ISLAMIC UNIVERSITY OF TECHNOLOGY (IUT)

Performance Comparison of Positive-Feedback Match-Line Sensing Schemes in High Speed Ternary Content Addressable Memory (TCAM)

A Thesis Presented to

The Academic Faculty

by

Md. Atik Foysal [102406]

Syed Muntasir Morshed [102425]

Md. Zahidul Anam [102426]

A Dissertation

Submitted in Partial Fulfillment of the Requirement for the

Bachelor of Science in Electrical and Electronic Engineering

Academic Year: 2013-2014

Department of Electrical and Electronic Engineering.

Islamic University of Technology (IUT)

A Subsidiary Organ of OIC

Dhaka, Bangladesh.

A Dissertation on,

Performance Comparison of Positive-Feedback Match-Line Sensing Schemes in High Speed Ternary Content Addressable Memory (TCAM)

Approved By

———————————————

Prof. Dr. Md. Shahid Ullah

Head of the Department

Department of Electrical and Electronic Engineering

Islamic University of Technology (IUT)

Gazipur-1704, Bangladesh.

Supervised by

———————————————

Dr. Syed Iftekhar Ali

Associate Professor

Department of Electrical and Electronic Engineering

Islamic University of Technology (IUT)

Gazipur-1704, Bangladesh.

Declaration of Authorship

This is to certify that the work presented in this thesis is the outcome of the analysis and investigation carried out by Md. Atik Foysal, Syed Muntasir Morshed and Md. Zahidul Anam under the supervision of Dr. Syed Iftekher Ali in the Department of Electrical and Electronic Engineering (EEE), IUT, Gazipur, Bangladesh. It is also declared that neither of the thesis nor any part of this thesis has been submitted anywhere else for any degree or diploma. Information derived from the published and unpublished work of others has been acknowledged in the text and a list of references is given.

Authors:

‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾

Md. Atik Foysal

Student ID: 102406

‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾

Syed Muntasir Morshed

Student ID: 102425

‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾

Md. Mahedi Anam

Student ID: 102426

TABLE OF CONTENTS

Topics

CHAPTER ONE

1   **INTRODUCTION**

CHAPTER TWO

2   **TERNARY CONTENT ADDRESSABLE MEMORY (TCAM)**

CHAPTER THREE

3   **MATCHLINE SENSING SCHEMES**

CHAPTER FOUR

4   **SIMULATION RESULTS AND COMPARISON**

CHAPTER FIVE

5   **CONCLUSION AND FUTURE WORK**

ACKNOWLEDGEMENTS

First and foremost we offer our sincerest gratitude to our supervisor, Dr. Syed Iftekhar Ali who has supported us throughout our thesis with his patience, motivation, enthusiasm and immense knowledge whilst allowing us the room to work in our own way.

Besides our advisor, we would also like to thank the department of Electrical and Electronic Engineering of the University for the Support throughout the thesis.

Last but not the least very special thanks goes to our senior brothers, friends and family, without their motivation and encouragement it would not be possible.

ABSTRACT

We survey different schemes in the design of TCAM. A TCAM is a memory that implements the lookup-table function in a single clock cycle using dedicated comparison circuitry. TCAMs are especially popular in network routers for packet forwarding and packet classification, but they are also beneficial in a variety of other applications that require high-speed table lookup. The main TCAM-design challenge is to reduce power consumption associated with the large amount of parallel active circuitry, without sacrificing speed or memory density. In this paper, we review TCAM-design techniques at the circuit level and at the architectural level. At the circuit level, we review low-power matchline sensing techniques and searchline driving approaches. At the architectural level we review four methods for reducing power consumption. In our thesis 16×16 bit TCAM is designed in 0.18µm CMOS. The proposed ML sense scheme reduces power consumption by minimizing search time and limiting voltage swing of MLs. In our simulation we used 1.8V supply voltage.

Keywords: Ternary Content Addressable Memory (TCAM). Matchline (ML), Complementary Metal-Oxide-Semiconductor (CMOS).

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

List	Meaning
ML	Match Line
DML	Dummy Match Line
MLSO	Match Line Sensing Output
DMLSO	Dummy Match Line Sensing Output
MLSA	Match Line Sensing Amplifier
MLRST	Match Line Reset
SL	Search Line
BL	Bit Line
MLEN	Match Line Enable
CR	Current Race
MD	Mismatch Dependent
SLPRE	Search Line Precharge
tx	Transistor

CHAPTER ONE

INTRODUCTION

Ternary Content Addressable Memory (TCAM) is a type of associative memory that offers ternary storage and supports partial data-matching. Each ternary bit can be a "0", a "1", or a "don't care" state. It is a key technology to enable the true power of the next-generation networking equipment and many lookup-intensive applications. TCAMs are hardware-based parallel lookup tables with bit-level masking capability, so gaining importance in high-speed intensive applications. For example, in the application where TCAM is doing the role of a co-processor in network processor. In this application, Network processing unit provide great value to metro network equipment by providing flexible and adaptable processing solution. However, memory intensive packet searches that serve purposes as: routing, access control and Quality of Service (QoS) can quickly– as consume power and bus bandwidth. In classification of network routers application, TCAM is often used in network routers where each address has two parts: Network address and Host address. Network address varies in size on subnet configuration whereas Host address occupies remaining bits. Routing is done by consulting routing table maintained by router which contains-Known destination network address and Information needed to route packets to destination. Using TCAM in the routing table, it makes lookup process very efficient where we store address using "X" for Host part of address. Therefore looking up destination address in TCAM immediately receives correct routing entry. Thus, modern systems are realizing that additional functionality must be achieved in innovative way – *co- processor,* that accelerate deep packet classification and forwarding in the next-generation networking equipment [1].

One of the interesting features of TCAM is: *parallel search,* due to which it has wide range of applications, and also got some advantages such as - Became basic building block of complex searching schemes, Searches every location in memory at once, which means, it doesn't give gap between searching of one cell and another cell in a whole TCAM cell array , Ordering of elements has got less importance because it doesn't include the ability of ordering parts of TCAM such as precharge unit, priority encoder unit etc. and Large indexing structures are avoided [1].

However, the high cost and power consumption are limiting their popularity and versatility. The increasing line-rates and growing deployment of IPv6 are demanding fast and wide TCAM S. Hence, the various techniques have been implemented to improve the performance of TCAM S by reducing the power considerably in the 180nm CMOS technology.

1.1 Motivation

- TCAM is the Key device to speed up packet forwarding in network routers.

- It also provides masking capability: 0, 1, and "don't care ".

- It also has longest prefix capability. It performs table lookup for searching longest prefix matches (i.e. IP routing tables).

- As maximum power is consumed in ML and SL, power consumption is an important issue.

- Area efficient sense amplifier designing is also a top priority.

1.2 Thesis Organization

In this thesis book the usefulness of TCAM circuits for high-speed search operation has been discussed. Variations, construction and operation of TCAM circuits will be presented in this book.

Discussion starts with the basic construction of TCAM from SRAM in chapter Two. Chapter Two also includes short discussion of SRAM operation, classification of TCAM cell, formation of TCAM word, TCAM array, and matchline structure with classification.

Our goal is to limit power consumption maintaining high search speed. Maximum power consumption occurs in matchline and searchline. In Chapter Three we focused on matchline power consumption. Based on this issue different matchline sensing schemes are discussed in this chapter.

In Chapter Four, our simulated results (using HSpice) and comparisons of different matchline sensing schemes are shown.

In Chapter Five, overall idea of this thesis is discussed in conclusion. Again future work is also discussed in this chapter.

CHAPTER TWO

TERNARY CONTENT ADDRESSABLE MEMORY (TCAM)

TCAM is Ternary Content Addressable Memory. It is a specialized type of data searching memory which can search its entire content in a single clock cycle. The significance of the term 'ternary' is that it can query and store for:

- High
- Low and
- Wildcard (Ternary) states.

It is the key device to speed up packet forwarding in internet routing. We know that IP (Internet Protocol) forwarding is the main objective of internet routers. Speed is the main performance parameter here. TCAM is probably the fastest device in this case.

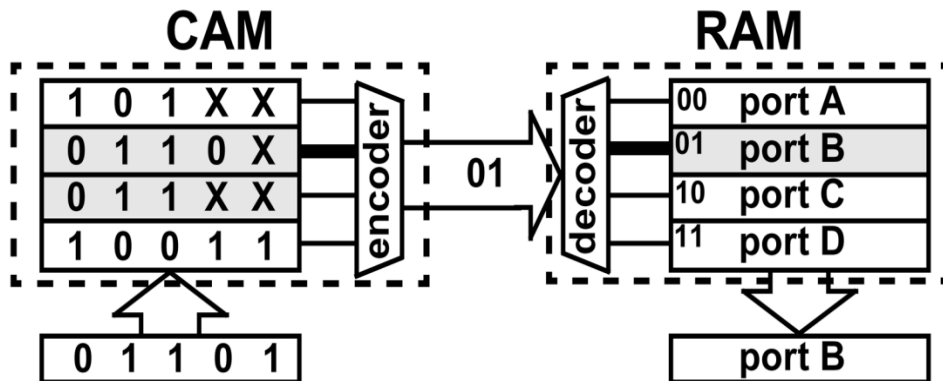| Entry no.	Address (Binary)	Output Port
1	101XX	A
2	0110X	B
3	011XX	C
4	10011	D

TABLE: 2.1 Example routing table

Fig. 2.1: CAM-based implementation of the routing table 2.1

Fig. 2.1 illustrates how a CAM accomplishes address lookup by implementing the routing table shown in Table I. On the left of Fig. 2.1, the packet destination-address of 01101 is the input to the CAM. As in the table, two locations match, with the (priority) encoder choosing the upper entry and generating the match location 01, which corresponds to the most-direct route. This match location is the input address to a RAM that contains a list of output ports, as depicted in Fig. 2.1. A RAM read operation outputs the port designation, port B, to which the incoming packet is forwarded. We can view the match location output of the CAM as a pointer that retrieves the associated word from the RAM. In the particular case of packet forwarding the associated word is the designation of the output port. This CAM/RAM system is a complete implementation of an address-lookup engine for packet forwarding. Farther examples of approaches that use CAMs for the purpose of sorting and searching are provided in [11], [12].

