Innovative Research and Applications in Next-Generation High Performance Computing

Qusay F. Hassan
Mansoura University, Egypt

A volume in the Advances in Systems Analysis, Software Engineering, and High Performance Computing (ASASEHPC) Book Series

Information Science REFERENCE
An Imprint of IGI Global

Published in the United States of America by
Information Science Reference (an imprint of IGI Global)
701 E. Chocolate Avenue
Hershey PA, USA 17033
Tel: 717-533-8845
Fax: 717-533-8661
E-mail: cust@igi-global.com
Web site: http://www.igi-global.com

Library of Congress Cataloging-in-Publication Data

Names: Hassan, Qusay F., 1982- editor.
Title: Innovative research and applications in next-generation high performance computing / Qusay F. Hassan, editor.
Description: Hershey : Information Science Reference, 2016. | Includes bibliographical references and index.
Identifiers: LCCN 2016007158| ISBN 9781522502876 (hardcover) | ISBN 9781522502883 (ebook)
Subjects: LCSH: High performance computing.
Classification: LCC QA76.88 .H38 2016 | DDC 004.1/1--dc23 LC record available at https://lccn.loc.gov/2016007158

This book is published in the IGI Global book series Advances in Systems Analysis, Software Engineering, and High Performance Computing (ASASEHPC) (ISSN: 2327-3453; eISSN: 2327-3461)

British Cataloguing in Publication Data
A Cataloguing in Publication record for this book is available from the British Library.

Chapter 1
Power Optimization Using Clock Gating and Power Gating: A Review

Arsalan Shahid
HITEC University, Pakistan

Muhammad Yasir Qadri
University of Essex, UK

Saad Arif
HITEC University, Pakistan

Saba Munawar
COMSATS Institute of Information Technology, Pakistan

ABSTRACT

The scaling of CMOS technology has continued due to ever increasing demand of greater performance with low power consumption. This demand has grown further by the portable and battery operated devices market. To meet the challenge of greater energy efficiency and performance, a number of power optimization techniques at processor and system components level are proposed by the research community such as clock gating, operand isolation, memory splitting, power gating, dynamic voltage and frequency scaling, etc. This chapter reviews advancements in the dynamic power optimization techniques like clock gating and power gating. This chapter also reviews some architectures and optimization techniques that have been developed for greater power reduction without any significant performance degradation or area cost.

INTRODUCTION

Recent advances in the field of computer architecture have marked power consumption as one of the major design constraints (Farkas, Jouppi, Ranganathan, & Tullsen, 2015). Power optimization is now a prerequisite for not only portable and mobile computing systems but also for advanced multicore platforms (Hager, Treibig, Habich, & Wellein, 2014). An energy-efficient design strives to deliver optimum throughput while minimizing power consumption (Hanumaiah & Vrudhula, 2014). Current research on energy-efficient processor architectures targets various abstraction levels, spanning from integration technology to algorithmic optimization. Integrated Chip (IC) manufacturing technologies

DOI: 10.4018/978-1-5225-0287-6.ch001

have largely improved during the past decades to cater the greater performance demands, while being energy efficient. Modern multiprocessor systems have billions of transistors with operating frequencies in gigahertz (GHz) range (Daud, Ahmad, & Lynn, 2014). Power optimization in such highly integrated systems is not just an option but a basic requirement. A microprocessor experiences both static and dynamic power consumption; the static power is mainly contributed by leakage current in the device, whereas the dynamic power consumption is a function of toggling frequency. A number of techniques have been developed to control the dynamic power of **C**omplementary **M**etal-**O**xide **S**emiconductor (CMOS) circuits in general, and processors in particular such as clock gating (CG), pre-computation, operand isolation, memory splitting, power gating, body biasing, dynamic voltage-frequency scaling, etc. (Arora et al., 2014; Shah & Ahir, 2013).

This chapter introduces and reviews recent advancements in *clock gating (CG)* and *power gating* (PG) techniques in details. The rest of this chapter is organized as follows: the next section presents a background of power optimization techniques and motivation to use clock gating and power gating among them. Then, a detailed review of clock gating techniques is presented in the next section. After that, power gating techniques to optimize power consumption in a microprocessor are described. Finally, the chapter concludes.

BACKGROUND

Reducing the power consumption has become one of very important research interest and a great challenge in various computing platform for the past decade. In this section we will discuss various techniques used for power optimization. Moreover, we will discuss the advantages of power gating and clock gating techniques over the other available techniques.

Power dissipation has emerged as an important factor in the design phase of a microprocessor. Careful and intelligent design is required at different levels of computer system to obtain optimal power performance. Therefore, it is very important to know the sources of energy consumption at different levels of memory hierarchies. Various energy models have been presented to understand the accurate power consumptions by integration with different cycle accurate simulators. One such example is energy models of multilevel cache memory presented by Qadri et al. (Qadri, M. Y., & McDonald-Maier, K. D., 2010). Energy models can also be subdivided in to three types, i.e., CPU level Energy models (Brooks, D. M. et al, 2000), complete system level models and interconnect level energy models. Hence, energy models do provide a very deep level of energy consumption analysis and result in to power optimization. Then comes the Dynamic Power Measurement (DPM) techniques, which can be classified into three subcategories:

1. CPU level DPM,
2. Complete system level DPM, and
3. Parallel system-level DPM.

These techniques target energy consumption reduction at run-time by selectively turning off or slowing down components when the systems are idle or serving light. DPM techniques can be applied in different ways and are applied at different levels; in accordance to which they lie in the three mentioned categories e.g. dynamic voltage scaling changes the processor's supply voltage and similarly frequency

scaling changes the operating frequency at runtime resulting into power optimization. Both of them lies in CPU-level DPM. Clock gating also comes in CPU level DPM. System and parallel level DPM are also further classified into two types:

1. Hardware based DPM, and
2. Software based DPM.

In this chapter, we discuss clock gating and power gating. Clock gating saves power by adding more logic to a circuit to prune the clock tree. Pruning the clock disables portions of the circuitry so that the flip-flops in them do not have to switch states. Clock gating is widely used because it is conceptually simple and a very small additional circuitry is needed to implement it. Moreover, the component can transit from an idle to an active state in only a few cycles. In power gating, reduction in power consumption is obtained by switching off the current to blocks of the circuit that are not in use. In addition to reducing stand-by or leakage power, power gating has the benefit of enabling Iddq testing. Iddq testing is a method for testing CMOS integrated circuits for the presence of manufacturing faults. The following section presents different clock gating techniques along with their design and efficiency.

CLOCK GATING

This technique is based on the concept of powering off the clock to the flip-flops or memory elements which are not taking part in the current processing; therefore, their unnecessary switching is stopped in order to save dynamic power that is required for their speculative toggling (Zhou, Peng, Hou, Wan, & Lin, 2014). This power reduction can also be achieved by minimizing the number of gates used in each stage and hence reducing the operating frequency at the cost of performance degradation.

From 20% to 40% of the total dynamic power consumptions can be saved so performance degradation can be neglected. (Qadri, Gujarathi, & McDonald-Maier, 2009). Therefore, the cost of performance degradation is bearable at this level of power saving and is based on a compromise between these two aspects.

Clock related power accounts for more than 60% of the total dynamic power requirements (Li, Wang, Choi, Park, & Chung, 2010). Clock gating (CG) is most recognized and commonly applied in number of systems. CG is further divided into three subcategories:

1. System level CG,
2. Combinational CG, and
3. Sequential CG (Donno, Ivaldi, Benini, & Macii, 2003).

System level CG requires information of architecture design and works when clock gating is required at system components level like caches, etc. Combinational CG is used at individual register level as shown in Figure 1, where a group or block of registers are gated in parallel through a single gate cell. Sequential CG (SCG) is applicable in components with serial chains of registers as in pipeline stages, etc. Combinational and Sequential CGs are used in industry for ultra-low power processors (Li et al., 2010).

Figure 1. Clock gating scheme
Adapted from Hsu & Lin, 2011.

