DEVELOPMENT OF A CMOS $I_{DDq}$ TESTING ENVIRONMENT

By

RICHARD J. BURGESS JR.

Bachelor of Science in Electrical Engineering

Worcester Polytechnic Institute

Worcester, Massachusetts

1988

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
MASTER OF SCIENCE
May, 1995
DEVELOPMENT OF A CMOS $I_{DQ}$

TESTING ENVIRONMENT

Thesis Approved:

[Signatures]

Thesis Adviser

Dean of the Graduate College
PREFACE

A majority of defects found in CMOS technology display elevated quiescent current magnitudes but still may pass functionality tests. By monitoring this power supply current, defect coverage can be elevated past the traditional stuck-at-fault coverage. This study provides a test methodology centered around current supply monitoring. By analyzing fabrication data, defect models, built-in current sensors, current and delay estimation, test set generation, and the QTAG standard, a technique is developed for CMOS integrated circuit testing. A built-in current sensor is presented, which through simulation, exhibits fast detection time. Novel techniques to enhance this time are also presented.

I would like to thank my advisor, Dr. Louis Johnson, for the encouragement and guided freedom he has given to me, while researching, throughout this project. Also, I wish to express my appreciation to my committee member, Dr. Chris Hutchens for his guidance and insight into testing and analog circuit design.

Special thanks to Dr. Carl Latino for serving on my advisory committee and Dr. James Baker, the former head of electrical engineering, who I feel was very instrumental in my acceptance in the Oklahoma State graduate program. Thanks also to Dr. Keith Baker of Phillips Semiconductor Netherlands Division who provided me with the information on the QTAG standard and Dr. R. D. McLeod of the University of Manitoba for providing me with information on his fabricated built-in current sensor.
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>( I_{DDQ} ) Basics</td>
<td>1</td>
</tr>
<tr>
<td>Built-in Current Sensing</td>
<td>4</td>
</tr>
<tr>
<td>A Testing Environment</td>
<td>5</td>
</tr>
<tr>
<td>II. CMOS FAULTS AND FAULT MODELS</td>
<td>7</td>
</tr>
<tr>
<td>Fabrication Process Accessment</td>
<td>7</td>
</tr>
<tr>
<td>Gate-Oxide Shorts</td>
<td>8</td>
</tr>
<tr>
<td>Floating Gate and Opens</td>
<td>9</td>
</tr>
<tr>
<td>Bridging Faults</td>
<td>12</td>
</tr>
<tr>
<td>III. RECENT BUILT-IN CURRENT SENSORS</td>
<td>14</td>
</tr>
<tr>
<td>Introduction</td>
<td>14</td>
</tr>
<tr>
<td>Maly and Patyra BICS</td>
<td>15</td>
</tr>
<tr>
<td>Shen et al. BICS</td>
<td>15</td>
</tr>
<tr>
<td>Hsue and Lin BICS</td>
<td>16</td>
</tr>
<tr>
<td>Singh and Hurst BICS</td>
<td>20</td>
</tr>
<tr>
<td>Brown and McLeod BICS</td>
<td>22</td>
</tr>
<tr>
<td>Integrating BICS</td>
<td>23</td>
</tr>
<tr>
<td>IV. PROPOSED BICS CIRCUIT</td>
<td>24</td>
</tr>
<tr>
<td>Introduction</td>
<td>24</td>
</tr>
<tr>
<td>( V_{DD} ) line Monitored Differential BICS</td>
<td>25</td>
</tr>
<tr>
<td>Small Signal Analysis</td>
<td>27</td>
</tr>
<tr>
<td>Comparison Performance</td>
<td>29</td>
</tr>
<tr>
<td>Current Profiling</td>
<td>32</td>
</tr>
<tr>
<td>Mismatch Effects and Minimization</td>
<td>34</td>
</tr>
<tr>
<td>V. CURRENT AND DELAYS IN CMOS CIRCUITS</td>
<td>35</td>
</tr>
<tr>
<td>Transient Current and Delays</td>
<td>35</td>
</tr>
<tr>
<td>( I_{DDQ} ) Estimation</td>
<td>36</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>29</td>
</tr>
</tbody>
</table>

Hurst Sensor Comparison
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>2</td>
</tr>
<tr>
<td>1.2</td>
<td>4</td>
</tr>
<tr>
<td>2.1</td>
<td>9</td>
</tr>
<tr>
<td>2.2</td>
<td>10</td>
</tr>
<tr>
<td>3.1</td>
<td>13</td>
</tr>
<tr>
<td>3.2</td>
<td>15</td>
</tr>
<tr>
<td>3.3</td>
<td>16</td>
</tr>
<tr>
<td>3.4</td>
<td>17</td>
</tr>
<tr>
<td>3.5</td>
<td>18</td>
</tr>
<tr>
<td>3.6</td>
<td>20</td>
</tr>
<tr>
<td>3.7</td>
<td>20</td>
</tr>
<tr>
<td>3.8</td>
<td>23</td>
</tr>
<tr>
<td>4.1</td>
<td>25</td>
</tr>
<tr>
<td>4.2</td>
<td>28</td>
</tr>
<tr>
<td>4.3</td>
<td>31</td>
</tr>
<tr>
<td>4.4</td>
<td>33</td>
</tr>
<tr>
<td>6.1</td>
<td>41</td>
</tr>
<tr>
<td>7.1</td>
<td>44</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

1.1 I_{DDq} Basics

I_{DDq} testing or Current Supply Monitoring (CSM) has increased in popularity as an effective technique for detecting defects in CMOS integrated circuits. These defects in many instances are undetectable by the classical stuck-at fault model. Thus, I_{DDq} testing is a powerful test technique to improve the fault coverage and to detect defects that are hard or impossible to detect using traditional voltage methods [1]. I_{DDq} testing is performed by measuring the power supply current that exists after the logic states have settled and the circuit is in a quiescent mode. In a full CMOS fault-free circuit this current is negligible since either the NMOS or PMOS transistor group is off and the output has reached a voltage rail. The circuit does contain some junction leakage current with magnitudes in the range of nanoamperes to tens of microamperes depending on circuit size. However, in the presence of various physical defects the I_{DDq} current will increase by a few orders of magnitude[2].

Even though the circuit shows elevated I_{DDq} the circuit can produce the correct logic output during functional testing. Figure 1.1 shows the I_{DDq} and the output voltage for a faulty Boundary Scan cell developed in [3]. Note that the I_{DDq} fault is excited only under certain input values. Some other defects may produce high I_{DDq} values always independent of the input. Magnitudes of I_{DDq} are functions of effective transistors resistances in the faulty current.
Figure 1.1  (A) $I_{DDq}$  (B) Output Voltage of Boundary Scan Cell [3]
As seen in Figure 1.1B, the correct output voltage is exhibited under all input patterns, however, rise and fall times are degraded. Since correct output levels are achieved, $I_{DDq}$ testing can be applied to truly know the "health" of the integrated circuit. Even though these circuits passes traditional functional testing, they have been shown to fail more frequently in life than normal IC's [4]. This raises reliability issues especially in critical applications like medical and aerospace.

