Characterization of Single-Electron Tunneling Transistors for Designing Low-Power Embedded Systems

Changyun Zhu, Student Member, IEEE, Zhengyu (Peter) Gu, Student Member, IEEE, Robert P. Dick, Member, IEEE, Li Shang, Member, IEEE, Robert Knobel, Member, IEEE

Abstract—Minimizing power consumption is vitally important in embedded system design; power consumption determines battery lifespan. Ultra-low-power designs may even permit embedded systems to operate without batteries by scavenging energy from the environment. Moreover, managing power dissipation is now a key factor in integrated circuit packaging and cooling. As a result, embedded system price, size, weight, and reliability are all strongly dependent on power dissipation.

Recent developments in nanoscale devices open new alternatives for low-power embedded system design. Among these, single-electron tunneling transistors (SETs) hold the promise of achieving the lowest power consumption. Unfortunately, most analysis of SETs has focused on single devices instead of architectures, making it difficult to determine whether they are appropriate for low-power embedded systems.

Evaluating the use of SETs in large-scale digital systems requires novel architectural and circuit design. SET-based design imposes numerous challenges resulting from low driving strength, relatively large static power consumption, and the presence of reliability problems resulting from random background charge effects. We propose a fault-tolerant, hybrid SET/C莫斯, reconfigurable architecture, named IceFlex, that can be tailored to specific requirements and allows trade-offs among power consumption, performance requirements, operation temperature, fabrication cost, and reliability. Using IceFlex as a testbed, we characterize the benefits and limitations of SETs in embedded system designs. In particular, we focus on the use of SETs in room-temperature ultra-low-power embedded systems such as wireless sensor network nodes. We also consider high-performance applications such as multimedia consumer electronics. We see this work as a first step in determining the potential of ultra-low-power embedded system design using SETs.

Index Terms—Single-electron tunneling transistor, Low-power design, Reconfigurable architecture, Embedded system, High-performance computing, Nanotechnology, Reliability.

I. INTRODUCTION

Energy consumption and thermal issues are now central issues in electronic system design. In high-performance applications, temperature affects integration density, performance, reliability, power consumption, and cost. For battery-powered embedded systems, power consumption determines system life time. Power consumption crises were historically solved by moving to new technologies that decreased energy per operation, allowing increases in density and eventually performance. Power and thermal concerns were primary motivations for replacing vacuum tubes with semiconductor devices in the 1960s and replacing bipolar junction transistors with CMOS in the 1990s. Although CMOS is the mainstream fabrication technology used today, as IC and system integration further increase, it will reach fabrication, power consumption, and thermal limits; it may soon be time for another transition to a dramatically different technology.

Device researchers have seen the coming challenges for CMOS devices and evaluated alternative technologies such as carbon nanotube transistors [1], nanowires [2], and single-electron tunneling transistors (SETs) [3]. The International Technology Roadmap for Semiconductors projects that SETs have the potential to achieve the lowest projected energy per switching event of any known device (1 × 10⁻¹⁸ J) [4]. However, their use poses unique architectural, circuit design, and fabrication challenges. For example, SETs are susceptible to reliability problems caused by random background offset charges. They have cyclic I–V curves (see Figure 2) that can complicate design but permit highly-efficient implementation of some useful logic functions that have proven inefficient using CMOS and threshold logic. Although the fabrication of SETs capable of operating at low temperatures is now common, feature sizes of only a few nanometers are required for room-temperature operation, making fabrication challenging.

A. Past Work

After their discovery in the 1980s [5, 6], there has been extensive research on fabrication, design, and modeling of SETs [3]. SET fabrication and use in high-sensitivity amplifiers at cryogenic temperatures has been the main research focus [7]. SETs and simple circuits with a variety of structures were proposed and fabricated using different methods and materials [8–10]. Recently, researchers have fabricated SETs that operate at room-temperature [11–13]. Various SET-based circuit applications, such as logic [14–17] and memory [18–20] have been developed. This work provides a promising start for SET circuit design. However, these articles did not provide an architectural evaluation. We do not claim to have improved

C. Zhu is with the Department of Electrical and Computer Engineering, Queen’s University, Kingston, ON K7L 3N6, Canada.
Z. Gu is with Synplicity Inc, Sunnyvale, CA 94086, U.S.A.
R. Dick is with the Department of Electrical Engineering and Computer Science, Northwestern University, Evanston, IL 60208, U.S.A.
L. Shang is with the Department of Electrical and Computer Engineering, University of Colorado at Boulder, Boulder, CO 80305, U.S.A.
R. Knobel is with the Department of Physics, Queen’s University, Kingston, ON K7L 3N6, Canada.
This work is supported in part by NSERC Discovery Grants #388694-01 and #298185-04, in part by the NSF under awards CNS-0347941, CNS-0721978, and CCF-0702761, and in part by the SRC under award 2007-HF-1593.

Acknowledgment

We thank the anonymous reviewers for their comments which improved this paper.
the performance of SET-based logic gates. Instead, we are the first to develop the modules necessary to support architectural design and synthesis and evaluate the architectural performance and power consumption implications of using SETs. They demonstrate orders of magnitude improvement in power consumption and energy efficiency compared to CMOS.

Research on SET modeling and simulation has been an active area. Monte Carlo simulation has been widely used to model SETs. SIMON [21] and MOSES [22] are the two most popular SET simulators. However, they are too slow for analysis of large circuits. Uchida et al. proposed an analytical SET model and incorporated it into SPICE [23]. Recently, Inokawa et al. extended this model to a more general form to include asymmetric SETs [24]. Mahapatra et al. propose a simulation framework for hybrid SET/CMOS circuit design and analysis [25]. Their model for SET behavior is similar to that of Uchida et al. These compact modeling techniques are efficient enough for use in SET circuit design and analysis and closely match Monte Carlo simulation results.

Significant challenges still remain for large-scale integration of SETs and for room-temperature operation. SETs that operate reliably at room temperature have critical dimensions of ∼1–10 nm. They are challenging to fabricate using current top-down lithographic techniques. However, several exciting advances make the evaluation of architectures for high-density logic based on SETs worthwhile. Scanning-probe microscopes can be used to create devices smaller than those using conventional lithography [11]. Continual progress has been made on bottom-up nano-fabrication techniques, where chemical techniques are used to make individual molecules with useful electronic properties. Molecular quantum dots [26] can display SET behavior. Larger structures, such as carbon nanotubes and nanowires, can act as SETs [10]. These bottom-up techniques can create structures supporting room-temperature SET operation. However, more research is needed in order to integrate individual devices into large-scale circuits. Very recent advances in graphene [27] devices show promise for SETs. Reliable methods for cooling to very low temperatures without supplies of liquid helium or nitrogen are also becoming more common [28]. For high-performance computing, the added complexity of operating at cryogenic temperatures may not be a limiting factor. Similarly, cryogenic temperatures are readily attained using passive methods in outer space.

B. Contributions

In this paper, we explore the potential use of SETs in low-power embedded systems. In order to take advantage of the power efficiency of SETs, it is critical to bring SET-based design to the system level, characterize the impacts of SETs on system design metrics, and evaluate the benefits and limitations of SETs. Our work starts from design space characterization of SET-based architectures. We evaluate the impacts of using SETs upon architectural, circuit-level, and device-level design, considering metrics such as energy efficiency, performance, reliability, maximum operating temperature, and ease of fabrication.

Based on our evaluation of the architectural and circuit-level features that can most effectively exploit the strengths of SETs while working within their limitations, we propose a fault-tolerant, reconfigurable, hybrid SET/CMOS based architecture called IceFlex. IceFlex is regular and cell-based. It is reconfigurable, permitting compensation for fabrication defects. It incorporates flexible, modular circuits to enable tolerance of run-time faults. In addition to compensating for the weaknesses of SETs, IceFlex exploits their strengths, e.g., we develop a two-SET design to implement Boolean functions that are not linearly separable.

We tailor IceFlex to both high-performance and battery-powered embedded systems and characterize its energy efficiency, performance, and power consumption by using it for a number of instruction processors and application-specific cores. Compared to CMOS-based designs, IceFlex improves energy efficiency by two orders of magnitude for both battery-powered and high-performance applications, while maintaining good performance. However, our results also indicate great challenges to the use of SET-based designs in portable embedded systems. Their use will either require advances in the compact cooling technologies or the fabrication of features with sizes approaching physical limits.