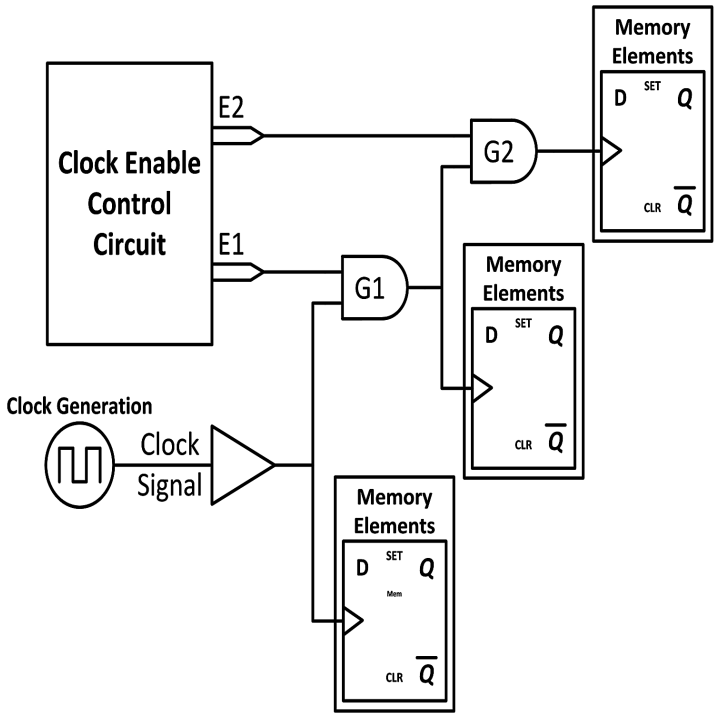


Figure 1 shows a scheme adapted for clock gating in a multi-staged CMOS circuit. Clock control logic block generates the *enable signals* "E1" and "E2" in order to enable clock for multi-staged circuit. "G1" and "G2" are the clock gates which block the clock signal to further stages of flip-flops depending on their usage requirement. Clock control logic decides the depth of stage up to which clock signal is supplied, and unnecessary flip-flops are powered off to avoid speculative toggling, and as a result power saving is achieved.

Huang et al. optimized total power consumption by efficiently designing clock gating circuit (Huang, Tu, & Li, 2012). Control logic used for clock gating has greater circuit area which is also responsible for increased dynamic power consumption. To avoid this, Integer Linear Programming (ILP) is used to optimize both the clock tree and circuit area of clock control logic. In this work, control logic circuit for clock gating is designed by reducing area requirements to 14.5% without loss of performance or increase in power consumption.

Johnson et al. designed a wire-speed power processor SoC (System on Chip) based on IBM 45nm SOI (Silicon on Insulator) technology for inline processing and filtering of data with 2.3GHz core frequency, 16 cores, and 64 threads (Johnson et al., 2010). This chip uses clock gating to achieve AC power saving up to 32% (40W). Brandt et al. developed a technique for identification of speculative or passive executions which are further routed to less power consuming parts through clock gating or operand isolation for low power consumptions (Brandt, Schneider, Ahuja, & Shukla, 2010).

Li et al. introduced Selective Sequential Clock Gating (SeSCG) technique to minimize power consumption in Multimedia Mobile Processor (MMP) designs (Li et al., 2010). SeSCG is used to select optimal sequential clock gating for MMP at Register Transfer (RT) level based on Wasting Toggle Rate

(WTR) analysis. The efficiency of this technique is checked on two real, industrial MMP implementations. Conventional SCG increased total power consumption by 4.77%, while SeSCG reduced it to 23.71% without significant performance degradation.

Hsu et al. described the optimized design of clock gating based on delay-matching technique of gated cells array as compared to the characteristics of type-matching clock gating design (Hsu & Lin, 2011). Their work showed that delay-matching clock gating is much better in achieving better slew and smaller latency with insignificantly changed clock skew and area overhead as compared to type-matching clock gating design. This clock skew is insensitive to process and operating corner variations. The original timing characteristics of the gated tree are preserved by introducing Engineering Change Order (ECO).

Figure 2 shows a clock gating design with type-matching in which all gates at a certain depth level of clock tree are the same. On the other hand, Figure 3 shows clock gating design based on delay-matching in which gated cells are replaced by buffers (inverters) with same timing characteristics.

Koppanalil et al. designed a 1.6GHz dual-core Cortex-A9 processor using 32nm CMOS bulk process (Koppanalil, Yeung, Driscoll, Householder, & Hawkins, 2011). This chip includes power saving schemes like *dynamic voltage-frequency scaling*, *power gating* and *clock gating* for operation at more than GHz frequency and low power. It has all these gates introduced in its all modules to greatly reduce static and dynamic power consumption in this ARM architecture. This Cortex A9 CPU uses Clock Tree Synthesis (CTS) flow for efficient clock distribution in complete tree with hierarchical clock gating schemes.

Figure 2. Type-matching clock gating scheme
Adapted from Hsu & Lin, 2011.

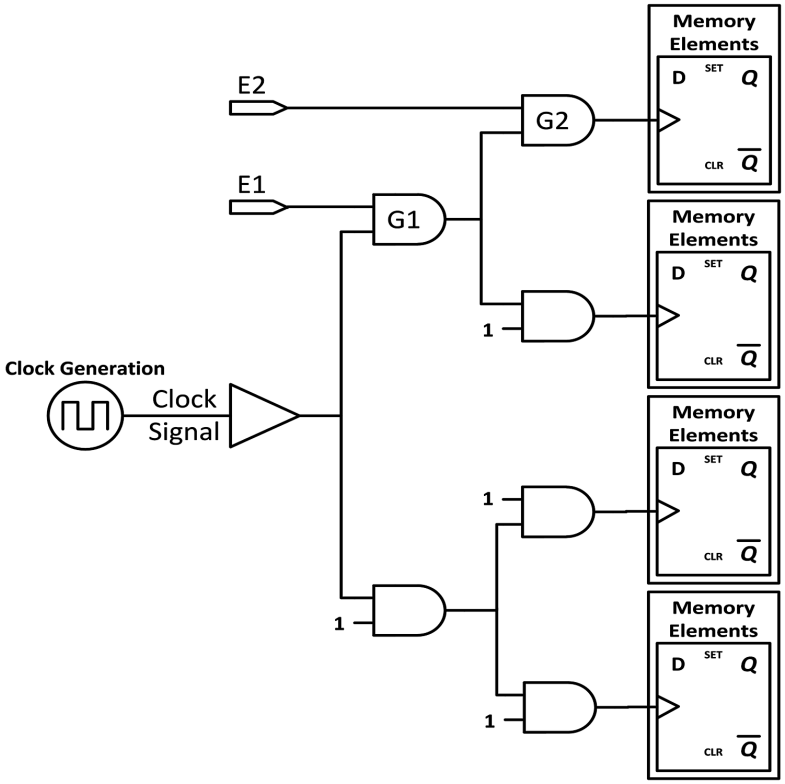
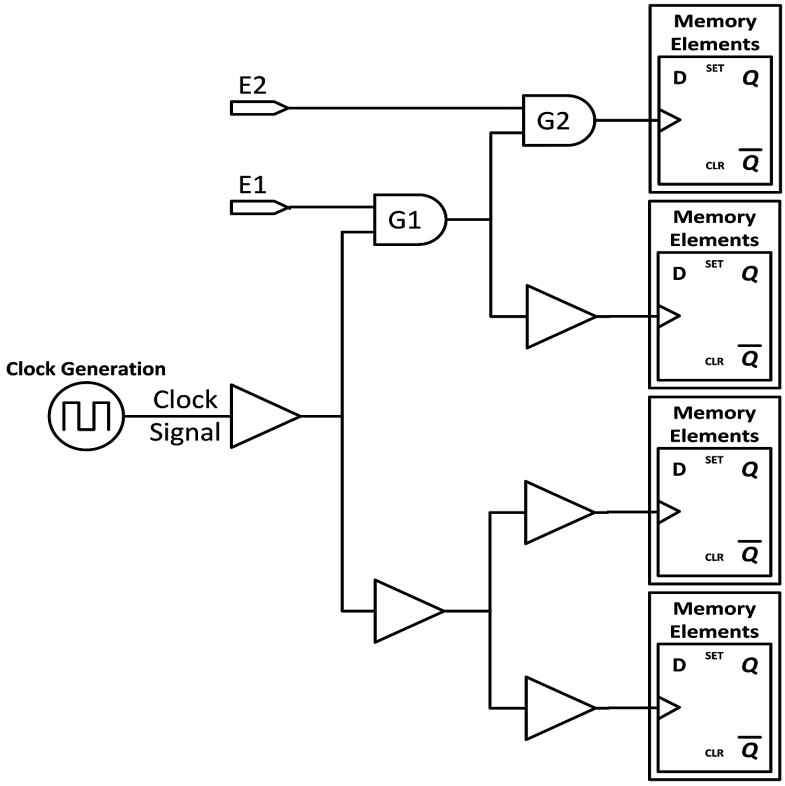


Figure 3. Delay-matching clock gating scheme
Adapted from Hsu & Lin, 2011.



