNR_{MIN})}{P_{LNA} + \gamma}$$

(12)

![Energy per bit and transceiver power distribution versus bandwidth for fixed link margin of 88 dB (e.g., $r \sim 25$ m by (3)).](image-url)
The BW dependent expression for minimum energy per bit is the sum of $P_{LNA}$, $P_{PA}$, and $P_{OH}$ divided by BW.

$$E_{\text{BIT,MIN}} = \frac{1}{BW} \left( P_{OH} + \frac{P_{LNA,\text{OPT}}^2}{\gamma} + P_{LNA,\text{OPT}} \right).$$  (13)

$P_{OH}$, $P_{PA}$, and $P_{LNA}$ are defined by (7), (10), and (12), respectively. $P_{OH}$, $P_{PA}$, $P_{LNA}$, and $E_{\text{BIT}}$ are plotted against BW in (Fig. 10) for a constant link margin of 88 dB (or $r_{\text{MAX}} = 25$ m with $n = 4$, $r_o = 1$ m, and a 900-MHz carrier). This tradeoff is most relevant for communication over a fixed range, as is the case when only one data path is available (Fig. 11, top). The values for $\gamma$, $e_{PA}$, $P_{OH}$, and $\text{SNR}_{\text{MIN}}$ are taken from the transceiver in [127]. According to (13), the benefits of increasing BW diminish as the BW dependent terms $P_{PA}$ and $P_{BB}$ exceed the fixed overhead $P_{OH,RF}$.

Equation (13) ignores the energetic cost of initializing the transceiver and synchronizing it with the network. Each time a mote wakes up to transmit or receive data, it must first enable all the necessary baseband analog circuits, lock its VCO to the correct center frequency with a phase locked loop (PLL), and synchronize with its neighbor(s) at both the MAC and the PHY layers. The details of the MAC-layer synchronization phase depend on the specific implementation [133]–[136], but in each case the goal is to make sure that one or more receivers is actively listening when a transmitter sends its packet. Whatever the approach, extra time is spent with mote radios on and no useful data flowing.

Once both receiver and transmitter are on and tuned to a channel, there is packet overhead that sets a minimum practical packet length. Packet overhead includes most of the following: the radio startup/training sequence, packet start symbol, packet length, addressing, encryption, and error detection overhead. The 802.15.4 standard requires 11 bytes for an ACK with no addressing or security. An acknowledged message with a one-byte payload, short addresses, and minimal message integrity sent in 802.15.4 would consist of a 20-byte packet separated from an 11-byte Ack by a standard-mandated 6-byte turnaround time between packet and ACK. This makes a total of 37 bytes of time that both radios need to be on to transmit a single byte payload, or less than 3% payload to packet efficiency. By using a maximum-length payload, the overall payload efficiency can be up to 76%.

During synchronization, most or all of the transceiver’s circuits are consuming power. Therefore, the initialization cost is only negligible if data packets are long enough such that the time spent sending data is much greater than the time spent in starting up, synchronization, and packet overhead.

Fig. 11. Top: lack of multiple paths imposes a range constraint on communication between nodes while bandwidth is flexible. Bottom: a dense network of nodes with interfering signals constrains available bandwidth, but link range remains flexible owing to redundancy of paths.
To incorporate the effect of transceiver startup time on the overall $E_{\text{BIT}}$ versus BW tradeoff, some knowledge of average number of data bits per transmission ($N_{\text{AVG}}$) and transceiver initialization, synchronization, and packet overhead time ($t_{\text{INIT}}$) is needed. For the purpose of illustration, we assume $N_{\text{AVG}} = 1000$ bits and $t_{\text{INIT}} = 1$ ms. Assuming the transceiver is consuming full power during synchronization, the energy cost per bit is then the product of total link power and $t_{\text{INIT}}$ divided by $N_{\text{AVG}}$

$$E_{\text{BIT,INIT}} = (P_{\text{OH}} + P_{\text{PA}} + P_{\text{LNA}}) \cdot \left( \frac{t_{\text{INIT}}}{N_{\text{AVG}}} \right). \quad (14)$$

The total energy per bit, including initialization and transmission, is the sum of (13) and (14). $E_{\text{BIT}}$ is minimized.

Fig. 12. Top: optimum ratio of PA to LNA power. Bottom: energy per bit per meter (EBIT-MTR) versus the sum of PA + LNA for 3 values of the path-loss exponent ($n$). Optimum link margin and range are labeled for each value of $n$. 
when the amount of energy spent during synchronization and data transmission are equal, or equivalently (see Fig. 10)

\[ BW_{OPT} = \frac{N_{AVG}}{T_{INIT}}. \]  

(15)

F. Optimal Link Margin and Range to Minimize \( E_{BIT-MTR} \)

Suppose we wish to send a set of data over a long distance through a dense network with many available paths (Fig. 11, bottom). From a global network energy perspective, should we send the data the entire distance in one hop, in several tiny hops to nearest neighbor motes, or is there an ideal link range somewhere in between? In dealing with this question, energy per bit per meter \( (E_{BIT-MTR}) \) is a more appropriate metric than \( E_{BIT} \).

