Self-Timed SRAM for Energy Harvesting Systems

Abdullah Baz, Delong Shang, Fei Xia, and Alex Yakovlev

Microelectronic System Design Group, School of EECE, Newcastle University
Newcastle upon Tyne, NE1 7RU, England, United Kingdom
{Abdullah.baz,delong.shang,fei.xia,alex.yakovlev}@ncl.ac.uk

Abstract. Portable digital systems tend to be not just low power but power efficient as they are powered by low batteries or energy harvesters. Energy harvesting systems tend to provide nondeterministic, rather than stable, power over time. Existing memory systems use delay elements to cope with the problems under different Vdds. However, this introduces huge penalties on performance, as the delay elements need to follow the worst case timing assumption under the worst environment. In this paper, the latency mismatch between memory cells and the corresponding controller using typical delay elements is investigated and found to be highly variable for different Vdd values. A Speed Independent (SI) SRAM memory is then developed which can help avoid such mismatch problems. It can also be used to replace typical delay lines for use in bundled-data memory banks. A 1Kb SI memory bank is implemented based on this method and analysed in terms of the latency and power consumption.

1 Introduction

With the wide advancement in such remote and mobile fields as wireless sensor based applications, microelectronic system design is becoming more energy conscious. This is mainly because of limited energy supply (scavenged energy or low battery) and excessive heat with associated thermal stress and device wear-out. At the same time, the high density of devices per die and the ability to operate with a high degree of parallelism, coupled with environmental variations, create almost permanent instability in voltage supply (cf. Vdd droop), making systems highly power variant. In the not so long past low power design was targeted merely at the reduction of capacitance, Vdd and switching activity, whilst maintaining the required system performance. In many current applications, the design objectives are changing to maximizing the performance within the dynamic power constrains from energy supply and consumption regimes. Such systems can no longer be simply regarded as low power systems, but rather as power adaptive or power resilient systems.

Normally, this kind of system has the following properties: 1) power efficient not just low power; 2) non-deterministic supply voltage (probably with known range, which tends to be low) variable over time. Recently a possible solution is proposed for this kind of system. It is a power elastic system which takes power and energy as dynamic resources [13]. For example, when power is not enough, some of the subsystems could either be powered off or be executed under lower power supplies (Vdds). When power is enough, systems can provide high performance. This means

R. van Leuken and G. Sicard (Eds.): PATMOS 2010, LNCS 6448, pp. 105–115, 2011.

that all tasks in a system are managed based on the power resources, performance requirements, and thermal constraints.

When systems are subjected to varying environmental conditions, with voltage and thermal fluctuations, timing tends to be the first issue affected. Most systems are still designed with global clocking and the design is often made overly pessimistic to avoid failures due to Vdd (timing) variations.

Along with the advent of the nanometre CMOS technology, the continuation of the scaling process is vital to the future development of the digital industries. The International Technology Roadmap for Semiconductors (ITRS) [1] predicts poorer scaling for wires than transistors in future technology nodes. This makes the above worst timing assumption even worse along with power supply voltage drooping [17].

Asynchronous techniques may provide solutions to all these problems. Unlike synchronous systems, asynchronous designs can completely remove global clocking. As a result, asynchronous designs may be more tolerant to timing variations.

The ITRS also predicts that asynchrony will increase with the complexity of on-chip systems. The power, design effort, and reliability cost of global clocks will also make increased asynchrony more attractive. Increasingly complex asynchronous systems or subsystems will thus become more prevalent in future VLSI systems.

In order to fully realize the potential of asynchrony in an environment of variable supply voltage and latencies, system memories may need to be asynchronous together with the computation parts. In this paper, we concentrate on asynchronous SRAM. Our main contributions include: analysing the behaviour of latency in SRAM memory systems under different Vdds, developing asynchronous SRAM memory, and proposing a new method to build delay elements for bundled SRAM memory. We develop a fully Speed Independent (SI) [16] SRAM cell and a bundled SRAM bank technology by using such SI SRAM cells as delay elements.

The remainder of the paper is organized as follows. Section 2 introduces existing asynchronous SRAM memory structures. Section 3 analyses the effects on the latency of the SRAM memory and its controller of different Vdds. Section 4 gives our asynchronous SRAM solutions and implementations, and proposes a new method to build SI delay elements for SRAM memory. Section 5 demonstrates a memory bank and the measurements in terms of latency, power consumption. Section 6 gives the conclusions and the future work.

2 Existing Asynchronous SRAM Memory

Several asynchronous SRAM methods have been reported [5,6,7,8,9].

In [5] a methodology was mostly developed for designing and verifying low power asynchronous SRAM. An SI SRAM cell was alluded to in [5]. This memory cell is different from the conventional six transistor cell [15] and provides the possibility of checking that the data has been stored in memory. The paper however does not explain how the cell needs to be controlled nor does it include a controller design.

[6,7,8,9] focus on asynchronous SRAM memory designs. [6] presents a four-phase handshake asynchronous SRAM design for self-timed systems. It proposes an SI circuit to realize completion detection of reading operations. However, the paper claims that completion detection is not suitable for writing operations. Because the