An additional advantage is that it has longest prefix capability which shows the most direct route to the output data.

As power consumption is higher reduction in consumption is an important issue here. Area efficient sense amplifier design is another issue.

2.1 TCAM BASICS

TCAM consists of cells each containing 2 SRAM cells. SRAM elaborates Static Random Access Memory. SRAM is a type of semiconductor memory that uses bi-stable latching circuitry to store each bit. The term "static" differentiates it from dynamic RAM (DRAM) which must be periodically refreshed. SRAM exhibits data remanence. It holds data as long as power is applied.

SRAM has 3 operations:

              (1) Hold

              (2) Write and
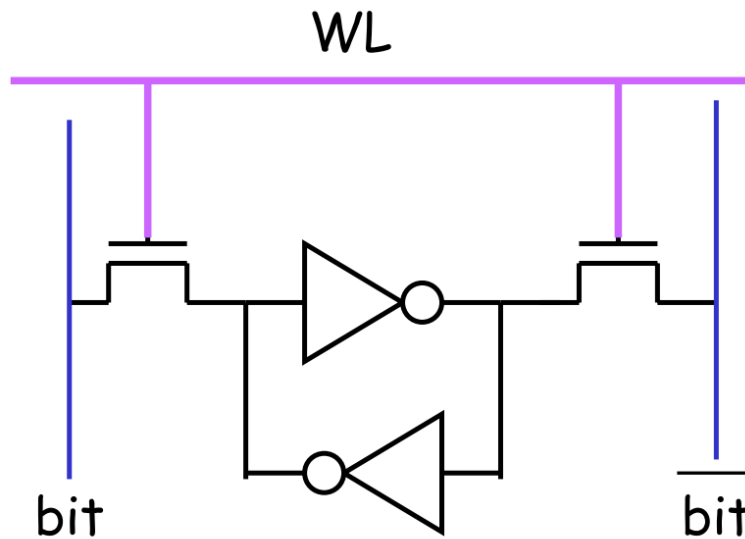
              (3) Read

Fig. 2.2: Basic 6T SRAM cell

- *Hold*

    – word line = 0, access transistors are OFF

    – data held in latch

- *Write*

    – word line = 1, access tx are ON

    – new data (voltage) applied to bit and bit_bar

    – data in latch overwritten with new value

- *Read*

    – word line = 1, access tx are ON

    – bit and bit_bar read by a sense amplifier

2.2   TCAM Cells

TCAM cell can be of two types:

                      1. NOR type cell and

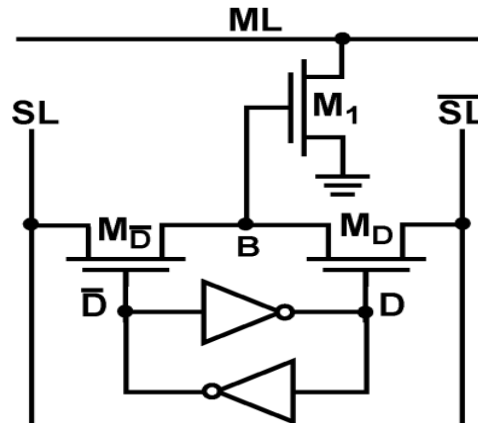   2. NAND type cell

2.2.1 NOR TYPE CELL

Fig. 2.3: 9-T NOR type CAM

The NOR cell implements the comparison between the complementary stored bit, D (and $\overline{D}$), and the complementary search data on the complementary searchline, SL (and $\overline{SL}$), using four comparison transistors, M_1 through M_4, which are all typically minimum-size to maintain high cell density. These transistors implement the pull-down path of a dynamic XNOR logic gate with inputs SL and D. Each pair of transistors, M_1/M_3 and M_2/M_4, forms a pull-down path from the matchline, ML, such that a mismatch of SL and D activates least one of the pull-down paths, connecting ML to ground. A match of SL and D disables both pull-down paths, disconnecting ML from ground. The NOR nature of this cell becomes clear when multiple cells are connected in parallel to form a CAM word by shorting the ML of each cell to the ML of adjacent cells. The pulldown paths connect in parallel resembling the pulldown path of a CMOS NOR logic gate. There is a match condition on a given ML only if every individual cell in the word has a match.

Modification for Ternary cell:

Another SRAM is added for wildcard operation. One bit D for left pull-down path and another bit $\overline{D}$ for left pull-down path. Ternary operation is done putting D and $\overline{D}$ to "1" (D and $\overline{D}$ is not necessarily complementary).As a result It disables both the pull-down paths and compels to show 'match' regardless of the inputs. Also allows searching for "X" by setting both SL and $\overline{SL}$ to logic "0"
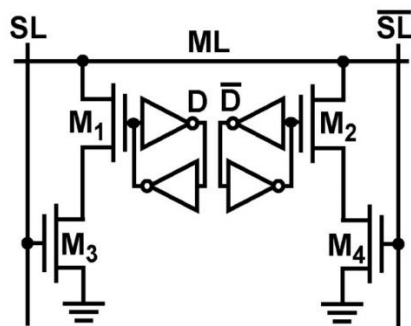
Fig. 2.4: Ternary core cell for NOR-type CAM [13], [14]

Stored value	Stored bit		Search bit	
	D	$\overline{D}$		
0	0	1	0	1
1	1	0	1	0
X	1	1	0	0

Table 2.2: Ternary encoding for NOR cell

2.2.2 NAND TYPE CELL
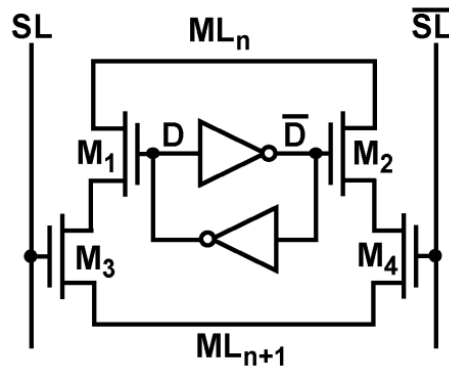
Fig. 2.5: 10-T NAND-type CAM

The NAND cell implements the comparison between the stored bit, D, and corresponding search data on the corresponding searchlines, (SL, $\overline{SL}$), using the three comparison transistors

M_1, M_D and $\overline{M_D}$, which are all typically minimum-size to maintain high cell density. We illustrate the bit-comparison operation of a NAND cell through an example. Consider the case of a match when $\overline{SL}=1$ and D=1. Pass transistor M_D is ON and passes the logic "1" on the SL to node B. Node B is the bit-match node which is logic "1" if there is a match in the cell. The logic "1" on node B turns ON transistor M_1. Note that M_1 is also turned ON in the other match case when SL=0 and D=0. In this case, M_D passes logic high to raise node B. The remaining cases, where SL≠D, result in a miss condition, and accordingly node B is logic "0" and the transistor M_1 is OFF. Node B is a pass-transistor implementation of the XNOR function SL®D. The NAND nature of this cell becomes clear when multiple NAND cells are serially connected. In this case, the ML_n and ML_{n+1} nodes are joined to form a word. A serial nMOS chain of all the transistors resembles the pull-down path of a CMOS NAND logic gate. A match condition for the entire word occurs only if every cell in a word is in the match condition.

Modification for Ternary cell:

For ternary operation an additional mask bit at node M is used. During ternary operation this mask bit is set to "1" which turns the M_{mask} ON and shows 'match' regardless of the value of D. Also allows to search for "X" setting both SL and $\overline{SL}$ to "1".

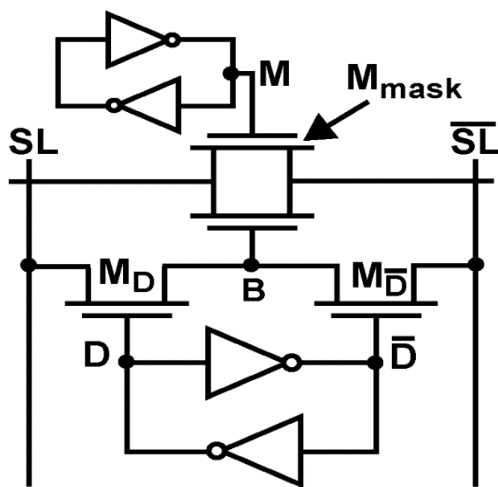Fig. 2.6: Ternary core cell for NAND-type CAM cell [13], [14]

Stored value	Stored bit		Search bit	
	D	M	SL	$\overline{SL}$
0	0	1	0	1
1	1	0	1	0
X	1	1	0	0
X	1	1	1	1

Table 2.3: Ternary encoding for NAND

NOR cell is usually preferred for TCAM as NAND cell has a major drawback. When searchlines are driven to supply voltage V_{DD} it provides a reduced logic "1" voltage at node B, which can reach only V_{DD} - V_{tn} due to NMOS. On the other hand NOR cell provides a full rail voltage to gates of all comparison transistors. Even in the worst case NOR-cell evaluation is faster than NAND-cell operation. So for high speed operation NOR cell is preferred. Also there is a potential charge sharing problem between ML nodes.

2.3 TCAM Array

Joining TCAM cells side by side forms TCAM Word. Usually of 36 to 144 bit and both NAND type and NOR type.