II. SET MODELING

In this section, we introduce the physical properties of SETs, and discuss SET analytical device modeling.

A. SET Basics

The operation of a single-electron tunneling device is governed by the Coulomb charging effect. As shown in Figure 1 a single-electron tunneling device consists of a nanometer-scale conductive island embedded in an insulating material. Electrons travel between the island, source (S), and drain (D) through thin insulating tunnel junctions. When an electron tunnels into the island, the overall electrostatic potential of the island increases by $e^2/C_S$, where $e$ is the elementary charge and $C_S$ is the island capacitance. For large devices, this change in potential is negligible due to the high island capacitance $C_S$. However, for nanometer-scale islands, $C_S$ is much smaller. As a result, the electrostatic energy change due to the addition or removal of a single electron can be larger than the thermal energy, particularly at low temperatures.

Changes to SET island potential results in an energy gap at the Fermi energy, preventing further electron tunneling. This phenomenon is called Coulomb blockade. It prevents current from flowing between source and drain ($I_{ds} = 0$), i.e., the SET is turned off. The Coulomb blockade effect can be overcome.
by changing the voltage of a conductor capacitively coupled to the island, thereby turning tunneling on and off. Although their transfer functions differ significantly from those of CMOS transistors, with careful circuit design, SETs can be used to realize logic functions using circuits analogous to CMOS, or using radically different design techniques [3].

As shown in Figure 1, a SET typically has four terminals. The source and drain terminals (S, D) serve as electron reservoirs. When the SET is turned on, electrons tunnel from one terminal, through the junction, to the conductive island. They then tunnel through the other junction to the other terminal, through the junction, to the conductive island. Each tunneling junction is modeled as a resistor $R_j$. They then tunnel through the other junction to the other terminal. The source and drain terminals (S, D) serve as electron reservoirs. When the SET is turned on, electrons tunnel from one terminal, through the junction, to the conductive island. They then tunnel through the other junction to the other terminal.


![Fig. 2. SET Coulomb oscillation ($C_g = 3.2 \text{ aF}$, $C_s = C_d = 1.0 \text{ aF}$, and $R_s = R_d = 10 \text{ MΩ}$).](image)

### Table I

**ISLAND SIZE ESTIMATION**

<table>
<thead>
<tr>
<th>Temperature (K)</th>
<th>$C_g = e^2/(10k_BT)$</th>
<th>Island capacitance (aF)</th>
<th>Island diameter (nm)</th>
<th>$C_g = e^2/(40k_BT)$</th>
<th>Island capacitance (aF)</th>
<th>Island diameter (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>40</td>
<td>4.65</td>
<td>52.48</td>
<td>1.16</td>
<td>13.12</td>
<td>4.65</td>
<td>52.48</td>
</tr>
<tr>
<td>77</td>
<td>2.41</td>
<td>27.26</td>
<td>0.60</td>
<td>6.82</td>
<td>2.41</td>
<td>27.26</td>
</tr>
<tr>
<td>103</td>
<td>1.80</td>
<td>20.38</td>
<td>0.45</td>
<td>5.10</td>
<td>1.80</td>
<td>20.38</td>
</tr>
<tr>
<td>120</td>
<td>1.55</td>
<td>17.49</td>
<td>0.39</td>
<td>4.37</td>
<td>1.55</td>
<td>17.49</td>
</tr>
<tr>
<td>200</td>
<td>0.93</td>
<td>10.50</td>
<td>0.23</td>
<td>2.62</td>
<td>0.93</td>
<td>10.50</td>
</tr>
<tr>
<td>250</td>
<td>0.74</td>
<td>8.40</td>
<td>0.19</td>
<td>2.10</td>
<td>0.74</td>
<td>8.40</td>
</tr>
<tr>
<td>300</td>
<td>0.62</td>
<td>7.00</td>
<td>0.15</td>
<td>1.75</td>
<td>0.62</td>
<td>7.00</td>
</tr>
</tbody>
</table>

Assuming disc capacitor model ($C_g = 8\epsilon r$). One side of island embedded in silicon dioxide. Other side exposed to Nitrogen.

In order to observe the Coulomb blockade effect, the following constraints must be satisfied.

1) Since thermal fluctuations can suppress the Coulomb Blockade effect, the electrostatic charging energy, $e^2/C_g$, must be much greater than $k_BT$, where $k_B$ is Boltzmann’s constant and $T$ is the temperature. In order to ensure reliability, $e^2/C_g > k_BT$ or the more conservative $e^2/C_g > 40k_BT$ constraint is enforced. These equations imply that the maximum allowed island capacitance is inversely proportional to temperature. At room temperature, an island capacitance below $1 \text{ aF}$ is required. Island capacitance is a function of island size. As shown in Table I, room-temperature operation requires an island size in nanometer range, making fabrication challenging. At present, the smallest island capacitance of a fabricated device is around $0.15 \text{ aF}$ [12].

2) To observe single-electron charging effects, electrons must be confined to the island, which requires that the junction resistance be higher than the quantum resistance, i.e., $R_S R_D > h/e^2$, $h/e^2 = 25.8 \text{ kΩ}$, where $h$ is Planck’s constant. Therefore, SETs have high resistances and low driving currents.

In order to operate voltage-state logic, SETs must exhibit voltage gain. The low-temperature voltage gain is equal to the gate capacitance divided by the sum of the junction capacitances: $G = C_G/(C_S + C_D)$. Achieving this gain requires low tunneling junction capacitances. It also requires close coupling of gate and island without a large increase in the total island capacitance. High gain has only been demonstrated for a few devices and has required operation at low temperatures [30, 31]. However, further advances in nanofabrication may overcome this limitation.

#### B. Random Background Charge Effects

Constant background charge effects have been a persistent problem for SETs. Charges near the SET island influence its equilibrium state [32]. Although the resulting voltage offsets can be compensated for with a biased second gate terminal, the required bias is unknown until fabrication. Worse yet, some devices are affected by random background charge effects, which result in run-time voltage fluctuations.

It is the tentative consensus of the research community that random background charge effects are caused by multiple, closely-spaced charge traps near the island, among which charge carriers tunnel. This produces run-time variation in gate bias, and may cause logic errors. Much work has been done to understand the nature and density of these defects [33-35]. Most SETs have been fabricated with aluminum islands. Some researchers have attempted to eliminate random background charge effects by fabricating SETs with alternative island materials such as silicon, based on the thesis that the use of non-crystalline, non-stoichiometric aluminum oxide junctions in conventional SETs leads to numerous charge-trapping defects. Silicon island based devices have high immunity to random background charge noise, with operation unchanged over several weeks [36]. However, random background charge effects remain the main source of run-time reliability problems for most SET designs. In this work, we describe a reconfigurable...
architecture that provides architectural resistance to the effects of random background charges.

C. SET Modeling

Circuit design involves extensive simulation. Despite their accuracy, Monte Carlo methods are too slow for large-scale circuit analysis. We build upon the SET analytical model developed by Inokawa et al. [24], which has been incorporated into SPICE. Combined with MOS transistor models, it provides an efficient and accurate simulation solution for hybrid SET/CMOS circuits. Inokawa’s model ignores random background charge effects and multi-gate effects. We incorporate these effects into the model.

The I–V characteristics of a SET with island charge equal to \( n \) or \( n + 1 \) electrons follow:

\[
I_{DS} = \frac{e}{4R_Cg_S} \times \frac{(1 - r^2)(V_{GS}^2 - \tilde{V}_{DS}^2) \sinh(\tilde{V}_{DS}/T)}{\sinh(V_{GS}/T) - (\tilde{V}_{DS} + rV_{GS}) \sinh(\tilde{V}_{DS}/T)}
\]

where

\[
\tilde{V}_{GS} = \frac{2\sum C_G V_{GS}}{e}, \quad \tilde{T} = \frac{2kBTC_G}{e^2} \quad (3)
\]

\[
r = \frac{R_D - R_S}{R_D + R_S}, \quad RT = \frac{2}{\frac{1}{R_S} + \frac{1}{R_D}} \quad (4)
\]

\[
C_S = C_D + \sum C_G_i \quad (5)
\]

In this model, \( \frac{\sum C_G_i V_{GS}}{e} \) models the Coulomb charging effects of the multiple gate terminals. \( \zeta \) is a real number that characterizes the random background charge effect.

