Design and Analysis of Hybrid NEMS-CMOS Circuits for Ultra Low-Power Applications

Hamed F. Dadgour and Kaustav Banerjee
Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106
e-mail: {hamed, kaustav}@ece.ucsb.edu

Abstract

Integration of nano-electro-mechanical switches (NEMS) with CMOS technology has been proposed to exploit both near zero-leakage characteristics of NEMS devices along with high ON current of CMOS transistors. The feasibility of integration of NEMS switches into a CMOS process is illustrated by a practical process flow. Moreover, co-design of hybrid NEMS-CMOS as low power dynamic OR gates, SRAM cells, and sleep transistors is explored. Simulation results indicate that such hybrid dynamic OR gates can achieve 60-80% lower switching power and almost zero leakage power consumption with minor delay penalty. However, the hybrid gate outperforms its CMOS counterpart both in terms of delay and switching power consumption with increase in fan-in beyond 12. Additionally, it is shown that the proposed hybrid SRAM cell can achieve almost 8X lower standby leakage power consumption with only minor noise margin and latency cost. Finally, application of NEMS devices as sleep transistors results in upto three orders of magnitude lower OFF current with negligible performance degradation as compared to CMOS sleep switches.

Categories and Subject Descriptors

B.7.1 [**Integrated Circuits**]: Types and Design Styles – *Advanced technologies.*

General Terms

Performance, Design, Reliability.

Keywords

Hybrid NEMS-CMOS technology, Leakage power, Low-power design, Nano-electromechanical switches, Suspended-gate MOSFET, VLSI.

1. Introduction

During the past two decades CMOS integrated circuits (ICs) have witnessed unprecedented improvements in their functionality and performance. This was primarily achieved by aggressive technology scaling, which resulted in device density and performance doubling roughly every *18* months as per Moore's law [1] while achieving a remarkable *25%* per year improvement in cost per chip function. As CMOS scaled from generation to generation, power dissipation increased proportionately to increasing transistor density and switching speeds. However, with the minimum feature size of the transistor entering the sub-100 nanometer regime, power dissipation is increasing ominously especially due to a substantial increase in the (sub-threshold) leakage power. Sub-threshold leakage power used to be insignificant for earlier generations of ICs but is becoming an increasing fraction of the total power [2]-[3]. The increase in sub-threshold leakage power arises due to the fact that power supply (V_{dd}) scaling necessitates threshold voltage (V_{th}) scaling. This trend, which is forecasted by the International Technology Roadmap for Semiconductors (ITRS) [4] is shown in **Figure 1**. Moreover, most leakage mechanisms are strongly temperature dependent. This strong coupling between temperature and leakage can cause further increase in total power dissipation [5].

The subthreshold leakage power is strongly influenced by the *subthreshold-swing* (S) of a device, which essentially indicates the amount of gate voltage reduction necessary to reduce the subthreshold current by one decade ($S = dV_{gs}/d\log I_d$) [6]. For bulk CMOS, the subthreshold swing has a fundamental lower limit of *60 mV/decade*.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
DAC 2007, June 4–8, 2007, San Diego, California, USA.
Copyright 2007 ACM 978-1-59593-627-1/07/0006 ...$5.00.

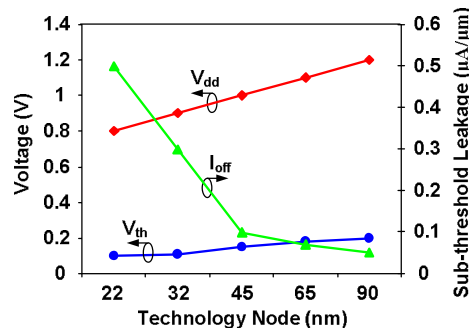

Figure 1. CMOS technology scaling trend and its impact on sub-threshold leakage current (Source: ITRS [4]).

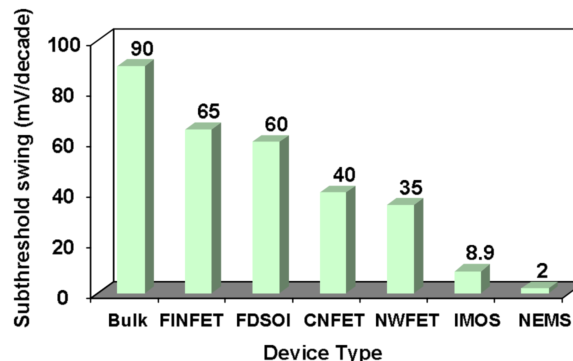

Figure 2. Minimum subthreshold swings reported in the literature for various emerging device structures. The value reported here for the carbon nanotube FET corresponds to the tunneling-CNFETs, which have the lowest subthreshold swing.

This fundamental limit for sub-threshold swing of CMOS devices, among other shortcomings, has directed researchers to explore potential alternatives for CMOS technology. Minimum reported subthreshold swings for various classical and non-classical devices are summarized in **Figure 2** [7]-[12]**.** It can be observed that all CMOS-based transistors (bulk, fully-depleted SOI (FDSOI) and FinFET) have subthreshold slopes higher than *60mV/decade*; whereas, tunneling carbon nano-tube transistors (T-CNFET), nano-wire based device (NW-FET) and impact-ionization based MOS (IMOS) show subthreshold slopes of *40*, *35*, and *8.9 mV/decade*, respectively. However, it has been experimentally shown that an electromechanical device exhibits astonishingly low sub-threshold slope of *2 mV/decade* [12]. The reason for extremely low leakage property of these devices will be discussed in Section 2.