Clock gates are inserted at different levels of clock tree for enhanced power saving. Architecture level blocks like CORE, NEON multimedia, signal processing accelerator, Data Engine (DE) and Snoop Control Unit (SCU) are gated with gates before them. Hence, blocking clock to these in case of idle conditions of processor. Power compiler has also introduced clock gates at register level to achieve enhanced clock gating and hence significantly reduces dynamic power consumptions.

Oliver et al. discussed the experimental results of power consumption measurements of ACTEL's Cyclone III based on 65nm CMOS process and Spartan 6 Field Programmable Gate Arrays (FPGAs) based on 45nm CMOS process (Oliver, Curto, Bouvier, Ramos, & Boemo, 2012). Power saving techniques such as clock gating, clock enable, and blocked inputs, etc., are used to reduce power consumption in standby and active modes. A comparison is performed based on the measurement results of both technologies, in both active and standby modes.

Experimental results showed the trend of power consumption with increasing number of flip-flops in Xilinx Spartan 6 for both active and standby modes. It is clearly evident from the comparison chart that clock gating is the best technique for power reduction in standby mode, and it also behaves well in active mode when compared to clock enable (Oliver et al., 2012). Further results showed the trend of power consumption with increasing number of flip-flops in Cyclone III for both active and standby modes. Here in this case, clock enable is more power saving technique as compared to clock gating. However, clock gating has proved to be the best power optimization technique in standby modes.

Clock gating can also be implemented at gate level rather than RT level. Han et al. developed gate level clock gating using combinational logic (Han & Shin, 2012). This is done by *matching factored forms* in existing combinational logic tree.

Experimental results showed that factored form matching has reduced number of gates used in combinational logic up to 25% as compared to another technique of Boolean division in which this reduction is observed to be only 10%. Consequently, this reduction in gates results in significant dynamic power saving.

Work has also been done on reduction of gate cells in combinational logic for clock gating. Chen et al. developed three phase clock gating technique to generate clock gating tree having minimum number of gate cells and buffers (W.-H. Chen, Chang, Hung, & Hsieh, 2012). Generation of this optimized clock tree is based on clustering and merging algorithms. The proposed scheme is tested on 2GHz Intel Xeon Linux system by using C programming language.

Dillen et al. applied clock gating for power saving in standby mode in Level Sensitive Scan Design (LSSD) with very low area overhead and is implemented on x86 and x64 based AMD microprocessor core called "Bulldozer" (Dillen, Priore, Horiuchi, Naffziger, & others, 2012). The authors of that work presented clock gating scheme for control of flip-flops' scanning operation. This clock gating circuit also provides window transparency. Figure 4 shows field-effect transistor (FET) based complex clock gating circuit with early and late clock gates. The delay between early and late clock signal determines the window transparency time which eventually reduces set-up time, and late arriving data is successfully written to memory elements.

Figure 4. Clock gate for LSSD
Adapted from Dillen et al., 2012.

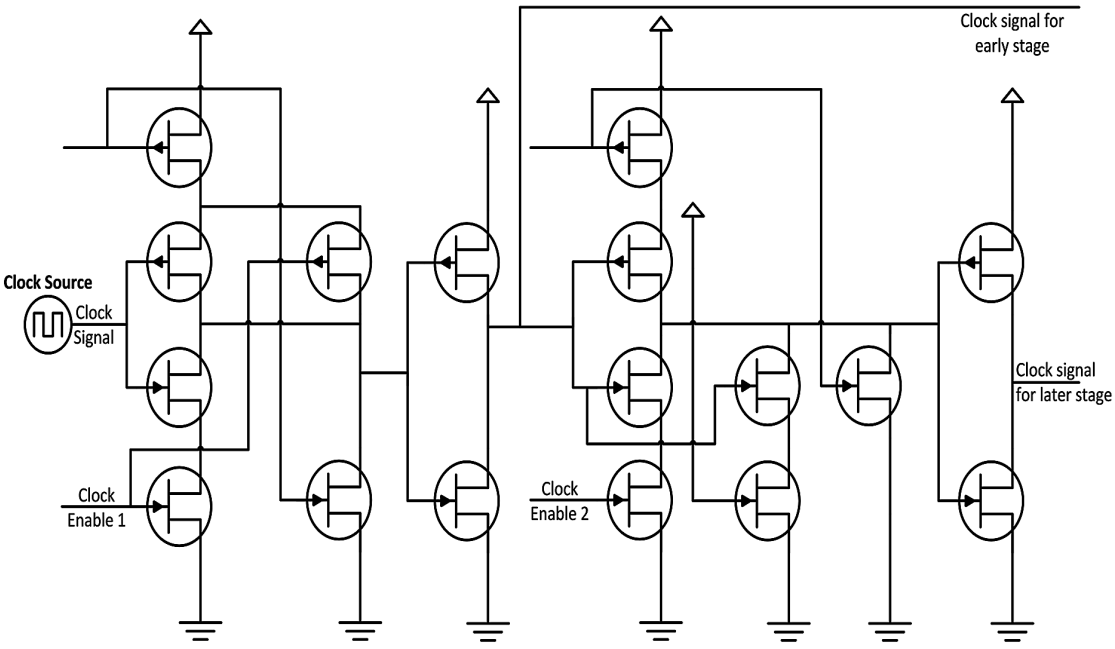
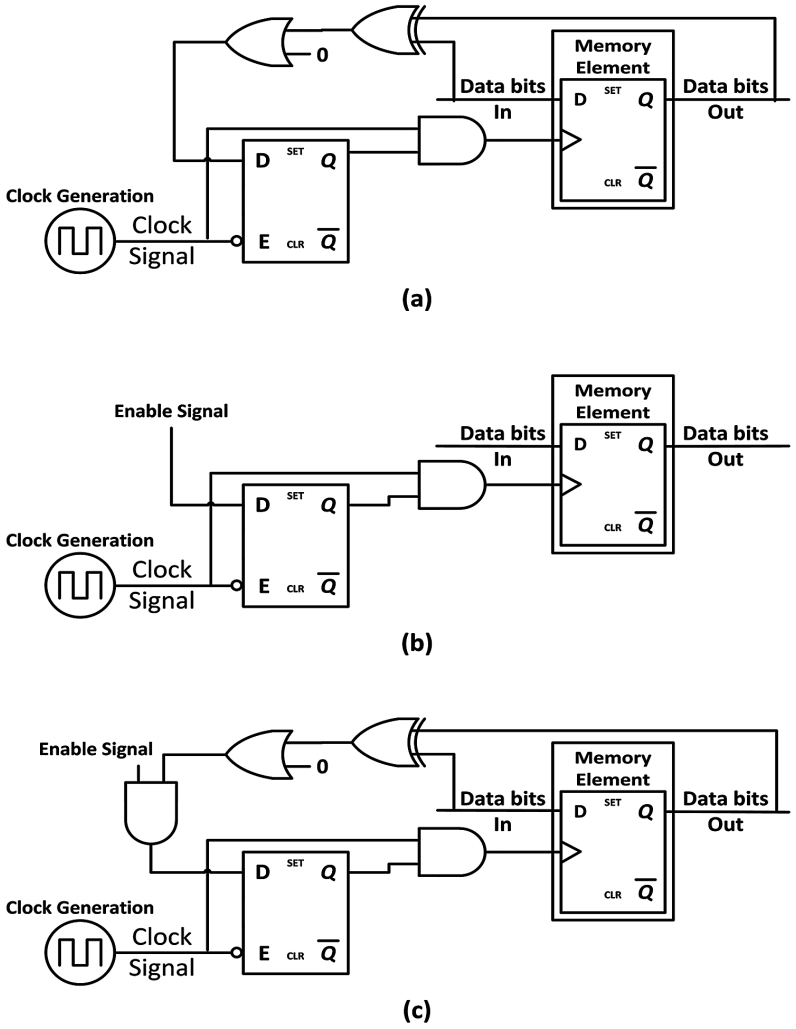