If path loss characteristics are known, we can find an optimum link range that will minimize the global network energy cost for data transport by minimizing \( E_{BIT-MTR} \). Since (13) relates \( E_{BIT} \) to both BW and \( r \), \( E_{BIT-MTR} \) can be obtained by simply dividing \( E_{BIT} \) by \( r \). \( E_{BIT-MTR} \) is plotted versus power with BW fixed at 1 MHz for three values of the path-loss exponent at the bottom of Fig. 12. This plot shows that there exists an optimum energy range and link margin for transporting data through a network that depends on path-loss conditions and transceiver characteristics. The optimum link margin \( (L_{M,OPT}) \) varies by only 11 dB for values of \( n \) from two to four and has the lowest value when the path-loss exponent is highest, implying shorter hops are preferred when path-loss is worst.

VII. DISCUSSION

It is clear that a system-on-chip wireless sensor node with an active power dissipation of less than 1 mW is not only possible, but likely to be commercialized. The performance possible in such a mote will be impressive, including secure wireless communication at hundreds of kilobits per second over distances of tens of meters, multihop mesh networking, onboard sensors, 10- to 16-bit ADCs, and a sensor datapath. Today’s commercially available software runs all motes in a mesh network at less than 1% radio duty cycle [26]. This implies average mote power consumption of between 1 and 10 \( \mu \)W. At these power levels, mote lifetimes above a decade will be possible with coin cell, or even button-cell batteries.

Near-term IC process scaling will reduce the area required for memory and digital circuits to below a square millimeter, but the analog and RF portions will not scale as readily. Radio transceivers are unlikely to shrink much in finer line width processes, as their area is determined more by the physics of inductors than the transistors that drive them. Unless integrated resonant LC tanks are abandoned, low-GHz radios are stuck around a square millimeter. Process scaling driven by purely high-speed digital constraints is unlikely to provide the low leakage necessary for submicrowatt operation, but other
applications will drive low-leakage options in fine-line width processes, and clever circuit design may solve the leakage problem even in standard processes.

MEMS technology is likely to play a role in the integration of a broader selection of sensors on chip. In addition, RF filters and frequency references for both real-time clocks and RF local oscillators are possible. Similarly, nanotechnology is likely to be added first in the area of sensors. Improvements in the stability of low-power real-time clocks, based on MEMS, nano, or any other technology, would have an immediate impact on mote-to-mote time synchronization and therefore power consumption. The integration of MEMS or nano could in principle reduce the size of radios well below a square millimeter, but these radios will face the same challenging RF environment as the radios that they replace, so millimeter, but these radios will face the same challenging

While in principle it is possible to integrate a battery, antenna, and timing reference into a single-chip mote with no external components, this is unlikely to be the most economical approach. Integration of all the components of a mote onto a single chip will involve making substantial sacrifices in performance. The efficiency of a millimeter-scale chip-based antenna will be lower than that of a well designed antenna external to the chip. Power scavenging and storage in a future integrated process will not match what is possible with optimized off-chip components. On the other hand, on-chip time-keeping and frequency references using MEMS or nano may ultimately rival or even exceed the performance of off-chip crystal references. Fig. 13 illustrates some possible incarnations of a wireless sensor mote, underscoring size, power, and performance tradeoffs of integration versus assembly.

For all of the performance and cost limitations of a true system-on-chip mote with no external components, surely at some point they will be produced, if only for academic research. When that is the case, then wafers full of completely functional motes will be formed in the final metal etch of a CMOS process, take their first photovoltaic breaths of life from the plasma’s glow, and start chatting with each other while waiting for wafer passivation and dicing.

REFERENCES


Vol. 94, No. 6, June 2006
Cook et al.: Soc Issues for RF Smart Dust


ABOUT THE AUTHORS

**Ben W. Cook** (Student Member, IEEE) received the B.E. degree from Vanderbilt University, Nashville, TN, in 2001. He is currently working toward the Ph.D. degree at the University of California, Berkeley. Since the summer of 2003, he has worked as a Design Engineer and Consultant for Dust Networks, Hayward, CA, where he has worked on ultralow-power transceivers for the 900-MHz and 2.4-GHz ISM bands. His research has focused on low-power, highly integrated hardware for wireless sensor networks, with a particular emphasis on RF transceivers.

**Steven Lanzisera** (Student Member, IEEE) received the B.S.E.E. degree from the University of Michigan, Ann Arbor, in 2002. He is currently working toward the Ph.D. degree at the University of California, Berkeley. He was an engineer with the Space Physics Research Laboratory at the University of Michigan from 1999 to 2002, where he worked on satellite integration and testing. He has held internships with Guidant Corporation and TRW Space Systems, respectively. His research has focused on low-power mixed signal IC design and RF time of flight ranging technologies.

**Kristofer S. J. Pister** received the B.A. degree in applied physics from the University of California, San Diego, in 1986 and the M.S. and Ph.D. degrees in electrical engineering from the University of California, Berkeley, in 1989 and 1992. From 1992 to 1997 he was an Assistant Professor of Electrical Engineering at the University of California, Los Angeles, where he helped developed the graduate microelectromechanical systems (MEMS) curriculum, and coined the term “Smart Dust.” Since 1996, he has been a Professor of Electrical Engineering and Computer Sciences at the University of California, Berkeley. In 2003 and 2004, he was on leave from the University of California, Berkeley, as CEO and then CTO of Dust Networks, Hayward, CA, a company he founded to commercialize wireless sensor networks. He has participated in many government science and technology programs, including the DARPA ISAT and Defense Science Study Groups, and he is currently a member of the Jasons. His research interests include MEMS, micro robotics, and low-power circuits.