IC foundries have utilized off-chip $I_{DDq}$ testing as part of production testing since the mid 70's [5-6]. These off-chip sensors are limited in speed due to the large time constant associated with the total chip capacitance. Other factors that degrade off-chip $I_{DDq}$ sensors include (1) the pulse width of the CMOS IC transient current, (2) the impedance loading of the tester, (3) current leakages into or out of the tester, and (4) the high noise environment of the load board [5]. Maly states [7], "the best rate of off-chip current testing still is, and will be, much slower than the rate which can be achieved by the normal 'voltage oriented' testing technique". A much higher rate can be achieved by a built-in current sensor (BICS), where the sensor is located on-chip. This falls into a self-checking philosophy which can be integrated with other built-in test strategies which is particularly useful in high density packing such as multichip modules (MCMs). With built-in sensors and appropriate partitioning of the circuit, an increase rate of testing as high as two orders of magnitude can be achieved [8]. In this thesis a new BICS is designed and an environment for using it is developed. Recently, however QTAG headed by Dr. Keith Baker at Phillips Research Laboratories have tried to develop a standard for current testing using off-chip current sensors. This will be dealt with in Chapter 7.
1.2 Built-in Current Sensing

Since the sensor is to monitor supply current there are two possible locations in relation to the circuit-under-test (CUT) as illustrated in Figure 1.2. Large circuits can be partitioned and have their own dedicated BICS. If the BICS is between GND and the CUT,

![Diagram of possible BICT connections to CUT](image)

the CUT ground node is at a virtual ground. Thus, there are two noisy supply lines. If the BICS is between Vdd and the CUT then there is one clean ground line but some of the leakage currents to ground will go undetected. The choice for on-chip sensor is usually between the CUT and GND and the choice for off-chip is usually between Vdd and CUT. The on-chip sensor causes an increase in parasitics which cause a performance degradation unless the BICS can be disconnected from the CUT after testing. This can be accomplished at the expense of extra pin counts.
Some characteristics of the BICS should be:

1. Testing speed should be fast
2. Minimal Performance degradation
3. Small silicon area
4. High current resolution
5. Minimal voltage drop due to the BICS
6. Ability to sink the high transient current

One through five are self-explanatory. Six is described below. Transient power supply current is the current when the logic levels are in transition. This results in a short-circuit current and the current required to charge and discharge the capacitance. The latter is usually the dominant term [9]. This current is shown in Figure 1.1. It is a function of clock waveform and frequency, timing skew, load capacitances, and input waveforms [10]. Peaks magnitudes can reach several hundred milliamperes. The sampling of supply current can be done any time after the current tail has reached a fixed low-value of current for the fault-free case [11]. The BICS must sink this current quickly or long delays times will result. However, the sensor must be sensitive enough to distinguish small current differences and determine when the current tail has ended. In order for this to occur, some form of clocking must be present which adds area. If this clocking is tied with the system clock, it could slow down the CUT. Some BICS do not use clocks but monitor through the transients. Thus, the output is a fault when the transient occurs and then the output changes in fault-free cases.

1.3 A Testing Environment

If, in the presence of a defect, different circuit behavior is exhibited than in fault-free case, then by testing for this, one can determine if a defect exists. However, to know what the defects and their characteristics are one must look at the fabrication process. To develop a
test set to test for these defects accurate fault models must be developed for these defects.

The traditional stuck-at fault model does not accurately depict most defects in CMOS circuits.

From accurate fault models, tests sets must be developed and then collapsed into smaller test sets if timing constraints are required. These test sets will then be applied to the circuit at the wafer, board, and system level. Thus, a testing environment encompasses the following:

1. Knowledge of defects and behavior
2. Accurate fault models of these defects
3. Minimum test sets for detection of these defects given a desired fault coverage
4. Sensors to determine the outcome of these test sets
5. Rules to facilitate this testing

In this thesis a new built-in current sensor is presented which attempts to combat previous reported BICS pitfalls and an environment in which to use this sensor is developed.

The rest of this thesis is organized as follows. Chapters 2 gives an overview of CMOS defects and their detectability by $I_{DDQ}$ testing. Also rules to facilitate current testing with be given. Chapter 3 describes previously reported BICS and their strengths and weaknesses. Chapter 4 presents the new BICS developed by these thesis. Also Current Profiling, which is used to increase the functionality and test speed of the BICS, is explained. Chapter 5 demonstrates a method to determine maximum transient current estimation and delays in CMOS circuits. This is needed since the BICS of this thesis is clocked and a determination of the dynamic current tail is necessary. Also a novel method to improve test time based on current profiling is developed. Chapter 6 reports previous test set generation methods for $I_{DDQ}$ testing. Chapter 7 explains the new QTAG standard being developed for $I_{DDQ}$ testing. Chapter 8 gives insight into the future direction for $I_{DDQ}$ testing.
Chapter 2

CMOS Faults and Fault Models

2.1 Fabrication Process Assessment

Any good fault model should accurately predict the behavior of the fault. It is widely accepted that the stuck-at model (a node stuck-at one or zero) for most CMOS defects does not [12-14]. Since faults are just behavioral properties of physical defects caused by the fab. process, the logical starting place for developing accurate fault models is at the fabrication level. Researchers at Carnegie-Mellon developed the idea of Inductive Fault Analysis [15] which consists of three steps:

1. Physical defect generation from statistical parameters of the fab. process.
2. Determination of circuit level faults caused by these defects.
3. Classification and ranking of these fault types.

In IFA, each spot on the wafer is characterized by a mask vector consisting of 1's and 0's. These correspond to each physical layer either being transparent or opaque, respectively. Defects are generated by changing the value of one variable in the vector, from statistical analysis, and interpreting its impact at the logic level. The faults are then classified and ranked according to their likelihood of occurrence. This is significant information because your test set generation should be geared toward detection of the most predominant fault. It was determined that bridging faults account for the most number of detects [15]. However as
transistor sizes shrink, gate-oxide shorts seem to be the major defect [16-18]. The rest of the chapter attempts to explain the Gate-oxide short (GOS), the floating gate and open, and the bridging fault models. Following this will be a set of rules to facilitate $I_{DDQ}$ testing.

### 2.2 Gate-Oxide Short

Gate-oxide shorts been have shown to have increased $I_{DDQ}$ levels but may not lose functionality [16-18]. GOS's can degrade with time and eventually cause failure. Thus, for reliability purposes $I_{DDQ}$ monitoring is necessary. A gate-oxide short is an electrical connection from the gate to the channel or source/drain regions. Thus, there is an undesired current path though the gate-oxide. GOS's can result from lithography defects or electrostatic discharge. Most GOS's are likely to be found between the gate and channel [18].

The circuit properties of GOS's were analyzed in [18]. It was found that for an NMOS a non-linear IV characteristic occurred for the pn junction of the n-doped polysilicon gate and the p-well short. Ohmic IV characteristics arose form shorts between the n-doped poly and the n-diffusion. For the PMOS non-linear IV characteristics arose form the n-doped gate to p-diffusion; However, n-doped gate and n-doped substrate did not produce ohmic but a non-linear relationship. This may be from the combined influence of channel inversion and the space charge between the inverted channel and substrate [18]. When the GOS is sensitized abnormal $I_{DDQ}$ levels results. In [16] a lot of CMOS 8085's with GOS's were evaluated. 14 out of 15 failed the $I_{DDQ}$ test but passed the functional test. These $I_{DDQ}$ levels were several orders of magnitude above the normal values.