Although NEMS devices have extremely low OFF current, a NEMS device does not offer as high ON current as CMOS transistors do. Therefore, for the first time, we propose the idea of *hybrid NEMS-CMOS technology* where near zero-leakage characteristics of NEMS can be combined with high ON current of CMOS transistors to simultaneously achieve ultra low-power and high performance operation. Hence, integrating NEMS devices into CMOS process can have significant implications for the IC industry.

To explore potential applications of hybrid NEMS-CMOS circuits, we have focused on three most critical applications for low power designs namely dynamic OR gate, SRAM cell, and sleep transistor design [2]. HSPICE simulations are carried out to compare NEMS-CMOS circuits against their pure-CMOS counterparts. The reported device model and characteristics in the literature have been used for

NEMS devices [13], and BSIM models [14] have been employed for CMOS transistors. Simulation results suggest that a hybrid dynamic OR gate can achieve *60-80%* lower switching power and near zero leakage power consumption with minor delay penalty and a hybrid SRAM cell can achieve *7.7X* lower standby leakage power consumption with only *14%* and *23%* penalty on noise margin and latency, respectively. Also, in the sleep transistor application, NEMS devices exhibit upto three orders of magnitude lower leakage current with minor performance degradation compared to the CMOS sleep transistors.

In summary, we have made two key contributions in this paper:

- Proposed the idea of hybrid NEMS-CMOS technology and a simplified process flow for its fabrication.
- Presented the design and analysis of low-power high-performance dynamic OR gate, SRAM cell and sleep transistors using hybrid NEMS-CMOS technology.

This paper is organized as follows: in Section 2 a brief introduction is provided on NEMS including its basic operation and modeling. Section 3 describes an integrated fabrication process for NEMS and CMOS devices. Section 4 includes proposed dynamic OR gate architecture and corresponding simulation results. In Section 5, a novel SRAM cell architecture is proposed and compared against existing works in literature. Section 6 discusses the application of NEMS devices as sleep transistors and finally concluding remarks are made in Section 7.

2. Nano-Electro-Mechanical Switches (NEMS)

A nano-electro-mechanical switch (NEMS), as suggested by its name, is a device where flow of current between source and drain terminals is controlled with both electrical as well as mechanical means. These devices are usually composed of a moveable beam which can deform in response to applied electrical bias; hence, aiding or blocking flow of electrical current.

2.1. Operation of a NEMFET Device

The first NEMFET device, in the form of a Suspended Gate MOSFET (SG-MOSFET), was proposed in [15] and then fabricated in [12]. A Scanning Electron Microscope (SEM) micrograph of the fabricated suspended gate MOS device is shown in **Figure 3**. In this device, the gate is physically suspended with four supporting arms (sides) over the channel area. That is, in the absence of any gate bias, an air-gap exists between the gate material and the gate oxide layer. Source and drain area are located at the top and bottom of the picture. When proper gate voltage is applied on the gate terminal, the suspended gate bends over from its original straight position, thereby touching the underlying oxide layer; thus attracting carriers of opposite charge and forming a conducting channel between source and drain in the underlying silicon substrate.

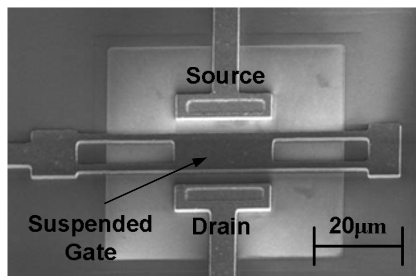

Figure 3. SEM picture of a SG-MOSFET switch with a 20μm long and 10μm wide suspended gate with four arms (Source: [12]).

Since the operation of SG-MOSFET involves both mechanical and electrical interactions, we explain the device behavior in ON and OFF states in more detail. **Figure 4** shows a cross-sectional view of the device in **Figure 3** as it can be seen from the side. In **Figure 4**, the suspended gate is connected to an anchor through a spring; however, it should be noted that there is no physical anchor or spring in the device structure and those concepts are solely used to model and demonstrate the movable suspended gate (which is essentially a bendable beam).

In **Figure 4(a)**, no bias voltage is applied to the device and the suspended gate is separated from the oxide layer by an air gap. Elastic force of suspended gate is in equilibrium with its weight force and there is no channel formed inside silicon (OFF state).

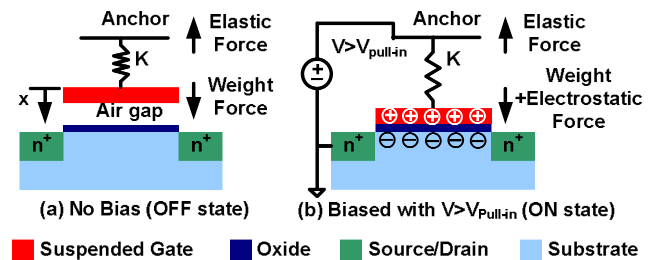

Figure 4. Basic operation of suspended gate NEMFET in ON and OFF states.

As we apply a positive gate voltage, shown in **Figure 4(b)**, positive charge appears on the suspended gate and a negatively charged layer is formed inside silicon. This creates an electrostatic force between opposite charges in the gate and silicon that causes the suspended gate to move downward in the x-direction. If the gate bias voltage is increased up to a threshold level called the pull-in voltage ($V_{pull\text{-}in}$) that is equivalent to the threshold voltage of standard CMOS transistors, the suspended gate ends up in a mechanically unstable equilibrium point. At this equilibrium point, sum of electrostatic and weight forces equals the maximum elastic force that can be supported by the spring or in other words, elasticity of suspended gate. Any further increase in the bias voltage beyond $V_{pull\text{-}in}$ pulls the suspended gate all the way down and the air gap vanishes. This situation is shown in **Figure 4(b)** where a conducting channel is formed in the Si-substrate and current can flow between source and drain (ON state).