Joining multiple words forms TCAM Array. An example of a 4-bit ×4-word array is shown in the figure. It stores the routing table in descending order of priority from bottom to top. Because of the mask bits, each entry represents a range. For example, the second highest priority entry 10XX means packets with destination addresses in the range 1000 to 1011 have to be routed to the encoder. Here search word 1011 matches with 00, 01, 10 matchlines. But as TCAM is implemented for Longest Prefix Match (LPM) it shows the most direct route to the output. So matchline 00 goes to the encoder.
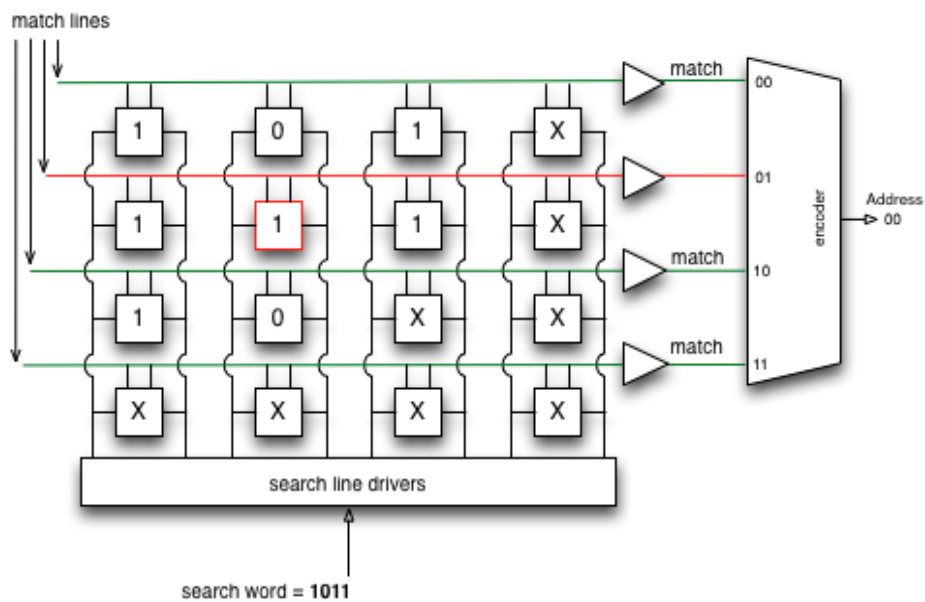
Fig. 2.7: A 4-bit × 4-word TCAM array

2.4 MATCHLINE STRUCTURE

Two key structures of TCAM are-

- Matchline
- Searchline

In this section Different matchline structures are discussed. As discussed previously, TCAM cells can be of two types-NOR cell, NAND cell. According to cell type Matchline also can be divided as follows-

- NOR Matchline
- NAND Matchline

2.4.1 NOR Matchline

Fig. 2.8: One TCAM data word consisting of n-bit NOR-type cells

NOR matchline is formed by connecting NOR cells in parallel. NOR search cycle has three phases:

1. Searchline Precharge
2. Matchline Precharge
3. Matchline Evaluation

1. *Searchline Precharge:*

At the first phase searchlines are precharged low. This is done to disable the pull-down paths in order to disconnect the matchline from ground.
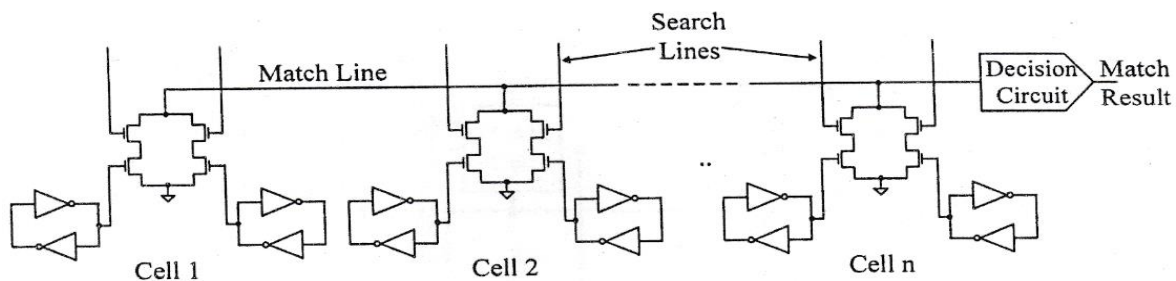
2. *Matchline Precharge:*

Matchlines are precharged by the M_{pre} transistors (not shown in the figure) to high.

3. *Matchline Evaluation*:

Matchline evaluation phase is triggered by driving the searchlines to search word values.

For match, ML (Matchline) voltage stays high as there is no discharge path to ground.

For mismatch, ML will find at least one path to discharge to ground.

The decision circuit (varies with ML sensing schemes) senses the voltage on ML nd generates a corresponding full-rail output match result. High speed operation (even in the worst case) is the major advantage of NOR matchline.
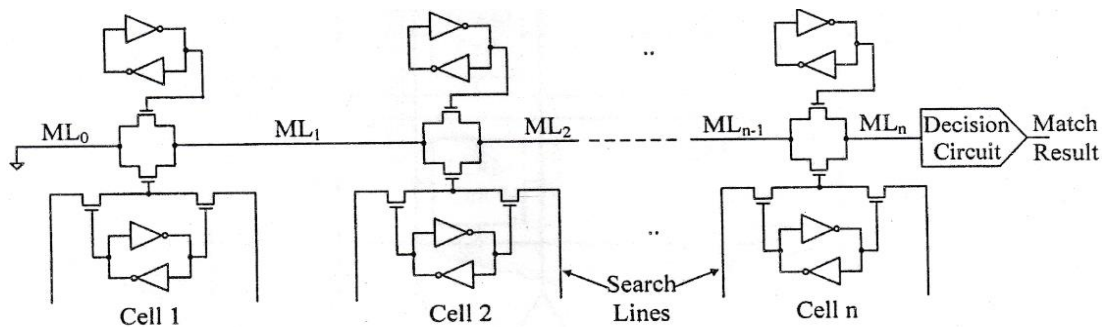
2.4.2 NAND Matchline

Fig. 2.9: One TCAM data word consisting of n-bit NAND-type cells.

NAND matchline is formed by cascading n-number of NAND cells.

Three phases of NAND search cycle is discussed below:

1. *Searchline Precharge*:

At the first phase searchlines are precharged low. This is done to disable the pull-down paths in order to disconnect the matchline from ground.

2. *Matchline Precharge:*

Matchlines are precharged by the M_{pre} transistors (not shown in the figure) to high. Evaluation transistors M_{eval} turns ON.

3. *Matchline Evaluation:*

Matchline evaluation phase is triggered by driving the searchlines to search word values.

For match, ML will find effective path to discharge to ground.

 For mismatch, At least one of the series NMOS transistors is OFF. So ML (Matchline) voltage stays high as there is no discharge path to ground.

The decision circuit (Sense amplifier MLSO) senses the voltage on ML and generates a corresponding full-rail output match result.

CHAPTER THREE

MATCHLINE SENSING SCHEMES

In the above discussion, it was assumed that the decision circuit produces high output for a full match and low output otherwise .The variation, constructions operations of different types of decision circuits will be the topic of discussion in this section.

Matchlines and searchlines are highly capacitive. Combined SL capacitance arising from large number of columns is also very large. During search, all the MLs and SLs are activated simultaneously. During match detection, switching of this highly capacitive lines cause huge power consumption. Energy efficient match detection or matchline sensing schemes try to reduce this power consumption.

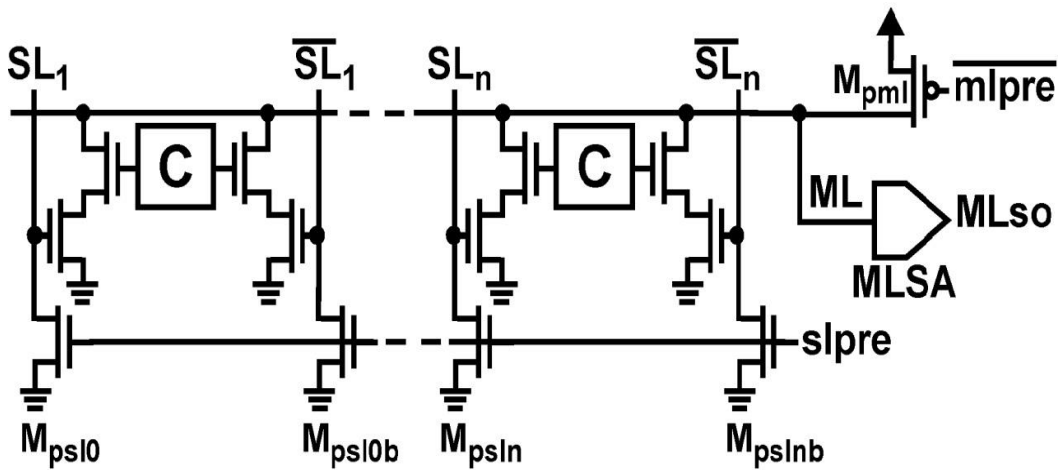## 3.1 CONVENTIONAL (PRECHARGE-HIGH) MATCHLINE SENSING SCHEME

Fig 3.1(a): The schematic with precharge circuitry for matchline sensing using the precharge-high scheme. [1]
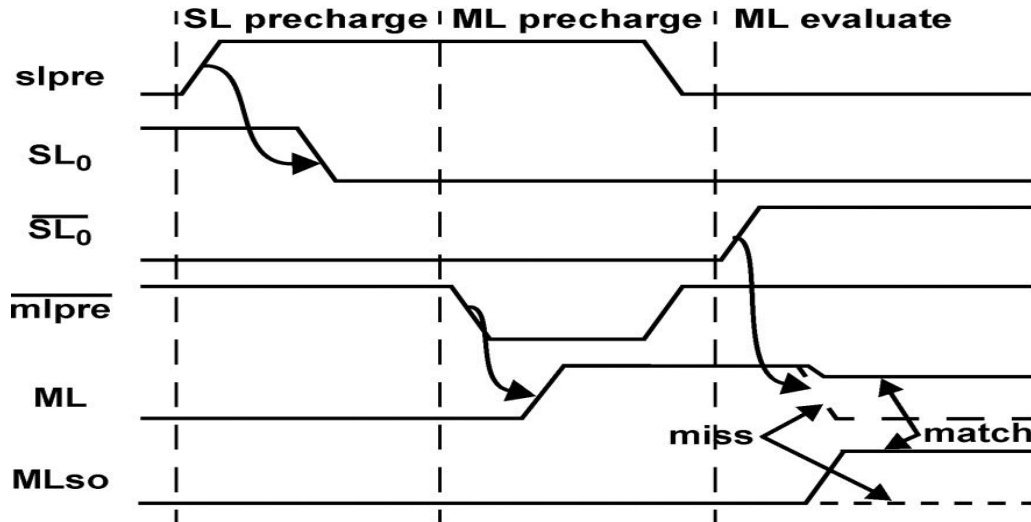
Fig 3.1(b): Corresponding timing diagram showing relative signal transitions. [1]

Basic Operation:

The basic scheme for sensing the state of the NOR matchline is first to precharge high the matchline and then evaluate by allowing the NOR cells to pull down the matchlines in the case of a miss, or leave the matchline high in the case of a match. Fig. 3.1(a) shows, in schematic form, an implementation of this matchline-sensing scheme. Fig. 3.1(b) shows the signal timing which is divided into three phases:

SL precharge:

The operation begins by asserting slpre to precharge the searchlines low, disconnecting all the pull down paths in the NOR cells.