In Figure 2.1 [14] a GOS circuit level fault model is shown. The model consists of three components. The first is a barrier(B) which models junction between the gate and
channel. This has a threshold voltage, a breakdown voltage, and forward and breakdown resistances. Rs is the contact resistance of the short. Two transistors are created where the barrier splits them. This is dependent on the location of the short. The two new transistors have beta’s that are dependent on the location of the barrier divided by the parallel combination of their beta’s values. From this model the $I_{DDQ}$ value is larger when the fault is located closer to the source. However, only for large values of Rs, will the defect be undetected by $I_{DDQ}$ testing. A circuit of three inverters [14] was created and once again, the GOS faulty circuit produced correct logic levels but exhibits an $I_{DDQ}$ 3 or 4 orders of magnitude higher than normal. Consult [16] for a more advanced GOS model.

2.3 Floating Gate and Opens

When a wanted connection is missing an open occurs. These opens have various analog and time dependent behavior [19-21]. An open at the output can cause sequential behavior in a CMOS gate. Some opens will not have elevated $I_{DDQ}$ levels. This is dependent on the steady-state values of the p and nMOS transistors. Thus, $I_{DDQ}$ levels may be observed after some time. This, time may be longer than the test set time. However, if the node was not driven out of it’s high impedance state while settling, the fault can be detected by $I_{DDQ}$ testing[19].
One specific type of open is an open in the poly gate path. This is known as the floating gate defect. Some floating gate transistors may turn on transistors that should be off. This can cause elevated $I_{DDQ}$ if it lies in a path from Vdd to GND. The floating gate voltage depends on the coupling capacitances of the transistor device and on the surrounding circuitry. Thus, the behavior of the transistor is dependent on these values. The following model for the floating gate with implications on $I_{DDQ}$ testing was developed in [14] and repeated below.

Figure 2.2 shows the transistor floating gate coupling capacitance, where
1. $C_{gs0}$ is the capacitance of gate-source overlap
2. $C_{gd}$ is the capacitance of gate-drain overlap
3. $C_{pb}$ is the capacitance of the poly-bulk. It depends on the length of poly path. Thus, is dependent on the location of the open.
4. $C_{mp}$ is the capacitance of the overlapped metal-poly.

![Figure 2.2 Capacitance associated with floating gate](image)

To estimate the induced gate voltage, the extrinsic capacitances surrounding the floating gate, the charges induced in the gate, and the gate oxide charge must be considered. If you find the charge equation at node G and set it equal to zero and include the charge induced at the gate by the intrinsic part of the transistor then $V_g$ is found below:

$$V_G = \frac{V_d C_{gd0} + V_m C_{mp} + Q_B + Q_I + Q_O}{C_{gd0} + C_{gs0} + C_{pb} + C_{mp}}$$
Where QB is the bulk depleted charge, QI is the channel inversion charge, and QO is the effective charge in the oxide. From the above equation, if \( C_{gdo} \) or \( C_{mp} \) increases \( V_g \) increases and if \( C_{pb} \) or \( C_{gso} \) increases \( V_g \) decreases. \( C_{pb} \) and \( C_{mp} \) are location dependent and if \( C_{pb} \) is large, the transistor may always be off regardless of \( C_{mp} \). Also, there is a linear relationship between the \( V_g \) and \( V_m \) and \( V_d \). Thus, the amount of current is dependent on the floating gate caps. and its operating region. If you take \( V_d=V_m=5v \) then, for the same \( C_{mp} \) and for higher values of \( C_{pb} \) the gate voltage and quiescent current decrease, for the same value of \( C_{pb} \) and higher \( C_{mp} \), the gate voltage and quiescent current increase. If \( C_{mp} \) doesn’t exist the model still holds with zero values for \( C_{mp} \).

For the floating gate to be \( I_{DDq} \) testable, the \( C_{mp} \) and \( C_{pb} \) must have value necessary to turn the transistor on. If this is the case the correct input pattern must be applied to the gate which cases a conduction path from \( V_{dd} \) to GND. Since, the gate voltage is highly dependent on the capacitances the expected fabrication parameters must be analyzed to ensure \( I_{DDq} \) testability.

Other opens were studied in [20-21]. It was found that signal coupling across breaks caused by electromigration appears to be directly related to the Fowler-Nordheim tunneling. For a floating input attached to a well or substrate though a reverse biased diode, the voltage will float to the substrate potential in an amount of time dependent on light and temperature.

2.4 Bridging Faults

Bridging faults are undesired electrical shorts between two or more lines. These can be at the transistor node, I/O of a logic gate, or to a power rail. These defects exhibit linear
or non-linear behavior. The ones with non-linear behavior are GOS's, soft pn junctions, and transistor punch through. Linear shorts have an impedance value which is dominated by the resistive term. Thus, the bridging fault can be modeled by a resistor. This is different than the typically assumed zero impedance. Thus, the defect behavior is determined by the relative impedances in the driving stages and the relative impedance of the bridge resistor which may be considerably different [22]. It has been shown in [22] that for resistive bridging to inputs of gates, an excessive $I_{\text{DDq}}$ can be measured up to about 50K while the stuck at fault model is only good until about 5K. For resistive transistor stuck-ons, these values are about the same however, a large delay is exhibited below 1K. For resistive stuck-open transistors, the fault is undetectable by both methods. However, above 100K an excessive delay and a fine $I_{\text{DDq}}$ are exhibited.

Depending on the circuit topology and the location of the bridge an anomalous reverse conduction can occur [2]. If the body of the transistor is not connected to the supply rail but connected to the source, then there is a chance, in the presence of a defect, that the drain voltage will be lower than the source voltage. Thus, the drain-substrate junction is forward-biased and a low resistance is formed between the drain and source through the substrate. This current can cause permanent failure of the device.

2.5 $I_{\text{DDq}}$ Testing Rules

There are circumstances in which $I_{\text{DDq}}$ will not give correct test results. These cases are usually due to some sort of charge sharing, floating nodes, or other undetectable faults masking other detectable faults. Also BiCMOS or any logic which has a high quiescent current and some sequential logic which has feedback control loops are two other topologies
that are undetectable by $I_{\text{DDq}}$ testing. Even though these circuits are not detectable by $I_{\text{DDq}}$, large portions of the circuit may be. Therefore, partitioning may be necessary. A set of design rules to ensure the proper use of $I_{\text{DDq}}$ testing is given below [23].

1. The gate and drain nodes of a transistor cannot be in the same transistor group. (A transistor group exists after partitioning the circuit under certain rules.)
2. During steady-state operation, no Vdd to GND connection may exist.
3. During steady-state operation, each output of a transistor group must be connected to Vdd or GND through a path of conducting transistors.
4. There are no control loops among transistor groups.
5. The bulk must be connected to a supply rail.
6. During testing, each primary input must be controlled by a strong power source whose current is also monitored.