This compact model is derived based on the steady-state master equation, which is not directly applicable to transient circuit analysis. However, when used in circuits, SETs are connected by metal wires. Based on existing fabrication processes, the capacitance of local interconnect is at least two orders of magnitude higher than SET island capacitance, thereby eliminating inter-SET Coulomb interaction. The independence of SETs enables the use of quasi-steady-state analysis [24, 37].

III. IceFlex: A Fault-Tolerant Hybrid SET/CMOS Reconfigurable Architecture

This section describes the design and analysis of IceFlex, the proposed low-power, fault-tolerant, reconfigurable, hybrid SET/CMOS architecture. The vast majority of devices in IceFlex are SETs, allowing extremely low power consumption. CMOS devices are sparingly used to improve the driving strength of global interconnect.

Our evaluation of the architectural constraints imposed by SETs led to four main conclusions.

1) Flawless fabrication will be challenging, especially for circuits that operate at room temperature. It is important to simplify fabrication and use post-fabrication adaptation to improve reliability.
2) An unpredictable subset of devices will be susceptible to random background offset charge effect noise: SET-based architectures should have the ability to tolerate run-time errors.
3) SETs have poor driving strength; this must be remedied, especially when driving global interconnect.
4) SETs have the ability to efficiently implement some functions that are inefficient using BJTs, CMOS logic, or threshold logic, e.g., non-linearly-separable functions. SET-based architectures should exploit such special properties.

A. SET Design Space Characterization

In order to characterize the benefits and limitations of SET circuits and architectures, we analyze the tradeoffs among the following metrics: temperature, performance, power consumption, reliability, and fabrication constraints. This study yields two design configurations, each of which is shown in Table II. One targets high-performance embedded applications such as multimedia consumer electronics and one targets ultra-low-power embedded applications such as wireless sensor networks.

1) Temperature: IceFlex was evaluated at seven temperature settings (see Table II). IceFlex is a hybrid SET/CMOS design; the temperature range starts at 40 K to permit reliable operation of the CMOS components. 77 K is achieved by liquid nitrogen cooling. 103 K is the average cloud top temperature. 120 K and below are defined as cryogenic. At 200 K, functional SET devices have been widely demonstrated in the literature. 250 K is a temperature that might be reached using a stacked Peltier heat pump. 300 K is room temperature.

2) Capacitance: To observe well-defined Coulomb blockade effects, electron charging energy must be higher than the thermal energy, i.e., \( \frac{e^2}{C_S} \geq 10k_BT \) or \( \frac{e^2}{C_D} \geq 40k_BT \), where \( k_B \) is Boltzmann’s constant and \( T \) is temperature. At room temperature, this constraint requires an island capacitance below 1 aF, making fabrication challenging but possible [12]. In order to operate voltage-state logic, SETs must exhibit voltage gain, which is equal to the gate capacitance divided by the sum of the junction capacitances: \( G = C_G/(C_S + C_D) \). Our results indicate that a gain of 1.5 is sufficient for use in digital logic. Targeting battery-powered systems, using \( C_S \leq e^2/(10k_BT) \), \( C_D \leq e^2/(40k_BT) \) and \( G = 1.5 \), the maximum allowed gate and junction capacitances are derived and shown in the “Low power, Capacitance” columns of Table II.

The maximal allowed capacitance decreases with increasing temperature. However, fabricating SETs with low gate capacitance is challenging. We assume the capacitances at 300 K are the minimum allowed. Given \( \frac{e^2}{C_S} \geq 10k_BT \), for high-performance applications, these minimal gate and junction capacitances are used at all the temperature settings and shown in the corresponding “High Performance, Capacitance” columns of Table II. Given \( \frac{e^2}{C_D} \geq 40k_BT \), which requires very low SET
capacitance at room temperature, $C_G = 0.09 \text{ aF}$. This makes fabrication very challenging. Due to fabrication concerns, for high-performance design, the capacitance and voltage are determined at the appropriate operation temperature, instead of room temperature.

3) Voltage: Consider a SET biased via a second gate, such that a $V_{GS}$ of zero places it in the middle of the positive voltage coefficient (PVC) region in Figure 2. In this case, the maximum range of current values can be traversed by letting $V_{GS}$ (i.e., $V_m$) vary in the range $[-e/(4C_G), e/(4C_G)]$. At all but the lowest temperatures, this range also provides near-optimal sensitivity to $V_{GS}$; we use this range. Once the range of $V_{GS}$ is known, a $V_{SS}$ of $-e/(4C_G)$ and a $V_{DD}$ of $e/(4C_G)$ naturally follow, shown in the “Voltage” columns of Table II. Note that a bias voltage applied via a second gate can be used to shift the zero $V_{GS}$ point from the PVC to negative voltage coefficient (NVC) region in Figure 2, permitting NMOS-like or PMOS-like behavior.

4) Junction Resistance: To observe single-electron charging effects, electrons must be confined to the island. This requires junction resistances that are much higher than the quantum resistance, i.e., $R_s, R_D \gg h/e^2$, $h/e^2 = 25.8k\Omega$, where $h$ is Planck’s constant. Therefore, SETs have high resistances and low driving currents. In this work, we pick two resistance settings: 100KΩ for high-performance applications and 10 MΩ for battery-powered systems, shown in the “Resist.” columns of Table II.

5) Reliability Implications: Researchers have pointed out the dangers posed by thermal noise as charging (state change) energy approaches thermal energy. We explicitly consider the effects of temperature on steady-state current during circuit analysis and its effects are reflected in our design decisions. We implicitly consider, and guard against, the effects of temperature-dependent shot noise by requiring charging energy to be a large multiple of the thermal energy. Designs with charging energies of both 10 and 40 times the thermal energy are evaluated in this paper (10$k_BT$ or 40$k_BT$). Researchers have reported device operation at each level but the $40k_BT$ requirement is more reliable. At charging energies over $10k_BT$, the model we use is accurate to within 4% of the time-dependent master equation [23, 38].

Random background charge effects [34, 35] are the main barrier to SET reliability. They are observed as $1/f$ noise on SET gate voltages, with some SETs susceptible and others immune. Several recent devices have shown improved immunity to this noise, as described in Section II-B. Currently, the distribution of random background offset charges can only be determined after fabrication [3]. Susceptible SETs may suffer transient errors infrequently, e.g., only once per day. In this work, we use architectural techniques to reduce the probability of failure using an entirely SET-based design. SETs are used in parallel to exploit the lack of SET-to-SET correlation in random background offset charge effects.

**B. IceFlex Design**

In this section, we present the architecture and circuit design of IceFlex. The microarchitecture of IceFlex is shown in Figure 3. IceFlex is a cell-based design. Each cell is a SET logic block (SELB) composed of the following components: (1) multi-gate SET-based reconfigurable look-up tables that can realize arbitrary $n$-input Boolean functions; (2) a SET-based arithmetic unit that allows efficient implementations of non-linearly separable arithmetic operations; (3) a SET-based reconfiguration memory array that caches multiple configuration contexts to support efficient run-time reconfiguration; (4) a multi-gate SET-based input switch fabric; and (5) SET registers. In addition, IceFlex includes SET threshold logic-based majority voting logic units, allowing a flexible solution to run-time reliability problems. In IceFlex, a multi-level on-chip interconnect fabric forms inter-SELB connections. Local connections rely on a custom-designed, SET-driven, variable-length, constant-latency interconnect. Using a constant-latency interconnect structure reduces power consumption and simplifies physical-level design automation, e.g., placement and routing. SETs have limited driving strength. Therefore, IceFlex uses hybrid SET/CMOS circuits to drive global interconnects.

We now explain each IceFlex component and discuss both circuit and architecture tradeoffs.

1) Multi-Gate SET Reconfigurable Lookup Table Component: Each SELB is equipped with $l$ sets of $n$-input reconfigurable look-up tables. Each look-up table can realize an arbitrary $n$-input Boolean function. The basic structure of the look-up table consists of an $m$-to-1 multi-gate SET multiplexer tree ($m = 2^n$), and an $m$-bit SET storage cell, which will be described in the next section.

