Solid-State Electronics 103 (2015) 19–29

Contents lists available at ScienceDirect

Solid-State Electronics

journal homepage: www.elsevier.com/locate/sse

On the effect of technology scaling on variation-resilient sub-threshold circuits

Nele Reynders *, Wim Dehaene

Dept. of Electrical Engineering (ESAT-MICAS), KU Leuven, Kasteelpark Arenberg 10, 3001 Leuven, Belgium

ARTICLE INFO

Article history:
Received 17 March 2014
Received in revised form 7 August 2014
Accepted 13 August 2014
Available online 10 October 2014

The review of this paper was arranged by Prof. A. Zaslavsky

Keywords:
Scaling
Ultra-low-voltage
Sub-threshold
Digital circuits
Variation-resilient
Minimum supply

ABSTRACT

This paper studies the impact that CMOS technology scaling has on circuits operating in the ultra-low-voltage region. Sub-threshold circuits are an attractive option for energy-constrained applications, but the influence of scaling on the energy consumption has not been studied thoroughly on on-chip ultra-low-voltage implementations. This paper aims to provide an answer to the benefits and disadvantages of scaling on such implementations. First, an equation to determine the minimum feasible supply voltage for digital circuits is derived. Out of this equation, a theoretical minimum as well as a practical minimum supply for a specific technology can be calculated. Second, a 16-bit Multiply-Accumulate Unit is selected as a test vehicle to study scaling effects. This test vehicle is designed, processed and fully measured in both a 90 nm and a 40 nm CMOS technology. An extensive comparison between the measurement results of both designs allows to clearly examine the different technology scaling trade-offs.

1. Introduction

CMOS scaling has been driven by the increased performance that is obtained for digital systems. However, many applications do not require such a high circuit speed. Instead, the energy consumption becomes the critical parameter. A wide range of these applications exists, e.g. sensor networks, RFID tags and biomedical signal processors. Ultra-low-voltage circuits can be a solution for such energy-constrained applications that are less stringent on speed requirements [1]. Lowering the supply voltage V_{dd} under or near the threshold voltage V_T enables large reductions in energy consumption, at the disadvantage of a simultaneous increase in circuit delay. Although few implementations only require operating frequencies in the kHz-range, to allow a more widespread use of ultra-low-voltage circuits, higher operating frequencies well within in the MHz-range are required.

The impact of CMOS technology scaling for digital circuits operating at the nominal supply has been extensively studied, through simulations, on-chip implementations and measurements. The influence of scaling on circuits operating in the weak inversion region has received some attention, but until now, this attention has been limited to device-level studies and circuit-level simulations.

Previous work investigated the effect of scaling on device-level and proposed different scaling strategies. Hanson et al. [2] performed a model study of sub-threshold transistors to investigate the implications of device scaling on sub-threshold operation from 90 nm down to 32 nm technology nodes. An alternative scaling strategy was proposed to help sub-threshold circuits to reliably scale to nanometer technologies. In [3], devices were redesigned specifically for sub-threshold operation. An optimized transistor structure to improve sub-threshold the circuit delay and the power delay product was proposed. The impact of technology scaling for nodes from 90 nm to 22 nm was examined in [4] and strategies for increasing the robustness of sub-threshold circuits were proposed.

Some prior works also performed simulations to examine the impact of technology scaling on ultra-low-voltage logic circuits, although this mostly consisted of simple circuits. In [5], simulations of a ring oscillator were used to validate an analytical approach for studying the effect of technology scaling and variability on performance of ultra-low-power integrated systems. The effects of process variations were exhaustively examined to study the sensitivity of a circuit in presence of these variations. Bol et al. [6] investigated the impact of technology scaling on sub-threshold logic in nodes from 0.25 μm to 32 nm CMOS. A circuit-level simulation of a benchmark 8-bit multiplier was used to study the scaling

* Corresponding author.
E-mail addresses: nele.reynders@esat.kuleuven.be (N. Reynders), wim.dehaene@esat.kuleuven.be (W. Dehaene).

http://dx.doi.org/10.1016/j.sse.2014.08.013
0038-1101/

effects, first using predictive technology models (PTM) and then validated by industrial models.

To the author's knowledge, only [7] presented measured results: two test chips were fabricated in a 130 nm and a 65 nm technology, consisting of a 1000-stage inverter chain and a 41-stage ring oscillator, both operating at ultra-low voltages. The measurements of these two simple circuits were used to validate a body biasing technique to adaptively balance the pMOS and nMOS transistors in strength. No papers have been published that present the design and measurements of a full digital system, implemented in different CMOS technology nodes. Moreover, a significant amount of the previous work did not use industrial models, but rather relied on PTMs (e.g. [2,4,5]). PTMs are reasonably accurate transistor models that benchmark future generations of technologies and are therefore a useful resource for early circuit design research [8]. Although PTM simulations are very suitable to investigate scaling trends, they do not provide the same value as designing with industrial models, followed by manufacturing and measuring the designed chip.

This paper aims to fill the hiatus between simulations and measured, confirmed results. An extensive digital circuit has been designed, processed and measured in both a 90 nm [9] and a 40 nm CMOS technology. The test vehicle that was used to study the effect of technology scaling on ultra-low-voltage circuits, is a 16-bit Multiply-Accumulate Unit (MAC). The MAC was chosen since it is a block that is very frequently used in DSP designs and because it is a complex block that includes feedback. Since the MAC is the critical component of DSP designs, it is also possible to design a processor that achieves similar ultra-low-voltage characteristics as this MAC with the same design principles [10]. Important to note is that the aim was to design variation-resilient circuits which are able to operate at both very low energy consumptions and $n \times 10$ MHz-speed to increase the industrial relevance of ultra-low-voltage circuits. To conclude, this paper focuses on providing a scaling analysis which is based on the design and measurements of a large ultra-low-voltage circuit.