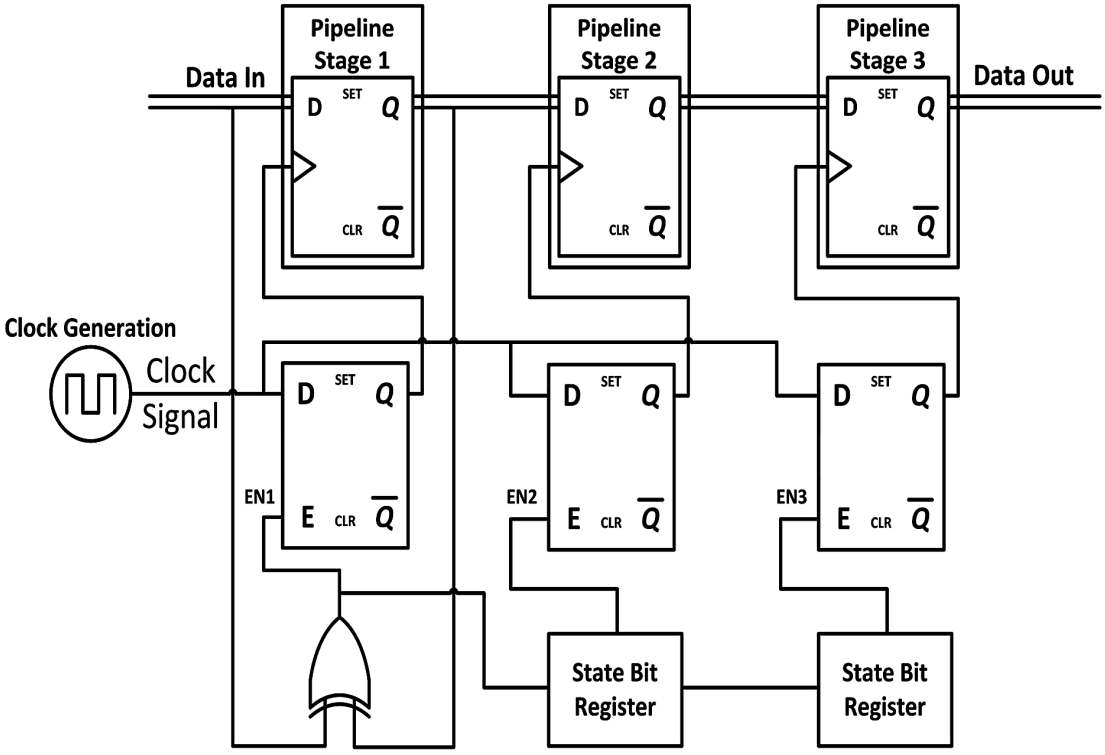
Figure 5. (a) BSCG, (b) LECG, (c) ECG
Adapted from Y. Zhang et al., 2012.



Shelar et al. developed a power optimizing clustering algorithm for power saving in local clock trees for testing on high speed 45nm microprocessor designs (Shelar, 2012). It makes clusters of memory elements and applies clock gating to them. Power reduction up to 14% is achieved by keeping slew and skew factors unchanged. Implementation of this algorithm has achieved up to 32% cost reduction than conventional clock tree synthesis techniques.

Zhang et al. implemented various gate level clock gating optimizations on TSMC 45nm CMOS technology for RT level in Very-large-scale integration (VLSI) designs with effective power reduction (Y. Zhang et al., 2012). *Enable signal* plays an important role in development of number of techniques. Bus Specific Clock Gating (BSCG), Threshold based Clock Gating (TCG) and proposed Optimized Bus Specific Clock Gating (OBSCG) are designed and tested with the absence of *enable signal*. Local Explicit Clock Gating (LECG), Enhanced Clock Gating (ECG), Wasting Toggle Rate (WTR) analysis based clock gating and Single Comparator based Clock Gating (SCCG) are tested when *enable signal* is

Figure 6. SCCG scheme
Adapted from Y. Zhang et al., 2012.



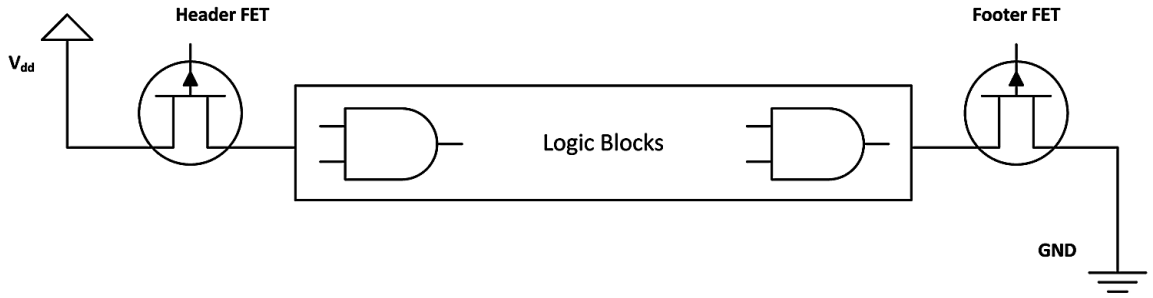
present. Furthermore, new power optimization techniques like *operand isolation* and *memory splitting* are also experimented and tested in this work for above mentioned CMOS technology. Figure 5 shows RT level clock gating schemes for BSCG, LECG and ECG applied on flip-flops. Figure 6 shows clock gating scheme for SCCG in which a single comparator is gating the clock signal to all the stages of pipeline stage registers. OBSCG scheme resulted in power savings up to 26.95% after testing on ISCAS89 bench circuits with Synopsys power compiler (Y. Zhang et al., 2012). Area overhead is increased up to 14.44% while performance degradation in terms of delay incurred is up to 5.77%.

Jotwani et al. designed an AMD x86-64 core implemented in 32nm SOI CMOS technology with a number of design changes and enhancements to improve power requirements which ranges from 2.5W to 25W and hence making the core more suitable for low power mobile devices and products (Jotwani et al., 2010). Enhanced design of clock gating architecture resulted in reduction of clock related power requirement which is now less than 10% of total dynamic power of the chip.

POWER GATING

This technique is based on the concept of powering off the supply voltage to the blocks of flip-flops or memory elements and other system components in idle conditions while system is in standby or sleep mode. Power gating is also applied in active mode for dynamic power reduction as well as static power

Figure 7. Power gating scheme with header/footer switches
Adapted from M. Y. Qadri et al., 2012.



reduction through blocking leakage currents and switching currents while not in use for required results computation. Power switches based on FETs etc. are introduced before and after clusters or logic blocks. Leakage current in active mode heavily takes part in dynamic power consumption which is up to 30% to 40% of total dynamic power consumptions (Seomun, Shin, & Shin, 2010). Figure 7 shows a typical design of power gating circuit with power enabling header and footer switches. State of these switches determines the power supply or blocking to logic block.

