Portland State University

PDXScholar

Dissertations and Theses

Dissertations and Theses

Winter 2-13-2013

Just-In-Time Power Gating of GasP Circuits

Prachi Gulab Padwal
Portland State University

Follow this and additional works at: https://pdxscholar.library.pdx.edu/open_access_etds

Part of the [Electrical and Computer Engineering Commons](#), and the [Other Computer Sciences Commons](#)

Let us know how access to this document benefits you.

Recommended Citation

Padwal, Prachi Gulab, "Just-In-Time Power Gating of GasP Circuits" (2013). *Dissertations and Theses*. Paper 211.

<https://doi.org/10.15760/etd.211>

This Thesis is brought to you for free and open access. It has been accepted for inclusion in Dissertations and Theses by an authorized administrator of PDXScholar. Please contact us if we can make this document more accessible: pdxscholar@pdx.edu.

Just-In-Time Power Gating of GasP Circuits

by

Prachi Gulab Padwal

A thesis submitted in partial fulfillment of the
requirements for the degree of

Master of Science
in

Electrical and Computer Engineering

Thesis Committee:
Xiaoyu Song, Chair
Ivan E. Sutherland
W. Robert Daasch

Portland State University
2013

© 2013 Prachi Gulab Padwal

Abstract

In modern integrated circuits, one way to reduce power consumption is to turn off power to parts of the circuit when those are idle. This method is called *power gating*. This thesis presents a state-preserving technique to achieve power savings in GasP family of asynchronous circuits by turning off the power when the circuit is idle.

The power control logic turns on the power in anticipation of the receiving data. The power control logic turns off the power when the stage is idle either because it is empty or because the pipeline is clogged. The low logical effort of GasP circuits makes just-in-time power gating possible on a stage-by-stage basis.

A new latch called *Lazy Latch* is introduced in this thesis. The lazy latch preserves its output and permits power gating of its larger transistors. The lazy latch is power efficient because it drives strongly only when necessary. A new latch called *Blended Latch* is proposed in this thesis which blends the advantages of the Conventional latches and the Lazy latches.

Performance of power gating is evaluated by comparing the power-gated pipeline against the non-power gated pipeline. Power savings achieved are dependent on the duty cycle of operation. The fact that just-in-time power gating achieves power savings after it is idle for a minimum of 10^6 cycles makes it useful in limited applications where a quick start is required after long idle times.

Acknowledgements

I would like to take this opportunity to thank all the wonderful people around me for their guidance and assistance during my graduate study at Portland State University. First and foremost, I would like to thank my research mentor Dr. Ivan E. Sutherland. This research would not have been possible without his constant guidance and motivation. His immense knowledge and research experience in the field of digital circuits not only assisted me in achieving my goal but also improve many aspects of the education like problem solving and writing skills. One of the most important things I learned from him is to THINK. I hope to inculcate the thinking process in every aspect of my life.

I would like to extend my gratitude to Professor Xiaoyu Song, my academic advisor, for his support and help during the period of my thesis. I would also like to thank Professor Rob Daasch for his valuable inputs in the process of this thesis and for serving on my M.S. Defense Examination Committee.

I would like to show my immense gratitude to Marly Roncken for being such a wonderful person she is. I admire her ability and efforts to make everyone around her feel comfortable. I would like to thank Willem Mallon for his wonderful idea of the reading group and all interesting sessions of general knowledge. I am indebted to my colleagues and friends Swetha, Navaneeth and Hoon for their assistance and fun times at Asynchronous Research Center (ARC). Special thanks to Rajesh and Chaitrali for their encouragement and wonderful times we spent together.

Finally, I owe my deepest gratitude to all the people closest to my heart. I would like to thank my parents for their constant efforts and blessings that encouraged me to pursue my graduate studies. I would like to thank my younger brother for keeping me grounded. I cannot thank Kapil enough for his constant support and great patience at all the times.

Contents

Abstract	i
Acknowledgements	iii
List of Figures	vi
1 Introduction: Need, Contribution and Organization	1
1.1 Need for Power Gating	1
1.2 Brief Description and Contribution of this Research	3
1.3 Organization of Thesis	4
2 Background	6
2.1 Two-phase bundled data convention	6
2.2 GasP circuits: Asynchronous Controllers	8
2.3 Power Gating	12
3 Latches	18
3.1 Conventional Latch	18
3.2 Lazy Latch	20
3.2.1 Lazy Latch: Design and Operation	21
3.2.2 Lazy Latch: Switch design details for 100 Latches	24
3.2.3 Lazy Latch: Fire Load	26
3.2.4 Drawbacks of Lazy Latch	27
3.3 Blended Latch	27

4	Methodology - Power Gating of the GasP Circuits	32
4.1	Power-GasP circuit	35
4.1.1	Turn ON the Power	36
4.1.2	Analysis of effect of step-up and amplification	38
4.1.3	Turn OFF the Power	41
4.1.4	Operation: Power-GasP circuit	41
4.2	Power-gating Block	43
5	Evaluation of Power-Gating Technique	47
5.1	Experimental Setup	47
5.1.1	Simulation Environment	47
5.1.2	Design Comparisons and Decisions	47
5.2	Why Just-In-Time	48
5.3	Results: Comparing Conventional Pipeline and Power-gated Pipeline	51
6	Conclusions	58
6.1	Contributions of this work	58
6.2	Power gating technique implementation requirement	59
6.3	Future Work	60
	References	63

List of Figures

1.1	Total chip power dissipation by ITRS	2
2.1	Two-phase bundled data convention	7
2.2	Two-phase bundled data convention(Waveform)	8
2.3	6/4 GasP circuit	9
2.4	6/4 GasP circuits(Waveform)	12
2.5	Block Diagram of power gating	13
3.1	Conventional Latch	19
3.2	Lazy Latch: Top level	21
3.3	Lazy Latch: Internal circuit diagram	22
3.4	Switch: Asymmetric gates for 100 latches.	25
3.5	Blended Latch: Circuit Diagram	28
3.6	Table 3.1. Output and fire loads	30
3.7	Table 3.2. Blend factor effect	31
4.1	Block Diagram: Power Gating of GasP circuit	34
4.2	Power-GasP circuit	36
4.3	Circuit Diagram of Power-gating Block	38
4.4	Operation: Power-GasP circuit	42
4.5	Circuit Diagram of Power-gating Block	44
5.1	Just-In-Time	50
5.2	Periods of Operation	51
5.3	Simulated circuit of the Conventional Pipeline	53

5.4	Simulated circuit of the Power-gated Pipeline	54
5.5	Comparison of table of currents	55

Introduction: Need, Contribution and Organization

1.1 Need for Power Gating

Continual scaling of process technologies significantly improves chip performance and increases density allowing more computations in a smaller area. However, technology scaling also leads to a significant increase in power consumption. Today battery life is a commonly cited requirement for daily electronic gadgets like cell phones and laptops. Also in servers and advanced computing machines, power consumption results in heat dissipation resulting in expensive cooling techniques [1]. All these factors have compelled design engineers to focus on low power design methodologies for digital circuits.

Power consumption in digital circuits can be classified into two main types: dynamic and static power consumption. Charging and discharging transistor capacitance and wire capacitance leads to dynamic power dissipation. Technology scaling requires reducing the supply voltage to avoid the destruction of transistors due to high electric field. Supply voltage scaling also saves significant amount of dynamic power but at the cost of performance. To maintain performance, threshold voltage should be scaled. However, a decrease in threshold voltage increases leakage. Threshold voltage and leakage have an exponential relationship and hence leakage power, commonly referred as static power, is becoming a larger fraction of the total power consumption. Leakage power was recorded to be 40% of the total power consumption for 130-nm CMOS technology [11]. Figure 1.1 shows total chip

power trend based on ITRS (2002) [2]. The plot shows that there is no significant increase in dynamic power with gate length or technology scaling. The dynamic power increase assumes a doubling of on-chip devices every two years. Subthreshold leakage and gate-oxide leakage constitute the static power. The two power plots for static power represent the 2002 ITRS projections normalized to those for 2001. The static power increases exponentially with decreasing gate length or technology scaling [3].

Figure 1.1: Total chip dynamic and static power dissipation trends from ITRS (2002)

Significant research has been done to propose various power reduction techniques. The leakage current can be reduced by using process- and circuit-level techniques [4]. At the process level, leakage reduction can be achieved by controlling the dimensions such as length, oxide thickness, junction depths, and doping profile in transistors. At the circuit level, threshold voltage and leakage current of transistors

can be controlled by controlling the voltages of different device terminals such as drain, source, gate and body. Major circuit level techniques include transistor stacking, multiple V_t , dynamic V_t , supply voltage scaling and power gating.

Power gating is the most effective way to achieve low-power performance [5]. When the circuit is idle or in sleep mode, power to the circuit is switched OFF by controlling an additional transistor called the sleep transistor. Because power is proportional to the square of the voltage supply, power gating the circuit reduces a significant amount of leakage power. Further, using high threshold voltage transistors as sleep transistors and low threshold voltage transistors for logic implementation provides additional power savings while maintaining performance.

The key to the successful power gating implementation is to keep the circuit ON, only as long as required to carry out its functions, and switch OFF power as soon as the circuit completes the work. The ultimate goal is to achieve maximum power saving by applying power gating to smaller parts of circuits without hindering performance.

1.2 Brief Description and Contribution of this Research

This thesis introduces a new Just-In-Time Power Gating technique for the GasP family of asynchronous circuits to achieve power savings. The technique introduced achieves power savings by reducing the subthreshold leakage in the idle state of the circuit. Additional transistors are inserted between the supply and the circuit. When the circuit is idle, the additional transistors cut-off the power to the circuit. Turning the power to the circuit OFF reduces the voltages at different nodes in the circuit. Subthreshold leakage in a transistor is dependent on the voltage across

it. Therefore, reduced node voltages reduce the subthreshold leakage. Turning the power to the circuit OFF to reduce leakage power is commonly referred as *power gating*.