2.2. Alternative NEMFET Implementation

An alternative approach for fabricating NEMS switches is based on cantilever- or carbon nanotube (CNT)- relays as shown in **Figure 5** [16]. The electro-mechanical principle of operation is similar; however, instead of a moveable gate, the conducting channel between source and drain is made moveable. In this structure, source terminal is connected to a conductive bendable cantilever (or a CNT), which is suspended over the gate terminal. In the off state, as shown in **Figure 5(a)**, there is no connection between the drain and source terminals. However, if sufficient voltage is applied between the gate and source, the cantilever can deform and make a connection between the source and drain terminals (**Figure 5(b)**).

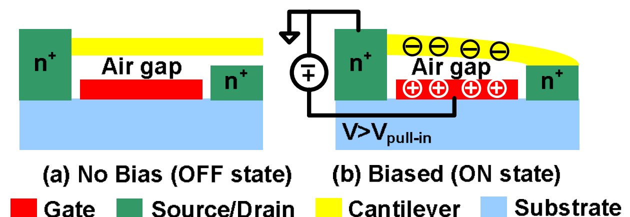

Figure 5. Cantilever or carbon nanotube based NEMFET in *ON* and *OFF* states.

2.3. Low Leakage Characteristics of NEMS

Unlike leakage current across source and drain terminals in solid state devices, the off-state leakage current in a NEMS device is determined only by the Brownian motion displacement current and vacuum tunneling currents [17]-[18]. Hence, if NEMS structures can be integrated with sub-65 nm CMOS, it can potentially revolutionize the IC industry by shifting the IC power trends to a new (and much lower) plane (much like what CMOS did to the BJT power trend).

While NEMS based on nanometer scale semiconductor structures (including Silicon, Silicon on Insulator, GaAs/AlGaAs systems, SiC on Si, Aluminum Nitride on Si as well as Carbon Nanotubes) have been reported [18]-[22], there is no report in the literature regarding co-design and integration of such NEMS structures with CMOS circuits and technology.

2.4. Modeling of NEMS Devices

Accurate modeling of NEMS devices is essential for performance comparison with existing CMOS technology. Precise modeling of NEMS devices is particularly complicated due to the interaction

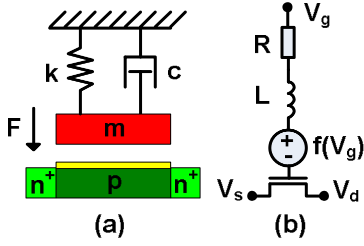

Figure 6. (a) Mechanical variables of a suspended-gate structure. (b) Equivalent electrical SPICE model of the suspended-gate device.

between electrical and mechanical components. However, using the electrical equivalents of mechanical variables, an entirely electrical SPICE model has been proposed in [23] for NEMS devices. In this model, each mechanical variable (shown in **Figure 6 (a)**) is replaced with its electrical equivalent, which leads to the SPICE model presented in **Figure 6 (b)**. In this model, the damping factor (c) of the suspended gate is modeled with a resistance (R) and mass of the moveable gate (m) is replaced with an inductance (L). Additionally, different forces (F), which pull the suspended gate toward the substrate, are modeled with a voltage-controlled source ($f(V_g)$). Since $f(V_g)$ is a complicated analytical function, a polynomial approximation is used through curve fitting [23]. According to this model, ON current of the NEMS device is lower than that of an identically-sized CMOS device because actual gate voltage for the NEMS transistor is smaller than V_g due to the voltage drop caused by $f(V_g)$ (**Figure 6 (b)**). This electrical model is then calibrated with reported data from actual NEMFET simulations [13]. The I_{ON} and I_{OFF} values, which were used for the calibration of the NEMS model, along with I_{ON} and I_{OFF} parameters of 90 nm CMOS devices are summarized in **Table 1**. In all simulations reported in the following sections, the 90 nm BSIM [14] models were used for the CMOS devices and the HSPICE model of **Figure 6 (b)** was used for the same-sized NEMS transistors.

Table 1. The I_{ON} and I_{OFF} values for NEMS and CMOS devices.

Device	I_{ON}	I_{OFF}
CMOS [4]	*1110 μA/μm*	*50 nA/μm*
NEMS [13]	*330 μA/μm*	*110 pA/μm*

3. Fabrication Feasibility of Hybrid NEMS-CMOS Circuits

A simplified fabrication process depicted in **Figure 7** is proposed to fabricate the necessary components of the NEMS switch along with CMOS transistors. In this figure, we combined fabrication steps of SG-MOSFET devices as reported in [12] with processing steps of the standard CMOS technology. The first step is to define polysilicon gates of CMOS devices (**Figure 7 (a)**) followed by a thermal oxidation and a nitride deposition to form an isolation bi-layer. Next step, shown in **Figure 7 (b),** is self-aligned definition of source and drain area for CMOS devices. It should be noted that since NEMFET employs a suspended gate, its active area can not be fabricated in a self-aligned fashion. Therefore, in the third step (**Figure 7 (c)**), phosphorous is implanted to define the source and drain regions of NEMS devices. Then the thick field oxide layer is formed (**Figure 7 (d)**).