Seomun et al. proposed an Active Mode Power Gating (AMPG) scheme to minimize dynamic power consumption because of leakage currents (Seomun et al., 2010). This scheme is a hybrid of power gating and clock gating integrated to actively reduce power consumption due to leakage currents up to 16% in comparison with conventional clock gated schemes, when tested on 45nm CMOS technology. A few constraints like functional, timing and current constraints are also addressed to ensure correct and fast working of the circuit. AMPG is based on power gating with clock gating controls. Figure 8 shows an AMPG scheme layout in which clock gating control signal controls the footer switches which in turn reduces the leakage current of combinational logic. If clock gating controller generates logic to turn off clock transmission, then "EN1" and "EN2" signals become zero, and hence stops clock signal to flip-flops and turns off their unnecessary switching. It also results into disabling of footer switches to stop the leakage currents of combinational logic circuit shown with gates connected by thick power wires. Performance degradation is kept minimized with 6% area overhead due to extra gates and footer switches and 30% wire length overhead due to interleaving wires for combinational logic circuitry, etc. Therefore, AMPG is proved to be a better power saving technique in standby mode as well as active mode for handheld devices like mobile phones and Personal Digital Assistant (PDAs), etc.

Ishihara et al. applied power gating scheme to asynchronous FPGA based on a single two input and one output Look-Up Table (LUT) level fine granularity (Ishihara, Hariyama, & Kameyama, 2011). The proposed scheme for FPGA is designed on ASPLA 90nm CMOS technology. A computer vision technique "Template Matching" is also tested on proposed FPGA design and it also have an added feature of detection of data arrival in order to wake up from sleep mode without delay incurred or unnecessary toggling power consumption. Experimental results of the proposed scheme in comparison with synchronous FPGA and asynchronous FPGA without power gating have showed power reduction up to 38% and 15% respectively at 85oC.

Figure 9 shows the control strategy adopted in proposed scheme. Here, input and output phases of Logic Block (LB) is compared through comparator in phase comparator module. If LB is in operation, then the comparator output will be one which controls the activity switch or FET (power gate) to be

Figure 8. Power gating scheme with clock gating control - AMPG
Adapted from Jun Seomun et al., 2010.

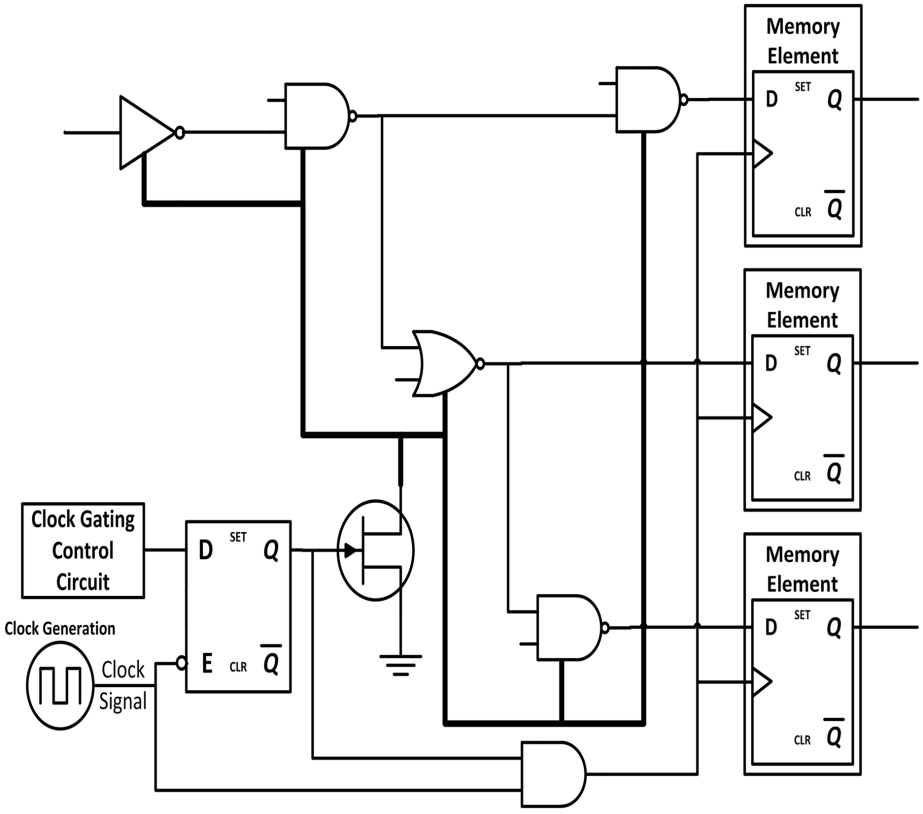
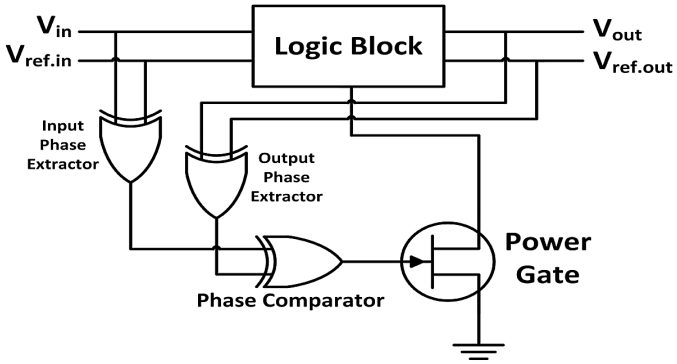
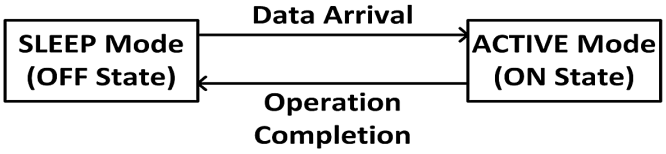


Figure 9. Autonomous power gating for synchronous FPGA (a) circuit, (b) control scheme



(a)



(b)

in ON state. This will keep LB working using the applied supply voltage, and if there is no activity or processing in LB, then the phase comparator will generate zero to stop the power gate in order to switch OFF the power supply. Figure 9(b) shows the adopted control scheme that whenever data arrives at LB, the power gate will switch the LB ON in active mode and after completion of operation, LB gets into sleep mode by getting power OFF signal from power gate. The provision of delay occurrence and power consumption in activating and deactivating power gate is more when compared to the saved power. therefore, to avoid this problem, new modified schemes are also discussed such as introducing an interval led checking of activity of LB, which powers OFF the gate if there is no activity observed on LB for a specific time interval, etc. (Ishihara et al., 2011).

Madan et al. have proposed algorithms based on guarded and predictive power gating for multicore settings at intra-core and inter-core level (Madan, Buyuktosunoglu, Bose, & Annavaram, 2012). Algorithms are developed to maintain proper balance of power gating at these levels in order to maximize the power savings at the cost of minimum performance degradation and area overhead. Power gating experimentation at inter-core level resulted in power savings up to 34% with average performance degradation of 4.8% in terms of response time; whereas, intra-core level power gating experimentation yielded maximum power saving up to 20% with a performance loss of 1%.

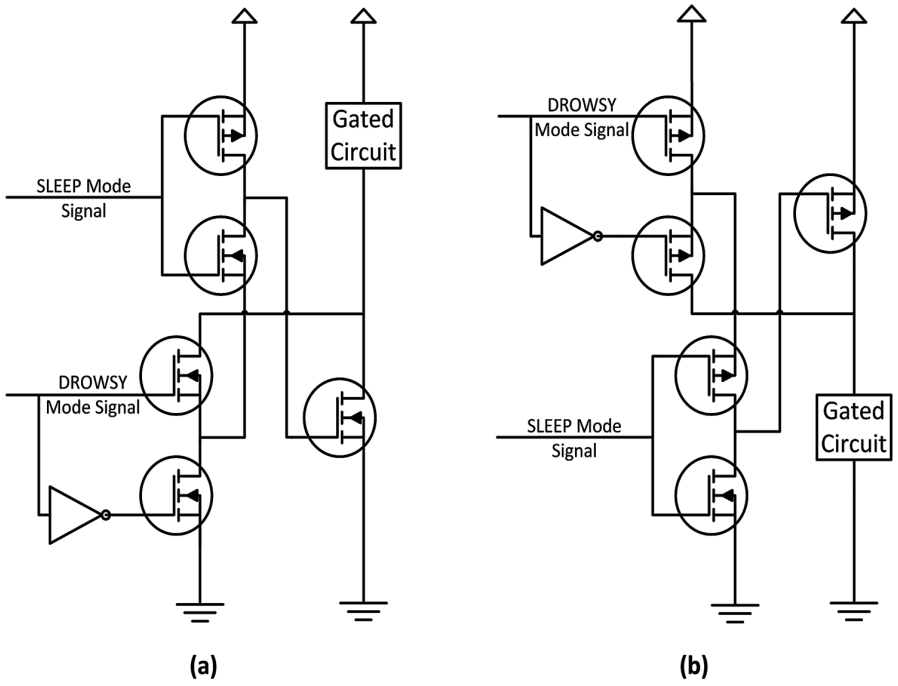
Chen et al. developed an ARM Cortex-M3 core based millimeter scale sensor system with battery and solar cells (G. K. Chen et al., 2010). Cortex-M3 and SRAM is power gated in order to reduce power consumptions during sleep mode and leakage current losses are reduced up to 20%. This power gating saves energy during sleep mode and provides long lasting battery life during measurement operations.