Individual stages in the pipeline are power gated when the stages are idle. Turning OFF the power may cause the loss of states in the pipeline. Keepers are added for each stage to avoid the loss of state. The stages are turned ON in anticipation of the incoming data. The power for a stage is restored just in time when the latches are driven. Benefits of turning the power OFF for individual stages of the pipeline are:

- Individual stages are turned OFF when the pipeline is EMPTY.
- Individual stages are turned OFF when the pipeline is FULL.
- The power is turned ON only when it is required.
- No loss of state when the power for the circuit is turned OFF.

1.3 Organization of Thesis

This thesis has been laid out in six chapters with different subsections. The next chapter gives a detailed description of the asynchronous controllers used in this research work: GasP circuits. It also describes the concepts of power gating and previous work in synchronous and asynchronous circuits. The third chapter describes the Conventional latch and introduces a new latch design called the Lazy latch. Further, a new latch called the Blended latch is also proposed in third chapter. The fourth chapter of this thesis describes the power gating technique developed for the GasP circuits. The fifth chapter of this thesis discusses different

design considerations and evaluates the power savings achieved from the developed techniques. This thesis concludes by summarizing the benefits and requirements of the power gating technique developed in this work.

Background

2.1 Two-phase bundled data convention

The GasP circuits use two-phase bundled data convention but with a single wire [6]. The block diagram for the two-phase bundled data convention is shown in Figure 2.1. The sender and receiver use transition signaling for communication, treating both rising and falling transitions as *events*. The sender puts valid data on the data wires and generates an event, either a rising or a falling transition, on the control wire called *request*. The receiver sees the request, accepts the data and produces an event on the control wire called *acknowledge*. This is the end of one transaction. The rising and falling events alternately follow each other for successive transactions. The two phases of this data convention are defined as: sender's active phase and receiver's active phase. A request signal on a control wire ends the sender's active phase. The data are free to change during sender's active phase. An acknowledge ends the receiver's active phase. The receiver can accept the data during the receiver's active phase. The bundled data part of the name indicates that the data wires and the request signal must be treated as a bundle. Data should reach the receiver and stabilize prior to the request signal reaching the receiver.

Figure 2.1: Block Diagram for a two-phase bundled data convention.

Figure 2.2 shows the waveform depicting two-phase bundled data convention. Every cycle consists of three events as shown in the figure with circled numbers.

1. Senders active phase where data may change.
2. Once correct values of data are established, the active phase of the sender comes to an end by the request event. After this event, sender must hold the data constant.
3. The receiver sees the request event, accepts the data and ends the active phase by sending an acknowledge event. This marks the end of one cycle. The next cycle follows by reversing the level on the request signal to report an event.

Figure 2.2: Waveforms depicting two-phase bundled data convention.

2.2 GasP circuits: Asynchronous Controllers

The GasP family of asynchronous circuits provides a control circuit for simple pipelines, for branching and joining pipelines and for join on demand through arbitration [7]. The GasP circuits follow the two-phase bundled data convention but with a single wire.

Figure 2.3 shows two GasP modules connected by a single wire W called the *state wire*. GasP utilizes the wire capacitance for storing states instead of flip-flops and latches used to store states conventionally. The transistors merely act as switches or paths to copy states from one storage location to other. In short, the state wires connecting the GasP modules act as storage whereas the GasP modules themselves act as switches controlling the flow of data.

Figure 2.3: Two stages of the 6/4 GasP circuits.

A GasP module switches when its predecessor stage has new data to proffer to the successor stage. The necessary condition for switching is: predecessor stage is FULL and successor stage is EMPTY. Various encoding can be used for FULL and EMPTY states. The encoding convention used in this document is: HI is FULL and LO is EMPTY.

The condition, predecessor is FULL and successor is EMPTY, is detected by an AND gate. When pred1 is HI and succ1 is LO, MODULE 1 is triggered and generates a fire signal, fire1, an ACTIVE high signal. A GasP module must achieve three things when it fires: (1) it must make data latches momentarily transparent, (2) it must declare its successor data wires FULL, and (3) it must declare its predecessor data wires EMPTY. So fire1 is followed by succ1 going HI and pred1 going LO. So the very condition that caused the fire signal is destroyed by the fire

signal. The fire signal remains HI for five gate delays. It is the responsibility of the latches to transfer the data in this time. The five gates (Gates 1, 2, 3, 4, 5 and Gates 1, 2, 3, 6, 7) form a ring oscillator whose oscillations are controlled by the NOR gate 1. Five stages are used in the ring oscillator to allow greater disparity between the sizes of the small and fast gates [8].

Operation of the GasP circuit relies on balancing the widths of the transistors in the circuit. The transistors are sized to make each logic gate operate in fixed time. The concept of logical effort [9] is used as the basis for designing the GasP circuits and all other circuits in this thesis. One might argue that fixed known time of completion hinders the very basis of the asynchronous circuits. But circuits designed to operate in fixed known time are logically simpler and faster than circuits that check completion every cycle [7]. GasP circuits and all the circuits in this thesis are designed for uniform gate delays. Hence counting the number of gates reflects timing delay.

The GasP circuit described in Figure 2.3 is a 6-4 GasP circuit. It takes six gate delays from predecessor FULL to successor FULL (Gates 1, 2, 3, 4, and 5 of MODULE 1 and 7 of MODULE 2). This is called *forward latency* shown by the blue arrow in the figure. Also, it takes four gate delays from successor EMPTY to predecessor EMPTY (Gates 1, 2, 3, 6 of Module 2). This is called *reverse latency* shown by the red arrow. Hence the name 6-4 GasP. More combinations of forward-reverse latencies, such as 2-8, 3-7, 4-6, 5-5, 7-3, 8-2, are possible. Note that the sum of forward and reverse latencies is ten. To make faster circuits it is convenient to have longer forward latencies. The reason is that it takes time to copy data forward through a latch but no time to move emptiness backwards. Also for 3-7,

5-5, 7-3 combinations require alternating encoding for FULL and EMPTY. Hence 6-4 GasP is the preferred GasP circuit and will be used for the work of this thesis.

The waveforms in Figure 2.4 depicts the operation of GasP circuits. Figure 2.4 shows that when the condition, pred1 is HI and succ1 is LO, MODULE 1, in Figure 2.4, is triggered. The fire signal, fire1, goes HI as shown by green curve in the figure. The fire signal remains HI for 5 gate delays. The forward latency for GasP is 6 gate delays. This is shown by the time interval between blue curve going HI followed by red curve going HI. The reverse latency is 4 gate delays. This is shown by the time interval between red curve going LO followed by blue curve going LO.

The master clear signal, not shown in the figure, establishes an initial condition. To sustain the state indefinitely keepers are required. The keepers are shown in dashed boxes in the figure. The GasP modules drives the state wires HI or LO for a brief period. It is the responsibility of the keepers to sustain those states. Keeper k1 should keep state wire W LO if it is EMPTY. Keeper k2 should keep state wire W HI if it is FULL. The keepers are split into two adjacent GasP modules so that each GasP module can control its half keeper when the modules drive the state wire to change its state. So keeper k1 is switches off by MODULE 1 when state wire W is driven HI whereas MODULE 2 switched off keeper k2 when state wire W is driven LO.

GasP circuits consists of all inverters and a simple NOR gate. These simple gates reduce the logical effort, making the GasP family very fast. The GasP circuits use a single wire to convey the state information making them power and area efficient.

The GasP family of asynchronous control circuits is used in the work presented in this thesis.

Figure 2.4: Waveforms depicting operation of 6/4 GasP circuits.

2.3 Power Gating

A significant portion of the total power consumption in high-performance digital circuits is due to leakage currents, like subthreshold conduction, junction leakage, and tunneling leakage through gate oxide [4], [11]. Power gating is the most effective technique to reduce subthreshold leakage. In power gating, either a PMOS transistor, called a header, or an NMOS transistor, called a footer, or both are used to connect the circuit block to the power rails. The NMOS transistors and PMOS transistors are commonly referred to as *sleep transistors*. The supply point of the circuit that connects to the external power supply is commonly referred to as *Virtual VDD* and *Virtual GND*. In power gating sleep transistors are inserted

between the virtual power supply of the circuit and the external power supply. When the circuit block enters idle mode, the sleep transistors are switched OFF to disconnect the external power supply from the virtual power supply of the circuit. Switching OFF the sleep transistors causes a voltage rise in virtual GND and drop in virtual VDD. The leakage current is dependent on the power supply and hence as the voltage drops at virtual VDD or rises at virtual GND, the leakage current also reduces. Hence power savings begin as soon as the sleep signal is asserted [12].

Figure 2.5: Block Diagram of overview of power gating.

Power gating reduces the leakage power by inserting sleep transistors in series with the circuit block. Sleep transistors are controlled by sleep signals generated by control blocks that synchronize with the idle times of the circuits. Sleep transistors should be wide enough to support large circuit loads. Extra power is consumed in the sleep transistors and control blocks generating the sleep signals. To achieve significant power savings, the leakage savings obtained from power gating should

exceed the overheads incurred due to power gating. Power gating also increases delay. Stacking the sleep transistor with the circuit reduces the drive current, lowering performance. To avoid performance loss, sleep transistor should be switched ON and OFF accurately. For additional power savings, high V_t transistors are used for sleep transistors to reduce the leakage in big transistors while the logic is implemented with low V_t transistors to maintain the performance [5].

Power gating can be implemented with either a PMOS transistor or an NMOS transistor or both. A PMOS transistor, also called a header switch, connects the VDD supply to virtual VDD of the circuit. PMOS transistors leak less than NMOS transistors of the same size and avoid ground bounce issues. The Negative Bias Temperature Instability (NBTI) effect increases the threshold voltage over time and makes PMOS transistors even less leaky [13]. The disadvantage of header switch is that a PMOS transistor is bigger in area and has lower drive current than an NMOS transistor of the same size. An NMOS transistor, also called a footer switch, connects external GND supply to virtual GND of the circuit. The advantage of an NMOS transistor is high drive current and smaller area compared to a PMOS transistor of the same size. But NMOS transistors leak more than PMOS transistors and have ground bounce issues [13].