The proposed multi-gate SET multiplexer tree differs from existing CMOS-based designs in the following way. A CMOS $m$-to-1 multiplexer tree requires $\lceil \log_2 m \rceil$ stages of transmission gates, plus buffers to meet the required driving strength. SETs may have multiple gate terminals. As described in Equation 5, the gate charging effect is a function of $\sum C_i V_{GS_i}$. Therefore, multiple control signals, e.g., the select signals for a multiplexer, can be supplied to a single SET, enabling a more compact circuit structure with better performance and power efficiency.

Figure 4 shows the proposed SET multi-gate multiplexer tree design. The basic building block is a $q$-to-1 multi-gate single-stage multiplexer, in which each of the $q$ paths consists of a single multi-gate SET controlled by $\lceil \log_2 q \rceil$ select signals. Using this design, the logic depth of a $n$-to-1 multiplexer tree reduces to $\lceil \log_2 m \rceil$ instead of $\lceil \log_2 m \rceil$. Figure 4 also shows a design case for $q = 4$. The output SET buffer is used to break long resistive path and improve the driving strength.

As described in Section II, thermal energy has significant impact on electron tunneling and the ratio of on to off currents, i.e., the ratio of the off to on resistance. This ratio decreases as the ratio of Coulomb charging energy ($e^2/(C)$) to thermal energy ($k_BT$) decreases. On the other hand, as the number of gate control signals per SET (hence the number of off paths connected in parallel) increases, the impact of the off paths on the circuit output increases. Consider, for the sake of example, the dual-gate 4-to-1 multiplexer design shown in Figure 4. The four logic inputs are 0001 and both select signals are logic one, i.e., $V_a = V_b = V$. Assume $C_a = C_b = C$. As shown in
TABLE II
DESIGN SPACE CHARACTERIZATION

<table>
<thead>
<tr>
<th>Temp. (K)</th>
<th>Low power</th>
<th>High performance</th>
<th>Low power</th>
<th>High performance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Capacitance (aF)</td>
<td>Resist. (MΩ)</td>
<td>Voltage (mV)</td>
<td>Resist. (kΩ)</td>
</tr>
<tr>
<td>40</td>
<td>2.78</td>
<td>0.93</td>
<td>14.36</td>
<td>10</td>
</tr>
<tr>
<td>77</td>
<td>1.45</td>
<td>0.48</td>
<td>27.65</td>
<td>10</td>
</tr>
<tr>
<td>103</td>
<td>1.08</td>
<td>0.36</td>
<td>36.99</td>
<td>10</td>
</tr>
<tr>
<td>200</td>
<td>0.56</td>
<td>0.19</td>
<td>71.82</td>
<td>10</td>
</tr>
<tr>
<td>250</td>
<td>0.45</td>
<td>0.15</td>
<td>89.77</td>
<td>10</td>
</tr>
<tr>
<td>300</td>
<td>0.37</td>
<td>0.12</td>
<td>107.70</td>
<td>10</td>
</tr>
</tbody>
</table>

Fig. 3. IceFlex microarchitecture.

Fig. 4. Multi-gate SET multiplexer tree.

the I–V curve on the right side of Figure 4 for the SET on path P3, the overall gate charge equals $2CV$. Therefore, the SET becomes fully conductive. For paths P1 and P2, the gate charges both equal $CV - CV = 0$, hence both switches are partially conductive. For path P0, even though the overall gate charge equals $-2CV$, at high temperature its resistance may still be within the same order of magnitude as that of path P3. Since the inputs of paths P0, P1 and P3 are all connected to logic zero (the worst-case scenario), these three parallel paths may reduce the output voltage, producing incorrect results.

In the high-performance setting, the same capacitance settings are used across the whole temperature range. Therefore, the ratio of Coulomb charging energy to thermal energy increases as the temperature decreases. Therefore, lower temperatures permit fewer multiplexer levels in the multiplexer tree, with more inputs to each individual multiplexer.

Detailed circuit analysis shows that, using the high-performance setting and $e^2/C_{\Sigma} \geq 10kBT$, the dual-gate design may be used at temperatures up to 200 K. At 250 K and 300 K, only the single-gate design is feasible. For the low-power setting, capacitance scaling maintains the same $e^2/C_{\Sigma} \geq kBT$ ratio. Therefore, the same design should be used for the whole temperature range. In addition, since both the low-power setting and the high-performance setting at room temperature use the same $e^2/C_{\Sigma} \geq 40kBT$ ratio, only the single-gate design is feasible for low-power, room-temperature operation. For the $e^2/C_{\Sigma} \geq 40kBT$ configurations of IceFlex, the dual-gate design may be used at all temperatures due to the increased charging energy.

2) SET Configuration Memory: In IceFlex, run-time reconfiguration is enabled by SET configuration memory, which consists of SET configuration cache and current configuration memory. In each SELB, the configuration cache stores multiple configurations. During run-time reconfiguration, one set of configuration bits stored in the configuration cache are placed into the current configuration memory to program SELB logic and interconnect. If $k$ copies of configuration sets are stored in the configuration cache, then the circuit can be reconfigured $k$ times during run-time execution without the need to access off-chip memory.

The left portion of Figure 5 shows the circuit structure of the configuration memory in IceFlex. The SET configuration
cache is the main on-chip configuration memory. Each storage cell consists of a dual-island SET [3]. A dual-island SET contains two capacitively-coupled SETs: a primary SET and a secondary SET. By controlling $V_{CG}$, electrons can tunnel through the control gate and charge the island of the secondary SET. The charge state of the secondary SET shifts the phase of the Coulomb oscillations of the primary gate, i.e., its conductivity condition shifts as a function of gate control voltage, $V_{GS}$. Therefore, under a certain $V_{GS}$, the primary SET is either conductive or open due to different island charges, representing either a logic one or logic zero.

In the configuration cache, selecting a configuration forms a short-circuit path between the pull-up resistor and SETs with a stored zero within the selected configuration set. The power consumption will be high if the configuration cache constantly controls the logic and interconnect. To minimize power consumption, separate on-chip memories are used to store the currently-used configuration.

We designed two types of SET-based on-chip storage to hold the current configuration. The first design is a dual-island based SET buffer. As shown in the SET SRAM portion of Figure 5, this buffer uses two opposite biasing voltages, $V_{G2}$ and $-V_{G2}$, and behaves like a complementary SET inverter. During run-time reconfiguration, for each dual-island SET, the corresponding configuration bit stored in the configuration cache updates the island charge of its secondary SET, hence the conductivity of its primary SET, thereby controlling the buffer output. The second design is a SET SRAM design, which is similar to CMOS SRAM.

3) Efficient SET Implementations of Non-Unate Functions and Implications for Arithmetic: SETs have the ability to support efficient implementation of some critical logic functions that have long frustrated designers using threshold logic, BJTs, and CMOS technologies. Most conventional transistors have either non-decreasing or non-increasing I–V curves. As a result, numerous devices are required to implement Boolean functions that are not unate, i.e., linearly separable. However, such functions are widely used, especially in digital arithmetic. The periodic nature of SET I–V curves can be exploited for efficient implementation of highly-useful non-unate functions such as exclusive-or.

The most efficient CMOS static pass-transistor logic design of a two-input exclusive-or gate in general use requires six transistors [39]. Moreover, it relies on strong input signals because it is not capable of signal restoration. A restoring version would require at least eight transistors. In contrast, it is possible to implement a two-transistor SET-based exclusive-or gate that is structurally equivalent to a CMOS inverter. In this design, each SET has two gates, each of which is connected to one of the exclusive-or inputs. The circuit structure for a SET-based $n$-input parity gate is shown to the right of Figure 5. This design is capable of signal restoration. Thanks to the periodic SET I–V curve, it is possible to directly determine whether the number of high inputs is odd or even. By appropriately adjusting the gate capacitances, the device can be adjusted such that switching a single gate will result in a 180° phase shift in the I–V curve (see Figure 2). Note that even or odd parity functions with additional inputs may be implemented using only two SETs. The number of inputs is bounded primarily by geometrical constraints on fabrication of additional gates.