Kang et al. proposed a software based power saving scheme based on integration of DVFS and power gating by minimizing the leakage currents and chip temperature (Kang, Kim, Yoo, & Kyung, 2010). Proposed scheme is tested with two software applications H.264 Decoder and Ray Tracing with benchmark program Equake and power reduction up to 19.4% to 27.2% is observed in comparison with conventional methods.

Pakbaznia et al. presented a tri-modal Multi Threshold CMOS (MTCMOS) based power gating scheme for power saving in active, drowsy and sleep modes (Pakbaznia & Pedram, 2012). Header and footer power switches are introduced for working in data retentive power gating, multi-drowsy mode structures and dynamic voltage scaling for reduction of dynamic power consumption. Experimental results on IBM 90nm technology with Synopsys power compiler shows power savings due to leakage current stopping is up to 50%, 71% and 91% in proposed drowsy, sleep and active mode circuits. Figure 10 shows the design of footer and header switches used for power gating in drowsy, sleep and active modes. When SLEEP signal is OFF then the gated circuit will be in ACTIVE state irrespective of the state of DROWSY signal. When SLEEP signal is ON and DROWSY signal is OFF then gated circuit will enter in SLEEP mode with power gating in action. When both SLEEP and DROWSY signals are ON then gated circuit will enter in DROWSY mode and hence enabling power saving at an intermediate level.

Yip et al. designed a reconfigurable SAR ADC whose power is scalable with resolution reconfiguration (Yip & Chandrakasan, 2013). Power gating is introduced in it which resulted in improved energy efficiency with reduced complexity and cost. This ADC has reconfigurable resolution of 5 to 10-bit and 0.4 to 1V. Exponential trend is achieved in power reduction with respect to resolution in DAC operation which accounts for 25% of total power consumptions. Leakage current losses are minimized by power gating and now total power is reduced up to 14% which was much higher than this without power gating.

Figure 10. (a) Footer power gate, (b) header power gate, for sleep drowsy and active modes
Adapted from Pakbaznia & Pedram, 2012.



Seok et al. investigated power optimization for ultra-low Vdd CMOS circuits during sleep modes by introducing power gating switches with minimal size (Seok, Hanson, Blaauw, & Sylvester, 2012). SPICE simulations of these proposed minimal sized power gating switches showed 125 times reduction in total power consumptions as compared to those achieved by conventional power saving techniques. Furthermore, this proposed scheme is evaluated and validated by taking measurements from ultra-low power microprocessor.

Jeong et al. introduced a new power optimization technique called Memory Access Power Gating (MAPG), which is used to save power during processor stalls without waiting for thread completion when a number of requests to memory are in progress and hence incurring latency to the system (Jeong, Kahng, Kang, Rosing, & Strong, 2012). Battery operated devices such as mobile phones are mostly prone to this type of problem where limited capacity of battery to be consumed efficiently for increased working span. This scheme has resulted in up to 38.07% power reduction. A counter based power gating is practically implemented resulting in up to 22.57% dynamic power reduction in cores without significant performance degradation and area overhead.

Takeda et al. proposed sleep mode control scheme for dynamic power saving due to leakage current and named it as Opt-Static (Takeda, Miwa, Usami, & Nakamura, 2012a). It uses sleep mode with certain depth level which is reconfigurable according to run-time requirements. Experimentation showed that average power savings by reducing leakage current with simple power gating and proposed multi-mode power gating are 34.4% and 44.1% respectively as compared to consumptions of non-power gated circuits. In various cases, multi-mode power gating resulted in more than 25% power savings in addition to

that of conventional power gating scheme. Proposed power gating scheme with depth level has achieved average power saving up to 93.8% as compared to that of ideal multi-mode power gating. Single depth level based power gating has achieved 88.1% power saving without significant degradation.

Wang et al. implemented high speed power gating for reduction of dynamic power consumption due to leakage currents in functional units of CPU MIPS 3000 based cores (Wang, Ohta, Ishii, Usami, & Amano, 2012). Three types of power gating are designed: *cell based*, *row based* and *ring based*. Experiments and simulation results have shown dynamic power savings up to 28% to 54% at 25°C with much smaller area overhead and design costs.

Chang et al. proposed a new power gating scheme based on balanced rush current for an enhanced active wakeup mode (Chang, Tso, Huang, & Yang, 2012). The proposed scheme is evaluated by implementing a 40-bit ALU on TSMC 0.18m CMOS technology and resulted in 10.23% reduction in wake-up time as compared to other power gating schemes without significant performance or power degradation.

Power optimization techniques for single core processors have been investigated in numerous research works. Some of these power saving schemes are also implemented in multi-core multi-threaded architectures. Zyuban et al. deployed power saving models in IBM POWER7 R processor by introducing design changes at microarchitecture, logic and circuits, etc. level (Zyuban et al., 2011). In another work, Roy et al. investigated power savings in multi-core processors having multi-threaded in-order cores by using power gating of register files with state retention capability (Roy, Ranganathan, & Katkoori, 2011). State retention is a powerful feature enabling register files to retain their data during memory stalls and sleep modes treated with power gating to reduce leakage power losses. A technique is proposed in control unit design of in-order cores with three multi-core configurations named as:

1. Coarse-Grained Multi-Threading (CGMT),
2. Fine-Grained Multi-Threading (FGMT), and
3. Simultaneous Multi-Threading (SMT).

SPEC 2000 integer benchmarks are used to evaluate the proposed technique with the help of simulations. Experiments on 8 core processor with 64 threads execution resulted in average power savings by reducing leakage current up to 42%, 7% and 8% in CGMT, FGMT and SMT configurations, respectively. Simulation results on integer register files showed that power saving is mainly dependent upon number of Thread Context (TC) running (Roy et al., 2011). In CGMT configuration, leakage power saving ranges between 0.9% to 2.9% for 2 TCs, and 22% to 42% for 8 TCs. In FGMT configuration, power saving lies between 0.8% to 2.02% for 3 TCs, and 3.09% to 7.8% for 8 TCs. In SMT configuration, power saving is in the range of 1.02% to 2.23% for 4 TCs and 2.97% to 7.27% for 8 TCs with degradation of 0.023% while each core is executing 8 threads in an 8 core processor.

Takeda et al. proposed a multi-depth based sleep control scheme, which decides the depth of sleep mode during runtime by keeping in view the idle duration, program load and temperature, etc. (Takeda, Miwa, Usami, & Nakamura, 2012b). This sleep control scheme is optimized for better leakage savings with stepwise depth selection of sleep mode. Body biasing technique is used here to effectively reduce leakage savings by putting system in sleep mode. Experiments of proposed scheme with floating point benchmarks of FPAlu at 1.0GHz, 75oC resulted in 43% net power saving by reducing leakage current.

Hong et al. proposed a power model for Graphics Processing Unit (GPU) named as Integrated Power and Performance modelling system (IPP) which saves power by considering core temperature increase and by optimizing number of cores to be used for a specific task according to power consumed by each

core (Hong & Kim, 2010). Power saving by power gating of cores in GPUs is also computed here. Experimentation and evaluation resulted in average power savings up to 10.99% in case of optimization of number of cores according to application demand. Power requirements are reduced up to 25.85% by introducing power gating before each core.