The sleep transistor can be implemented either in *coarse-grain* or *fine-grain* power gating style [13]. In fine-grained implementation, a sleep transistor is allotted to a single standard cell. The advantage of the fine-grain sleep transistor implementation is that the virtual power nets are short. But fine-grain implementation also results in significant increase in area. In coarse-grained implementation, a sleep transistor is allotted to large circuit block consisting of multiple standard cells.

The advantage of coarse-grained sleep transistor implementation is less area overhead. However, the entire block must be idle before it can be power gated. Hence the leakage savings achieved vary depending on the idle times of the circuit blocks.

Power gating requires inserting sleep transistors, limiting the area and power efficiency of the circuit. The sleep transistor should be sized to provide sufficient current to the latches and the combinational circuits. A single sleep transistor can be assigned for a latch and a combinational block. When the latches are transparent, only the latches draw current from the virtual power supply while the combinational circuit is idle. Once the data are stabilized at the output of the latches, latches stop drawing current from the virtual power supply and combinational circuit starts the computation and hence draws current from the virtual power supply. Either the latch or the combinational block draws current from the virtual power supply, but not both. Hence, the sleep transistor should be sized to provide supply to the larger of the two loads, latches or combinational circuits.

Synchronous and asynchronous circuits require different techniques for implementing power gating. The control blocks used for power gating need to be efficient in area and power consumption. The circuit blocks should be activated and deactivated quickly to minimize the performance impact. Also, the idle times for the circuit blocks should be long enough to compensate for the additional power incurred by power gating. A micro-architectural method in which the number of cycles for which a block is idle is counted is discussed in [12] [14]. A block is power gated when the number of idle cycles exceeds the number of cycles required to compensate for additional power entailed by power gating. A Clock gating signal can also be utilized to decide when to power gate the circuit blocks [11], [15]. When

data stored in flip-flops is not updated, clock to the flip-flops is gated to reduce dynamic power. The combinational blocks feeding the flip-flops are not required to compute new data while flip-flops are not updated. This idle period can be detected and the leakage current can be reduced by forcing the combinational blocks into power gating mode. Asynchronous controllers can generate the sleep signal locally and hence provide an easier fine-grained power gating implementation.

Power gating in synchronous circuits mostly relies on the clock gating signals or on the count of the number of idle cycles. The transition of the power gated circuit between sleep and active modes may disturb the clocking synchronization. Hence, implementation of fine grained power gating in synchronous circuits is very difficult. On the other hand, asynchronous circuits use a handshaking protocol to implement data synchronization and so asynchronous controllers have inherent attribute to determine when the circuits are active and inactive. Hence fine-grained power gating is easier to implement because the sleep signal is generated locally.

A technique to power gate the combinational circuit in an asynchronous pipeline is discussed in [16]. The request signal is used as the sleep signal to switch ON the sleep transistor and to quickly wake the circuit up, from sleep mode. The control circuit for controlling the power gating consists of a PMOS transistor and an NMOS transistor as sleep transistors. A simple control circuit consisting of an XOR gate to achieve power savings in an asynchronous MOS Current-Mode Logic (MCML) pipeline is discussed in [17]. Both the methods discussed above power gate the combinational logic leaving latches fully powered. Hence latches continue to leak when idle and curtail the power savings.

A fine-grained power gating for asynchronous pipelines, to power gate combinational blocks as well as latches, is discussed in [18]. Power gating is implemented based on adjacent state monitoring. The control circuit for power gating includes two OR gates. However, if the latch feeding the combinational block is opaque and the receiver latch is transparent then the combinational block will be in active mode. Unless the feeding latch has keepers that maintain the states, input signals to the combinational block will be floating leading to short circuit current in the combinational block. Hence, while power gating combinational blocks and latches; care should be taken not to leave the output of the latch floating.

Power gating techniques for asynchronous pipelines that are discussed above achieve power savings when the pipeline is mostly empty. As the occupancy increases, sleep transistors of more stages remain ON. Hence, the more the occupancy of the pipeline, the less are the power savings achieved. The work presented in this thesis power gates pipeline stages in lower occupancy as well as higher occupancy states. Also, latches used in this thesis work have keepers in latches that avoid output of latches from floating at any time thus maintaining the states.

3

Latches

An asynchronous pipeline consists of data latches and data controllers. The controllers assert enable signals that guide data through the latches. The pipeline designed in this thesis consists of GasP modules as the data controllers. The data latches are configured to receive the data and the fire signal. The latches used in this thesis are transparent for the HI duration of the fire signal.

This thesis covers three latches: Conventional latch, Lazy latch and Blended latch. The Conventional latch and the Lazy latch are used in the circuits of this thesis. The Blended latch is a latch design proposed based on the advantages and disadvantages of the Conventional latch and the Lazy latch. The Blended latch is a blend of both the Conventional latch and the Lazy latch.

3.1 Conventional Latch

The Conventional latch consists of three parts in the following order: (1) Switch (2) Keeper (3) Amplifier. Figure 3.1 shows the circuit diagram of a Conventional latch. As shown in the figure, *data* is the latch input, *out* is the latch output and *fire* is an input signal from the corresponding GasP module. The switch consists of inverter I and pass transistors N1 and N2. The fire signal from the GasP module controls the pass transistors. When the fire signal is high, the pass transistors transfer the input to the output. The amplifier consists of an inverter with wide transistors to drive the large fan-out. The keeper consists of a back to back inverter.

Figure 3.1: Circuit diagram of the Conventional latch

To achieve maximum power savings during power gating, amplifier transistors can be power gated. In the event of power gating, however, the latch output will be lost. The keeper keeps the state but fails to maintain the latch output because the amplifier is power gated. Using power gating with the Conventional latch requires retention logic to retain the state once the power is restored. Additional startup time is associated with the retention logic. Also a power gated latch stage may provide undefined input to the next stage. Undefined inputs may cause issues like short circuit current in the next stage. Isolation logic is required to avoid providing undefined inputs to the next stage.

The fire signal driving the Conventional latch has to overcome the opposition of the keeper while driving the amplifier. Assume the latch has to drive a load of D_{conv} and a step-up of s for the entire design. D_{conv} is the capacitance of the next stage, including the gate capacitance of the next stage and the wire capacitance. Step-up is a combination of logical effort and electrical effort. With a load of D_{conv} and step-up of s , amplifier transistors are sized as $\frac{D_{conv}}{s}$. The pass transistor driving

the amplifier transistors are sized $\frac{D_{conv}}{s^2}$. The load offered to the fire driver by the two NMOS pass transistors is,

$$F_{conv} = \frac{2}{3} \times \frac{D_{conv}}{s^2} \quad (3.1)$$

The Conventional latch is insensitive to the noise coupled to its output wire. The amplifier isolates the noise from the keeper. In spite of power gating and noise, the keeper always maintains the state, though not at the latch output.

3.2 Lazy Latch

In this thesis a new latch design called the Lazy latch is introduced. The Lazy latch drives its output only when required and therefore avoids wasting power. The Conventional latch is described in Section 3.1. The arrangement of the parts of the Conventional latch is: Switch, Keeper and Amplifier. For every fire signal, the latch drives the output. The keeper acts as the memory element and keeps the last data to pass the latch. Therefore, one of the transistors in the amplifier is always ON causing full capacity leakage in the other transistor. Also, the keeper imposes a load on the fire signal. The keeper in the Conventional latch is sized to drive the big amplifier transistors. Therefore, the fire driver is heavily loaded. The Lazy latch overcomes all the drawbacks by moving the keeper after the amplifier and turning OFF the amplifier transistors when they are unneeded.

The Lazy latch drives the big amplifier transistors only if necessary [19]. The latch parts in the Lazy latch are similar to the original latch. The latch parts are, however, rearranged as follows: Switch, Amplifier and Keeper. The switch has two outputs that separately drive the big amplifier transistors. Therefore, the wide

NMOS and PMOS transistors in the amplifier can be turned OFF individually when not required. Turning OFF both the transistors in the amplifier avoids the full capacity leakage in either transistor. A feedback signal from the output feeds the switch. If the data remains stable for more than one cycle, then the feedback signal forces the latch to stop driving the output. The keeper maintains the state when the amplifier stops driving the output. Also, moving the keeper after the latches, omits the loading of the fire driver by the keeper.

3.2.1 Lazy Latch: Design and Operation

Figure 3.2 shows the top level circuit diagram of the Lazy latch. The Lazy latch has two inputs, Data from previous pipeline stage, and fire signal from the corresponding GasP module. The Lazy latch in the figure consists of three parts in following order parts: (1) Switch and (2) Amplifier (3) Keeper. The switch inputs indicate when the latch should be made transparent. Latches may have a large fan-out and wire capacitance. Therefore, the PMOS and NMOS transistors in the amplifier are wide to drive the large wire capacitance quickly.

Figure 3.2: Lazy Latch: Top level circuit diagram

Figure 3.3: Lazy Latch: Internal circuit diagram

Figure 3.3 shows the internal circuit of the switch and the amplifier of the Lazy latch. A switch consists of a 3-input NAND gate and a 3-input NOR gate. The amplifier consists of a wide PMOS transistor and a wide NMOS transistor, keeper and a chain of inverters to carry the output state back to the switch. The latch output is inverted and fed back to the switch, where it participates in determining when the latch should drive its output. The latch design can be viewed as two branches (1) data - NAND - Amplifier PMOS (P1) - out and (2) data - NOR - Amplifier NMOS (N1) - out. Data values of 1 are copied to the output through the upper branch whereas data values of 0 are copied to the output through the lower branch. Consider the following conditions of input and output data values and corresponding latch operation:

1. Latch input = 1, latch output = 0: Because initially the latch output is 0, the storeb signal is 1. Hence the output of the NOR gate is 0 turning OFF the lower branch. However, the output of the NAND gate is 0 thus turning

ON the PMOS P1 transistor and copying a 1 to the latch output. Therefore data values of 1 at the latch input are copied to the latch output via the upper branch.