Rule 1 excludes the possibility of self-control inside a transistor group. Rule 3 disallows charge sharing or retention. Rule 4 excludes the possibility of a loop regenerating to a value because it does not have an input source. Thus, no faulty $I_{\text{DDq}}$ will flow. Rule 5 excludes anomalous reverse conduction. Rule 6 excludes faults in the driving circuitry. Remember in built-in current sensing the large current in the drivers is not sensed. Therefore, testing time decreases. However, bridging faults that exist between two primary inputs are not $I_{\text{DDq}}$ testable unless the driver current is monitored. Bridging fault between primary inputs and other nodes are $I_{\text{DDq}}$ testable. Whatever, topology the device under test uses, with careful layout and partitioning, large portions of the circuit are still $I_{\text{DDq}}$ testable.
Chapter 3

Recent Built-In Sensors

3.1 Introduction

A list of guidelines for high-performance built-in current sensors was given in Chapter 1. It must sink large transients currents but, be able to turn smaller $I_{ddq}$ magnitudes into detectable voltage levels. The voltage across the BICS must be small such that the voltage swing of the DUT is large enough not to degrade noise margins below an appropriate level. It must be semi-insensitive to the virtual ground line capacitance otherwise large time-constants increase the test speed above an appropriate level. These and other factors will be considered when examining BICS architectures. The basic idea of a BICS is shown in Figure 3.1. The rest of this chapter presents BICS in the recent literature. It is by no means an exhaustive list and the order in which they are presented is not chronological.

![Figure 3.1 Generic BICS](image)

Figure 3.1 Generic BICS
3.2 Maly & Patyra BICS [24]

Figure 3.2 shows the circuit diagram for the Maly and Patyra BICS. The M1, M2 mirror sets the tail current in the M4-M7 differential amplifier. The input to the diff. amp comes from a Vref and the voltage drop across the Bipolar device. The base voltage of the BJT is set by a circuit breaker. Thus, when the current in the DUT is large and the Virtual GND line voltage rises above a certain level, the BJT turns off. The BJT turns on when the virtual GND line voltage is below a certain level and the current flows through the diode configured BJT. This circuit is based on the same idea as one used previously by Nigh and Maly [25] to test an ALU except that the differential amplifier was replaced with a cross-coupled latch and the circuit breaker consisted of less control logic.

![Diagram](image)

**Figure 3.2** Maly & Patyra BICS

3.3 Shen et al. BICS [26-27]

Figure 3.3 shows the circuit diagram for the Shen et al. BICS. The diode, which is made of the source and a pwell of an NMOS, sinks the large transient current. Thus, the voltage drop during peak currents will not exceed .65 volts. The rest of the circuit is similar to a bit line sense amp. This BICS is clocked with a two-phase non-overlapping clock. The
clock isolates the amplifier from the input node in the transient state and prevents the DUT from being affected by the interior signals of the amplifier. During CLK1, the voltages pass through to the latch. During CLK2, the latch is triggered and the output settles depending on the difference in the input voltages. The clock must be non-overlapping because an error may result if the PMOS controlled by CLK2 conducts before the voltage values on the inputs of the cross coupled inverters are isolated from the DUT and Iref.

The BICS has been reported to have a 2ns detecting time [26]. Figure 3.4 verifies this in SPICE. The detection time is defined as the time CLK1 is high. This is also known as the sampling period (CLK1=1) which only has to be for 2ns. This is the time needed after the input nodes on the cross-coupled inverter has settled. However, the sampling period cannot start until the transient has died out. This is governed by the amount of capacitance on the ground line and the resistance of the NMOS. Therefore, the total operational frequency depends on the virtual ground line discharge time, the sampling time, the window between non-overlapping clocks, and the evaluation period (CLK2=0). The time, however, is reduced due to pipelining. The test is run so that the output value displays the results of the previous test vector. To further increase the operational frequency the on chip clock should be
synchronized with CLK2; Therefore, the transient will have a longer time to die out.

A fabricated chip was designed with this BICS [27]. Static and dynamic tests were performed. In performing the static test an input vector was applied to the pins and how fast CLK1 and CLK2 could be clocked and still get correct results was determined. This is not a true test since the transients have already died out. In the dynamic tests the inputs were
changed every clock cycle. It was shown that the BICS operating speed was only 2Mhz. This is a very different result from the claimed 2ns detecting time or the 33Mhz operational frequency of the static test. The test chip had six implanted defects. Only four could be detected by the BICS. The two defects that passed did not produce a high enough Iddq to be detectable. Also false alarms happen when the Iref was taken to be less the 23uA.

This BICS is highly dependent on the virtual GND line capacitance. The circuit could be partitioned so that an ideal ground line capacitance can be found. However, other strategies should be explored. Also, it was reported that the DUT speed degraded by 15% due to the voltage drop across the BICS [27].

3.4 Hsue and Lin BICS [28]

Figure 3.5 shows the circuit diagram for the Hsue/Lin BICS. Unlike the Shen et al. BICS this one does not require a clock. The diode under the DUT is used as a current sink and a voltage-to-current sensor. The other diode is used to provide a reference voltage so that a difference between the faulty and the reference I_{DDq} is measurable. M4 and M5 will set the gate voltage depending on the desired I_{DDq} level in order for this current to flow through M6 and M13. M8 and M7 form a current source inverter. Thus, depending on the gate voltage of M8, node 17 will be set to a certain value which will be the input to the M9,M10, M14-M16, diff. amp. Node 20 should stay constant depending on the I_{DDq} limit. Node 23 is the output of the diff. amp and the input of the current sink inverter M17-M18. The output, node 24, will then be amplified by two push-pull inverter(not shown). M17 and M18 stabilize the quiescent voltage at node 24 regardless of the supply voltage changes. The drain of M17 will
follow the change in supply voltage. Thus, the $V_{gs}$ of M17 is stable. The output of the BICS produces a pass/fail 1/0 output. The speed of the sensor to produce a non-fault output when the current is less than the $I_{DDq}$ limit, determines the speed at which $I_{DDq}$ testing can be performed. Since this circuit is not clocked the transients are compared. Thus, the output will show a failure until the transient dies out and a change to a pass occurs if the circuit is fault-free. Figure 3.6 show the SPICE results for this BICS. The delay is 125ns. This matches the 100ns reported. The difference is probably due to the different sensitivities of the push-pull inverter. They also have the diff. amp biased around the inverter trip voltage. It is pertinent to point out that the value of the $I_{DDq}$ limit sets the speed of the BICS. If the value is lower it will take longer for the virtual ground line to discharge. Finally a negative pull down voltage of -.5volts is supplied, so that the virtual GND node will be zero volts which in turn will not slow down the DUT. This circuit too is highly dependent on the GND line capacitance.

3.6 Singh and Hurst BICS [29-30]

Figure 3.7 shows the abbreviated circuit diagram for the Singh/Hurst BICS. This
Figure 3.6 SPICE Results for the delay of the Hsue/Lin BICS

Figure 3.7 Singh and Hurst BICS
sensor uses the same basic idea as the Shen et al. BICS, however, it is now in the form of a dual-input sensor. The purpose of which is to cancel out the capacitive discharge currents of each DUT's virtual GND line capacitance. Thus, $I_{\text{DDq}}$ measurement can be taken before the virtual GND line capacitance has fully discharged. Note, however, the transient only has to die down enough for the minimum delta V difference of the comparator to work. This means that the virtual GND line capacitance and the length of the transients currents must closely match. This can be achieved by adding additional GND line capacitance and extra switching gates respectively. This BICS is clocked and it uses the same non-overlapping clock as Shen et al. Testing is carried out as follows. $I_{\text{ref1}}$ is turned on and $I_{\text{ref2}}$ is off. Vectors are then applied to both DUT's to detect faulty $I_{\text{DDq}}$ level in DUT2. The opposite is done for testing of DUT1. Testing speed was measured as 3 times as fast as the Shen et al. BICS with the same DUT at a GND Line cap. of 15p [29].