In SET-based architectures, we propose the use of fast carry chains based on the proposed exclusive-or (sum) computation logic. We have found that this design is approximately 75% more energy-efficient and 25% faster than a design based on a conventional CMOS-style exclusive-or sum implementation, when both are implemented using SETs. This design style is impossible for threshold logic, BJTs, and CMOS technologies. Note that carry-out logic is equivalent to 2-out-of-3 majority vote logic.

4) Reconfigurable Interconnect Network: IceFlex consists of a variety of reconfigurable interconnect resources, including SET local interconnects, hybrid SET/CMOS global interconnects, and SET switch fabric.

Interconnect consumes a substantial proportion of total power consumption in IceFlex: its power efficiency is important. For SET-based interconnect, the static power consumption dominates due to the impact of thermal energy on device conductance, especially at high temperatures. In addition, static power consumption increases with wireload because maintaining unchanged communication latency with higher wireload requires lower junction resistance. In contrast, the dynamic power consumption of SETs is low due to the low SET gate capacitance and low voltage swing. For hybrid SET/CMOS-based interconnect, SETs are only used to drive CMOS buffers, which in turn drive wires. In this case, SETs with low driving strength, hence high junction resistance, are allowed. Compared to SETs, CMOS has lower static power consumption but higher capacitance and dynamic power consumption. Therefore, dynamic power dominates in the hybrid SET/CMOS-based design. Circuit analysis shows that, given the same performance constraint, SET-based design is more energy-efficient for local interconnect and the hybrid SET/CMOS design is more energy-efficient for global interconnect.

In IceFlex, local interconnects driven directly by SET buffers support communication between nearby SELBs. Three types of local interconnects are supported: single length, double length, and hex length. The proposed SET local interconnect design guarantees a constant latency across different routing lengths. Consider, for the sake of example, a local communication architecture in which the maximum interconnect delay is constrained and the longest interconnect is appropriately buffered to meet this constraint. In this case, it would be possible to similarly drive shorter interconnects, thereby decreasing their delays, relative to that of the longest interconnect. It would also be possible to reduce the driving strength on shorter interconnects to reduce power consumption and produce a local interconnect architecture in which all interconnects have uniform delay. We propose the second design because it improves interconnect power efficiency and also simplifies placement and routing during physical design.

The proposed SET local interconnect is designed as follows. A SET buffer with minimal driving strength (hence high junction resistance) is first determined. Next, for local interconnects with different routing lengths, minimal driving
strength SET buffers are connected in parallel to meet driving strength requirements imposed by performance constraints. The main motivation for using parallel SET buffers is that SET junction resistance cannot be reduced arbitrarily \((R_D, R_S \gg h/e^2)\). Using homogeneous SET buffers in parallel instead of heterogeneous SET buffers may also simplify fabrication.

Remote connections introduce the high capacitive loads of long metal wires. To address the driving strength problem of SET-only circuits, we have designed hybrid SET/CMOS interface circuitry to drive global interconnect. Figure 6 shows the circuit structure, which contains two complementary SET inverters and two CMOS inverters. A SELB output is first fed to the input of SET inverter SINV1. SINV1 drives the CMOS inverter, CINV1. Unlike the SET logic used inside SELBs, SINV1 uses a low-resistance design to improve driving strength. Fortunately, it is possible to achieve sufficient driving strength with a single SET. Since the voltage range of SET logic is much smaller than that of CMOS logic, the output signal of SINV1 is within the switching range of the CMOS inverter. Since both MOS transistors are conductive within the switching region, short-circuit power is high. To solve the short-circuit power consumption problem, CINV1. Thus, \(e/C_G\) can match the output signal range of CMOS inverter CINT2. Although source–island and drain–island junctions must be short to permit tunneling, there is no such bound on gate–island separation.

In IceFlex, each SELB is equipped with a reconfigurable input switch fabric that selects the connections among local and global interconnects. The input switch fabric is implemented using multi-gate SET multiplexor tree, similar to that in the reconfigurable look-up table described in Section III-B1.

5) Design and Modeling of IceFlex Majority Voting Logic: Although researchers are making progress on reducing the severity of noise resulting from random background offset charge effects, it may continue to pose run-time noise problems in the future. Even if this problem can be entirely solved, resistance to run-time faults may be useful in SETs, e.g., to allow resistance to Alpha particle induced faults or other single event upsets. IceFlex incorporates support for hierarchical spatial redundancy to improve fault tolerance. Although much of the literature predicts the need for fault-tolerant architectures in nanoelectronics, the level of fault tolerance is currently unknown. Therefore, we consider the results for a number of possible SET failure rates and in the presence of three fault-tolerance configurations.

Other researchers have proposed a number of architectural techniques to support reliable computation using nanoscale electronics that are susceptible to fabrication-time and run-time faults. Dehon described the use of structural redundancy and programming-time defect-aware configuration in a carbon nanotube and silicon nanowire based programmable logic array architecture [40]. Goldstein et al. describe the use of a defect map that is generated during post-fabrication testing to avoid the use of faulty devices [41]. Bahar et al. present a method of expressing logic circuits using Markov Random Fields, permitting Boolean functions to be computed using devices susceptible to potentially-frequent transient faults [42]. We think it likely that the random background charge problem will ultimately be dealt with by a combination of improved fabrication technology, post-fabrication testing to identify and avoid a subset of the affected SETs, and run-time fault-tolerance via conventional structural redundancy or recent advances in probabilistic computation. IceFlex provides for regular structural redundancy and run-time error correction.
TABLE III
IMPACT OF MAJORITY VOTE LOGIC ON SELB FAULT PROBABILITY

<table>
<thead>
<tr>
<th>SET fault probability Majority vote inputs</th>
<th>1/1,000</th>
<th>1/10,000</th>
<th>1/100,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw fail prob.</td>
<td>6.20E-2</td>
<td>6.20E-2</td>
<td>6.20E-2</td>
</tr>
<tr>
<td>Best prob.</td>
<td>1.11E-2</td>
<td>2.17E-3</td>
<td>4.45E-4</td>
</tr>
<tr>
<td>SET MVL prob.</td>
<td>1.11E-2</td>
<td>2.18E-3</td>
<td>4.57E-4</td>
</tr>
</tbody>
</table>

We now consider the fault model for IceFlex SELBs. Every path from SELB input to output contains 64 SETs. In the third row of Table III, we show the SELB raw failure probabilities, i.e., the probability of a SELB producing an incorrect output. SELB failure probability is a function of the SET fault probability, for which Table III shows three values. Likharev estimates the long-term density of background offset charge susceptible SETs [3]. We follow his assumptions but correct a typographical error in that article, arriving at one susceptible SET in 10,000. The resulting 1/f noise produces long-duration failure periods. Therefore, in this analysis, we (conservatively) assume that susceptible devices consistently fail. In reality, errors may not be consistent. We also consider the higher SET fault probability of 1/1,000 and the lower fault probability of 1/100,000. Advances in fabrication and detection of most SETs susceptible to random background offset charge effects by post-fabrication testing may permit reduction in run-time SET fault probability.

We have considered the effect of using no MVL (Raw fail prob.), fault-free MVL (Best prob.), and SET MVL. Using a given reliability configuration, it is not possible for MVL-based designs to produce lower SELB fault probabilities than those shown in the Best prob. row. SET MVLs are constructed from multi-gate SETs. We focus on the three-input SET MVL design to simplify depiction; the five-, input-, and seven-input SET MVL follows an analogous design style. This circuit has identical structure to the parity gate shown in the SET parity circuit portion of Figure 5. However, the separation of gates and island are adjusted such that the circuit traverses only 1/2 Coulomb oscillation period during use. The SET pull-up gates are separated sufficiently to require the majority of the gates to be high. The converse is true of the pull-down gates. For each SET depicted in the figure, four SETs are used in parallel in order to permit the failure of one SET while still producing correct results. We have computed the delay of the SET MVL by considering the worst-case scenario, in which a path that is 3/5 or 4/7 closed has a faulty driver SET and a path that is 2/4 or 3/7 closed has no faulty SETs.