2. Latch input = 0, latch output = 1: The storeb signal is 0 turning OFF the upper branch. The lower branch gets the inverted fire signal and so all the inputs to the NOR gate are 0. Hence the output of the NOR gate is 1 which turns ON the amplifier NMOS N1 transistor and copying data values of 0 to the output. In this case the lower branch conducts to copy the latch input to the output.
3. Latch input = latch output: If the data value at the latch input and latch output are same, then both the branches are turned OFF.

The keeper in the latch drives the output when the amplifier transistors in the latch stop driving the latch output. When the data at the latch input and output match, the amplifier transistors are turned OFF. The keeper then takes over and preserves the state. It will be seen in later sections that when the power to the latches is turned OFF, the amplifier transistors are turned OFF. The keeper preserves the state of the latch in the event of power gating.

If the amplifier stops driving the latch output, the keeper drives the output. Therefore, when the latch output = latch input, the amplifier transistors are turned OFF and keeper takes over to keep the state. As will be seen in the later section, if the power to the Conventional latch is turned OFF, the output state is lost. However, in the Lazy latch, if the power is turned OFF then the keeper maintains the state of the latch.

The output signal passes through three inverters before going back to the switch of the latch. This completes a ring with five inverters. Five inverters in a ring relax the sizing constraints [8]. The duration of five inverters allow sufficient time for the output to attain the supply rail before it is driven by the keeper. The fire signal to the latch may repeat every 10 gate delays. Hence, adding five gate delays in feedback loop in the latch avoids garbling the data value captured by the latch.

3.2.2 Lazy Latch: Switch design details for 100 Latches

The NAND and the NOR gates in the switch are asymmetric gates to favor the fire signal input at the expense of the other inputs - data and storeb. The latch should quickly transfer the input data to the output as soon as the fire signal drives the latch. Therefore transistors in both the NAND and the NOR gates are modified to favor the fire input. As shown in Figure 3.4, NAND gate and NOR gate drive PMOS transistor and NMOS transistor respectively, both of same drive strength. PMOS transistor offers twice as much load as the NMOS transistor. Therefore the transistors in the NAND gate are twice as wide as the transistors in the NOR gate.

Figure 3.4: Asymmetric gates scaled for 100 latches. (1) Asymmetric 3-input NAND gate (2) Asymmetric 3-input NOR gate

Figure 3.4 shows the circuit diagram for the asymmetric NAND gate and asymmetric NOR gate. The NAND gate should quickly turn on the PMOS transistor in the amplifier when the fire signal drives the latch. To avoid loading the amplifiers in the GasP module, the gate capacitance offered by the transistor N1 should be small. However, to maintain the drive required for the wide PMOS transistor, transistors N2 and N3 driven by inputs data and storeb should be large. The NAND gate is responsible for copying data 1 to the latch output. However, if the latch input data are 0, then the NOR gate copies it to the output. While copying a 0 the PMOS transistor should be turned off quickly to avoid any short circuit current when the NMOS transistor is turned on by the NOR gate. Hence the transistor P2 driven by the data input is wide. Once the input data are transferred to the output, the amplifying transistors are turned off and the keeper maintains the

data value. If the last value copied was a 1 and the input value is stable at 1, then storeb signal is equal to 0 thus turning off the PMOS transistor in the amplifier. However, the PMOS transistor can be turned off slowly without drawing much current, thus lowering the effort on the storeb signal. The switching of the fire signal switches the pull up network in the NAND gate every cycle. This can be avoided by having small P1 transistor driven by the fire input. The fire and the storeb signal drive the PMOS transistors with a step-up of 100. The step-up of 100 is a random choice.

For an asymmetric NOR gate, the pull-up transistors are sized to provide fast drive current to quickly turn on the amplifying NMOS transistor while reducing the effort on the fire signal. The transistor driven by the input data in the pull-down network is sized large to quickly turn off the amplifying NMOS transistor whereas transistors driven by inputs storeb and fire are sized small to reduce the efforts on those inputs. A low skewed inverter is used to prefer falling edge so that NOR gate switches almost at the same time as the NAND gate.

3.2.3 Lazy Latch: Fire Load

The fire signal driving the Lazy latch has to drive the asymmetric gates in the switch. The keeper is moved after the amplifier and is turned OFF when the amplifier is turned ON. Therefore, the fire signal has to drive just the switch. Assume the latch has to drive a load of D_{lazy} and a step-up of s for the entire design. With a load of D_{lazy} and step-up of s , amplifier transistors are sized as $\frac{D_{lazy}}{s}$. The amplifier transistors are driven by separate signals. NAND gate driving the PMOS amplifier transistor is sized $\left(\frac{2}{3} * \frac{D_{lazy}}{s^2}\right)$. NOR gate driving

the NMOS amplifier transistor is sized $\left(\frac{1}{3} * \frac{D_{lazy}}{s^2}\right)$. The total load As seen in previous section, NAND and NOR gate in switch are asymmetric gates offering minimum load to the fire driver. The pull-up networks offer negligible load on the fire driver. An approximate load offered by the switch of the the Lazy latch to the fire driver is,

$$F_{lazy} = \frac{1}{3} \times \frac{D_{lazy}}{s^2} \quad (3.2)$$

3.2.4 Drawbacks of Lazy Latch

The keeper in the Lazy latch is moved after the amplifier. In the event of power gating, the keeper maintains the state. However, the latch output drives large fan-out at a large distance. A large coupling noise is associated with the latch output. The noise may override the data kept by the keepers resulting in loss of data. A new latch called the Blended latch is proposed in next section that overcomes the drawbacks of the Lazy latch.

3.3 Blended Latch

The Blended latch is a blend of the Conventional latch and the Lazy latch. The Blended latch combines advantages of both the Conventional latch and the Lazy latch. The Conventional latch fails to maintain the state in the event of power gating. The keeper in the Conventional latch loads the fire driver because the fire driver has to overcome the opposition from the keeper. However, the Conventional latch is insensitive to noise. The Lazy latch can maintain the state in the event of power gating. Keeper in the Lazy latch is moved after the amplifiers and keeper is turned OFF when the amplifiers are turned ON. The fire driver has to drive the switch gates which are highly skewed to reduce the load on the fire signal. However,

keeper of the Lazy latch is sensitive to the noise at the latch output. The Blended latch has the advantage of the Lazy latch of maintaining the state in the event of power gating while making the keeper insensitive to the noise. The Blended latch offers intermediate of loads offered by the Conventional latch and the Lazy latch. This thesis proposes a Blended latch but the power gating technique introduced in this thesis is evaluated only for the Lazy latch.

Fig 3.5 shows the circuit diagram of the Blended latch. The figure shows the Conventional latch and the Lazy latch blended together. The amplifier and the switch of the Lazy latch can power gated whereas the switch, keeper and amplifier can powered from true power supply. Only one keeper is required to maintain the state and therefore the keeper of the Lazy latch can be omitted. Therefore, the Blended latch maintains the state in the event of power gating as well as noise.

Figure 3.5: Circuit Diagram of the Blended Latch.

The Blended latch blends advantages of the Conventional latch and the Lazy latch. The advantage gained from the Conventional latch and the Lazy latch depends on the amount of contribution of each of the two latches. The contribution of the latches can be determined by a new factor called the *blend factor* α . The blend factor defines the advantages gained from the Conventional latch and the Lazy latch.

Let D_{blend} be the output load offered to the Blended latch. The output load represents the gate capacitance of the next stage and the wire capacitance. The Conventional latch part can provide α parts of D_{blend} whereas the Lazy latch can provide $(1-\alpha)$ parts of D_{blend} . Let D_{conv} be the load offered to the Conventional latch. Let D_{lazy} be load offered to the Lazy latch. The load offered by the Conventional latch to the fire driver from equation 3.1 is,

$$F_{conv} = \frac{2}{3} \times \frac{D_{conv}}{s^2} \quad (3.3)$$

The load offered by the Lazy latch to the fire driver from equation 3.2 is,

$$F_{lazy} = \frac{1}{3} \times \frac{D_{lazy}}{s^2} \quad (3.4)$$

Let F_{blend} be the load offered by the Blended latch to the fire driver. The load offered by the Blended latch is a combination of load offered by the Conventional latch and the Lazy latch.

$$F_{blend} = F_{conv} + F_{lazy} \quad (3.5)$$

$$= \left(\frac{2}{3} \times \frac{D_{conv}}{s^2} \right) + \left(\frac{1}{3} \times \frac{D_{lazy}}{s^2} \right) \quad (3.6)$$

$$= \left(\frac{2}{3} \times \frac{\alpha D_{blend}}{s^2} \right) + \left(\frac{1}{3} \times \frac{(1 - \alpha) D_{blend}}{s^2} \right) \quad (3.7)$$

$$= \left(\frac{(1 + \alpha)}{3} \right) \left(\frac{D_{blend}}{s^2} \right) \quad (3.8)$$

Table in Figure 3.6 summarizes the output loads and fire loads for Conventional latch, Lazy latch and Blended latch.

	Conventional Latch	Lazy Latch	Blended Latch
Output load	$D_{conv} = \alpha D_{blend}$	$D_{lazy} = (1 - \alpha) D_{blend}$	$D_{blend} = D_{conv} + D_{lazy}$
Fire load	$F_{conv} = \left(\frac{2}{3} \right) \left(\frac{D_{conv}}{s^2} \right)$	$F_{lazy} = \left(\frac{1}{3} \right) \left(\frac{D_{lazy}}{s^2} \right)$	$F_{blend} = F_{conv} + F_{lazy}$ $= \left(\frac{1 + \alpha}{3} \right) \left(\frac{D_{blend}}{s^2} \right)$

Figure 3.6: Table showing Output load and Fire load offered by the Conventional, the Lazy and the Blended latch.