This dual BICS architecture seems to be a novel solution since transients cancel. If pulse widths between DUT's increase, then the BICS detection capabilities decrease except for small GND line caps. Also, the tolerance for the difference in GND line capacitance goes down as the testing speed increases.

3.7 Brown and McLeod BICS [35]  

Figure 3.8 shows the Brown/McLeod BICS. It consists of a control loop, a reference generator, and a differential amplifier in the cross-coupled latch. The voltage drop in this BICS occurs at the Vdd line so as to not degrade the noise margins near GND. The circuit works as follows. Diode D1 provides current to the DUT. The current necessary will create a voltage drop across the diode. This voltage is feed into a current conveyor made up of M2
and M1 and the cascode mirror. The common base (M9) sets a maximum level and provides a better ac response. As the properties of the current conveyor state, the voltage drop across D1 with be the same as D2 and the current out of D2 will equal the current out of D1 that does not go to the DUT. Thus, the voltage drop across D1 sets up a current which is mirrored to M5 and M12 and then again mirrored to M7. This current is dependent on the I_{DDq} level. This current and the reference current go through to the cross-coupled sense-amp latch for comparison. Note, with the many mirrors and the scaled diodes, precise matching is necessary as to not mask the faulty I_{DDq}. No mention of this was discussed.

Simulation results were performed with a load circuit of approximately 4500 minimum width transistors. It was shown that the time from the I_{DDq} of the DUT to fall below 40u to the time the current in the control loop fell below 40u was 28ns. This added to a detection time of 36ns, bringing the total time to 80ns. Thus, 80ns after the DUT is being tracked by the control loop a decision can be made about the I_{DDq} level. Although a performance

![Diagram of Brown and McLeod BICS](image-url)

Figure 3.8 Brown and McLeod BICS

22
comparison in [31] shows that it performs better than Shen et al., a differential architecture such as Singh/Hurst BICS would improve the overall delay in evaluating the faulty $I_{DDq}$.

### 3.8 Integrating BICS

The final three BICS reviewed integrate the charge from the $I_{DDq}$ in some way. Miura and Kinoshita [32] integrate the dynamic and static currents and compare it to a reference. This is achieved by turning the monitored current into a voltage and depending on the time it is above a certain level an integral part of the circuitry is working. This circuit requires precise transient behavior for each applied vector.

Rubio et al. BICS [33] integrates only the $I_{DDq}$. The current is fed into a capacitor. If the voltage that accumulates is above the threshold voltage of a transistor, a node gets discharged else it stays the same. Segura et al. BICS [34] claims to be able to detect for opens that are not detectable by $I_{DDq}$ testing. This is achieved by integrating the total charge from both the static and dynamic current. However, as the transient current increases the reference capacitor gets large and small differences in $I_{DDq}$ levels will get masked.
Chapter 4

Proposed BICS Circuit

4.1 Introduction

When designing the proposed BICS, the advantages and performance limiting effects of the recent BICS's in Chapter 3 were taken into consideration. It was decided that a differential architecture[28-29] has the potential for a faster $I_{DDq}$ testing speed due to capacitive discharge current cancellation. Note, a differential BICS (Figure 3.7) can produce a faster result then a single sided BICS when there is no fault. A single sided BICS, where the $I_{DDq}$ is compared to a reference, is faster if there is a fault because the $I_{DDq}$ is already above the reference. However, when there is no fault it must discharge the virtual node capacitance until the $I_{DDq}$ is below the reference. This is another instance where the $I_{DDq}$ reference sets the speed of the BICS. When there is a fault in the differential BICS, since BICS both sides have an $I_{DDq}$, one faulty and one not, and the fault free side has a reference current added to it, if the difference between the fault-free and faulty $I_{DDq}$ level is close to the reference current level, the voltages on the virtual GND capacitances will not diverge away from each other quickly. This may allow the single sided BICS to perform better in some situations and is dependent on the faulty $I_{DDq}$ current and the magnitude of the virtual node capacitance. The thesis proposes a novel solution to this problem by current profiling which is discussed in more detail in Section 4.2 - 4.4.
4.2 Vdd line Monitored Differential BICS

The BICS proposed in this thesis is shown in Figure 4.1. It adopts the differential architecture proposed in [29-30] and the Vdd line monitoring in [31]. It works as follows: diodes D1 and D2 provide the current to DUT1 and DUT2, respectively. (As stated in Chapter 3, DUT1 and DUT2 are divided pieces of the total DUT that ideally match in with respect to the virtual Vdd line capacitance.) Whatever, the magnitude of the transient and static currents are, a proportional voltage will be setup at node 2 and 12. A diode is used so as to limit the voltage to node 2 and 12 to about 4.3 volts, for a 5 volt supply, regardless of the magnitude of the transient current. After the transient has died, the node voltages on 2 and 12 will assume a voltage that is proportional to the I_ddq level of the DUT. These voltages are fed into the sources of M2 and M11 which form a common-gate differential pair. A proportional current flows out the drain of M2 and M1 and is fed into a sense amp with a cross-coupled differential mirror [35] used as source degeneration. Transistor M12 is not

![Figure 4.1 Proposed BICS](image-url)
necessary and has been removed. An amplified difference of node 2 and node 12 is now set up across node 7 and node 8. The initial difference is given below:

\[
\frac{V_8 - V_7}{V_2 - V_{12}} = \frac{g_{m1}}{gm_p - 2g_{o11}}
\]

Where \(g_{m2}\) is the small signal transconductance of M2 and \(g_{o11}\) is the drain-source conductance of M11 biased in the triode region and \(gm_p\) is the tranconductance of M8. M11 is then turned off and the cross coupled latch uses positive feedback to amplify this difference again and provide a power rail voltage output. M11, M3-M10, M12 are similar to a clamped bit-line sensed amplifier developed in [36] except M3-M6 in the BICS circuit were replaced by two transistors biased in the triode region.

As a brief example, if one wanted to test DUT1, the reference current flowing out of node 12 will be turned on and its magnitude will be some \(I_{DDq}\) limit. The reference out of node 2 will be off. Tests vectors would be applied to both DUT1 and DUT2. The transients currents pulled by DUT1 and DUT2 will probably clamp the node voltage at 2 and 12 at about 4.3 volts. After the transients die the node voltage at 12 will diverge to a value proportional to the \(I_{DDq}\) reference level. Node 2 will diverge to 5V if the \(I_{DDq}\) is ideally zero. If the \(I_{DDq}\) of DUT1 is above the \(I_{DDq}\) value, it will diverge to a voltage proportional to the faulty current. As one can see, the speed at which these voltage diverge away from each other is dependent on the reference current for the fault-free case or dependent on the difference between the faulty \(I_{DDq}\) minus the reference current for the faulty case. Of course the total speed that they both diverge from 4.3 volts is dependent on the virtual Vdd line.
If there is a fault, node 2 will be less than node 12. Thus, more current will flow in M1 than in M2. One observation is that the body effect helps in this situation. Since the source voltage of M2 is more negative, its threshold voltage is less than that of M1. Thus, M2's current is less than that of M1's not only by the difference in source voltage but also the difference due to the threshold degradation. More current is fed into node 6 than node 5. This ac current will flow into the lower impedances of M1 and M2 and will set up a voltage difference across node 8 and 7.