As shown in Table III, it is possible for a seven-input SET-only MVL with redundant SELBs to reduce the failure rate to 1/8,500,000, given a SET fault probability of 1/10,000, or 1/830,000,000, given a SET fault probability of 1/100,000. Given recent trends in noise-resistant SET design and fabrication, it seems likely that a less aggressive fault tolerance configuration will be necessary in the future (see Section II-B).

If a method of rapidly determining which SETs are susceptible to random background charge effects is ever developed, these effects can be avoided in the same way that fabrication defects are avoided: via the use of a regular computation structure in which operations are mapped only to fault-free devices. There has been some promising work on this topic, in which illumination is used to produce ions, accelerating the onset of random background charge effects [43].

IV. EXPERIMENTAL RESULTS

In this section, we evaluate the suitability of using SETs in low-power embedded system design. We start from the microarchitecture characterization of IceFlex. IceFlex is then used as a testbed to characterize the benefits and limitations of SETs for both high-performance and battery-powered embedded application.

A. Characterization of the IceFlex Architecture

Following the design parameters shown in Table II, we evaluate the performance and power consumption of IceFlex using HSPICE. For SET circuitry, the SPICE model and device parameters are described in Section II-C. For CMOS logic and metal wire, we use the 22 nm Berkeley BSIM4 predictive technology model, which models the impact of temperature on MOS devices. We analyzed designs adhering to the $C_S = e^2/(40k_BT)$ constraint. We also analyzed designs with the less conservative $C_S = e^2/(10k_BT)$ constraint, although space constraints force us to omit detailed results tables for this setting. A low-power setting (targeting megahertz-range frequencies) and a high-performance setting (targeting gigahertz-range frequencies), are considered.

Tables IV and V summarize the performance and power characterization of the logic components and interconnect fabric of the $C_S = e^2/(40k_BT)$ version of IceFlex, including multi-gate SET reconfigurable lookup table (LUT)$^1$ SET register (Register), SET and CMOS four-out-of-seven majority voting logic (MVL), multi-gate (MG) and CMOS-style (CS) exclusive-or, (CO) carry-out logic, and SET local interconnect (Single, Double, and Hex), hybrid SET/CMOS global interconnect (Global) and SET input switch fabric (ISF). From these results, we make the following observations.

First, IceFlex has high energy efficiency, good performance, and high flexibility in terms of performance and energy efficiency tradeoff. At the low-power setting, the power consumptions of SET-based logic components and local interconnect fabric are nano-Watts. The hybrid SET/CMOS global interconnect has the highest power consumption. This is a result of the high capacitance of global wires and high power consumption of the CMOS buffers. All components in the low-power version of IceFlex still have latencies in the range of nanoseconds. SETs have high junction resistance and low driving strength. Using the high-performance setting, by scaling the SET junction resistance down to 100 kΩ, the latencies of the SET-based logic and local interconnect fabric are consistently lower than 100 ps. Even though reducing

$^1$To allow comparison with Xilinx FPGAs, a 16-to-1 setting is used.
resistance results in a 100× increase in power, as demonstrated in [Section IV-B] the overall energy efficiency of IceF lex is still orders of magnitude higher than that of CMOS-based solutions.

Second, these results demonstrate the impact of temperature on SET performance and power consumption – as the temperature increases, performance increases and the power efficiency decreases. This is a result of the impact of thermal energy on tunneling events and therefore circuit behavior, which is described in Section II. The number of electrons with sufficient energy to overcome the Coulomb blockade effect increases with temperature, thereby increasing tunneling rate, performance, and power consumption.

The $C_S = e^2/(40k_BT)$ setting enables greater resistance to shot noise than the $C_S = e^2/(10k_BT)$ setting. However, it also imposes performance and power consumption penalties. For SET circuitry, the required supply voltage is inversely proportional to gate capacitance. Compared to the $C_S = e^2/(10k_BT)$ setting, $C_S = e^2/(40k_BT)$ requires a further reduction of SET gate capacitance and an increase in supply voltage. Note that the driven capacitance of a SET gate design is feasible at temperatures greater than 250 K. From 200 K to 250 K, both latency and power consumption increase. In addition, when using the same design, we observe that both the circuit performance and power consumption increase with temperature. The same trend was described in Section IV-A. Using the low-power design of IceFlex, only the single-gate design is feasible at temperatures greater than 250 K due to high static currents at these temperatures. As a result, circuit power consumption is increased at high temperatures. From 200 K to 250 K, both latency and power consumption increase. In addition, when using the same design, we observe that both the circuit performance and power consumption increase with temperature. The same trend was described in Section IV-A. Using the low-power design of IceFlex, only the single-gate design is feasible at temperatures greater than 250 K due to high static currents at these temperatures. As a result, circuit power consumption is increased at high temperatures. From 200 K to 250 K, both latency and power consumption increase. In addition, when using the same design, we observe that both the circuit performance and power consumption increase with temperature. The same trend was described in Section IV-A. Using the low-power design of IceFlex, only the single-gate design is feasible at temperatures greater than 250 K due to high static currents at these temperatures. As a result, circuit power consumption is increased at high temperatures. From 200 K to 250 K, both latency and power consumption increase. In addition, when using the same design, we observe that both the circuit performance and power consumption increase with temperature. The same trend was described in Section IV-A. Using the low-power design of IceFlex, only the single-gate design is feasible at temperatures greater than 250 K due to high static currents at these temperatures. As a result, circuit power consumption is increased at high temperatures. From 200 K to 250 K, both latency and power consumption increase. In addition, when using the same design, we observe that both the circuit performance and power consumption increase with temperature. The same trend was described in Section IV-A. Using the low-power design of IceFlex, only the single-gate design is feasible at temperatures greater than 250 K due to high static currents at these temperatures. As a result, circuit power consumption is increased at high temperatures. From 200 K to 250 K, both latency and power consumption increase. In addition, when using the same design, we observe that both the circuit performance and power consumption increase with temperature. The same trend was described in Section IV-A. Using the low-power design of IceFlex, only the single-gate design is feasible at temperatures greater than 250 K due to high static currents at these temperatures. As a result, circuit power consumption is increased at high temperatures.

### TABLE IV

**CHARACTERIZATION OF ICEFLEX MICROARCHITECTURE FOR $C_S = e^2/(40k_BT)$**

<table>
<thead>
<tr>
<th>Latency (ns)</th>
<th>Low Power</th>
<th>High Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>40 K</td>
<td>77 K</td>
</tr>
<tr>
<td>LUT</td>
<td>10.04</td>
<td>7.86</td>
</tr>
<tr>
<td>Register</td>
<td>1.42</td>
<td>1.09</td>
</tr>
<tr>
<td>Logic</td>
<td>1.52</td>
<td>1.31</td>
</tr>
<tr>
<td>Arithmetic</td>
<td>1.52</td>
<td>1.31</td>
</tr>
</tbody>
</table>

### TABLE V

**CHARACTERIZATION OF ICEFLEX INTERCONNECT FABRIC FOR $C_S = e^2/(40k_BT)$**

<table>
<thead>
<tr>
<th>Latency (ns)</th>
<th>Low Power</th>
<th>High Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>40 K</td>
<td>77 K</td>
</tr>
<tr>
<td>ISF</td>
<td>0.96</td>
<td>0.73</td>
</tr>
<tr>
<td>Single</td>
<td>0.74</td>
<td>0.66</td>
</tr>
<tr>
<td>Double</td>
<td>0.75</td>
<td>0.67</td>
</tr>
<tr>
<td>Hex</td>
<td>0.69</td>
<td>0.64</td>
</tr>
<tr>
<td>Global</td>
<td>0.96</td>
<td>0.73</td>
</tr>
<tr>
<td>ISF</td>
<td>0.21</td>
<td>0.18</td>
</tr>
<tr>
<td>Single</td>
<td>0.00</td>
<td>0.02</td>
</tr>
<tr>
<td>Double</td>
<td>0.00</td>
<td>0.02</td>
</tr>
<tr>
<td>Hex</td>
<td>0.00</td>
<td>0.02</td>
</tr>
<tr>
<td>Global</td>
<td>0.00</td>
<td>0.02</td>
</tr>
</tbody>
</table>

### 2) Power and Performance of Interconnect Design:

Power consumption, performance, and the tradeoff between them are of central importance in interconnect design. We considered both SET-only and SET/CMOS hybrid interconnect driver designs. The relative static power benefit of the SET/CMOS hybrid design over the SET-only design increases as the wireload increases. This is mainly due to an increase in the static power consumption of the SET-only design as more SET
buffers are used to meet the driving strength requirements. The SET-only design has superior power efficiency. As the wire length increases, the proportion of capacitance contributed by CMOS buffer gates becomes less significant relative to wire capacitance. Therefore, compared to the SET-only design, the dynamic power consumption of the SET/CMOS hybrid design also improves, but is still inferior to that of the SET-only design. At 300 K, for both the $C_\Sigma = e^2/(40k_BT)$ and $C_\Sigma = e^2/(10k_BT)$ settings, we found that SET-only designs had better energy efficiencies for wires shorter than approximately 1 mm, and SET/CMOS hybrid designs were better for longer wires. As temperature increases, the thermal energy impact increases. As a result, the static power consumption of SETs increases. Therefore, the wire length at which the SET/CMOS design begins to outperform the SET-only design decreases as temperature increases.