ML precharge:

With the pull down paths disconnected, the operation continues by asserting to precharge the matchline high. Once the matchline is high, both slpre and are de-asserted.

ML evaluation:

Evaluation begins by placing the search word on the searchlines. If there is at least one single-bit miss on the matchline, a path (or multiple paths) to ground will discharge the matchline, ML, indicating a miss for the entire word, which is output on the MLSA sense-output node, called

MLSO. If all bits on the matchline match, the matchline will remain high indicating a match for the entire word.

Drawbacks:

All the matchlines are precharged to V_{DD}. As number of mismatch of words is lower than match word and those mismatched ML voltage is discharged to ground it causes large amount of energy to be wasted [1].

The comparison circuit suffers from charge sharing problem. It depend on whether the storage bits are connected to the top transistors or bottom transistors in the pull-down path.

(a)   (b)
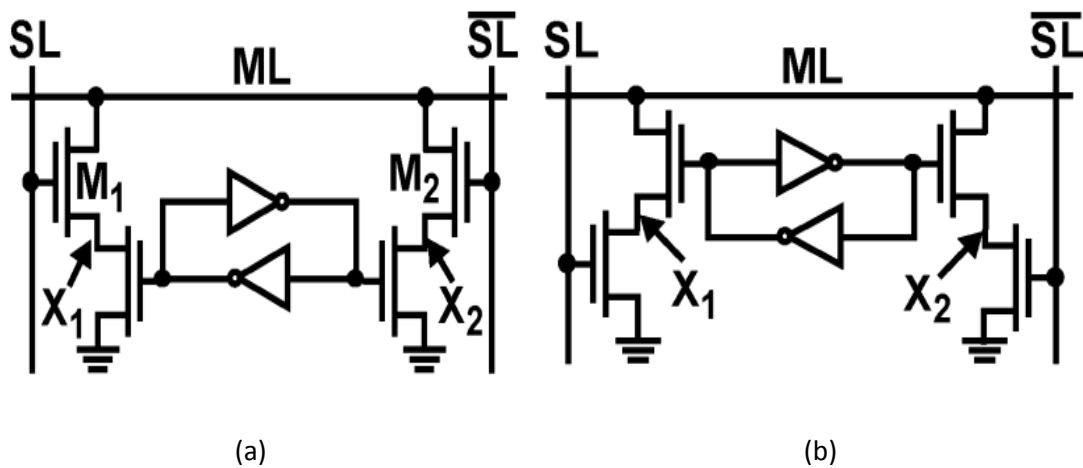
Fig 3.2: Two possible configurations for the NOR cell: (a) the stored bit is connected to the bottom transistors of the pulldown pair, and (b) the stored bit is connected to the top transistors of the pulldown pair. [1]

In the above figure transistors are connected to the bottom transistors. Charge sharing occurs between ML and nodes X_1, X_2. After the matchline precharge phase during searchline precharge one of the searchline is turned ON causing ML voltage V_{ML} to drop.

To solve this from this charge sharing problem configuration shown in Fig. 3.2(b) can be used where search bits are connected to the top transistors. Since stored bit is constant during a search operation, charge sharing is eliminated [1].

3.2 SELECTIVE-PRECHARGE SCHEME

Fig 3.3: Simple implementation of the selective-precharge matchline technique.

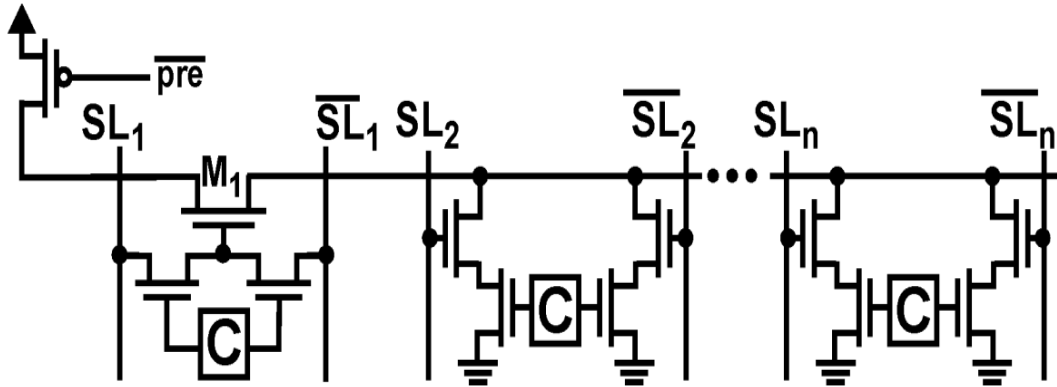In the conventional scheme, same amount of energy is required for each matchline regardless of the data pattern and match/mismatch. Selective precharge scheme can provide a good saving in the power consumption. Except the worst case scenario selective precharge scheme is the most common method used to save power on matchline. In this scheme only a subset of TCAM cells in the entire word is compared first and further requirement of comparison is decided accordingly.

Fig. 3.3 is a simplified schematic of an example of selective precharge similar to that presented in [15]. The example uses the first bit for the initial search and the remaining bits for the remaining search. To maintain speed, the implementation modifies the precharge part of the precharge-high scheme conventional scheme. The ML is precharged through the transistor M_1, which is controlled by the NAND CAM cell and turned on only if there is a match in the first CAM bit. The remaining cells are NOR cells.ML of the NOR cells must be pre-discharged (circuitry not shown) to ground to maintain correct operation in the case that the previous search left the matchline high due to a match. Thus, one implementation of selective precharge is to use this mixed NAND/NOR matchline structure.

Drawbacks:

- Initial matching draws higher power

- Not applicable for non-uniform data distribution

- Initial match bits can be identical

3.3 CURRENT RACE SCHEME

To overcome the problems of conventional ML sensing schemes Current Race (CR) scheme is proposed by [4]. CR scheme is the most versatile scheme. Many authors [5], [10], [7], [8], [9], [2], [3] have used this technique as a basis for the construction of their own techniques in an attempt to have better performance than CR This scheme precharges the ML low and evaluates the ML state by charging the ML with a current I_{ML} supplied by a current source. The precharge signal, MLRST, starts the search cycle by precharging the ML low. Since the ML is precharged low, the scheme concurrently charges the SL to their search data values, eliminating the need for a separate SL precharge [1].

The main difference between CR scheme and conventional scheme is that CR scheme is a precharge-low scheme while conventional scheme is precharge-high. Instead of charging all MLs to high, CR scheme pre-discharges all MLs to low. During match evaluation, all MLs are charged towards high. Matched MLs charge quickly to large voltage (due to NOR-type cells) while mismatched MLs have much lower voltage due to presence of discharging paths. Precharge-low technique has another advantage, in conventional scheme, the SLs need to be discharged to zero voltage during precharge phase to ensure that MLs remain disconnected from ground. During evaluation, SLs are loaded with the actual search key. In CR scheme, since MLs are pre-discharged, SLs need not to be switched to zero. This reduced SL switching activity results in around 50% saving of SL energy. Furthermore, CR scheme eliminates unnecessary ML charging by using replica control (dummy word). Figure 3.4 shows the CR scheme.
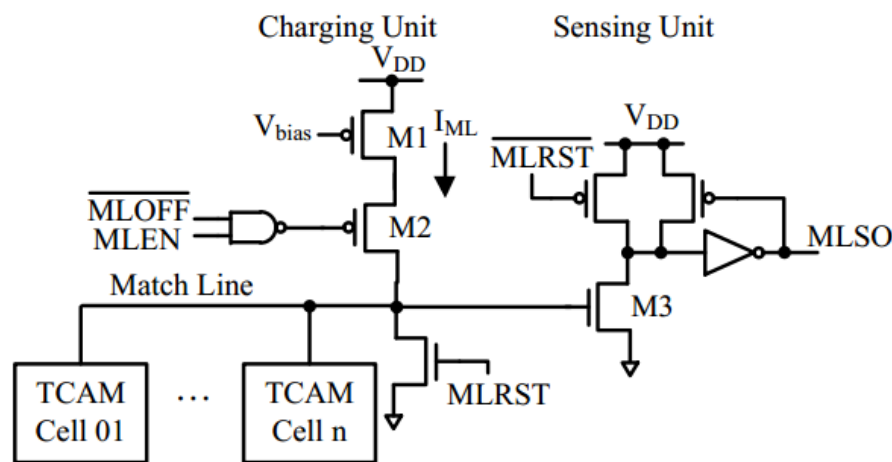

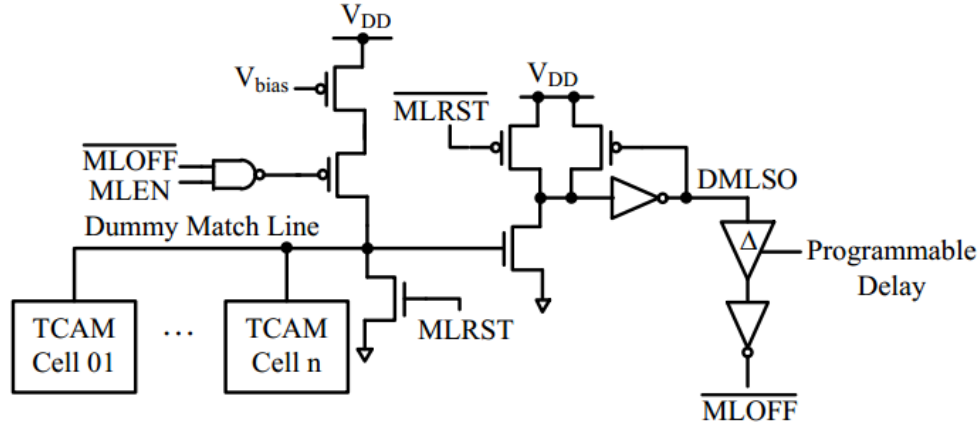**Fig. 3.4(a)**. Word in Current Race Scheme, [4]