4.3 Small Signal Analysis

The differential pair M1 and M2, as stated previously, sets up the initial condition between node 8 and 7. The gain from node 2 to node 5 is just $g_{m2}$ divided by $g_{m10}$. The gain from node 5 to node 8 is just $g_{m10}$ times $R_{11 \text{trans}}$. Multiplying these two gains the initial condition is found (previously given in Chapter 4.2). The rest of this chapter focuses on the cross-coupled inverter.

Transistors M3-M11 make up the cross coupled inverter. Transistors M3-M6 can just be modeled as a resistor ($R_{\text{out}}$) when a differential signal is impressed across node 5 and 6. The circuit is now just a regular cross-coupled inverter with high resistances between the nFET sources and GND. The cross coupled inverter has a reset time associated with turning on M11, and a regeneration time associated with turning off M11. Since resetting occurs during the transient and the onset of capacitive discharge, sufficient time for resetting is allowed. Thus, a minimum size transistor can be used in for M11. The will not be the case for fast resetting times.
Figure 4.2 shows the small signal model used to calculate the loop gain. The circuit is in the regeneration mode, M11 is off, which disconnects node 7 and 8. To find the 'reset' small signal model, M11 should be modeled as a resistor between node 7 and 8.

Looking at node 8, there is a pMOS and an nMOS with source degeneration to GND. Thus, the equivalent transconductance is just $g_{mp}$ (small signal transconductance of M6) and the equivalent resistance is just $R_{op}$ (small signal drain source resistance of M7). Finding the current equations at node 7 and 8, setting them equal, and solving for $V_8/V_7$ gives the loop gain found below:

$$GBP = \frac{g_{mp}}{C_s}$$

Where $C_s = C_8 = C_7$. The regeneration time at which the latch comes out of its metastable state is inversely proportional to the gain-bandwidth product (GBP) [37]. To increase the GBP a minimum length pFET should be used. Note, the bandwidth of a regular cross-coupled inverter is larger than this one with source degeneration in the nFET. However, transistors
M3-M6 are used to remove the latched voltage outputs at node 8 and 7 from influencing the voltages at nodes 2 and 12 because of lambda effects.

If a time domain analysis is performed on Figure 4.2, the time for node 7 and 8 to reach a logic level given an initial difference is:

\[ T_{\text{reg}} = \frac{C}{g_{np}} \ln \frac{V_{\text{Logic Level}}}{V_{IC}} \]

The initial condition was derived in 4.2. If the initial condition is small the regeneration time increases. This may be unacceptable for certain testing frequencies. This non-regenerative nature forms the basis of current profiling.

4.4 Comparison Performance

This section compares the proposed BICS against some of the recent BICS presented in Chapter 3. It attempts to recreate the testing environment that was decried by other designs. The first comparison will be made against the Singh/Hurst BICS.[29-30] In Hurst thesis [29], he compared his BICS to Shen et al.[26-27] by measuring the time to detect a no-fault. A transient of 5mA that lasted for 16ns with 1ns rise and full times and \( I_{DQ} \)

<table>
<thead>
<tr>
<th></th>
<th>2.5pF</th>
<th>10pF</th>
<th>25pF</th>
<th>50pF</th>
<th>100pF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed</td>
<td>45MHz</td>
<td>43MHz</td>
<td>40MHz</td>
<td>37MHz</td>
<td>35MHz</td>
</tr>
<tr>
<td>Hurst [29]</td>
<td>40MHz</td>
<td>40MHz</td>
<td>37MHz</td>
<td>37MHz</td>
<td>30MHz</td>
</tr>
<tr>
<td>Shen [29]</td>
<td>25MHz</td>
<td>17MHz</td>
<td>10MHz</td>
<td>5MHz</td>
<td>2MHz</td>
</tr>
</tbody>
</table>

Table 1 Hurst Sensor Comparison
reference of 80ua will be assumed. The point where a constant latch sensitivity was achieved was measured and 4ns was added to allow for sampling and regeneration. Table I shows the results of varying the Vdd cap. Shen et al. is included even though truly, the comparison should be done at half the value of Shen et al. cap since the other BICS are differential.

The Hurst BICS is very close in speed to the proposed BICS. His speed is proportional to the difference between RC time constant composed of a GND line cap. and a transistor in the linear region. The proposed BICS speed is proportional to the difference between the RC time constant composed of a Vdd line cap. and a linear transistor.

Note, that the frequency reported in Table I is for a given transient. Thus, the DUT sets the frequency of BICS operation. More performance results for different transients for the proposed BICS is given in 4.5. Also, the comparison in Table I is for the fault free case. As discussed before the faulty case is were the differential BICS can have increased delay.

The proposed BICS is now compared to the Brown/McLeod [31] BICS under the conditions in [31]. A Vdd line cap of 94pF and a reference current of 40uA were used. The transient waveform used is shown in Figure 4.3. On the last transient a faulty $I_{DDQ}$ of 50u was applied. The Brown/McLeod BICS needed 80ns before a decision could be made. The proposed BICS needed only 56ns before a measurement can occur (MCO) as shown in Figure 4.7. This was only possible for the faulty case when the reference current was timed so that it was applied after the transients died out otherwise the other BICS showed a better detection time.

If a reference current is pulled out of the diode on one side and no current is pulled out of the diode on the other side, the diode voltages diverge from each other faster. If a
reference current is pulled out of the diode on one side and a faulty \( I_{DDq} \) is pulled out of the diode on the other side, that have around the same magnitude, the voltages across the diode will diverge away from each other slowly. This slow divergence time is more pronounced for large \( V_{dd} \) line capacitance. Thus, to combat this problem timing the reference current could be used, where the reference current is not applied until after the transients die out. This method fails if the faulty \( I_{DDq} \) is just below the reference current.

![Graph showing voltage and current over time](image)

Figure 4.3 Proposed BICS SPICE Results under a 94pF cap and 100mA transient
When the latch does not regenerate, testing does not have to be slowed down until regeneration occurs. The fact that it does not regenerate in a certain time interval means that the reference is close in magnitude to the $I_{DDq}$ level. Current profiling can be used to detect these times.

4.5 Current Profiling

In current profiling, the M11 is switched on and off constantly throughout the transient and after it dies out. Using the equation for regeneration time in Chapter 4.3, a time limit can be decided based upon the $I_{DDq}$ difference set. For example, if a detectable difference between the DUT's of 10μA is desired, then regenerate, if under do not. Now find out, based upon 10μA, what the initial condition will be and the time it takes for regeneration. This will be the clock period. If no regeneration occurs then the $I_{DDq}$ difference is below 10μA. In this situation there is no need to use a reference current. Since both DUT's current should be the same, the latch should never regenerate within the time interval. A reference current could be used to determine how much of a difference does exist if there is a fault. Figure 4.4 shows the $I_{DDq}$ current of both DUT's and the latch output. This output is a current signature for the applied input vector. The signature would be different if a reference current was used. In current profiling the reset time needs to be as fast as the regeneration time so that more sampling can occur. Thus, a minimum geometry M11 transistor may not be the best solution.