Equation 3.8 shows that the blend factor determines the load offered by the Blended latch to the fire driver. The blend factor also determines the noise sensitivity and the power gating effectiveness. The blend factor can have a value between 0 and 1. If $\alpha = 0$, the Blended latch has complete characteristics of the Lazy latch with no properties of the Conventional latch. This latch is very effective with power gating but most sensitive to noise. If $\alpha = 1$, the Blended latch has complete characteristics of the Conventional latch with no properties of the Lazy latch. For the intermediate values of α , the Blended latch shows the combined

properties of the Conventional latch and the Lazy latch depending on the value α . As the value of α increase from zero to one, the noise sensitivity improves but the power gating effectiveness decreases. Table in Figure 3.7 summarizes the effect of value of α on the Blended latch.

	$\alpha = 0$	$\alpha = 0.25$	$\alpha = 0.5$	$\alpha = 0.75$	$\alpha = 1$
Fire load	$\left(\frac{4}{12}\right)\left(\frac{D_{blend}}{s^2}\right)$	$\left(\frac{5}{12}\right)\left(\frac{D_{blend}}{s^2}\right)$	$\left(\frac{6}{12}\right)\left(\frac{D_{blend}}{s^2}\right)$	$\left(\frac{7}{12}\right)\left(\frac{D_{blend}}{s^2}\right)$	$\left(\frac{8}{12}\right)\left(\frac{D_{blend}}{s^2}\right)$
Noise sensitivity	Highest	Much	Medium	Little	Insensitive
Power gating	100%	75%	50%	25%	State loss
	Lazy Latch	Blended Latch			Conventional Latch

Figure 3.7: Table showing the effect of blend factor α .

The Blended latch has advantages of both the Conventional latch and the Lazy latch. The power gating effectiveness ranges between full effectiveness offered by the Lazy latch to none of the Conventional latch. Similarly, noise insensitivity ranges between minimum offered by the Lazy latch and maximum offered by the Conventional latch. The characteristics of the Blended latch change with the blend factor α .

Methodology - Power Gating of the GasP Circuits

This thesis introduces a new fine-grained power gating technique designed to achieve idle-state power savings in the GasP family of asynchronous circuits. In the GasP family of asynchronous circuits, pipeline stages are controlled by GasP circuits. Using the new fine grained technique described in this thesis, we can cut-off power to idle pipeline stages by asserting necessary sleep signals and forcing the stages into sleep mode. The technique implemented in this thesis inserts a PMOS transistor called a header, and a NMOS transistor called a footer, as sleep transistors. The sleep transistors are turned ON in anticipation of incoming data. Fine-grained power gating controls the power for individual stages of the pipeline. In [16], [18], individual stages of the pipeline are turned OFF when the pipeline is empty. The technique introduced in this thesis turns the power OFF not only when the pipeline is EMPTY but also when it is FULL. The traditional power gating turns the power OFF for combinational circuits. In this technique, sleep transistors are shared between a latch and the combinational circuit associated with it. The power is turned OFF for the combinational circuit as well as the latches. Sharing the sleep transistors reduces the area overhead for power gating.

Each stage of the pipeline can be viewed as three separate blocks namely,

1. Power-gating Block: This block consists of the sleep transistors (header and footer) to turn the power ON and OFF.
2. Power-GasP circuit: The Power-GasP circuit sequences the data through the

latches. This circuit also includes the logic to generate the sleep signal.

3. Latches and Combinational circuit: The latches and combinational circuit to be power gated.

Figure 4.1 shows the block diagram of a power gated pipeline. Each stage of this power gated pipeline consists of 100 latches that drive 100 bits of data. A Power-GasP circuit controls each of those 100 latches. The Power-GasP circuit has the logic required to generate the sleep signal. The sleep signal is referred in this thesis as the *power-control* signal. The power-control signal drives the sleep transistors in the power-gating block. A power-gating block provides virtual power supplies, V_{vdd} and V_{gnd} . V_{vdd} and V_{gnd} are the power supplies for the latches and the combinational circuits of that stage. The Power-GasP circuit is powered from regular power supply.

Figure 4.1: Block diagram of Power Gating of Power-GasP circuits

4.1 Power-GasP circuit

This thesis modifies the traditional GasP circuit (refer to section 2.2) to include power gating logic and we call it the Power-GasP circuit. The Power-GasP circuit generates the power-control signal as well as the state wires and the fire signal. The power-control signal slowly turns OFF the power to the latches and the combinational circuit but quickly turns ON the power to avoid any data loss.

Figure 4.2 shows the circuit diagram of a Power-GasP circuit. Keepers have been omitted for clarity. Every Power-GasP circuit drives 100 latches and consists of two state wires, namely, a Data-state wire and a Power-state wire.

1. Data-state wire: The data-state wire indicates the presence or absence of data between the stages of Power-GasP circuits. The data-state wire is same as the state wires of the traditional GasP circuit. When the predecessor state wire is HI or FULL *and* the successor state wire is LO or EMPTY, the Power-GasP circuit is said to fire. The fire signal allows the transfer of data from one stage to another. The data-state wire is driven by the predecessor driver from one end and the successor driver from the other end.
2. Power-state wire: For power gating we add a second state wire, the *power-state wire*. When the power state wire is HI, power should be applied to the latches and the combinational logic associated with the latches. When the power state wire is LO, power may be interrupted and the sleep transistors may be turned off to allow the gated power supply to decay. The powerOn driver C drives the power-state wire HI from one end; whereas the powerOff driver D of the next stage drives the power-state wire LO from the other end.

Figure 4.2: Circuit Diagram of Power-GasP circuit

4.1.1 Turn ON the Power

The power for the latches should be turned on in anticipation of receiving data. When the predecessor stage proffers new data or is FULL and the successor stage is ready to accept the new data or is EMPTY, the Power-GasP circuit fires. An AND function, AND1, detects the FULL-EMPTY condition. The FULL-EMPTY condition should also turn on power ON for that stage. A parallel but separate AND function, AND2, drives the power-state wire. The power-state wire goes HI

four gate delays before the fire signal. The keepers keep the power-state wire HI until the data gets accepted by the successor stage. Once the data are passed on to the successor stage, the power for the current stage can be turned off. The power can be turned off even before the successor stage fires. The powerOff driver D of the successor stage drives the power-state wire LO. A LO power state wire indicates that the power is no longer required and may be turned OFF.

The power-control signal drives the sleep transistors instead of the power-state wires because of the inability of the power-state wires to drive big loads. The sleep transistors are large because they have to supply current for the latches and the combinational circuits. These large sleep transistors should be turned ON quickly. But the power-state wire is a low current signal and fails to provide the required drive for the sleep transistors. Therefore, another signal called the power-control signal, actually turns ON or OFF the sleep transistors. The power-control signal which drives large sleep transistors is a heavy current signal. A third AND function, AND3 drives the power-control signal. The AND3 function detects the FULL-EMPTY condition similar to the data-state wire and the power-state wire. The AND3 function detects the required condition early in the generation of fire signal. The AND3 function allows power-control signal to turn ON the sleep transistors and keep the power ready just in time before the fire signal drives the latches.

The data-state wire carries less current than the power-control signal but more than the power-state wire. Traditional GasP circuits place only the load of the AND function that generates fire signal on the data state wire. However, in the Power-GasP circuit, three AND functions load the data-state wire, namely, (i)

AND1 to generate fire signal, (ii) AND2 to generate power-state wire signal and (iii) AND3 to generate power-control signal. Of the three AND functions, the AND3 function drives the large sleep transistors and hence offers the largest load. The drivers for the data state wires drive the AND3 function as well. Therefore, drivers driving the data state wires in the Power-GasP circuit are comparable to the drivers driving the power control signal. Adding more amplification on the power control signal reduces the load offered by the sleep transistors. Figure 4.2 shows the added amplifiers. To match the delay of the power control signal and fire signal, an equal number of amplifiers are added on the fire signal. Added amplifiers cause a small delay in data but reduce the load on data-state wires significantly. The next subsection analyzes the effect of the increase in amplification on the data-state wire driver.

4.1.2 Analysis of effect of step-up and amplification

Figure 4.3: Circuit Diagram of Power-gating Block (s is the step-up)

Figure 4.2 refers to the Power-GasP circuit which was described in the earlier section. Figure 4.3 refers to the power-gating block which consists of the sleep

transistors. Detailed description of the power-gating block is provided in the later section.

Let us examine the effect of amplification of gates on the drivers of the Power-GasP circuit. Assume the sleep transistors with width X . The power-control signal directly drives the NMOS sleep transistor in the power-gating block. The power-control signal is inverted and applied to the PMOS sleep transistor in the power-gating block. PMOS sleep transistors drives the Vvdd and NMOS sleep transistor drives the Vgnd. The inverter and the PMOS sleep transistor in the power-gating block are the amplifiers between power-control signal and Vvdd.

1. Let us examine the effect of amplification of two gates. The two gates correspond to the inverter and the PMOS sleep transistor in the power-gating block. The power-control signal from the AND3 function is amplified twice in the power-gating block to drive Vvdd. Similarly, there are two gates of amplification between AND1 function and the fire signal. Assume a step-up of 3.

- Load seen by power-control signal = Load offered by the NMOS sleep transistor and inverter driving the PMOS sleep transistor.

$$= \frac{X}{3} + \frac{2 \times X}{9} = \frac{5 \times X}{9} \quad (4.1)$$

- Drive of the powerOn driver E

$$= \frac{\frac{5 \times X}{9}}{9} = \frac{5 \times X}{27} \quad (4.2)$$

- Load offered by the powerOn driver E to the data-state wires

$$= \frac{4}{3} \times \frac{5 \times X}{27} = \frac{20 \times X}{81} \quad (4.3)$$

2. Let us examine the effect of amplification of four gates. The four gates correspond to the inverter and the PMOS sleep transistor in the power-gating block and the two inverters driving the power-control signal. Similarly, there are four amplification gates between the AND1 function and the fire signal.