Fig. 3.4(b). Dummy word in Current Race Scheme, [4]

MLSA consists of a charging unit and sensing unit. The operation starts by resetting any voltage on ML (VML) and MLSA outputs (MLSOs and DMLSO) to zero by MLRST signal. Then, asserting signal MLEN turns on transistor M2 causing I_{ML} to flow. The match line capacitance (C_{ML}) starts charging. If the word is fully matched, ML can charge up to the threshold voltage ofM3 turning it on. Therefore, MLSO can become high. If the word is not matched, V_{ML} is small since ML has discharging path(s) to ground. M3 remains OFF and MLSO remains zero. The dummy word is always matched and its MLSA always produces high DMLSO. A delayed and inverted version of DMLSO, i.e., MLOFF is used to turn off M2 transistors in all words. This eliminates unnecessary charging (and energy consumption) of MLs. The dummy word also minimizes the effects of process variations since it is situated close to the regular words and goes through the same process variations as those words. The programmable delay after DMLSO ensures that all the matched MLs get 35 sufficient time to charge up to the threshold voltage of M3 if the dummy word detects the match earlier (due to process variations). V_{bias} controls I_{ML} and hence controls the speed and energy consumption of the match detection process. In CR scheme, NOR cell comparison circuit is preferred over the comparison circuit. Since search data bits are same along a column, C_{ML} remains the same for all MLs in a search. This ensures good matching between MLs and prevents sensing error due to capacitance variation. Current race scheme supplies the same initial current (on the rising edge of MLEN) to both mismatched and matched MLs. But, since matched MLs charge to higher voltage, I_{ML} gradually decreases in matched MLs. On the contrary, mismatched MLs have lower resistance path(s) to ground. The equivalent resistance of the ML pulldown path decreases with increasing number of mismatches. Therefore, I_{ML} increases with increasing number of mismatches. Since most of the MLs are mismatched, large currents to mismatched MLs cause significant wastage of energy. This problem can be solved by supplying smaller currents to mismatched MLs.

This sensing scheme saves power in two ways. First, by cutting the current when DML reaches the threshold voltage of the ML sense amplifier, the voltage swing of all MLs is limited to 960 mV (~VDD/2) [5]. This reduces the ML power dissipation by a factor of two when compared to a full swing ML sensing scheme. Second, by precharging the MLs to ground (mismatch state), the SLs do not need to be reset between consecutive searches, hence minimizing SL switching activity and thereby reducing SL power dissipation by a factor of two [4].

3.4 POSITIVE FEEDBACK MATCHLINE SENSING SCHEME

The conventional CR-MLSA charges all MLs with the same magnitude of current for both matched and mismatched cases. But, since matched MLs charge to higher voltage, I_{ML} gradually decreases in matched MLs. On the other hand, mismatched MLs have lower resistive path to ground. The equivalent resistance of the ML decreases with increasing number of mismatches. Therefore, I_{ML} increases with increasing number of mismatches. Since most of the MLs are mismatched, significant number of energy is wasted due to large current in these mismatched MLs. The proposed Positive Feedback Matchline Sensing Scheme solves this problem by supplying smaller current to mismatched MLs.

3.4.1 MISMATCH-DEPENDENT POWER ALLOCATION SCHEME

Mismatch-Dependent Power Allocation Scheme is proposed by [3]. Here feedback unit has been added to the basic CR MLSA to detect mismatched MLs during charging and reduce current to these MLs. Fig. 3.5 shows the proposed Mismatch-Dependent Power Allocation Scheme.
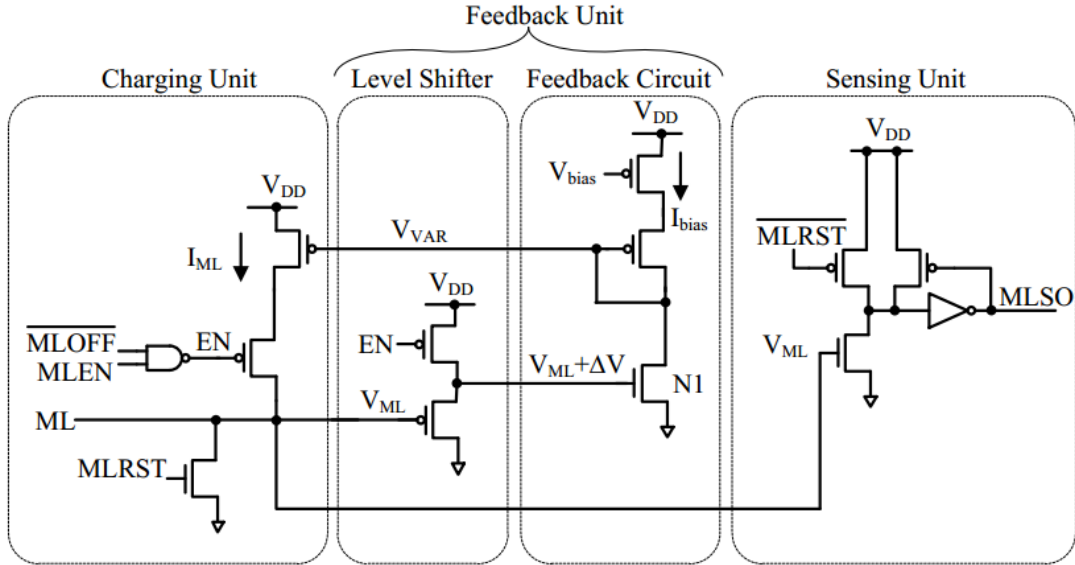
Fig. 3.5. MLSA in Mismatch-Dependent Power Allocation Scheme. Dummy word isn't shown here, [3], [9]

The feedback unit contains a level shifter and a feedback circuit to implement feedback. In speed optimized setting, the MLSO, V_{ML} and V_{VAR} are pre-discharged to ground to first. MLEN starts the charging of all MLs. Initially both V_{ML} and V_{VAR} are increased by current I_{ML} and I_{bias} respectively. As V_{VAR} increases I_{ML} decreases. But with increasing V_{ML}, the level shifter output also increases.

V_{ML} in a matched ML increases faster than that in mismatched ML. the level sifter output becomes significantly high in a matched ML to turn on N1 and V_{VAR} starts to discharge. As a result I_{ML} for matched ML increases again.

For a mismatched ML, V_{ML} rises slowly and to a lower value which depends on the number of mismatches presents in word. So for mismatched MLs N1 may get weakly turned on or it may remain off depending on the number of mismatches. This results small or no reduction in V_{VAR}. As a result I_{ML} in a mismatched ML keeps decreasing.

So matched MLs get higher current than mismatched MLs. Same dummy word in CR MLSA is used to generate MLOFF signal to control the ML charging time. This scheme shows significant energy reduction over CR scheme. But the level shifter and feedback unit consume considerable amount of energy. In case of a mismatch N1 sometimes remains weakly turn on which means there is a conducting path between VDD and ground. So for a large number of mismatches, V_{ML} is small since ML has multiple path to ground. This causes a static power consumption. Also capacitive loading on the ML increases due to large transistor size in level shifter. Again transistor count in this MLSA is higher than CR MLSA. Finally feedback action is very sensitive to V_{bias} and transistor sizes. So the transistor sizes need to be tuned for better performance.

3.4.2   ACTIVE FEEDBACK SCHEME

To overcome the problems of MD scheme Active Feedback ML sensing scheme is proposed. The proposed scheme does not consume any static power and it outperforms the MD-MLSA in speed, energy, area and robustness [2]. Fig. 3.6 shows the proposed MLSA with active feedback.
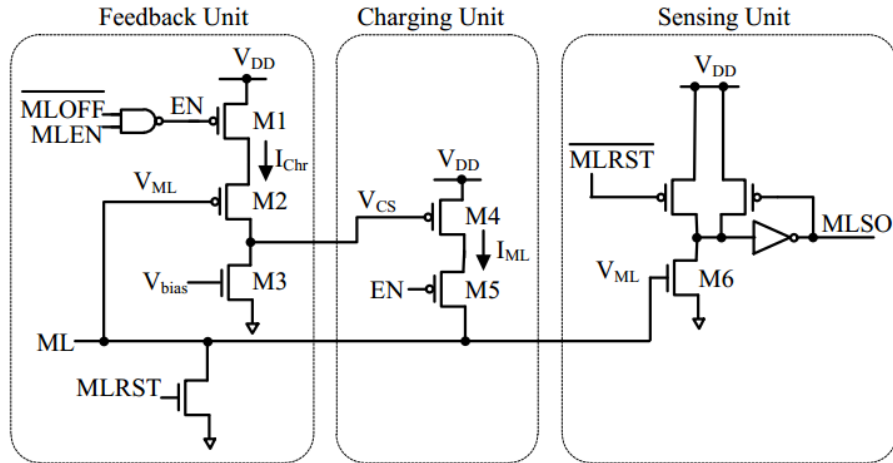

Fig. 3.6. MLSA in Active Feedback Scheme. Dummy word isn't shown here, [2], [10]

In active feedback scheme V_{ML} and MLSO are reset to zero by MLRST signal. Search starts by asserting MLEN. M1 and M2 turns on making gate capacitance of M4 to charge. Initially I_{ML} and I_{Chr} are same. So, V_{cs} and V_{ML} increase with time. As a result I_{ML} and I_{Chr} decrease, this makes M2 to turn off. M3 is always on due to dc V_{bias}.

In matched MLs V_{CS} discharges through M3. That makes M4 to conduct and I_{ML} to increase again.

In mismatched MLs, V_{ML} rises slowly. This keeps M2 on and makes Vcs to increase which makes M4 to turn off. As a result I_{ML} decreases.

So matched MLs get higher current than mismatched MLs. Energy consumption in mismatched words become small. V_{BIAS} and MLSA transistor sizes need to be tuned for better performance. As M3 is always on energy is wasted due to current flow through it in matched MLs. So further energy consumption reduction is needed.