The next step is growth of a sacrificial layer which finally will be removed for the gate release. Both cured polymide and polysilicon can serve as a sacrificial layer to form the gap between the gate oxide and the suspended gate **(Figure 7 (e))**. Polymide can be patterned through a photolithography process or by dry oxygen etching and the polysilicon uses a SF6 plasma etching. Both options require a two-step Chemical Mechanical Polishing (CMP) to flatten the device. It is important to note that the modeling study presented in [13] showed the need to form gaps of a few nanometers to optimize the V_{on} and V_{off} values. Hence, dry etching might be a more suitable alternative for obtaining nm-order gaps. High-quality poly-Si, the most commonly used material for surface micromachined MEMS compatible with CMOS technology, cannot be used for the sacrificial layer, since it requires high-temperature annealing for stress reduction and dopant activation [19].

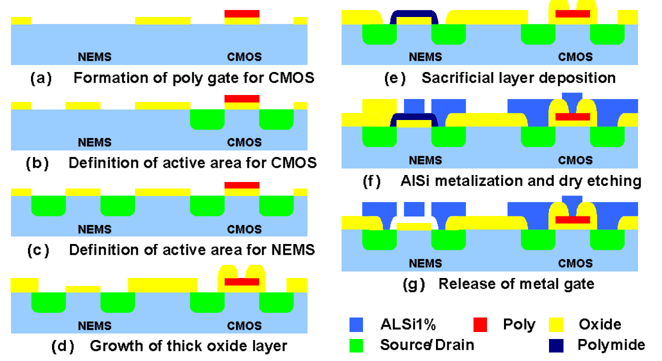

Figure 7. Simplified process flow for fabrication of hybrid NEMS-CMOS circuits.

Thus, after patterning any of these sacrificial materials, a layer of AlSi can be sputtered to build the mechanical component and then patterned with chlorine plasma etch (**Figure 7 (f)**). Finally, the structure can be released with an isotropic dry oxygen plasma (for polymide) or SF6 plasma (for polysilicon) (**Figure 7 (g)**).

4. Application of Hybrid NEMS-CMOS in Dynamic Circuits

Due to near zero leakage characteristics of NEMFETs, these devices are ideal for low-power applications. However, since NEMFET devices have relatively low ON current characteristics compared to their CMOS counterparts, it is more desirable to build hybrid NEMS-CMOS circuits to combine good characteristics of each type of devices. Therefore, for the first time, implementation of hybrid NEMS-CMOS circuits is proposed. In this section, we discuss limitations of pure CMOS dynamic OR gates. Next sub-section includes details of our proposed dynamic OR gate, which is then followed by HSPICE simulation results.

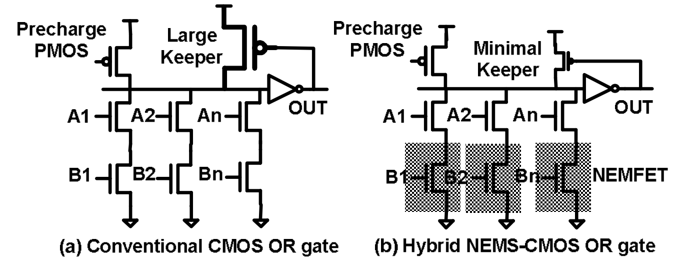

Figure 8. Conventional and proposed dynamic OR gates.

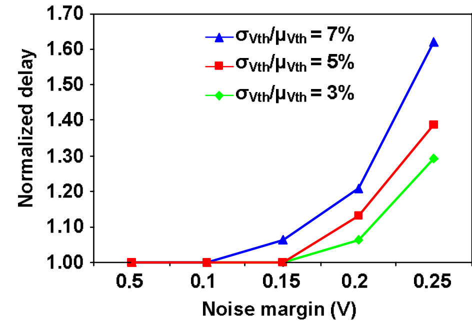

Figure 9. Normalized delay of an 8-input dynamic OR gate versus its noise margin under different levels of process variation.

4.1. Limitations of Pure CMOS Dynamic OR Gates

Dynamic implementation of wide fan-in OR-gates offers low latency, because it dose not require a PMOS transistor stack unlike their static CMOS counterparts. However, the major disadvantage of dynamic gates is their low noise margin, which conventionally is improved by employing a small PMOS keeper (**Figure 8 (a)**). But,

under increasing process variation and higher leakage levels, keeper size must be increased to meet the noise margin criterion, which on the other hand, degrades performance of dynamic gates.

The trade-off between noise margin and performance is shown graphically in **Figure 9** for an 8-input dynamic OR gate, where Y- and X-axis represent normalized worst-case delay and noise margin in volts, respectively. Three curves show different level of process variation measured in terms of standard deviation of threshold voltage (σ_{Vth}) as percentage of its nominal value (μ_{Vth}) [24].

4.2. Hybrid NEMS-CMOS Dynamic Gates

To combine low leakage characteristics of NEMFET devices and high ON current of MOSFET devices, a novel dynamic gate architecture is presented as shown in (**Figure 8 (b)**). In this architecture, NEMFETs are placed in series below the standard NMOS devices in the pull down network. Since NEMS devices are known to have much lower sub-threshold leakage than their CMOS counterparts, this will tremendously reduce leakage current of the pull down network. As a result, the size of the keeper can be made to be minimal. Since contention between keeper circuit and pull down network causes increased delay and power consumption, considerable improvement is expected due to minimization of keeper circuit. On the other hand, due to low ON current of NEMS devices, pull down network of this architecture poses higher ON resistance compared to the conventional all-CMOS pull down circuits. Hence, performance of the proposed gate is comparable to that of an equal-sized conventional dynamic gate.

A major advantage of the proposed circuit is its superior low power operation. Switching power consumption of the new gate is much lower than conventional dynamic gates due to minimal contention between keeper and pull down network. Moreover, leakage power consumption of the new gate is much smaller due to almost zero-leakage characteristics of NEMS switches.