Section 2 gives an overview of previous theoretical studies to find the fundamental limit for the lowest feasible supply voltage $V_{dd,min}$ and provides a new, practical equation to determine $V_{dd,min}$ for a specific technology. Section 3 gives a comparison between the used technologies, while Section 4 covers the design of the 16-bit MAC. Section 5 explains in detail the transistor-level implementation of the different logic components. The design changes that were necessary for the technology scaling are also addressed. In Section 6, the implementation of the timing is described, with a specific focus on the different design decisions necessary for both technologies. Section 7 presents the results obtained from the measurements of both chips.

2. Theory of sub-threshold operation

Before explaining the design and implementation of the MAC test case in the ultra-low-voltage region, this section first provides a practical expression that estimates the minimum possible supply voltage that can be expected for a certain technology. Previous research has focused on theoretically finding the fundamental limit for the lowest operating voltage for CMOS technologies. Already in 1972, Swanson and Meindl [11] studied the minimum usable supply of an inverter, with the requirement that the inverter should have sufficient maximum gain at $V_{dd}/2$ to be usable in a digital circuit. Based on measurements in a technology available at that time, the authors estimated that the minimum usable V_{dd} would have a value of about $8kT/q$ (where k is the Boltzmann constant, T the absolute temperature and q the electrical charge of an electron), or 207 mV at 300 K. In 2001, Bryant et al. [12] proposed another theoretical limit of the lowest operable supply. To achieve this V_{dd}, the nMOS and pMOS off-currents must be equalized. Following this requirement, the ideal supply limit of $4kT/q$ is proposed, which is 103 mV at 300 K.

These are all theoretical limits that predict the lowest possible supply voltage of CMOS digital circuits. However, they are not practical limits that take into account the specific details of the technology at hand. Therefore, we have derived a practical limit for the minimum feasible supply from the equations listed below. The basic equation for the current flow in the weak inversion region consists of an exponential relationship with V_{GS}:

$$I_{DS} = I_0 \cdot \exp\left(\frac{V_{GS} - V_T}{n \cdot V_{th}}\right)\left(1 - \exp\left(\frac{-V_{DS}}{V_{th}}\right)\right) \tag{1}$$

where V_{th} is the thermal voltage ($V_{th} = kT/q = 26$ mV at room temperature 300 K), n is a process-dependent parameter and I_0 is the current when $V_{GS} = V_T$. I_0 is dependent on process and device geometry [13,14]. The third term incorporates the current roll-off, and only has an influence when V_{DS} drops to within a few multiples of V_{th}. Taking the ratio of the on-current I_{on} at $V_{GS} = V_{dd}$ and the off-current I_{off} at $V_{GS} = 0$ gives:

$$\frac{I_{on}}{I_{off}} = \frac{I_0 \cdot \exp\left(\frac{V_{dd} - V_T}{n \cdot V_{th}}\right)}{I_0 \cdot \exp\left(\frac{-V_T}{n \cdot V_{th}}\right)} = \exp\left(\frac{V_{dd}}{n \cdot V_{th}}\right) \tag{2}$$

In both I_{on} and I_{off}, $V_{DS} = V_{dd}$ and therefore V_{DS} dependencies can be omitted. In (2), a direct relationship between the variables V_{dd}, I_{on} and I_{off} is obtained since V_{th} is fixed for a certain temperature and n is fixed for a certain technology. An equation for the supply voltage can be derived:

$$V_{dd} = \ln\left(\frac{I_{on}}{I_{off}}\right) \cdot n \cdot V_{th} \tag{3}$$

The value of n is affected by depletion region characteristics [15] and is equal to 1 for an ideal transistor, but unfortunately larger than 1 for actual devices. It is typically in the range of 1.3–1.7 for CMOS processes [14]. Since it is difficult to accurately determine n for a certain technology, the link with the so-called sub-threshold slope S_S will be made. The sub-threshold slope is defined by the amount by which V_{GS} must be increased in order for the weak inversion current I_{DS} to be increased by one order of magnitude. It is expressed in mV/decade[15]:

$$S_S = n \cdot V_{th} \cdot \ln(10) \tag{4}$$

Substituting n in (3) by using (4), results in:

$$V_{dd} = \frac{\ln\left(\frac{I_{on}}{I_{off}}\right) \cdot S_S \cdot V_{th}}{V_{th} \cdot \ln(10)} = \log_{10}\left(\frac{I_{on}}{I_{off}}\right) \cdot S_S \tag{5}$$

This result shows that for a certain CMOS technology (and thus for a certain S_S), the minimum supply V_{dd} is only dependent on the minimum I_{on}/I_{off} current ratio. This equation makes it possible to derive a practical as well as a theoretical limit for the minimum feasible supply voltage for a circuit operating in the weak inversion region. The V_{dd} dependence of the I_{on}/I_{off} ratio is logical: the lower V_{dd}, the lower I_{on} will be obtained, and the lower the current ratio will become. From experience, a fair minimum value for the I_{on}/I_{off} current ratio is 50. A lower value of the current ratio becomes problematic, since the circuit robustness in the presence of variations will be compromised. A theoretical limit for the minimum supply voltage can be found through the theoretical lower bound of the sub-threshold slope S_S. In the ideal case, n is equal to 1 and therefore the minimum S_S is equal to 60 mV/decade at room temperature. The theoretical $V_{dd,min}$ can then be calculated to be 101 mV.

However, although devices with an ideal sub-threshold slope are optimal for sub-threshold applications [16], typical S_S values for a bulk CMOS process range from 70 to 120 mV/decade [17],