Table V illustrate two interesting trends for global interconnect. The power consumption of both the low-power and the high-performance $C_\Sigma \leq e^2/(40k_BT)$ hybrid SET/CMOS designs decrease with increasing temperature. At low temperatures, the output voltage ranges and driving currents for the SETs are small, increasing CMOS buffer static power consumption.

3) Performance and Power Characterization of SET Non-Unate Logic: SETs support the efficient implementation of some non-unate arithmetic functions. We evaluate the power consumption and performance of an exclusive-or gate, a non-unate Boolean function widely used in arithmetic logic, e.g., in addition and multiplication. We compared the two different implementations described in Section III-B3, the proposed SET-based design and the CMOS-style SET implementation.

Figure 8 shows the power and performance characterization of these two designs for the low-power and $C_\Sigma = e^2/(40k_BT)$ settings. These results demonstrate the superior power consumption and performance of this design style, which is not possible using BJTs, CMOS, or threshold logic. Compared to the CMOS-style SET implementation, the design that exploits the periodic I–V curve of SETs achieves the latency and power consumption reductions indicated in Table VI, i.e., approximately a 25% reduction in latency and 75% reduction in energy consumption.

B. Characterization of High-Performance and Battery-Powered Embedded Applications

This section characterizes the performance and power consumption of IceFlex when used to implement numerous general-purpose and application-specific processor cores. We evaluate the suitability of IceFlex for use in both portable battery-powered and high-performance embedded systems by determining is performance and energy efficiency when used to implement the processor cores described below. We have divided the cores into battery-powered and high-performance categories for the convenience of the reader.

Battery-powered: AES (Rijndael) IP core (AES), AT-Mega103 microcontroller (AVR), coordinate rotation computer (CORDIC), ECC core (ECC), 32-bit IEEE 754 floating-point unit (FPU), Reed–Solomon encoder (RS), USB 2.0 function (USB), and video compression systems (VC).

High-performance: Power-efficient RISC CPU (ARM7), synchronous / DLX core (ASPIDA DLX), five-stage pipeline RISC CPU (Jan RISC), entire SPARC V8 processor (LEON2 SPARC), RISC CPU (Microblaze), MIPS I clone (minMIPS), MIPS processor (MIPS) supporting most MIPS I opcodes (Plasma), MIPS I integer only clone (UCore), and MIPS I clone (YACC).

The Xilinx Virtex-II XC2V2000 FPGA is used as a base case for comparison. Each application is synthesized with Xilinx ISE to determine the number of required LUTs, maximum frequency, and power consumption, using a switching probability of 10% [44] and a 65 nm feature size. Then, we scale the FPGA synthesis results into a 22 nm process based on HSPICE predictive technology model simulation results for the two technologies [45]. We used FPGA synthesis software to estimate the number of IceFlex SELBs required. 16-entry Virtex-II LUTs were used due to their functional (but not structural) similarity to IceFlex SELBs. For each design, the maximum frequency for IceFlex was determined by multiplying the number of SELBs along the longest combinational path by the delay of an IceFlex SELB plus the delay of a local interconnect. IceFlex power consumption was computed by taking the sum of the power consumptions of all components.
at the maximum operating frequency. Note that, since Xilinx ISE does not report use of global interconnect for any of the processors we synthesized, we exclude the hybrid global interconnect from IceFlex power analysis. In designs that use primarily local interconnect (i.e., single, double, and hex interconnect), the reported power consumption results will be accurate. However, for designs in which global hybrid SET-CMOS interconnect dominates, the power consumption may approach that of global interconnect in a corresponding 22 nm CMOS design.

Table VII shows the operating frequencies and energy efficiency in Joules per clock cycle of the CMOS FPGA and IceFlex variants for each benchmark application. As described in Section III-A5, recent progress in fabrication is reducing the severity of the random background charge problem. Due to space constraints, we show the characteristics of non-spatially-redundant versions of IceFlex.

We have also analyzed the performance and power consumption of a configuration with seven-way SELB spatial redundancy. The MVL uses internal fine-grained four-way SET parallelism. The MVL fault-tolerance hardware delays are added to the delay of each SELB stage. The fault-tolerant versions of the processors have similar maximum frequencies to those listed in Table VII. However, they only improve on the energy efficiency of 22 nm CMOS FPGAs by averages of 8.64 × (battery-powered) and 10.55 × (high-performance).

1) Ultra-Low-Power Applications: The data in Table VII indicate that the non-redundant, room temperature, low-power version of IceFlex is suitable for use in applications such as sensor network nodes, if they can be fabricated with sufficiently small island capacitances. In the following analysis, we shall focus on the AVR core, which is representative of a commonly-used sensor network node processor. Alkaline AA batteries typically have 2,800 mAh of energy and nominal operating voltages of 1.5 V, i.e., they can deliver approximately 15,000 J. Using the conservative $C_S \leq e^2/(40k_BT)$ constraint, a low-power IceFlex AVR implementation running at 4 MHz consumes approximately 200 µW, permitting it to run for 20 years on one AA battery, i.e., longer than the shelf life of most such batteries. When the less conservative $C_S \leq e^2/(10k_BT)$ constraint is used, the average energy consumption improvements increase to 95.60 × (non-redundant battery powered), 115.65 × (non-redundant high performance), 12.27 × (redundant battery powered), and 15.27 × (redundant high performance).

This power consumption is also low enough to permit an AVR processor to operate on energy scavenged from the environment. If we assume an energy scavenging volume of 5 cm$^3$ and use Roundy’s power densities of 4 µW/cm$^3$ for indoor solar energy, 200 µW/cm$^3$ for vibrations, 10 µW/cm$^3$ for daily temperature variation, and 0.003 µW/cm$^3$ for acoustic noise at 75 dB [46], we find that one sensor network node is capable of scavenging enough energy to power an IceFlex AVR processor running at the maximum clock frequency from vibrations or daily temperature variation, at 3.7 MHz from indoor solar energy, and at 2.8 kHz from 75 dB acoustic noise. However, SET circuits that operate at room temperature and adhere to the $C_S \leq e^2/(40k_BT)$ constraint will rely on features with sizes approaching (but not crossing) physical limits. Although the use of SETs in battery-powered applications has potential, it depends on the solution of formidable fabrication challenges or the development of compact, low-power cooling methods.

2) Energy-Efficient High-Performance Applications: We can draw the following general conclusions from Table VII. For a wide range of processor cores, the SET-based IceFlex architecture is capable of achieving energy efficiencies two orders of magnitude better than 22 nm CMOS-based FPGAs. Peak frequencies ranging from 200 MHz to 2 GHz are maintained for all processors.

One might expect the high-performance version of IceFlex to consistently achieve higher frequency but lower energy efficiency than the low-power version of IceFlex. However, its energy efficiency is typically better, as well. Operating at higher frequencies can permit reduced static energy consumption, and therefore better energy efficiency, especially at room temperature where static power consumption is high (see Figure 2). Therefore, for SET-based architectures that are operated at room temperature and have low performance requirements, it will generally be more energy efficient to operate the device at high frequency and periodically enter a power-gated sleep mode than to continuously operate at a low frequency.