4.3. Simulation Results

Proposed dynamic OR circuit was simulated using HSPICE to investigate its potential advantages over pure CMOS counterpart. In **Figure 10**, simulation results for two 8-input dynamic OR gates, one realized using only CMOS devices (**Figure 8 (a)**) and the other with hybrid NEMS-CMOS architecture (**Figure 8 (b)**), are shown. Switching power and the worst case delay of two gates are plotted on the two vertical Y-axes for different fan-out values (plotted on X-axis). It can be observed that hybrid NEMS-CMOS gate shows *10%* and *20%* higher delay when gate is loaded with fan-out of 1 and 5, respectively. However, hybrid NEMS-CMOS gate consumes up to *60%-80%* less switching power compared to CMOS gate at the same fan-out level.

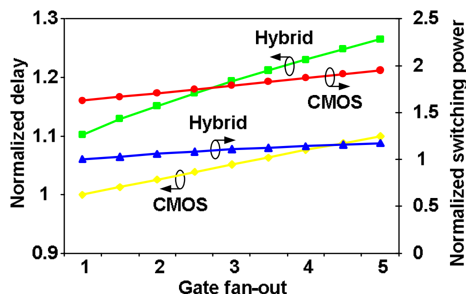

Figure 10. Normalized switching power consumption and the worst case delay of two 8-input hybrid NEMS-CMOS and CMOS dynamic OR gate for different fan-out levels. Switching power and delay values are normalized with respect to the power of the hybrid gate and delay of the CMOS gate with fan-out of one, respectively.

In **Figure 11**, simulation results are shown for the two dynamic OR gates discussed above with different number of inputs (fan-in). Similar to the previous figure, power consumption and delay of gates are plotted on the vertical axes; whereas, X-axis shows fan-in of the gates. For small gates (*4* and *8*-input) CMOS circuits exhibit lower delay; however, power consumption of proposed hybrid gate is much lower than the conventional dynamic gate. It is interesting to note that as we move toward larger fan-in (*12* and *16*-input), the hybrid gate outperforms its CMOS counterpart both in terms of delay and switching power consumption.

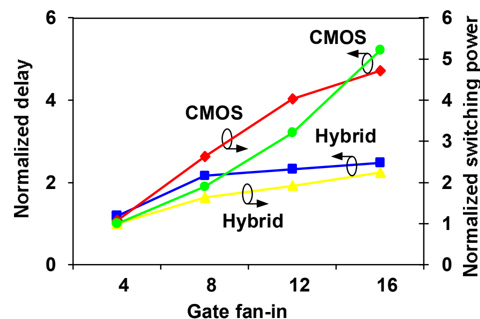

Figure 11. Normalized switching power consumption and the worst case delay of hybrid NEMS-CMOS and CMOS dynamic OR gate for different fan-in levels. Fan-out is kept constant at 3. Switching power values are normalized with respect to the power and delay value of the hybrid gate with fan-in of four.

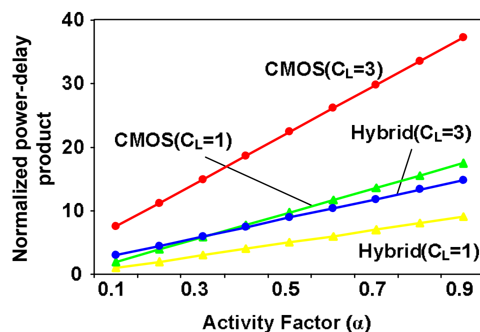

Figure 12. Power-delay product metric comparison for hybrid NEMS-CMOS and pure CMOS dynamic OR gates with different output capacitances.

To capture contributions of delay and power consumption on overall performance of these two types of gates, power-delay product metric was calculated. In order to incorporate impact of both idle state leakage and switching power consumption, the total power consumption of the circuits were evaluated using **Equation 1**. Here, α denotes the activity factor of the dynamic circuits, P_L denotes leakage power of each gate, P_S is the switching power consumption and D is the worst case delay of each gate.

$$P.D. = ((1-\alpha).P_L + \alpha.P_S).D \quad (1)$$

In **Figure 12**, the power-delay product metric is plotted for the two dynamic gates with two different output load values of $C_L=1$ and $C_L=3$, which denote fan-outs of *1* and *3*, respectively. In this figure, activity factor (α) is varied from *0* to *1,* which is plotted on the X-axis. As can be easily observed, the proposed hybrid architecture strongly surpasses the CMOS gate in terms of power-delay product in both cases.

5. Application of Hybrid NEMS-CMOS in SRAM Cell Design

SRAM cell design is known to be one of the most challenging tasks of low power designers [3]. In this section, we introduce a novel SRAM architecture based on the proposed NEMS-CMOS technology. Also, we compare the proposed SRAM cell against various existing low power SRAM cells and the conventional SRAM circuit.

5.1. Leakage Issue in Pure CMOS SRAM Cells

Performance of modern microprocessors is strongly dependent on the size of their on-chip cache memories. The cache memories are composed of arrays of SRAM cells (**Figure 13 (a)**). As technology scales down, more and more SRAM cells can be integrated in the same area. However, due to increased leakage current of transistors, the relative fraction of leakage power consumption (as compared to the fraction of switching power) is increasing. Moreover, the probability of read failures (toggling of stored value during read operation) and read latency (delay between accessing the cell and sensing voltage change on bit lines) also degrades with scaling.