- Load seen by power-control signal = Load offered by the NMOS sleep transistor and inverter driving the PMOS sleep transistor.

$$= \frac{X}{3} + \frac{2 \times X}{9} \approx \frac{2 \times X}{3} \quad (4.4)$$

- Load seen by the powerOn driver E (two gates of amplification between powerOn driver E and inverter for the PMOS sleep transistor)

$$= \frac{2 \times X}{27} \quad (4.5)$$

- Drive of the powerOn driver E

$$= \frac{\frac{2 \times X}{27}}{3} = \frac{2 \times X}{81} \quad (4.6)$$

- Load offered by the powerOn driver E to the data-state wires

$$= \frac{4}{3} \times \frac{2 \times X}{81} = \frac{8 \times X}{243} \quad (4.7)$$

Thus the size of the drivers driving the data-state wires with four amplifier gates is reduced to $2/15^{th}$ of the size of the drivers with two amplifier gates. Therefore, adding amplifiers reduce driver sizes in the Power-GasP circuits.

4.1.3 Turn OFF the Power

Turning ON the sleep transistors is urgent; however, turning them OFF can be a slow process. The power for a stage can be turned off once the data are transferred to the successor stage. Therefore, irrespective of the readiness of the successor stage to accept the data, the power to the current stage can be turned OFF. However, there is no guarantee how long the data will stay at a stage before it is accepted by the next stage. Therefore, to avoid floating of the data, the latches and combinational circuits should have keepers to maintain the output states. Lazy latch and Blended latch design introduced earlier maintain the state in the event of power gating. Therefore, once every stage transfers the data, the power to that stage can be turned OFF, by pulling the power-control signal LO, irrespective of the condition of the successor stage. Thus, turning OFF the individual stages helps in saving more power.

4.1.4 Operation: Power-GasP circuit

Figure 4.4 shows two successive Power-GasP circuits. The two Power-GasP circuit stages, 1 and 2 share the state wires between them. The Power-GasP circuit of stage1 pulls the shared state wires HI whereas the Power-GasP circuit of stage2 pulls them LO. The timing differences in the operation of the two Power-GasP circuits avoid pulling the state wires HI and LO at the same time.

Figure 4.4: Operation of Power-GasP circuit

When `pred1` is HI and `succ1` is LO, the AND functions in `stage1` are pulled HI at the same time. Thus, power-state wire goes HI immediately. This state wire indicates only the need for power and does not directly drive the sleep transistors. The power-control1 goes HI after two gates of amplification. Two gate delays after the power-control1 goes HI, `fire1` signal goes HI to make the latches transparent. The `succ2` goes HI and indicates the presence of new data to the Power-GasP circuit of `stage2`. The NAND gate in the Power-GasP circuits has two inputs namely, the data-state wire and `done`, a completion signal from the combinational circuit. The completion signal can either be a `done` signal from the combinational circuit or delay matching element. In this thesis the `done` signal is connected to the power supply `Vdd` because of the absence of the combinational circuit. That is why, when the `succ2` goes HI, `NAND2` immediately goes LO. Therefore, powerOff driver `D2` in the Power-GasP circuit of `stage2` pulls the power-state wire LO. A LO power-state wire indicates that `stage1` does not need power anymore. However the power to the `stage1` is not turned off abruptly. Instead when the powerOff driver `D2` from `stage2` pulls the power-state wire LO, powerOff driver `F1` in the Power-GasP circuit of `stage1` slowly pulls down the power-control1 wire. Thus the power to the `stage1` is turned off slowly and irrespective of the `fire` signal of `stage2`.

4.2 Power-gating Block

The power-gating block, controlled by the power-control signal from the Power-GasP circuit, provides power for the latches and the combinational circuits. The power-gating block, shown in Figure 4.5, consists of the sleep transistors. The power gating technique discussed in this thesis inserts two sleep transistors: a PMOS transistor, the header, and an NMOS transistor, the footer. The drain of

the PMOS transistor or header acts as the new power supply, called *Virtual Vdd* or *Vvdd*. The drain of the NMOS transistor or footer acts as the new ground, called *Virtual Gnd* or *Vgnd*. The latches and the combinational circuits are powered from the new supplies, Vvdd and Vgnd.

Figure 4.5 shows the circuit diagram for the power-gating block. The power-gating block consists of a PMOS transistor or a header, an NMOS transistor or a footer and a LO skewed inverter. The power-control signal which is generated in the Power-GasP circuit, is the input for the power-gating block. When the power-control signal is high, the sleep transistors are turned ON quickly whereas when the power-control signal is low, the sleep transistors are turned OFF slowly. The power-control signal directly drives the NMOS sleep transistor whereas a LO-skewed inverter inverts the power-control signal to drive the PMOS sleep transistor. The LO-skewed inverter quickly pulls the output low to turn ON the PMOS sleep transistor. However, the LO-skewed inverter slowly turns OFF the PMOS transistor. Therefore, virtual power supply Vvdd is turned OFF slowly.

Figure 4.5: Circuit Diagram of Power-gating Block

The widths of the sleep transistors depend on the latches and the combinational circuits. The sleep transistors should be wide enough to supply the maximum current drawn by either:

1. The latches and the load driven by the latches, or
2. The combinational circuits and the load driven by the combinational circuits

The latches and the combinational circuit associated with the latches draw current from the power supplies at different times. Hence the sleep transistors can be sized to supply maximum current required to charge the sources of the transistors in latches or combinational circuits and the load offered by the circuit they drive.

The sleep transistors drive the source capacitances of the transistors in the latches or the combinational circuits and the gate capacitances of the load that they drive. Additionally they also drive the wire capacitance. To represent the true power supply on a chip, by-pass capacitors are added on the V_{dd} and V_{gnd} . The by-pass capacitors are shown in Figure 4.5. The more the capacitance, the more is the current supplying ability of the power supply. However, the circuit draws more energy to charge all the capacitances on the virtual power supply. But this capacitance supports the Virtual V_{dd} . Hence the wires carrying the supply signals are made wider than the rest of the wires in the circuit. The capacitance of the gated power supply wires and the leakage drain on them set a time constant at which the delivered power voltage will decay when the power is turned OFF. This natural decay time should be longer than several cycles of normal operation. Artificially hastening the drain of charge already on the local power supply wires would merely waste energy.

The power-gating block consists of simple gates that turn ON or OFF the virtual power supplies, therefore reducing the leakage in the transistors of the latches and the combinational circuit. A power-control signal, generated in the Power-GasP circuit, decides when the power supply must act. The power-control signal turns ON the sleep transistor only when needed. Individual power-control signals control individual power-gating blocks, thereby allowing to turn OFF the power supplies when the pipeline is EMPTY as well as FULL.

Evaluation of Power-Gating Technique

5.1 Experimental Setup

5.1.1 Simulation Environment

This thesis uses an EDA tool - Electric VLSI Design System, to draw schematics and to do integrated circuit layouts. This thesis uses Hspice - a circuit simulator, to simulate the circuits built in Electric. The experiments in this thesis demonstrate power-gating technique with the available 180nm technology. This thesis uses BSIM3.1 transistor model.

5.1.2 Design Comparisons and Decisions

- This thesis compares power consumption of two designs:
 1. Design1 - Conventional Pipeline: The Conventional pipeline is a non-power gated pipeline with *Conventional latches* controlled by the traditional GasP circuits. Every stage of the pipeline consists of one hundred Conventional latches.
 2. Design2 - Power-gated Pipeline: The Power-gated pipeline is a pipeline with *Lazy latches* which are power gated and controlled by the Power-GasP circuits. Every stage of the pipeline consists of one hundred Lazy latches. The Power-GasP circuit is powered from true power supplies Vdd and Gnd.
- A single Conventional latch is scaled to represent one hundred Conventional

latches. Similarly, a single Lazy latch is scaled to represent one hundred Lazy latches. The transistors in the traditional GasP circuit, Power-GasP circuit and power-gating block are sized to drive a group of hundred Conventional latches and Lazy latches respectively.

- Conventional latches and Lazy latches drive the next stage of the pipeline and the wire capacitances. Extra capacitors are added between the stages to account for a step-up of the design.
- Wire capacitances were extracted from the rough layouts of the circuits. The sizes of the transistors in the schematics were modified to account for the wire capacitances.
- The two pipelines are operated at 1GHz. The data changes with a frequency of 500MHz. A current controlled voltage source (CCVS) measures the current entering a stage of the pipeline.

5.2 Why Just-In-Time

When the power of a stage is turned OFF, the voltage at the virtual supply V_{dd} decays and V_{gnd} rises. To avoid transferring ambiguous data, the voltages at V_{dd} and V_{gnd} should be charged back to supply rails before the fire signal makes the latches transparent. The power-gating technique implemented in this thesis powers V_{dd} and V_{gnd} just-in-time to avoid transferring ambiguous data to the next stage.

If the power to a stage is turned OFF, the voltages at V_{dd} and V_{gnd} decay until the leakage currents in the pull-up and pull-down transistors of the latches match.

Once the leakage currents match the voltages at V_{dd} and V_{gnd} remain constant. The voltages at virtual supplies are different if different data values are latched before the power is turned OFF. Different data values force different voltages across the transistors. If the last data value to be latched before the power was turned OFF were 0, the voltage at V_{dd} decays to a certain value. This decay in voltage is more than the decay in the voltage at V_{dd} if the last data value latched at the latch output were 1. Similarly, if the last data value to be latched before the power was turned OFF were 1, the voltage at V_{gnd} rises to a certain value. This rise in the voltage is more than the rise in the voltage at V_{gnd} if the last data at the latch output were 0.