3.4.3   RESISTIVE FEEDBACK SCHEME

Fig. 3.7 shows the proposed MLSA with resistive shielding. It uses an nMOS transistor (M3) in the triode region to decouple the ML and its MLSA. The N3 channel resistance shields the sensing point [SP in Fig. 8.5] from the highly capacitive ML. This way, the current source (I_{Chr}) can be sized down to save power without sacrificing the sensing speed. Note that due to the body effect and the decreasing gate-to-source voltage (V_{GSN3}) the N3 channel resistance increases when the ML voltage is rising up [2]. The N3 channel resistance depends on the number of mismatch bits.
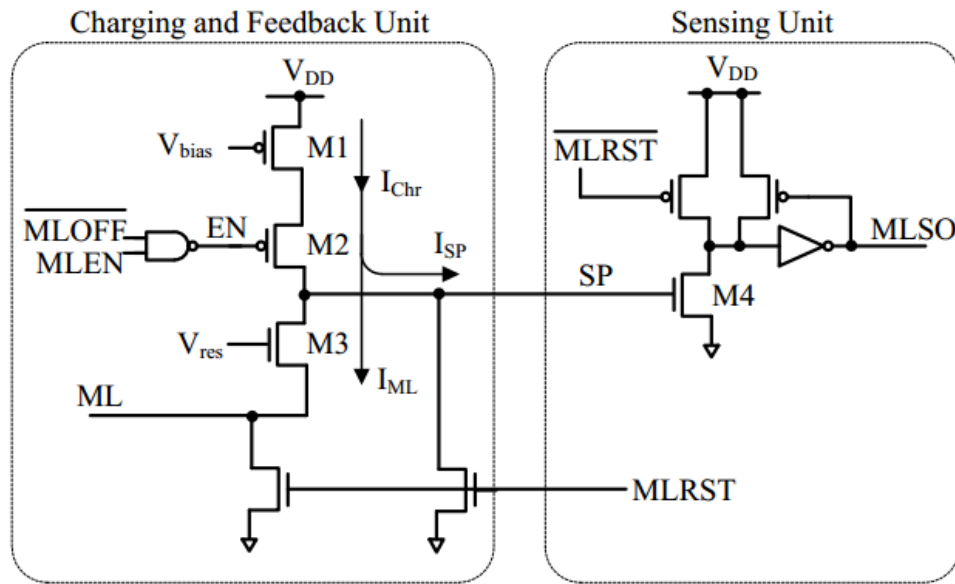

Fig. 3.7. MLSA in Resistive Feedback Scheme. Dummy word isn't shown here, [2], [10]

Voltages of ML, sensing point (sp) and MLSO are reset to zero by MLRST signal. The MLEN signal (shown in fig. 8.5) enables the MLSA by activating EN thus turning on M2. V_{BIAS} controls the total charging current I_{Chr}. Initially IML and IChr are same in all MLs. V_{ML} is increased by I_{ML}, so drain to source voltage of M3 decreases. Therefore I_{ML} decreases. Reduction in I_{ML} causes increase in I_{SP}. Increase in I_{SP} causes increase in V_{SP}. This, in turn, increases drain to source voltage of M3 again which increases I_{ML} again. As a result feedback continues between ML and SP.

In case of a match V_{ML} rises quickly because there is no discharging path to ground. Due to feedback action V_{SP} also increases rapidly. Since capacitance at node SP is much lower than that of ML increase of V_{SP} is much larger than V_{ML}.

In case of a mismatch V_{ML} rises slowly. Therefore V_{SP} also rises slowly. V_{SP} in a matched ML reaches threshold voltage much quicker than that in a mismatch ML and produce a high output at MLSO. V_{SP} of mismatched words do not get the sufficient time to charge up to the sensing threshold voltage. Therefore MLSOs remain low.

This technique of ML sensing is extremely simple in construction. Two additional transistors are required than CR MLSA. Faster sensing of the dummy word also reduces energy consumption because the ML current sources are shut down sooner. The charging current of mismatched ML is less affected by the M3 resistance because it has a larger and a weaker body effect than that of a matched ML. Energy and delay can be further reduced by decreasing V_{res} (shown in fig. 8.5) [2]. The sensitivity of the feedback action to process variation is also low since the feedback is weak. There are two analog control voltages which has a negative effect. A combination of small V_{res} and I_{Chr} large may reduce the voltage margin, causing a false match. The effectiveness of this scheme is further reinforced by the fact that a reduction in V_{res} decreases the EDP more rapidly than the voltage margin. In addition, for a small V_{res}, a reduction in I_{Chr} improves the voltage margin significantly without making much difference in the EDP [2].

In terms of match detection speed and energy, this technique has the potential to excel other existing feedback techniques.

CHAPTER FOUR

SIMULATION AND COMPARISON

The main purpose of this thesis is to compare the performances of different ML sensing schemes. Performance largely depends on the search speed that means the search time and noise margin that means voltage margin. The advantage of low search time is that search speed increases which is one of the main concerns in designing TCAM. Again the advantage of higher voltage margin is that if noise is added then it will not show a false match.

We used 16×16 array size and 0.18μm technology. We simulated these schemes in HSpice simulation software.

We performed simulations and the simulation waveforms are shown here

4.1 CURRENT RACE SCHEME

Simulated search time in CR scheme is 5.0446ns which is measured by subtracting the starting time of a matched MLSO (in figure DMLSO) and starting time of MLRST signal. Fig. 4.1(a) shows the search time of current race scheme.
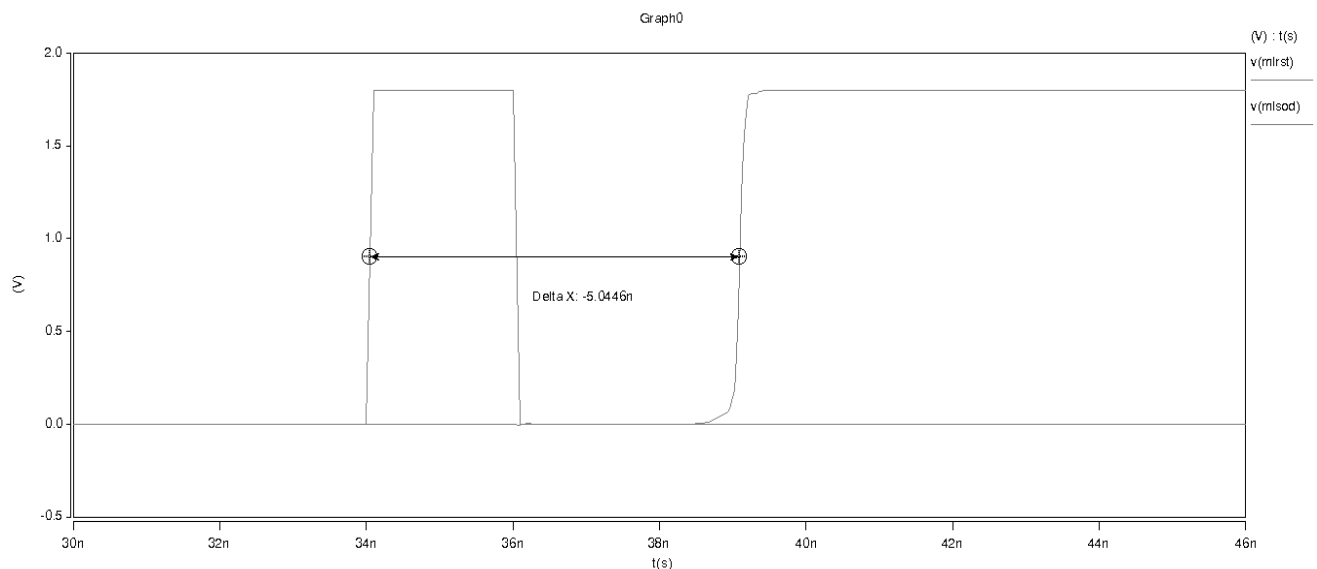

Fig. 4.1(a). Search time measured between V_{MLRST} and V_{MLSOD} at 0.9V (50% of V_{in}) in CR scheme

Again simulated voltage margin in CR scheme is 1.112225V which is measured by subtracting the amplitude of crossing between matched ML (ml1) and MLSO (mlso1) and maximum magnitude of 1 bit mismatched ML (ml3). Fig. 4.1(b) shows the voltage margin of current race scheme.
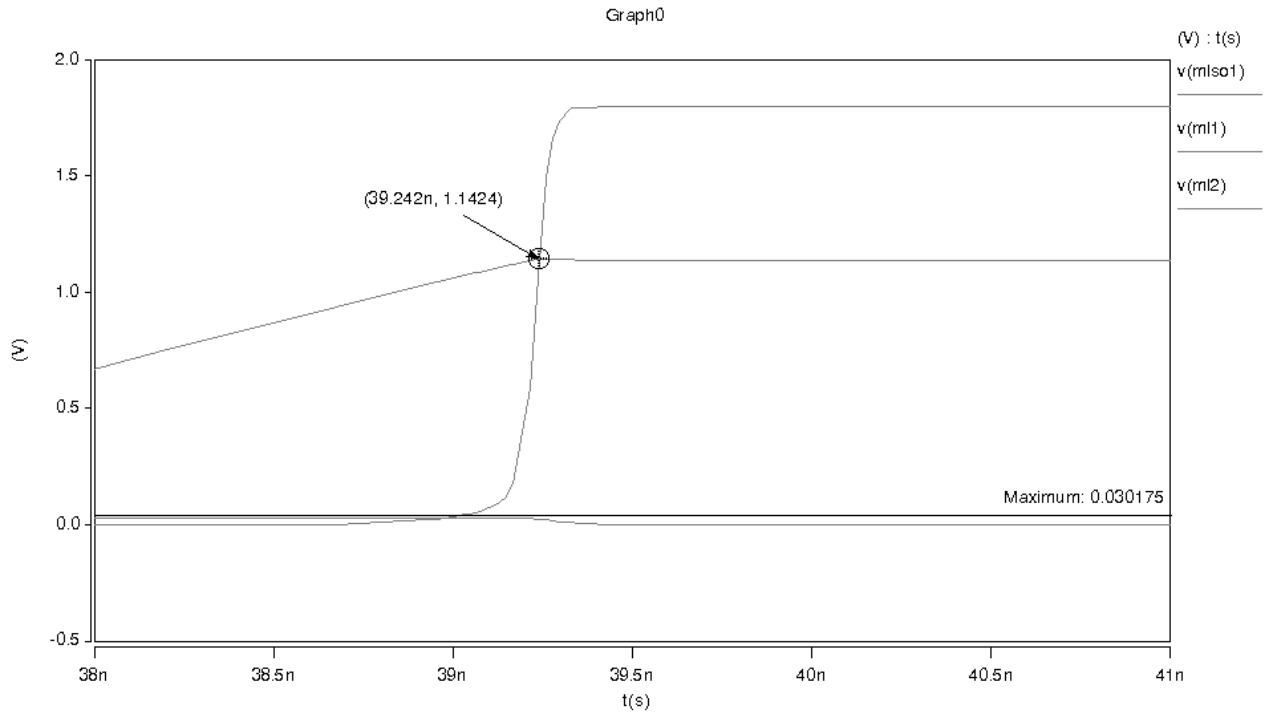