Read stability degrades because current of the access transistor *AR* (in saturation region) increases at a higher pace than the current of the pull down transistor *NR* (in linear region) as shown in **Figure 13(a)**. Also, the higher leakage current of OFF access transistors (in other cells that are connected to the BLB), makes it tougher for the access transistors to create the necessary voltage difference for sense amplifiers. Therefore, read access delay is also adversely affected by leakage. Hence, using low-leakage devices can potentially lead to better performance and stability as well as lower power consumption [2].

5.2. Low-Leakage CMOS SRAM Cells

Low leakage SRAM cell architectures must be able to reduce the leakage power consumption; however, they also must have minimal impact on the read and write latency of cache memory. Recently, a dual-V_t SRAM cell has been proposed in [25] where both high- and low-V_t transistors were employed to reduce leakage current at the cost of lower cell stability and performance (**Figure 13 (b)**). Also, an asymmetrical SRAM cell architecture has been proposed using dual threshold voltage technology [26] (**Figure 13 (c)**). It is argued that since data stored in cache are more likely to be *zeros* than *ones*, high V_t devices can be used to reduce the leakage in *zero* storing state and low V_t transistors can be used for the *one* state.

It should be noted that all existing low-power SRAM cells achieve low-leakage characteristics with some noise margin and latency cost. In the following sub-section, we propose employing hybrid NEMS-CMOS structures and show that although the stability and performance loss in our proposed architecture is comparable to those of other approaches, superior low leakage characteristics can be achieved with hybrid cells (**Figure 13 (d)**).

5.3. Hybrid NEMS-CMOS SRAM Cells

In this section, a hybrid NEMS-CMOS SRAM cell is proposed as shown in **Figure 13 (d)**. Compared to the conventional cell, NMOS pull-down transistors (*NR* and *NL*) along with the PMOS pull-up devices (*PL* and *PR*) are also replaced with their NEMS counterparts. Since NEMS devices have very low OFF current, NEMS transistors draw minimal leakage current from power supplies; however, low ON current of NEMS devices degrade the stability of the cell against read failure errors. It should be noted that replacing access transistors (*AR* and *AL*) with NEMS devices is not a good idea because of their huge impact on latency.

Another alternative to the proposed architecture is to only replace the PMOS pull-up transistors (*PL* and *PR* in **Figure 13 (d))** with NEMS devices. Since, PMOS devices are OFF during the read operation; low ON current of PMOS NEMS devices does not affect the read latency. However, in this case, leakage power saving is not as much as that of the proposed version due to the leaky NMOS devices. Different trade-offs of employing hybrid NEMS-CMOS in SRAM cells is discussed in the following sub-section.

5.4. Simulations and Comparisons

Different SRAM cell structures of **Figure 13** are simulated using HSPICE. As in the previous section, for NEMS devices, I-V characteristics were taken from [13]. Three major metrics are used for comparison: (a) Static Noise Margin (SNM), (b) Read latency and (c) Standby leakage power consumption of SRAM cells. Four different SRAM cell architectures of **Figure 13** are examined. In the simulation results, we call architectures of **Figure 13 (a)**, **(b)**, **(c)** and **(d)** as "Conv.", "Dual V_t", "Asym.", and "Hybrid", respectively. In **Figure 14,** butterfly curves are shown for different SRAM cells and measured values for SNM are reported. It can be observed that the proposed hybrid architecture exhibits *14%* lower noise margin as compared to that of the conventional cell; however, its noise margin is slightly higher than those of the other two low-leakage SRAM cell architectures.

In **Figure 15,** the read latency and standby leakage current of various SRAM cells are compared. It should be noted that the read latency and leakage values are normalized to those values of the conventional SRAM cell for ease of comparison. Also, since the asymmetric cell shows different read latencies for zero and one values, an average value is used in this graph. As can be observed, all three low-leakage SRAM cells have higher latency compared to the conventional design. The proposed hybrid design has *23%* higher read latency. However, it is possible to further reduce the latency of the

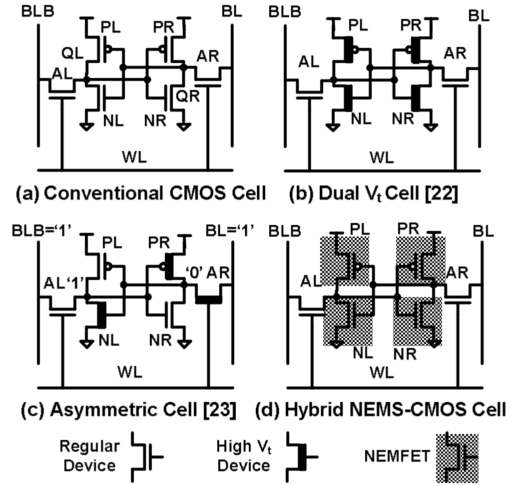

Figure 13. Different SRAM cell architectures in the literature: (a) conventional, (b) dual V_t [25], (c) asymmetric [26], and (d) proposed hybrid NEMS-CMOS SRAM cell. *AR*, *AL*, *PR*, *PL*, *NR*, and *NL* refer to different transistors; *QR* and *QL* refer to the output node of the inverter located on right and left side of the SRAM cell, respectively.

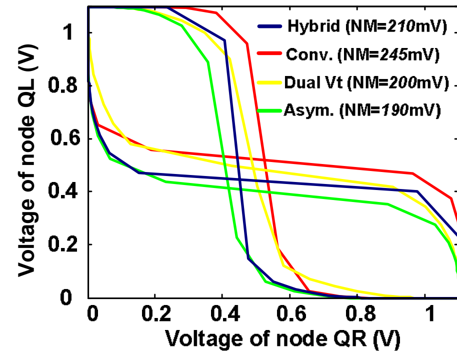

Figure 14. SRAM butterfly curves used for measuring SNM values for different SRAM cell architectures of Figure 13.