Current profiling can also detect mismatch between the Vdd. cap and may detect transient mismatch. For the detectable current values given certain comparator characteristics see Appendix A. If there is a mismatch between the Vdd line caps or transients of the two devices to a certain percentage, the latch will regenerate when not expected to. Test time will
still have to increase because a faulty $I_{DDq}$ may get masked due to different capacitive discharge currents. Therefore it is still necessary to keep these fairly close.

Current Profiling is a novel solution to increase test time and to wisely equate the situation of non-regeneration into observable information. From current profiling a library of current signatures can be constructed based upon simulation or a known good die. Then
deviations from these signatures could mean failure, Vdd cap mismatch, or even a delay fault.

4.6 Mismatch Effects

Any mismatch in the transconductance, conductance, parasitic capacitance, threshold voltage of the transistor's or the load capacitance causes undesirable effects on the metastability [38]. The equation for regeneration time is now dependent on mismatch as well as the initial condition. Careless layout can result in mismatch due to fabrication gradients in the different layers. In the proposed BICS matching transistor pair include M1-M2, M8-M9, Mr-Mr, and M3-M6. Layout techniques for better matching are summarized below [39-40]:

1. Transistors should be laid out in Fingers
2. Matched pairs can be common centroided
3. Further matching occurs with interdigitizing
4. Dummy transistors added so both sides of match transistors see the same thing to prevent etching errors.
5. Match transistor should have same orientation.

Since close to minimum geometry is used for M9-Mr it is hard to layout in fingers or common-centriod, however the transistor could be made larger to facilitate this. M3-M6 could be interdigitized because, above a certain degree of mismatch a negative impedance can occur [35] (See Appendix B). M1-M2 should be common centroided. All transistors should have the same orientation and etching effect should be minimized.

As for the mismatch in the Vdd cap, once the DUT is split in two a good extractor should be used to determine the two Vdd caps. The difference should be added to the small one under the same conditions as the greater Vdd cap. This means that perimeter and area ratio should match [41] and the effects of etching should be minimized by allowing both Vdd lines to see the same thing.
Chapter 5

Current and Delays in CMOS Circuits

5.1 Transient Current and Delays

Since the proposed BICS presented in Chapter four is clocked, delays must be evaluated. Even when current profiling is employed, the delay until the transient dies out must be known. Then, an additional delay needs to be added before the application of the next test vector. Maximum currents need to be known for the design of the sinking diodes of the BICS. Since the DUT has to be split in two parts with matching transients, determination of currents dictate the best optimal split.

Many techniques exist for current and delay estimation [42-45]. In [42] a survey of power estimation techniques is presented. Some techniques require knowledge of input vectors and some do not. One based on probability waveforms is Crest. Crest assigns a probability to every node based on the node being high or low. This is very similar to assigning values of controllability and observability to each node. These values are then used to estimate the current pulse and delay.

In [44-45] each individual gate is turned into an equivalent inverter. Rules apply for collapsing a gate into an inverter for different series and parallel combinations of transistors. All parasitic capacitances are lumped in the model and the charging positive maximum, discharging positive maximum, and discharging negative maximum are found. This seems to
be the best method. A speed improvement for delay and maximum current approach 4 orders of magnitude over HSPICE and 3 orders of magnitude for current waveform estimation. The accuracy is within 10 and 12% [44]. After these equivalent inverters are derived from gates, delay estimation is achieved by governing rules. These are based on input transition times, input delay times and effective width calculation. Whatever method is used, it must be accurate and computational fast.

5.2 $I_{DDq}$ Estimation

The $I_{DDq}$ is composed of junction and leakage currents. A technique to find its magnitude is explained [46]. The amount of reverse biased junction leakage currents is proportional to the area of all reversed p-n junctions (well-sub, drain/source - well) in the circuit. The drain/source component to $I_{DDq}$ can be estimated below:

1. Compute the relationship between the drain leakage current and drain area
2. Determine actual value using test structure.
3. Compute the total current using the formula below:

$$I_L = I_s \sum_{i=1}^{M} [W_i K_i + \pi X_j (W_i K_i) + \frac{1}{6} \pi X_j^2]$$

$I_s$ is the area per micron saturation current; $M$ is the number of drain regions; $W_i$ is the width; $K_i$ is the length of the drain; and $X_j$ is the junction depth. The same procedure can be used for determining the well junction leakage currents. The subthreshold leakage current is small. However, if charge is present in the gate oxide the subthreshold current may become an essential part of the total $I_{DDq}$. The subthreshold leakage current is proportional to the number of channel squares.
After estimating the $I_{dBq}$ predicted level, partitioning can be employed as to not exceed a certain magnitude. In an ideal differential BICS these currents should cancel. However, if a significant leakage current mismatch occurs an extra reference current on the side with the lower leakage current may be added.
Chapter 7

QTAG Standard

7.1 Introduction

The QTAG (Quality Test Action Group) was formed at the 1993 International Test Conference. It was created to define a standard for off-chip \( I_{DDQ} \) monitoring on test fixtures in a production environment. New developments in the proposal for the standard can be found in [56-63]. Recent off-chip current sensors [64-65] can be integrated with the standard. The goals set out by QTAG are given below [62]:

1. A definition of a "minimal pin" configuration for a monitor whose specific design may differ for any manufacturer. That such a monitor allows any digital test system to drive and receive information appropriate for \( I_{DDQ}/I_{SSQ} \) testing.

2. A defined pin configuration that requires packaging in a minimal number of package types to cover most application needs.

3. A standard ATE interface definition and physical size that allows all common test fixtures to be used: DUT board, probe-card and contactors.

4. A standard that defines both \( V_{PD} \) and \( V_{SS} \) line monitor configurations.

5. A standard that would recommend or define multiple monitors and power pin configurations.

6. Definition of a Monitor Description Format (MDF) that would allow different monitor
implementations to easily be driven by ATE hardware and software.

7. A standard that allows tester selection of a variable $I_{DDQ}/I_{SSQ}$ threshold.

8. A standard that allows robust physical construction and reliable operation in a stressful ESD environment.

Figure 7.1 shows a small QTAG monitor on a test fixture (probe card).

![Diagram of QTAG Monitor on a Probe Card](image)

**Figure 7.1** QTAG Monitor on a Probe Card [56]

Four different types of monitors are proposed: 1) A threshold monitor with a low pin count and a pass fail output, 2) A measurement monitor which beside a pass fail output can provide
a digital value for an analog $I_{DDQ}$, 3) An analog monitor which beside a pass fail output returns an analog voltage derived from the $I_{DDQ}$, 4) A modeless monitor which is not part of the QTAG standard. It uses a analog pin to set the $I_{DDQ}$ threshold instead of the serial interface. There is both little foot and Big foot monitors. The little foot is an 8 pin monitor with pin descriptions below: [63]

1) VDD_MON Pin - This is the Vdd supply for the QTAG $I_{DDQ}$ Monitor.
2) VSS Pin - This is the Vss supply for the QTAG $I_{DDQ}$ Monitor.
3) VDD_PSU Pin - This is the VDD from the tester which needs to be applied to the DUT.
4) VDD_DUT - Pin - This is the VDD output of the QTAG Monitor to the DUT.
5) MODE Pin - The mode pin is used to switch between the control and monitor modes of the QTAG monitor. Control Mode is used to configure the QTAG monitor and monitor mode is used to measure the quiescent current values.
6) CLK Pin - In control mode, the rising edge of the CLK signal is used to clock serial data in and out of the QTAG Monitor. In monitor mode, the rising edge of the CLK signal is used to indicate a measurement should be made.
7) DI Pin - In Control Mode, this pin is used to load serial data. In Monitor Mode, this pin is used to read the quiescent current measurement result which is present at the DO output pin of the previous monitor in the chain.
8) DO Pin - In Control Mode, this pin is used to read out serial data. In Monitor Mode, this pin is used to feed the result of the quiescent current measurement back to the tester.