Fig. 4.1(b). Voltage margin measured between crossing of matched ML (ml1) and MLSO (mlso1) and maximum magnitude of 1 bit mismatched ML (ml3) in CR scheme

4.1   MISMATCH-DEENDENT POWER ALLOCATION SCHEME

Simulated search time in MD scheme is 2.5927ns which is measured by subtracting the starting time of a matched MLSO (in figure DMLSO) and starting time of MLRST signal. Fig. 4.2(a) shows the search time of mismatch-dependent power allocation scheme.
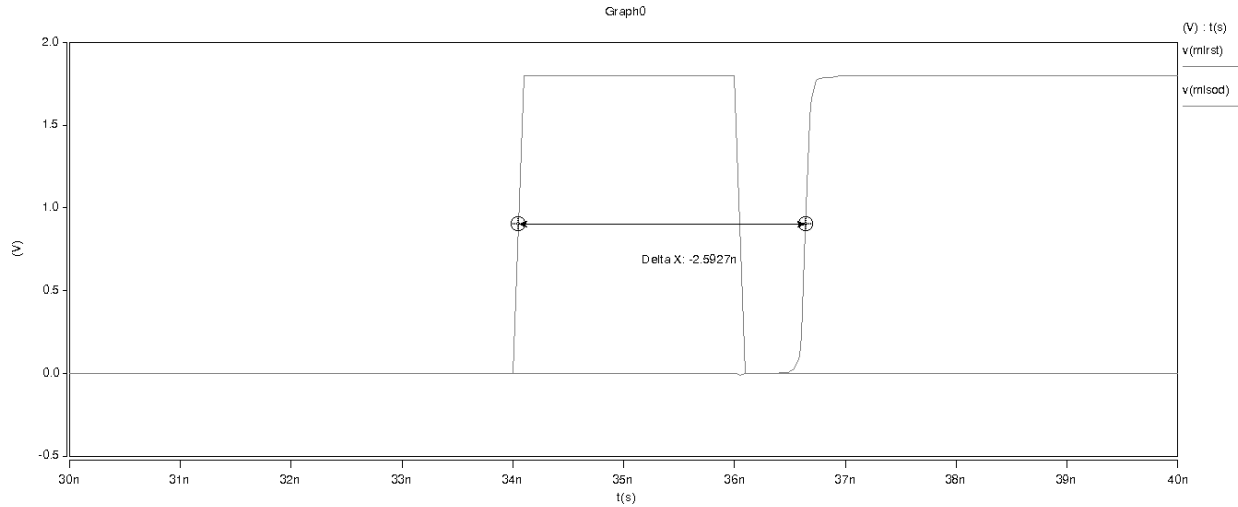
Fig. 4.2(a). Search time measured between V_{MLRST} and V_{MLSOD} at 0.9V (50% of V_{in}) in MD scheme

Again simulated voltage margin in MD scheme is 0.81093V which is measured by subtracting the amplitude of crossing between matched ML (ml1) and MLSO (mlso1) and maximum magnitude of 1 bit mismatched ML (ml3). Fig. 4.2(b) shows the voltage margin of mismatch-dependent power allocation scheme.
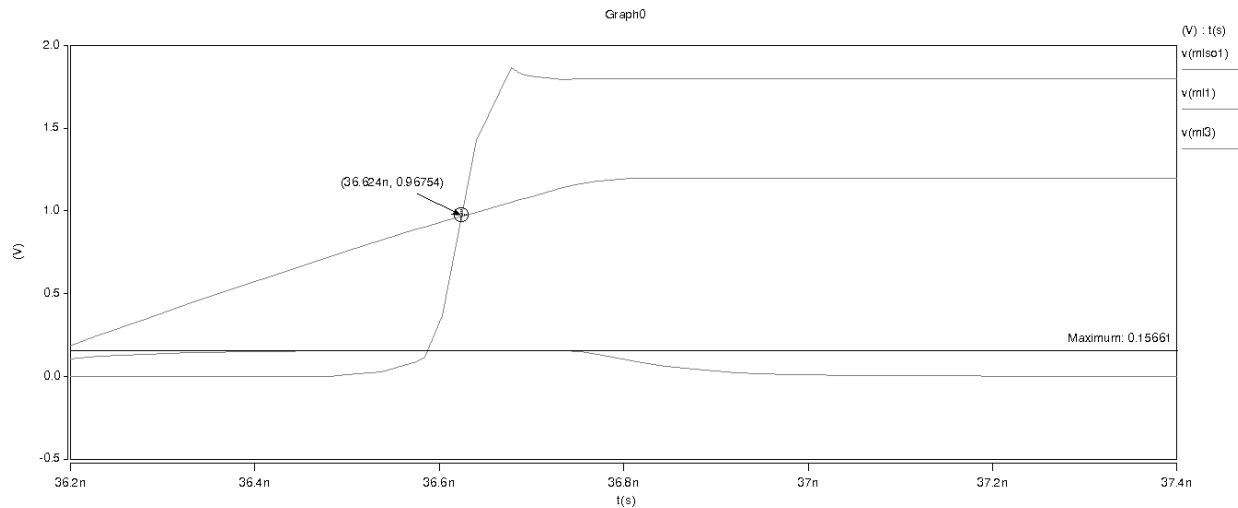

Fig. 4.2(b). Voltage margin measured between crossing of matched ML (ml1) and MLSO (mlso1) and maximum magnitude of 1 bit mismatched ML (ml3) in MD scheme

4.2 ACTIVE FEEDBACK SCHEME

Simulated search time in AF scheme is 2.6361ns which is measured by subtracting the starting time of a matched MLSO (in figure DMLSO) and starting time of MLRST signal. Fig. 4.3(a) shows the search time of active feedback scheme.
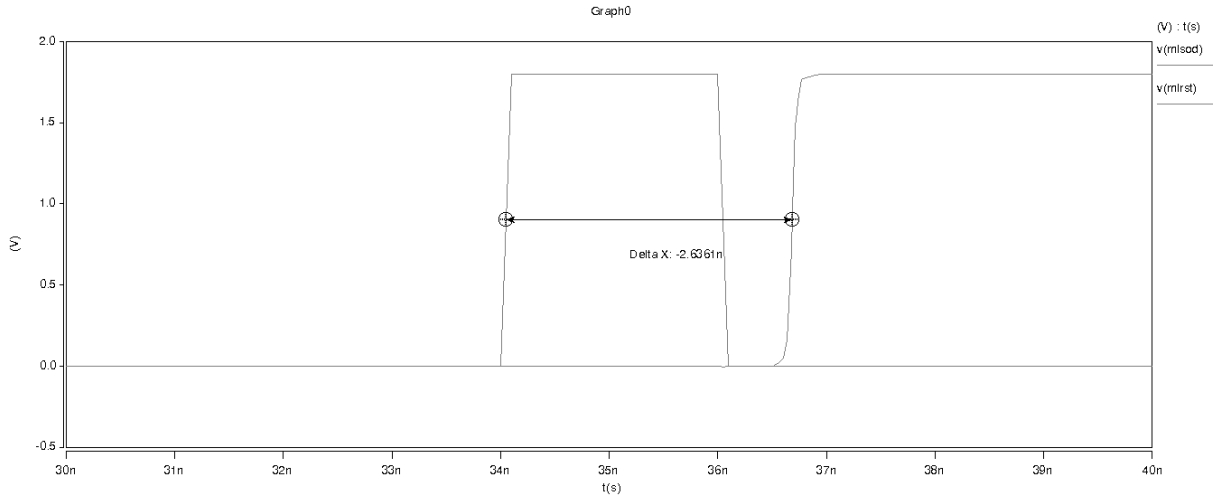

Fig. 4.3(a). Search time measured between V_{MLRST} and V_{MLSOD} at 0.9V (50% of V_{in}) in AF scheme

Again simulated voltage margin in AF scheme is 0.859028V which is measured by subtracting the amplitude of crossing between matched ML (ml1) and MLSO (mlso1) and maximum magnitude of 1 bit mismatched ML (ml3). Fig. 4.3(b) shows the voltage margin of active feedback scheme.
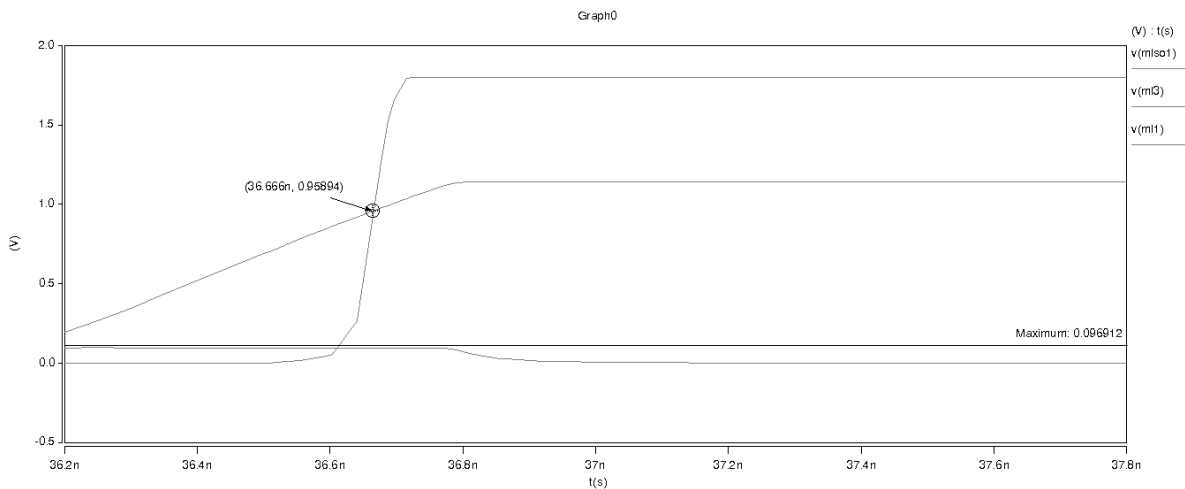