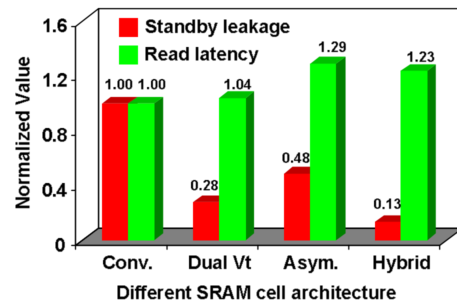

Figure 15. Comparison between conventional, existing low-leakage and proposed SRAM cell architectures.

hybrid cell via proper transistor and circuit optimization. Finally, from a leakage perspective, the hybrid SRAM cell outperforms all other circuits and has almost *8X* lower leakage power consumption compared to the conventional SRAM cell.

6. Application of NEMS Devices as Sleep Transistors

Low leakage characteristics of NEMS devices can also be utilized in designing more efficient sleep transistors. Sleep transistors refer to switches which are placed between the power supply and the circuit to reduce leakage current [27] (**Figure 16**). During the normal operation,

the sleep transistor is turned on to connect the circuit to power supply and it is turned off to reduce leakage current during the idol periods when circuit is not active. There are two different types of sleep transistors; header-type (**Figure 16(a)**), which is composed of a PMOS device and is placed between the actual V_{dd} and the circuit or, it can be a footer-type (**Figure 16(b)**) where an NMOS device is placed between the actual ground and the circuit. Sleep transistors can also be categorized as fine- and coarse-grain types. In the case of fine-grain sleep transistor (**Figure 16(c)**), each digital gate is separated from the power supply independently by one sleep device; however, in the coarse-grain implementation (**Figure 16(d)**), one sleep transistor is responsible for separation of a block of digital gates.

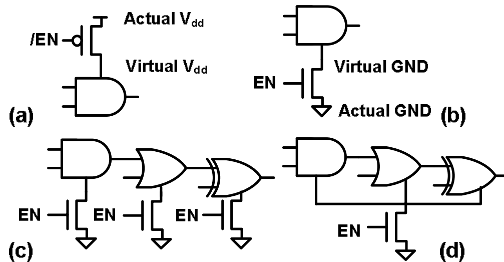

Figure 16. Different types of sleep transistors: (a) a header-type, (b) a footer-type, (c) fine-grain, and (d) coarse-grain.

Since NEMS devices demonstrate ultra low leakage characteristics, they are an ideal choice as sleep transistors. There are three important features of each sleep transistor [27]. First, it must have low sub-threshold leakage current; second, it must have low ON state resistance so that the voltage difference between virtual and actual power supply node (**Figure 16**) remains minimal. Third, its area should be preferably small compared to the rest of the circuit. A comparison between ON state resistance and OFF state leakage of NEMS and CMOS based sleep transistors for different device areas are shown in **Figure 17**. It can be observed that NEMS-based sleep device outperforms CMOS-based sleep transistor in terms of off-state leakage current. However, it can also be noticed that NEMS devices exhibit higher on resistance for equal area size, due to lower drain current of NEMS devices. It can be observed from **Figure 17** that for larger sleep transistors, the difference between the ON resistance of NEMS and CMOS sleep transistors become minimal. Hence, the NEMS sleep transistors can be sized up to offer up to three orders of magnitude lower leakage current with negligible performance degradation compared to that of their CMOS counterparts.

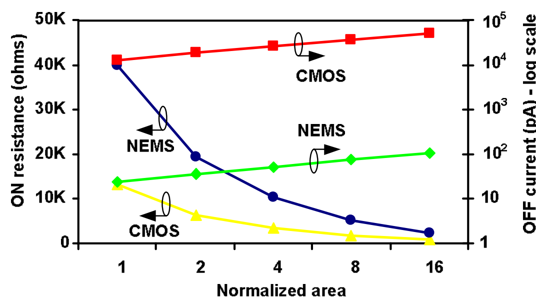

Figure 17. Comparison between on resistance and off current of NEMS and CMOS sleep transistors. All areas are normalized w.r.t the area of a CMOS device with W/L=5 in 90 nm technology.

7. Conclusions

Integration of nano-electro-mechanical switches (NEMS) with CMOS technology has been proposed for the first time, which combines near-zero leakage characteristics of NEMS switches with high ON current of CMOS transistors in order to achieve ultra low-power and high-performance operation. A simplified process flow has been outlined for integration of NEMS devices in CMOS technology. Possible low power applications of the proposed hybrid technology have been explored through design of hybrid NEMS-CMOS dynamic OR gates and SRAM cells at the 90 nm node. Simulation results indicate that such dynamic gates can achieve *60-80%* lower switching power consumption and almost zero leakage power consumption with minor delay penalty. Most importantly, the hybrid gate outperforms its CMOS counterpart both in terms of delay and switching power consumption with increase in fan-in beyond *12*. Additionally, it is shown that a hybrid SRAM cell can achieve nearly *8X* lower standby leakage power consumption with only minor noise margin and latency cost, which can be further reduced with adequate device/circuit optimization. Finally, as sleep transistors, NEMS devices offer up to three orders of magnitude lower OFF current with negligible performance degradation compared to their CMOS counterparts.