In high-performance applications for which parallel computation is appropriate, improved energy efficiency can be traded for improved performance with the same energy budget. For example, given a power budget of 125 mW and $C_S \leq e^2/(40k_BT)$, one could use one LEON2 SPARC implemented with an FPGA and running at 85 MHz or 5 LEON2 SPARCs implemented with the high-performance variant of IceFlex and operating at 1,025 MHz. This implies an overall performance 60 × higher than that of the FPGA version. Taken to its logical extreme, assuming a power budget of 100 W and one instruction per cycle, one could execute 4.8 Terra IPS. These numbers are intended to give the reader some indication of the potential

### Table VII

<table>
<thead>
<tr>
<th>Benchmarks</th>
<th>FPGA 22 nm CMOS technology*</th>
<th>Battery-powered</th>
<th>High-performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freq (GHz)</td>
<td>Energy (fJ/cycle)</td>
<td>Freq (MHz)</td>
<td>Energy (J/cycle)</td>
</tr>
<tr>
<td>ARM7</td>
<td>26.3 2.96e-09</td>
<td>2.0 5.47e-11</td>
<td>224.0 4.79e-11</td>
</tr>
<tr>
<td>ASPIDA DLX</td>
<td>125.7 8.86e-10</td>
<td>11.5 6.37e-12</td>
<td>1333.3 5.58e-12</td>
</tr>
<tr>
<td>Jam RISC</td>
<td>95.9 8.92e-10</td>
<td>12.8 3.65e-12</td>
<td>1481.5 3.19e-12</td>
</tr>
<tr>
<td>LEON2 SPARC</td>
<td>85.9 1.88e-09</td>
<td>8.8 2.39e-11</td>
<td>1025.6 2.99e-11</td>
</tr>
<tr>
<td>Microblaze RISC</td>
<td>18.1 7.28e-10</td>
<td>16.4 2.01e-12</td>
<td>1900.8 1.76e-12</td>
</tr>
<tr>
<td>miniMIPS</td>
<td>88.0 4.87e-10</td>
<td>9.6 9.78e-12</td>
<td>1111.1 8.56e-12</td>
</tr>
<tr>
<td>MIPS</td>
<td>50.4 2.10e-12</td>
<td>10.5 3.43e-12</td>
<td>1212.1 3.80e-12</td>
</tr>
<tr>
<td>Plasma</td>
<td>75.4 1.13e-09</td>
<td>8.8 9.61e-12</td>
<td>1025.6 6.05e-12</td>
</tr>
<tr>
<td>UCore</td>
<td>136.4 8.19e-10</td>
<td>12.8 5.45e-12</td>
<td>1481.5 4.78e-12</td>
</tr>
<tr>
<td>YACC</td>
<td>72.1 1.86e-09</td>
<td>19.2 3.08e-12</td>
<td>2222.2 2.69e-12</td>
</tr>
<tr>
<td>AES</td>
<td>205.3 3.43e-10</td>
<td>28.7 2.34e-12</td>
<td>3333.3 2.05e-12</td>
</tr>
<tr>
<td>AVR</td>
<td>71.9 2.67e-10</td>
<td>9.6 5.34e-12</td>
<td>1111.1 4.67e-12</td>
</tr>
<tr>
<td>CORDIC</td>
<td>271.8 1.37e-10</td>
<td>114.9 2.05e-12</td>
<td>1333.3 1.79e-13</td>
</tr>
<tr>
<td>ECC</td>
<td>39.1 4.91e-10</td>
<td>11.5 6.92e-12</td>
<td>1333.3 6.05e-12</td>
</tr>
<tr>
<td>FPU</td>
<td>28.4 1.00e-09</td>
<td>2.6 8.02e-11</td>
<td>296.3 7.02e-11</td>
</tr>
<tr>
<td>RS</td>
<td>496.7 1.28e-11</td>
<td>57.5 4.61e-14</td>
<td>6666.7 4.05e-14</td>
</tr>
<tr>
<td>USB</td>
<td>171.6 3.24e-10</td>
<td>38.3 1.53e-12</td>
<td>4444.4 1.34e-12</td>
</tr>
<tr>
<td>VC</td>
<td>114.6 1.24e-09</td>
<td>23.0 1.04e-11</td>
<td>2666.8 9.10e-12</td>
</tr>
</tbody>
</table>

An energy improvement 68.58 × 78.40 ×
to improve performance given a power budget. In practice some of this performance will be lost due to parallelization inefficiency and off-chip communication latency. A similar comparison can be used for the MIPS processor, for which IceFlex permits a 268× improvement in energy efficiency compared with an FPGA implementation.

V. Conclusions

In this article, we have analyzed the impact of using SETs in architecture and circuit design; proposed IceFlex, a fault-tolerant, reconfigurable, hybrid SET/CMOS architecture for use in high-performance and battery-powered embedded systems; and evaluated the energy efficiency, power consumption, and performance of IceFlex in these applications. Our results indicate that using SETs for computation poses many design challenges, some of which can be solved with the proposed architecture and circuit design techniques. In addition, we find that SETs have unique properties that permit significant improvements in circuit efficiency when compared with BJT, CMOS, and threshold logic based design. In summary, we find that SETs have unique properties that permit significant improvements in circuit efficiency when compared with BJT, CMOS, and threshold logic based design. In conclusion, we have evaluated the energy efficiency, power consumption, and performance of IceFlex in these applications. Our results indicate that using SETs for computation poses many design challenges, some of which can be solved with the proposed architecture and circuit design techniques.

REFERENCES


**Changyun Zhu** (S’06) received his B.E. and M.E. degrees from Tsinghua University in 2002 and 2005. He is currently a Ph.D. student at Queen’s University’s Department of Electrical and Computer Engineering. His research interests include computer-aided design of integrated circuits, reliability modeling and optimization, and design for nanotechnologies.

**Zhenyu (Peter) Gu** (S’04) received his Ph.D. degree from Northwestern University in 2007, and B.S. and M.S. degrees from Fudan University, China in 2000 and 2003, respectively. Gu has published in the areas of behavioral synthesis and thermal analysis of integrated circuits. He is currently with Synopsys Inc., CA.

Robert P. Dick Robert Dick (S’95-M’02) is an Assistant Professor of Electrical Engineering and Computer Science, Northwestern University. He received his Ph.D. degree from Princeton University and his B.S. degree from Clarkson University. He worked as a Visiting Professor at Tsinghua University’s Department of Electronic Engineering and as a Visiting Researcher at NEC Labs America. Robert received an NSF CAREER award and won his department’s Best Teacher of the Year award in 2004. His technology won a Computerworld Horizon Award and his paper was selected by DATE as one of the 30 most influential in the past 10 years in 2007. He served as a technical program subcommittee chair for CODES-ISSS and is an Associate Editor of IEEE Transactions on VLSI Systems and serves on the technical program committees of several embedded systems and CAD/VLSI conferences.

**Li Shang** Li Shang (S’99-M’04) is an Assistant Professor at the Department of Electrical and Computer Engineering, University of Colorado at Boulder. Before that, he was with the Department of Electrical and Computer Engineering, Queen’s University. He received his Ph.D. degree from Princeton University, and his B.E. degree with honors from Tsinghua University. His work on hybrid SET/CMOS reconfigurable architecture was nominated for the Best Paper Award at DAC 2007. His work on thermal-aware incremental design flow was nominated for the Best Paper Award at ASP-DAC 2006. His work on temperature-aware on-chip network has been selected for publication in MICRO Top Picks 2006. He also won the Best Paper Award at PDCS 2002. He is currently serving as an Associate Editor of IEEE Transactions on VLSI Systems and serves on the technical program committees of several design automation conferences. He won his department’s Best Teaching Award in 2006. He is the Walter F. Light Scholar.

Robert G. Knobel Robert Knobel received his MSc degree in 1995 from the University of British Columbia and his PhD degree from the Pennsylvania State University in 2000, both in physics, studying high temperature superconductors and diluted magnetic semiconductors, respectively. From 2000-2003 he was a postdoctoral scholar at the University of California, Santa Barbara working with Andrew Cleland on nanomechanical systems and single electron transistors. Since 2003 he has been an Assistant Professor in the Department of Physics, Engineering Physics and Astronomy at Queen’s University, where he studies the quantum limits of measurement in nanoscale devices.