The monitors must be designed to cope with transient currents in the ampere range. These transients must pass through the monitor without introducing large voltage drops.

Besides the test monitor location, type, and pin specifications the standard deals with commands, Monitor Description format, and other issues. For an in-depth discussion of these topics consult [55-65]. When completed the standard will resemble the Boundary Scan
standard[66] in the sense that its pin count, mode of operation, and interfacing will be
described and deviating from that will not be supported by the standard. Another similarity
is that Boundary Scan has it own VHDL based language [67] and the I_{DDQ} standard is
developing one [57].

Current Profiling and the BICS developed in this thesis can be used with the standard.
With modification the BICS can become a little foot sensor. Since it is differential, two chips
will have to be tested in parallel or on chip will have to be tested against a simulated reference
current. Using Current Profiling with the standard can turn a threshold monitor into a
measurement monitor. This is achieved by varying the reference until regeneration will not
occur.
Chapter 8

Future Direction and Conclusions

8.1 Future Testing Methodologies

Since testing is fabrication and layout driven, it should proceed as follows; With any new process, IFA should be employed. From this, defect models should be constructed. Test set generation should then be geared to detect these defects and real metrics should be used when discussing the test set coverage. I_{DDQ} is almost necessary since traditional voltage methods can not detect a majority of CMOS defects. Even though at present date, I_{DDQ} is only being investigated in the testing environment at Motorola[68], it needs to be employed.

When the QTAG standard is finalized on-chip current sensors will probably be phased out. This reason being, the area and pin count overhead are not economic in a consumer electronics market [64]. However, the Built-in current sensor proposed in this thesis, with modification, can be made into a QTAG I_{DDQ} monitor. Current profiling can easily be integrated with the standard, and has merit as a mixed-mode defect detection solution. In mixed-mode IC's, in which both analog and digital circuits are found on the same chip, the only way to I_{DDQ} test the chip is portioning. If current profiling is allowed from transient to steady state the power supply signature will allow a judgement to be made about the chip's defects. This idea of power supply signature analysis has been researched previously [69-73]. In this method the power supply is ramped and the current drawn from the supply is
monitored and a signature is created. The benefits of this is that analog, digital, and mixed mode circuits can be tested, no area or pin count overhead is necessary, and only one test vector is applied. The problem is that in large circuits small current deviation caused by some defects may get masked. However, this method should be further researched and integrated with the QTAG Standard.

8.2 Conclusions

This thesis has developed a test environment for CMOS integrated circuits. The environment includes fabrication extraction of defects, defect modeling, $I_{\text{DDq}}$ testing, and test pattern generation. A built-in current sensor was proposed which exhibits simulated test speeds over previous BICS. Current profiling was presented and was shown to increase test speeds. Since the BICS is timed and dependent on the DUT split circuits, a method to estimate delay and currents was explained.

Future research on topics in this thesis should include: 1) The fabricated BICS, 2) The modifications and problems of Current profiling on a fabricated chip, 3) Test Pattern generation with the defect models presented, 4) Integrating Current Profiling or Power Supply Ramping with the QTAG Standard.
REFERENCES


pp. 72-77.


39. Laker, K. & Sansen, W. *Design and Analysis of Analog Integrated Circuits*


52. Fritzemeier, R. Hawkins, C. Soden, J. "CMOS IC Fault Models, Physical Defect


55. Bruls, E., Arnout, G., Richardson, A. "A Practical defect Coverage Metric for Digital CMOS." QTAG documentation in PS format available via baker@prl.philips.nl.


57. Baker, K., Waayers, T., Verstraelen, M., Hales, A. "Plug & Play $I_{DDQ}$ Monitoring with QTAG." QTAG documentation in PS format available via baker@prl.philips.nl.


60. Baker, K. Bratt, A. Welbers, A. "Provisional Specification for IDUNA-2" QTAG documentation in PS format available via baker@prl.philips.nl.

61. Baker, K. "Requirements Specification for the QTAG Monitor Description Format (MDF)" QTAG documentation in PS format available via baker@prl.philips.nl.


63. Hales, A. "Texas Instruments Draft Proposal for the QTAG Off-Chip IDDQ/ISSQ Monitor Standard." QTAG documentation in PS format available via baker@prl.philips.nl.


68. Personal Communication. Ed Flanagan. Probe Manager Motorola, Autsin, TX.


Appendix A  Cross Coupled Mirror Analysis

A.1 Cross Coupled Mirrors Calculations

Figure A.1 shows the cross-coupled mirrors used as an active load in the proposed BICS. The current equations are given below:

\[ i = -g_{m1}V + Vg_o + Vg_{m2} \]

\[ -i = g_{m3} - Vg_o - Vg_{m4} \]

Solving for the transconductance yields:

\[ \frac{i}{v} = g_{eqv} = g_o + \frac{g_{m2} - g_{m1}}{2} + \frac{g_{m4} - g_{m3}}{2} \]

A negative resistance will result if the above equation becomes negative. This can happen in the presence of mismatch. A method to reduce this effect is shown in Figure A.2.
By adding a current source in parallel, the effective transconductance is reduced. The differences in the transconductance equation contribute less of a negating effect. An observation, found below, gives the allowable mismatch.

\[
\frac{1}{\text{%of mismatch}} = \frac{g_m}{2g_o}
\]

Figure A.2  Method to Reduce Mismatch Effects
VITA

Richard Burgess Jr.

Candidate for the Degree of

Master of Science

Thesis: DEVELOPMENT OF A CMOS I_{DDq} TESTING ENVIRONMENT

Major Field: Electrical Engineering

Biographical:

Personal Data: Born in Worcester, Massachusetts, October 10, 1970, the son of Richard and Lenore Burgess.

Education: Graduated from Holy Name Central Catholic High School, Worcester, Massachusetts in May, 1988; received the Bachelor of Science degree in Electrical Engineering from Worcester Polytechnic Institute, Worcester, Massachusetts in December 1992. Completed the requirements for the Master of Science degree with a major in Electrical Engineering at Oklahoma State University in May, 1995. Plan to pursue a Ph.D degree at the University of Virginia starting in Fall 95 under an awarded fellowship.

Experience: Employed by Oklahoma State University, Department of Electrical Engineering as a research and teaching assistant; Co-op at Sikorsky Aircraft Trumbull, CT. as a simulation engineer. Summer positions at Allegro Microsystems Worcester, MA and Norton Company Worcester, MA as a test engineer. Temp. employment at Applix Co. Westboro, MA and Simplex Time Recorder Gardner, MA in the Unix and DOS quality assurance dept. Currently Serving as Associate Editor of IEEE Potentials.