In this thesis, one hundred latches for sequencing one hundred bits of data are grouped together. A power-gating block provides power to the hundred latches. The voltage at V_{dd} and V_{gnd} , after the power is turned OFF, is dependent on the data value at each latch output. The voltage at the V_{dd} decays to the lowest voltage value if all the latch outputs are 0. Similarly, the voltage at the V_{gnd} rises to the maximum value if all the latch outputs are 1. The time required in charging V_{dd} and discharging V_{gnd} back to the power supply rail depends on the settled voltages at V_{dd} and V_{gnd} . The worst case time required is if all the latch outputs are either 1 or 0. The charging of V_{dd} is more critical than the discharging of V_{gnd} . The NMOS sleep transistor is turned ON one gate delay before the fire signal drives the latches transparent (refer to section 4.2). One gate delay provides sufficient period to discharge the accumulated charge on V_{gnd} . However, the PMOS sleep transistor is turned ON just as the fire signal drives the latches transparent. Therefore, the PMOS sleep transistor should charge V_{dd} just in time as fire signal drives the latches transparent.

Figure 5.1: Just-In-Time

Sequence - 1:Incoming data 2:pred signal HI in anticipation of new data 3:Vvdd is charged to the rail 4:Vvdd is charged up just in time as the fire signal drives the latch to reliable transfer the data to latch output

Figure 5.1 shows the waveforms of a simulated power-gated circuit when 0 was the last data value to be latched before the power was turned OFF. Therefore, Vvdd decays more than Vgnd rises. The figure shows that Vvdd decays to 1.6V whereas Vgnd almost remains almost zero. In anticipation of new incoming data, the GasP stage pulls the predecessor state wire HI. If the required condition for the GasP module to fire, predecessor state wire is HI and successor state wire is LO, is satisfied, the power to the stage is turned ON. The virtual power supplies are charged before the fire signal makes the latches transparent. When the latches are made transparent, the data at the input is passed reliably to the latch output.

5.3 Results: Comparing Conventional Pipeline and Power-gated Pipeline

This section provides performance evaluation of the power-gating technique introduced in this thesis. This section compares the pipelines with the Conventional latches that are not power gated and the Lazy latches that are power gated.

The operation of a stage of a Conventional pipeline can be divided in two periods,

- Idle: In this period, the pipeline stage is in a quiescent state either because it is empty or waiting for a vacancy. This period lacks data activity. Leakage current is the only current consumption.
- Active: In this period, the pipeline stage deals with valid data bursts. Data bursts causes activity in the stage. The current drawn from the circuits depends on the data activity.

Figure 5.2: Periods of operation in a Power-gated pipeline.

The waveforms in Figure 5.2 are an example to represent the periods of operation of a Power-gated pipeline. The blue curve represents the voltage across the virtual

power supply Vvdd. I_{ds} is the current drawn by a stage of the Power-gated pipeline. The operation of a Power-gated pipeline can be divided in four periods,

- Idle: This period is same as the idle period in the Conventional pipeline. Constant leakage current I_{STEADY} is the only source of current consumption. Virtual power supply Vvdd has a constant voltage of V_{STEADY} .
- Startup: A power-gated circuit requires a startup period to turn ON the power supply ready for the incoming data. The startup period is a short duration period during which a pipeline stage draws a large amount of current to charge all the capacitances of that stage.
- Active: This period is same as the active period in the Conventional Pipeline. The current drawn depends on the data activity.
- Transient: If idle periods are detected, the power to the stage is turned OFF. The transient period is the time taken by the leakage currents in the pull-up and pull-down networks to match. After the transient period, the circuit enters a stable state.

The startup and transient periods are absent in the Conventional pipeline. The startup and transient periods are associated with turning ON and OFF the power. The transient period in the Figure 5.2 is visible. However, practically, the transient period continues for a long time, of the order of 10^6 cycles of operation. This value corresponds to I_{on}/I_{off} ratio of 180nm technology. The resource limitations prohibit the measurement of the currents in idle period following the transient period. Hspice assumes the initial inactivity as the idle period and provides idle current measurements.

The power gating technique is validated by measuring the currents for different periods of operation for Conventional and Power-gated pipelines. Figure 5.3 shows a stage of the Conventional pipeline. The current drawn by individual parts of pipelines are measured by the current controlled voltage sources (CCVS). Sum is the total current drawn by a stage of the Conventional pipeline. $I(\text{convlatch})$ and $I(\text{gasp})$ are the currents drawn by the Conventional latch and the GasP circuit for a stage. Similarly, Figure 5.4 shows a stage of the Power-gated pipeline. A stage of Power-gated pipeline consists of three parts - power-gating block, the Lazy latch and the Power-GasP circuit. Sum is the total current current drawn by a stage of the Power-gated pipeline. $I(\text{total})$ is the total current drawn by all the three parts of a stage. $I(\text{lazylatch})$ and $I(\text{powergasp})$ are the currents drawn by the Lazy latch and the Power-GasP circuit.

Figure 5.3: A stage of a simulated circuit of the Conventional Pipeline.

Figure 5.4: A stage of a simulated circuit of the Power-gated Pipeline.

The table in Figure 5.5 shows the currents measured in active and idle periods for different parts of a stage of the pipelines. The current drawn by the Power-GasP circuit and the power-gating block are summed together in a column. The current drawn during the startup period charges the bypass and wire capacitance. The charge stored in the capacitance supplies the required current during the active period. Therefore, the column for startup period is absent. Transient period is an unpredictable period where the currents are constantly changing because of the transition from the active to idle periods. During the transient period, the charge stored in the capacitance is discharged and no additional current is drawn. Therefore, the column for transient period is absent. In the table, compare current values with same colors in a single row. The last column in the table shows the ratio of the total active and idle currents of the Power-gated pipeline and the Conventional pipeline.

Currents	Non-power gated pipeline			Power gated pipeline		
	Conventional Latches	GasP	Total T_{np}	Lazy Latches	GasP + Power Gating	Total T_{pg}
Active Current (mA)	20	4	24	16	4	20
Idle Current (mA)	$5*10^{-6}$	$1*10^{-6}$	$6*10^{-6}$	$1.8*10^{-6}$	$3.2*10^{-6}$	$5*10^{-6}$
						T_{pg}/T_{np}
						0.83
						0.83

Figure 5.5: Active and Idle currents of Conventional and Power-gated pipeline.

The table indicates the currents drawn by different parts of the pipelines for different periods of operation. In the active and idle periods, the Lazy latch draws 20% less current than the Conventional latch. The Lazy latch turns ON only when necessary and therefore, during active periods the Lazy latch draws less current than the Conventional latch. The Lazy latch offers half the load to the Power-GasP circuit as offered by the Conventional latch to the traditional GasP circuit. Therefore, in spite of the additional drivers in the Power-GasP circuit, the Power-GasP circuit draws same current as the traditional GasP circuit during the active periods. Overall in the active period, the Power-gated pipeline draws 16% less current than the Conventional pipeline. During the idle period, the Lazy latch is power gated and the Conventional latch is powered from true power supply. Therefore, the Lazy latch draws 64% less current than the Conventional latch. However during the idle period, the Power-Gasp circuit and the power-gating block draw more leakage current than the traditional GasP circuit in the Conventional pipeline. The wide sleep transistors in the power-gating block and the additional drivers in the Power-GasP circuit add to the leakage current in the Power-gated pipeline. Therefore, during the idle period the Power-gated pipeline draws 68% less current than the Conventional pipeline. Overall in the idle period, the Power-gated pipeline draws 16% less current than the Conventional pipeline.

The power consumed by the circuit is given by $P = VI$, where V is voltage in volts(V) and I is current in Amperes(A). For the designs in this thesis, the supply voltage is 1.8V. The currents are measured for a stage of the pipeline. The product gives the power consumed by one stage of the pipeline. This thesis compares the Conventional pipeline and the Power-gated pipeline based on the power consumed by a single stage of the pipelines.

In the active period 16% power savings are achieved. The Lazy latch introduced in this thesis contributes to 16% power savings during active period. The idle periods have to be very long to save more energy than the energy consumed during the active period. The energy savings achieved during idle periods depend on their duration. The energy savings achieved increases with the increase in the idle period. The idle current is of the order 10^6 less than the active current. Therefore, the pipeline stage should be idle for 10^6 cycles before achieving considerable energy savings.

The power gating technique introduced in this thesis is evaluated by comparing the Power-gated pipeline against the Conventional pipeline. The Power-gated pipeline consumes 16% less power than the Conventional pipeline during active period. The power saving achieved during active period are because of the Lazy latch. Power-gated pipeline also consumes 16% less power than the Conventional pipeline. The power savings achieved during idle period are because of the power gating technique. Ideally the power saving technique requires additional wide sleep transistors and control circuit to control the sleep transistors. The additional gates increase the power consumption during the active period. The Lazy latch introduced in this thesis compensates for the additional gates required for power gating. Therefore, 16% power savings are achieved during active period as well. The power savings achieved during idle period is proportional to the duration of the idle period. To achieve considerable power savings, more than those obtained during the active period, a pipeline stage should be idle for at least 10^6 cycles of operation.

Conclusions

6.1 Contributions of this work

This thesis presents for the first time a fine-grained power gating technique for the GasP family of asynchronous circuits. Fine-grained power gating can control power for each stage in the pipeline. The sleep transistors for each stage control the power to that stage and its combinational circuit. The power to the stage is turned ON in anticipation of receiving data. The power is turned OFF as long as the stage is idle either because it is empty or because the pipeline is clogged. The control circuit for implementing power gating is simple. This thesis introduces a new state wire called power-state wire, in addition to the traditional data state wire. The two state wires determine the requirement of power for each stage. A new power-control signal actually turns ON and OFF the sleep transistors to power the stage.