Singh et al. presented a power gating scheme with multiple sleep modes for improved leakage reduction in combinational circuits (Singh, Agarwal, Sylvester, & Nowka, 2007). Every sleep mode has unique power and performance level and multi-mode power gating which allows for improved power savings for processors. After simulations, this multilevel sleep mode power gating resulted in 17% extra reduction in leakage losses as compared to single mode power gating results. State retentive mode resulted in 19% reduction of leakage losses with active data retention.

Zhang et al. proposed a new power gating scheme with further improvements in comparison with multiple sleep mode power gating as described in an earlier work (Singh et al., 2007). Analysis of previous scheme of (Singh et al., 2007) resulted in high sensitivity for process changes effecting manufacturing and application level up to only two intermediate sleep levels (Z. Zhang, Kavousianos, Chakrabarty, & Tsiatouhas, 2011). This new multilevel power gating scheme is designed to overcome the drawbacks of previous scheme and allows more than two intermediate sleep levels resulting in increased power savings with lesser complexity and area overhead cost.

Annavaram et al. highlighted few challenges in designing an optimal power management system like negative outcomes and increased performance degradation in terms of response time and area overhead (Annavaram, 2011). Guarded power gating models are proposed in this work to overcome above mentioned challenges. Two proposed schemes are:

1. Idleness-Triggered Per-core Power Gating (IdlePG), and
2. Utilization-based Per-core Power Gating (UtilPG).

IdlePG is based on monitoring of idle state duration in all cores of processor which is very effective in power saving but incurs delay in wakeup duration of cores. So idleness duration threshold should be carefully chosen in this case in order to reduce degradation and negative outcomes. Experimentation showed performance degradation up to 8.4% if wakeup latency is increased up to 100ms. UtilPG is based on sampling system utilization repeatedly in a specific time period and based on this utilization, minimum number of cores required to be active are switched on. This scheme is sensitive to both parameters power saving and system response latency. A maximum performance degradation of 11% is observed with longest utilization sampling time period of 1s during experimentation. Both power gating schemes were tested against a few examples of different workloads for failure analysis, and a guard mechanism is proposed in order to make these schemes fail safe and effective in power saving in all types of workloads and circumstances (Annavaram, 2011).

CONCLUSION

This chapter has reviewed the advancements in power optimization techniques (i.e. clock gating and power gating) being used in a number of systems like multi-core CPUs, CMOS circuits and FPGAs, etc. These techniques have proven to be optimal in achieving static and dynamic power savings without significant performance degradation or additional area overheads. A number of other optimization tech-

niques such as operand isolation, pre-computation, memory splitting and dynamic voltage-frequency scaling are also being developed and implemented with integration of these two techniques to generate hybrid power saving schemes.

Clock gating and power gating proved to be efficient resulting in up to 30% to 40% total dynamic power reduction. These are being used in IBM POWER7 R, ARM (Cortex A9, Cortex M3, Bulldozer), Intel (Xeon), IA and in various other in-order multi-threaded and multi-core processors, and Multi-Mini-Processor (MMPs). Furthermore, power saving techniques for FPGAs such as RT level implementation and gate-based architectures (e.g. type-matching and delay-matching) were discussed.

The power savings schemes discussed in this chapter are widely being used in system-level components like cores, caches, pipelines, buses and memories, etc. They are also being used in advanced multi-core architectures to deliver high grade performance with minimal static and dynamic power consumptions and area overheads.

REFERENCES

Annavaram, M. (2011). A case for guarded power gating for multi-core processors. *Proceedings of the 2011 IEEE 17th International Symposium on High Performance Computer Architecture (HPCA)* (pp. 291–300).

Arora, M., Manne, S., Eckert, Y., Paul, I., Jayasena, N., & Tullsen, D. (2014). A comparison of core power gating strategies implemented in modern hardware. Proceedings of the 2014 ACM international conference on Measurement and modeling of computer systems (pp. 559–560). ACM. doi:doi:10.1145/2591971.2592017 doi:10.1145/2591971.2592017

Brandt, J., Schneider, K., Ahuja, S., & Shukla, S. K. (2010). The model checking view to clock Sgating and operand isolation. *Proceedings of the 2010 10th International Conference on Application of Concurrency to System Design (ACSD)* (pp. 181–190). IEEE.

Chang, C.-Y., Tso, P.-C., Huang, C.-H., & Yang, P.-H. (2012). A fast wake-up power gating technique with inducing a balanced rush current. *Proceedings of the 2012 IEEE International Symposium on Circuits and Systems (ISCAS)* (pp. 3086–3089). IEEE. doi:doi:10.1109/ISCAS.2012.6271972 doi:10.1109/ISCAS.2012.6271972

Chen, G. K., Fojtik, M., Kim, D., Fick, D., Park, J., Seok, M., . . . Blaauw, D. (2010). Millimeter-scale nearly perpetual sensor system with stacked battery and solar cells. Proceedings of ISSCC (Vol. 10, pp. 288–289). doi:doi:10.1109/ISSCC.2010.5433921 doi:10.1109/ISSCC.2010.5433921

Chen, W.-H., Chang, H.-H., Hung, J.-H., & Hsieh, T.-M. (2012). Clock tree construction using gated clock cloning. *Proceedings of the 2012 4th Asia Symposium on Quality Electronic Design (ASQED)* (pp. 54–58). IEEE.

Daud, S., Ahmad, R. B., & Lynn, O. B. (2014). CPU Power Prediction on Modern Multicore Embedded Processor. *International journal of Computer Science and Mobile Computing,* 3(2), 709–715. IJCSMC.

Dillen, S. J., Priore, D., Horiuchi, A. K., & Naffziger, S. D. (2012, September). Design and implementation of soft-edge flip-flops for x86-64 AMD microprocessor modules. *Proceedings of theCustom Integrated Circuits Conference CICC'12* (pp. 1-4). IEEE. doi:doi:10.1109/CICC.2012.6330707 doi:10.1109/CICC.2012.6330707

Donno, M., Ivaldi, A., Benini, L., & Macii, E. (2003). Clock-tree power optimization based on RTL clock-gating. Proceedings of the Design Automation Conference (pp. 622–627). IEEE.

Farkas, R. K. K., Jouppi, N. P., Ranganathan, P., & Tullsen, D. M. (2015). A Multi-Core Approach to Addressing the Energy-Complexity Problem in Microprocessors.*Proceedings of Workshop on Complexity-Effective Design*.

Qadri, M. Y., & McDonald-Maier, K. D. (2010, March). Analytical Evaluation of Energy and Throughput for Multilevel Caches. *Proceedings of the 2010 12th International Conference on Computer Modelling and Simulation (UKSim)* (pp. 598-603). IEEE. doi:10.1109/UKSIM.2010.114

Brooks, D. M., Bose, P., Schuster, S. E., Jacobson, H., Kudva, P. N., Buyuktosunoglu, A., & Cook, P. W. (2000). Power-aware microarchitecture: Design and modeling challenges for next-generation microprocessors. *Micro*, *20*(6), 26–44. doi:10.1109/UKSIM.2010.114

Hager, G., Treibig, J., Habich, J., & Wellein, G. (2014). Exploring performance and power properties of modern multi-core chips via simple machine models. *Concurrency and Computation*.

Han, I., & Shin, Y. (2012). Synthesis of clock gating logic through factored form matching. *Proceedings of the 2012 IEEE International Conference on IC Design & Technology (ICICDT)* (pp. 1–4). IEEE. doi:doi:10.1109/ICICDT.2012.6232835 doi:10.1109/ICICDT.2012.6232835

Hanumaiah, V., & Vrudhula, S. (2014). Energy-efficient operation of multicore processors by DVFS, task migration, and active cooling. *IEEE Transactions on Computers*, *63*(2), 349–360. doi:10.1109/TC.2012.213

Hong, S., & Kim, H. (2010). An integrated GPU power and performance model. In *ACM SIGARCH Computer Architecture News* (Vol. 38, pp. 280–289). ACM. doi:10.1145/1815961.1815998

Hsu, S.-J., & Lin, R.-B. (2011). Clock gating optimization with delay-matching. Proceedings of the Design, Automation & Test in Europe Conference & Exhibition (pp. 1–6). IEEE.