References

[1] G.E. Moore, "Cramming more components onto integrated circuits," Electronics, Vol. 38, pp. 114 - 117, 1965.
[2] V. De and S. Borkar, "Technology and design challenges for low power and high performance microprocessors," Proc. ISLPED, 1999, pp. 163 –168.
[3] S. Borkar, T. Karnik and V. De, "Design and reliability challenges in nanometer technologies," DAC, 2004, pp. 75.
[4] International Technology Roadmap for Semiconductors (ITRS), http://public.itrs.net.
[5] K. Banerjee, S-C. Lin, A. Keshavarzi, S. Narendra and V. De, "A self-consistent junction temperature estimation methodology for nanometer scale ICs with implications for performance and thermal management," Tech. Digest, IEDM, 2003, pp. 887-890.
[6] Y. Taur and T. Ning, Fundamentals of Modern VLSI Devices, Cambridge University Press, 1998.
[7] J. Appenzeller, Y. M. Lin, J. Knoch, and Ph. Avouris, "Band-to-Band Tunneling in Carbon Nanotube Field-Effect Transistors", Phys. Rev. Letter, Vol. 93, 196805, 2004.
[8] J. Appenzeller, L. Yu-Ming, J. Knoch, C. Zhihong, and P. Avouris, "Comparing carbon nanotube transistors - the ideal choice: a novel tunneling device design", IEEE Transactions on Electron Devices, Vol. 52, pp. 2568-2576, 2005.
[9] D. J. Wouters, J. P. Colinge, and H. E. Maes, "Subthreshold slope in thin-film SOI MOSFETs," IEEE Trans. Electron Devices, Vol. 37, pp. 2022–2033, 1990.
[10] H. C. Lin, M. H. Lee, C. J. Su, and S. W. Shen, "Fabrication and Characterization of Nanowire Transistors With Solid-Phase Crystallized Poly-Si Channels", IEEE Trans. Electron Devices, Vol. 53, pp. 2471-2477, 2006.
[11] K. Gopalakrishnan, P. B. Griffin and J. D. Plummer, "I-MOS: A Novel Semiconductor Device with a Subthreshold Slope Lower than kT/q," Tech. Digest, IEDM 2002, pp. 289-292.
[12] N. Abele et al., "Suspended-gate MOSFET: bringing new MEMS functionality into solid-state MOS transistor," IEDM Tech. Digest. 2005, pp. 479- 481.
[13] H. Kam et al., "A new nano-electro-mechanical field effect transistor (NEMFET) design for low-power electronics," IEDM Tech. Digest, 2005, pp. 463- 466.
[14] http://www-device.eecs.berkeley.edu/~ptm/mosfet.html.
[15] A. M. Ionescu, V. Pott, R. Fritschi, K. Banerjee, M. J. Declercq, Ph. Renaud, C. Hibert, Ph. Fluckiger and G-A. Racine, "Modeling and Design of a Low-Voltage SOI Suspended-Gate MOSFET (SG-MOSFET) with a Metal-Over-Gate-Architecture," IEEE International Symposium on Quality Electronic Design, 2002, pp. 496-501.
[16] J. M. Kinaret, T. Nord and S. Viefers, "A carbon nanotube based nanorelay," Applied Physics Letters, Vol. 82, pp. 1287-1289, 2003.
[17] T. Rueckes, K. Kim, E. Joselevich, G. Y. Tseng, C. L. Cheung, and C. M. Lieber, "Carbon Nanotube-Based Nonvolatile Random Access Memory for Molecular Computing", Science, Vol. 289, pp. 94-97, 2000.
[18] M. Dequesnes, S. V. Rotkin, N. R. Aluru, "Calculation of pull-in voltages for carbon nanotube-based nanoelectromechanical switches", Nanotechnology, Vol. 13, pp. 120-131, 2002.
[19] H. Takeuchi et al., "Thermal budget limits of quarter-micrometer foundry CMOS for post-processing MEMS devices," IEEE Trans. on Electron Devices, Vol. 52, no.9, pp. 2081- 2086, 2005.
[20] K. L. Ekinci and M. L. Roukes, "Nanoelectromechanical systems," Review of Scientific Instruments, 76, 2005.
[21] J. E. Jang et al., "Nanoelectromechanical switches with vertically aligned carbon nanotubes," Applied Physics Letters, 87, 163114, 2005.
[22] N. Abele et al., "Ultra-Low Voltage MEMS Resonator Based on RSG-MOSFET," 19th IEEE International Conference on MEMS, 2006, pp. 882- 885.
[23] V. Pott, A. Ionescu, R. Fritschi, C. Hibert, P. Fluckiger, G. Racine, M. Declercq, P. Renaud, A. Rusu, D. Dobrescu, and L. Dobrescu, "The suspended-gate MOSFET (SG-MOSFET): A modeling outlook for the design of RF MEMS switches and tunable capacitors," International Semiconductor Conference, 2001, pp. 137-140.
[24] H. F. Dadgour, R. V. Joshi and K. Banerjee, "A Novel Variation-Aware Low-Power Keeper Architecture for Wide Fan-in Dynamic Gates," DAC 2006, pp. 991-996.
[25] F. Hamzaoglu, Y. Ye, A. Keshavarzi, K. Zhang, S. Narendra, S. Borkar, M. Stan, and V. De, "Dual Vt-SRAM cells with full-swing single-ended bit line sensing for high-performance on-chip cache in 0.13µm technology generation," Proc. IEEE International Symposium on Low Power Electronics and Design, 2000, pp. 15–19.
[26] N. Azizi, F. N. Najm and A. Moshovos, "Low-leakage asymmetric-cell SRAM", IEEE Trans. on Very Large Scale Integration Systems, Vol. 11, pp. 701-715, 2003.
[27] S. Kaijian and D. Howard, "Challenges in sleep transistor design and implementation in low-power designs," DAC 2006, pp. 113-116.