The Conventional latch avoids the need to save the state in the event of power gating. A new latch design - the Lazy latch is introduced in this thesis which keeps the state in the event of power gating. The Lazy latch drives its output only when necessary. The Lazy latch offers reduced power consumption in active and idle periods of operation. The Conventional latch is insensitive to noise whereas the Lazy latch may be sensitive to noise. Based on advantages and disadvantages of the Conventional and Lazy latches, a new latch design - the Blended latch is proposed in this design.

Power gating is evaluated for the Lazy latch. This thesis compares a Power-gated pipeline consisting of the power-gated Lazy latches and a Conventional pipeline consisting of non-power gated Conventional latches. The reduced load offered by the Lazy latch and turning OFF of the amplifier transistors when unnecessary achieves 16% power savings during active period. If no work is done for a long time then energy is saved. Energy savings achieved during idle period depends on the duration of the inactivity. The longer the inactive period the more the energy savings achieved. The idle periods of the order of 10^6 cycles are required to achieve energy savings more than those obtained during active periods.

6.2 Power gating technique implementation requirement

A new library of gates is required to implement power gating. The existing library is powered from true power supply Vdd and Gnd. For power gating, the gates in the latches should be powered from virtual power supplies - Vvdd and Vgnd. Modeling new gates requires building a new library.

The power gating technique introduced in this thesis may be applied to latches and combinational circuits. The power is turned OFF if the latches and the combinational circuits are idle. If the power is turned OFF, the states of the circuits are lost. To avoid the loss of the state, keepers are required. The Lazy latch introduced in this thesis maintains the state in the event of power gating. All the combinational circuits will need keepers. If the combinational circuits of a stage don't have keepers then in order to avoid loss of state, the stage cannot be power gated. Therefore, keepers are the requirements to maintain the state if the stages are to be power gated when they are full.

6.3 Future Work

1. Power multiple stages from single sleep transistor: In the thesis, each stage is powered from a PMOS sleep transistor and an NMOS sleep transistors. Dedicating sleep transistors to each individual stage increases the circuit area. Powering multiple stages from the same sleep transistor can reduce the circuit area. The design of the GasP circuits prevent two consecutive stages from firing at the same time. Even three consecutive stages draw current from the power supply at different times. In a group of three stages, if the second stage fires, the first stage fires 4 gate delays later while the third stage fires 6 gate delays after the second stage fires. Therefore, a group of three stages can be powered from a dedicated PMOS sleep transistor and a dedicated NMOS sleep transistor. Reducing the number of sleep transistors will reduce the circuit area. However, a group of three stages power gated together will require changes in the control circuit. The control circuit should turn ON the sleep transistors in anticipation of the receiving data at the first stage. The power can be turned OFF only after the received data is processed by the last stage in the group.
2. Gradually turn ON multiple sleep transistors: This thesis uses a single wide sleep transistor to power gate a stage of a pipeline. Multiple sleep transistors of smaller widths should be used to get more effective power savings. Multiple transistors of smaller widths leak less than a single wide transistor. When the power is turned ON, the sudden inrush of current causes IR drop, Ldi/dt noise and electromigration of the power bus [20]. The sudden inrush of current can be avoided by gradually turning ON the multiple sleep transistors.

When the power of a stage is turned ON, all the circuit parts don't need current at the same time. Therefore, a pipeline stage can be powered from multiple sleep transistors of smaller widths. The multiple sleep transistors can be turned ON gradually with the current requirements.

3. Slowly turn ON the power: Implementing power gating requires additional drivers and sleep transistors. The Just-In-Time power supply requires turning ON the power supply quickly. The just-in-time circuit requires wide sleep transistors and wider transistors in the additional drivers. These wide transistors draw more current during active periods and leak more during idle periods. The effect of power gating takes longer time to overcome the additional cost of implementing just-in-time. Instead of just-in-time the stages can be turned ON slowly. Turning ON the stages slowly will require smaller transistors in power-gating block and drivers. Smaller drivers mean slowly turning ON the power supply to the stages. Slowly turning ON the power supply will be more effective in achieving considerable power savings.
4. Evaluate the technique for recent technologies: The gate sizes are getting smaller with the recent technologies. Reduction in the gate sizes need reduction in the supply voltages and threshold voltages as well. Reducing the threshold voltage increases the leakage current. The power gating technique introduced in this thesis can be evaluated for newer technologies. The newer technologies can be modeled with the existing 180nm technology. The transistors of 180nm technology can model newer technologies by adjusting the I_{on}/I_{off} ratio. The I_{on}/I_{off} ratio can be changed by adjusting the body bias [20]. Adjusting the body bias changes the threshold voltage of the

transistor. Subthreshold current is exponentially dependent on the threshold voltage. Thus, a transistor model modeling recent technologies can be implemented by adjusting the body bias.

5. Gate leakage - Another leakage problem: Finally, in recent technologies thickness of the oxide is constantly decreasing. Gate leakage occurs when carriers tunnel through a thin gate dielectric when a voltage is applied across the gate [20]. Gate leakage can be reduced by stacking the transistors in series as well as arranging the inputs. More new methods should be introduced to deal with the problem of gate leakage.

References

- [1] A. Chandrakasan and R. W. Brodersen, *Low-Power CMOS Design*. Wiley-IEEE Press, January 1998.
- [2] International Technology Roadmap for Semiconductors 2002 Update. [Online]. Available: <http://www.itrs.net>
- [3] N. Kim, T. Austin, D. Baauw, T. Mudge, K. Flautner, J. Hu, M. Irwin, M. Kandemir, and V. Narayanan, “Leakage current: Moore’s law meets static power,” *Computer*, vol. 36, no. 12, pp. 68 – 75, dec. 2003.
- [4] K. Roy, S. Mukhopadhyay, and H. Mahmoodi-Meimand, “Leakage current mechanisms and leakage reduction techniques in deep-submicrometer cmos circuits,” *Proceedings of the IEEE*, vol. 91, no. 2, pp. 305 – 327, feb 2003.
- [5] S. Mutoh, T. Douseki, Y. Matsuya, T. Aoki, S. Shigematsu, and J. Yamada, “1-v power supply high-speed digital circuit technology with multithreshold-voltage cmos,” *Solid-State Circuits, IEEE Journal of*, vol. 30, no. 8, pp. 847 – 854, aug 1995.
- [6] I. E. Sutherland, “Micropipelines,” *Commun. ACM*, vol. 32, no. 6, pp. 720–738, Jun. 1989. [Online]. Available: <http://doi.acm.org/10.1145/63526.63532>
- [7] I. Sutherland and S. Fairbanks, “Gasp: a minimal fifo control,” in *Asynchronous Circuits and Systems, 2001. ASYNC 2001. Seventh International Symposium on*, 2001, pp. 46 –53.
- [8] N. Jamadagni, “Five-stage ring oscillator, designers delight!” aug. 2009.

- [9] B. S. Ivan Sutherland and D. Harris, *Logical Effort Designing Fast CMOS Circuits*. Morgan Kaufmann, February 1999.
- [10] A. Peeters, F. te Beest, M. de Wit, and W. Mallon, “Click elements: An implementation style for data-driven compilation,” in *Asynchronous Circuits and Systems (ASYNC)*, 2010 IEEE Symposium on, may 2010, pp. 3–14.
- [11] J. Tschanz, S. Narendra, Y. Ye, B. Bloechel, S. Borkar, and V. De, “Dynamic sleep transistor and body bias for active leakage power control of microprocessors,” *Solid-State Circuits, IEEE Journal of*, vol. 38, no. 11, pp. 1838–1845, nov. 2003.
- [12] Z. Hu, A. Buyuktosunoglu, V. Srinivasan, V. Zyuban, H. Jacobson, and P. Bose, “Microarchitectural techniques for power gating of execution units,” in *Low Power Electronics and Design, 2004. ISLPED '04. Proceedings of the 2004 International Symposium on*, aug. 2004, pp. 32–37.
- [13] K. Shi and D. Howard, “Sleep transistor design and implementation - simple concepts yet challenges to be optimum,” in *VLSI Design, Automation and Test, 2006 International Symposium on*, april 2006, pp. 1–4.
- [14] A. Youssef, M. Anis, and M. Elmasry, “A comparative study between static and dynamic sleep signal generation techniques for leakage tolerant designs,” *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, vol. 16, no. 9, pp. 1114–1126, sept. 2008.
- [15] K. Usami and N. Ohkubo, “A design approach for fine-grained run-time power gating using locally extracted sleep signals,” in *Computer Design, 2006. ICCD 2006. International Conference on*, oct. 2006, pp. 155–161.

- [16] T. Lin, K.-S. Chong, B.-H. Gwee, and J. Chang, “Fine-grained power gating for leakage and short-circuit power reduction by using asynchronous-logic,” in *Circuits and Systems, 2009. ISCAS 2009. IEEE International Symposium on*, may 2009, pp. 3162 –3165.
- [17] T. W. Kwan and M. Shams, “Design of high-performance power-aware asynchronous pipelined circuits in mos current-mode logic,” in *Asynchronous Circuits and Systems, 2005. ASYNC 2005. Proceedings. 11th IEEE International Symposium on*, march 2005, pp. 23 – 32.
- [18] T. Kawano, N. Onizawa, A. Matsumoto, and T. Hanyu, “Adjacent-state monitoring based fine-grained power-gating scheme for a low-power asynchronous pipelined system,” in *Circuits and Systems (ISCAS), 2011 IEEE International Symposium on*, may 2011, pp. 2067 –2070.
- [19] B. Nikolic, V. Oklobdzija, V. Stojanovic, W. Jia, J. K.-S. Chiu, and M. Ming-Tak Leung, “Improved sense-amplifier-based flip-flop: design and measurements,” *Solid-State Circuits, IEEE Journal of*, vol. 35, no. 6, pp. 876 –884, june 2000.
- [20] N. Weste and D. Harris, *CMOS VLSI Design: A Circuits and Systems Perspective*. Addison-Wesley, March 2012.