Huang, S.-H., Tu, W.-P., & Li, B.-H. (2012). High-level synthesis for minimum-area low-power clock gating. *Journal of Information Science and Engineering*, *28*(5), 971–988.

Ishihara, S., Hariyama, M., & Kameyama, M. (2011). A low-power FPGA based on autonomous fine-grain power gating. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, *19*(8), 1394–1406.

Jeong, K., Kahng, A. B., Kang, S., Rosing, T. S., & Strong, R. (2012). MAPG: Memory access power gating.*Proceedings of the Conference on Design, Automation and Test in Europe* (pp. 1054–1059). EDA Consortium.

Johnson, C., Allen, D.H., Brown, J., Vanderwiel, S., Hoover, R., Achilles, H., Cher, CY., May, G.A., Franke, H., Xenidis, J. & Basso, C. (2010). 5.5 A Wire-Speed PowerTM Processor: 2.3 GHz 45nm SOI with 16 Cores and 64 Threads.

Jotwani, R., Sundaram, S., Kosonocky, S., Schaefer, A., Andrade, V., Constant, G., & Naffziger, S. (2010). An x86-64 core implemented in 32nm SOI CMOS. *Proceedings of the 2010 IEEE International Solid-State Circuits Conference-(ISSCC)*. IEEE.

Kang, K., Kim, J., Yoo, S., & Kyung, C.-M. (2010). Temperature-aware integrated DVFS and power gating for executing tasks with runtime distribution. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, *29*(9), 1381–1394. doi:10.1109/TCAD.2010.2059290

Koppanalil, J., Yeung, G., Driscoll, D. O., Householder, S., & Hawkins, C. (2011). A 1.6 GHz dual-core ARM Cortex A9 implementation on a low power high-K metal gate 32nm process. Proceedings of the 2011 International Symposium on VLSI Design, Automation and Test (VLSI-DAT) (pp. 1–4). IEEE.

Li, L., Wang, W., Choi, K., Park, S., & Chung, M.-K. (2010). SeSCG: Selective sequential clock gating for ultra-low-power multimedia mobile processor design. *Proceedings of the 2010 IEEE International Conference on Electro/Information Technology (EIT)* (pp. 1–6). IEEE.

Madan, N., Buyuktosunoglu, A., Bose, P., & Annavaram, M. (2012). Guarded power gating in a multi-core setting. In *Computer Architecture* (pp. 198–210). Springer.

Oliver, J. P., Curto, J., Bouvier, D., Ramos, M., & Boemo, E. (2012). Clock gating and clock enable for FPGA power reduction. *Proceedings of the 2012 VIII Southern Conference on Programmable Logic (SPL)* (pp. 1–5). IEEE. doi:doi:10.1109/SPL.2012.6211782 doi:10.1109/SPL.2012.6211782

Pakbaznia, E., & Pedram, M. (2012). Design of a tri-modal multi-threshold CMOS switch with application to data retentive power gating. *IEEE Transactions on Very Large Scale Integration Systems*, *20*(2), 380–385.

Qadri, M. Y., Gujarathi, H. S., & McDonald-Maier, K. D. (2009). Low Power Processor Architectures and Contemporary Techniques for Power Optimization–A Review. *Journal of Computers*, *4*(10), 927–942. doi:10.4304/jcp.4.10.927-942

Roy, S., Ranganathan, N., & Katkoori, S. (2011). State-retentive power gating of register files in multicore processors featuring multithreaded in-order cores. *IEEE Transactions on Computers*, *60*(11), 1547–1560. doi:10.1109/TC.2010.249

Seok, M., Hanson, S., Blaauw, D., & Sylvester, D. (2012). Sleep Mode Analysis and Optimization with Minimal-Sized Power Gating Switch for Ultra-Low Operation. *IEEE Transactions on Very Large Scale Integration Systems*, *20*(4), 605–615.

Seomun, J., Shin, I., & Shin, Y. (2010). Synthesis and implementation of active mode power gating circuits. *Proceedings of the 2010 47th ACM/IEEE Design Automation Conference (DAC)* (pp. 487–492). IEEE. doi:doi:10.1145/1837274.1837395 doi:10.1145/1837274.1837395

Shah, N. A., & Ahir, D. H. (2013). Power Consumption Reduction using Microarchitecture Techniques for Modern Multicore Microprocessors. *Current Trends in Information Technology*, *3*(3), 16–20.

Shelar, R. S. (2012). A fast and near-optimal clustering algorithm for low-power clock tree synthesis. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, *31*(11), 1781–1786.

Singh, H., Agarwal, K., Sylvester, D., & Nowka, K. J. (2007). Enhanced leakage reduction techniques using intermediate strength power gating. *IEEE Transactions on Very Large Scale Integration Systems*, *15*(11), 1215–1224.

Takeda, S., Miwa, S., Usami, K., & Nakamura, H. (2012a). Efficient leakage power saving by sleep depth controlling for Multi-mode Power Gating. *Proceedings of the 2012 13th International Symposium on Quality Electronic Design (ISQED)* (pp. 625–632). IEEE. doi:doi:10.1109/ISQED.2012.6187558 doi:10.1109/ISQED.2012.6187558

Takeda, S., Miwa, S., Usami, K., & Nakamura, H. (2012b). Stepwise sleep depth control for run-time leakage power saving.*Proceedings of the great lakes symposium on VLSI* (pp. 233–238). ACM. doi:10.1145/2206781.2206838

Wang, W., Ohta, Y., Ishii, Y., Usami, K., & Amano, H. (2012). Trade-off analysis of fine-grained power gating methods for functional units in a CPU. Proceedings of IEEE 2012 Cool Chips XV (COOL Chips), (pp. 1–3). IEEE. doi:doi:10.1109/COOLChips.2012.6216587 doi:10.1109/COOLChips.2012.6216587

Yip, M., & Chandrakasan, A. P. (2013). A resolution-reconfigurable 5-to-10-bit 0.4-to-1 V power scalable SAR ADC for sensor applications. *IEEE Journal of Solid-State Circuits*, *48*(6), 1453–1464. doi:10.1109/JSSC.2013.2254551

Zhang, Y., Tong, Q., Li, L., Wang, W., Choi, K., Jang, J. E., . . . Ahn, S. Y. (2012). Automatic Register Transfer level CAD tool design for advanced clock gating and low power schemes. *Proceedings of the 2012 International SoC Design Conference (ISOCC)* (pp. 21–24). IEEE.

Zhang, Z., Kavousianos, X., Chakrabarty, K., & Tsiatouhas, Y. (2011). A robust and reconfigurable multi-mode power gating architecture. *Proceedings of the 2011 24th International Conference on VLSI Design (VLSI Design)* (pp. 280–285). IEEE. doi:doi:10.1109/VLSID.2011.29 doi:10.1109/VLSID.2011.29

Zhou, Y., Peng, X., Hou, L., Wan, P., & Lin, P. (2014). Clock gating-A power optimization technique for smart card. *Proceedings of the 2014 12th IEEE International Conference on Solid-State and Integrated Circuit Technology (ICSICT)* (pp. 1–3). IEEE. doi:doi:10.1109/ICSICT.2014.7021594 doi:10.1109/ICSICT.2014.7021594

Zyuban, V., Friedrich, J., Gonzalez, C. J., Rao, R., Brown, M. D., Ziegler, M. M., & Culp, J. A. (2011). Power optimization methodology for the IBM POWER7 microprocessor. *IBM Journal of Research and Development*, *55*(3), 7–11. doi:10.1147/JRD.2011.2110410

KEY TERMS AND DEFINITIONS

Clock Gating: A Technique of reducing power consumption by addition of extra logic in the circuit; resulting in pruning the clock tree.

Logic Circuit: A basic building block of a real world computing platform.

Logic Gate: A basic building block of digital system having specific number of inputs and a outputs based on a certain logic.

Multi-Core System: A processor system composed up of multiple central processing units in a form of a single computing package for reading and executing instructions.

Power Gating: A technique to use integrated circuit design for reducing the power consumption by switching off the blocks of circuit which are not being used.

Power Optimization: The process of reducing the power consumption of a digital electronic system by the use of electronic design automation tools.