Fig. 4.3(b). Voltage margin measured between crossing of matched ML (ml1) and MLSO (mlso1) and maximum magnitude of 1 bit mismatched ML (ml3) in AF scheme

4.3 RESISTIVE FEEDBACK SCHEME

Simulated search time in RF scheme is 2.5266ns which is measured by subtracting the starting time of a matched MLSO (in figure DMLSO) and starting time of MLRST signal. Fig. 4.4(a) shows the search time of resistive feedback scheme.
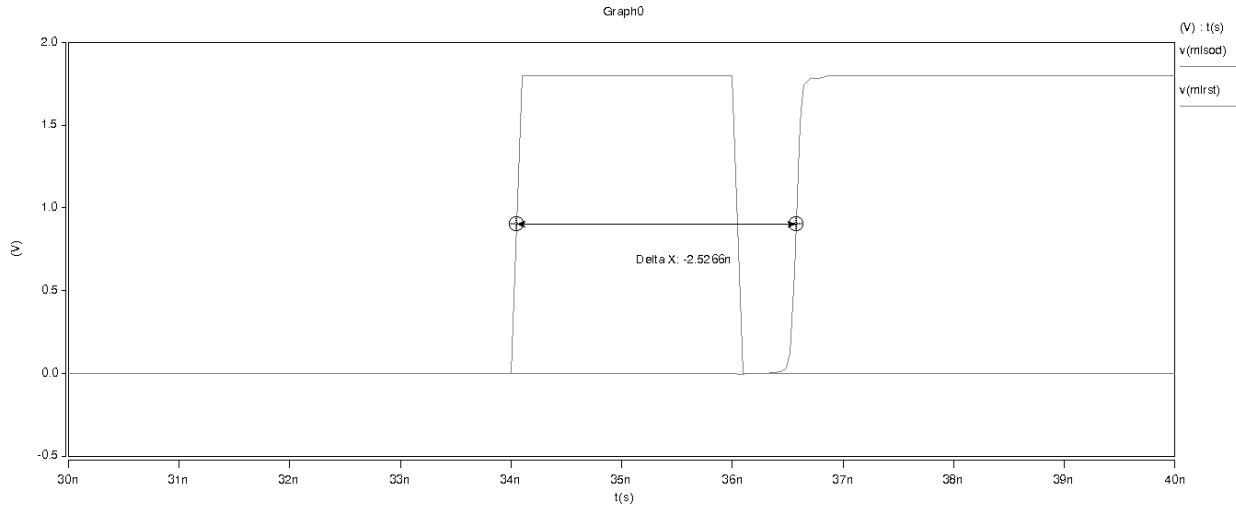

Fig. 4.4(a). Search time measured between V_{MLRST} and V_{MLSOD} at 0.9V (50% of V_{in}) in RF scheme

Again simulated voltage margin in RF scheme is 0.67165V which is measured by subtracting the amplitude of crossing between matched ML (ml1) and MLSO (mlso1) and maximum magnitude of 1 bit mismatched ML (ml3). Fig. 4.4(b) shows the voltage margin of resistive feedback scheme.
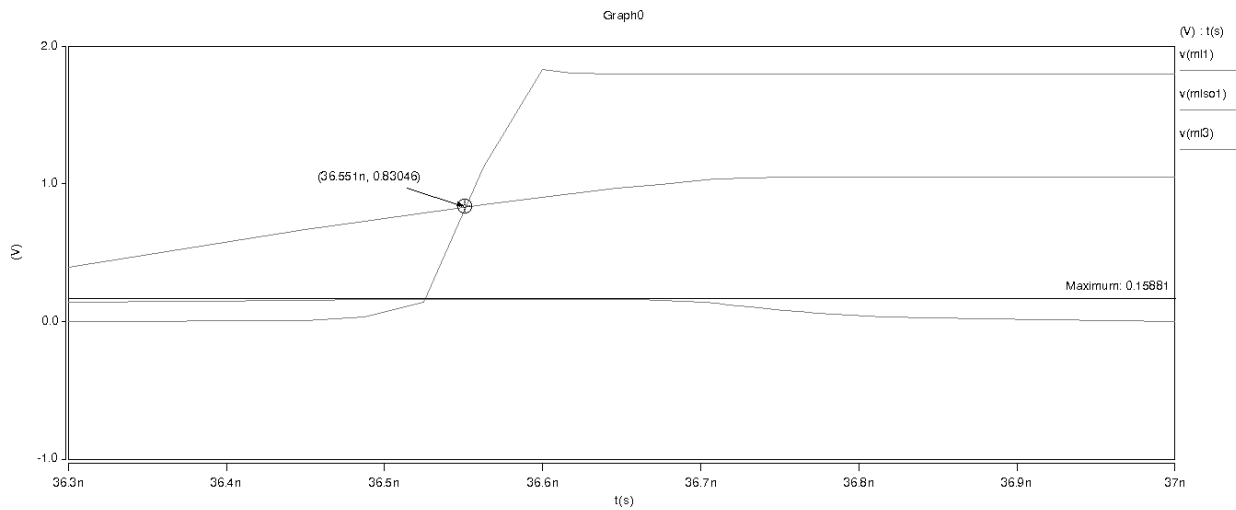

Fig. 4.4(b). Voltage margin measured between crossing of matched ML (ml1) and MLSO (mlso1) and maximum magnitude of 1 bit mismatched ML (ml3) in RF scheme

4.4   Comparison Table

We performed simulations and the simulation results (search time and Voltage margin) are shown in table 4.1

Schemes	Search Time (ns)	Voltage Margin (V)
Current Race(CR)	5.0446	1.112225
Mismatch Dependent(MD)	2.5927	0.81093
Active Feed-back(AF)	2.6361	0.859028
Resistive Feedback(RF)	2.5266	0.67165

Table 4.1. Comparison among different schemes

.

From the above table we can see that Resistive Feedback scheme has the lowest search time. So among these schemes RF scheme is with the best speed performance.

Again Current Race Scheme has the highest voltage margin. So among these schemes CR scheme is with the best noise immunity performance. But, its search time is too high (almost double than that of other positive feedback schemes). As a result its overall efficiency is compromised. Other than CR scheme Active Feedback scheme has higher voltage margin.

It is very difficult to get an optimum value of both speed and noise performances. It's a trade off as if we want to have better speed then voltage margin have to compromise and vice versa. So for better speed performance RF scheme can be used and for noisy environment AF scheme can be used.

CHAPTER FIVE
CONCLUSION AND FUTURE WORK

5.1 Conclusion

The proposed positive feedback schemes improve search speed at the cost of slight degradation of voltage margin. These MLSAs use the positive feedback to reduce the power consumption in TCAMs [3].

MD scheme allocates power to match decisions based on the number of mismatched bits in each TCAM word. With allocating less power to mismatched MLs and with most MLs being in this category, this scheme results in a considerable power reduction [3].

Again AF and RF schemes require less transistors than the MD scheme. The active-feedback MLSA improves energy savings and voltage margin without consuming any static power. Energy measurement results of the two MLSAs show reductions of 56% and 48%, respectively, over the conventional CR-MLSA [2].

5.2 Future Work

Our future works will be focused on these areas:

- Simulation and comparison of energy consumption among the matchline sensing schemes discussed in this thesis

- We will also try to increase the array size so that we can compare larger data. Larger sized array will take more simulation time than smaller sized array.

- We will also try to simulate layout designing.

- Tuning different sensitive transistors.

REFERENCES

[1] K. Pagiamtzis and A. Sheikholeslami, "Content-addressable memory (CAM) circuits and architectures: a tutorial and survey," IEEE Journal of Solid-State Circuits, vol. 41, no. 3, pp. 712-727, March 2006.

[2] N. Mohan, W. Fung, D. Wright, and M. Sachdev, "A low-power ternary CAM with positive-feedback match-line sense amplifiers," IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 56, no. 3, pp. 566-573, March 2009.

[3] I. Arsovski and A. Sheikholeslami, "Amismatch-dependent power allocation technique for match-line sensing in content-addressable memories," IEEE Journal of Solid-State Circuits, vol. 38, no. 11, pp. 1958-1966, Nov 2003.

[4] I. Arsovski, T. Chandler, and A. Sheikholeslami, "A ternary contentaddressable memory (TCAM) based on 4T static storage and including a current-race sensing scheme," IEEE Journal of Solid-State Circuits, vol. 38, no. 1, pp. 155–158, Jan 2003.

[5] N. Mohan and M. Sachdev, "Low-capacitance and charge-shared match lines for low-energy high-performance TCAMs," IEEE Journal of Solid-State Circuits, vol. 42, no. 9, pp. 2054-2060, Sept 2007.

[6] Banit Agrawal, Trimorthy Sherwood, "Modeling TCAM Power for Next Generation Network Devices," In Proceedings of IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS-2006).

[7] K. Pagiamtzis and A. Sheikholeslami, "Pipelined match-linesand hierarchical search-lines for low-power content-addressable memories," in Proceedings of the IEEE Custom Integrated Circuits Conference (CICC), 2003, pp. 383–386.

[8] K. Pagiamtzis and A. Sheikholeslami, "A low-power content-addressable memory (CAM) using pipelined hierarchical search scheme," IEEE Journal of Solid-State Circuits, vol. 39, no. 9, pp. 1512–1519, Sep 2004.

[9] I. Arsovski and A. Sheikholeslami, "A current-saving match-line sensing scheme for content-addressable memories,"in Digest of Technical Papers of IEEE International Solid-State Circuits Conference (ISSCC), 2003, pp. 304–305.

[10] N. Mohan, "Low-power high-performance ternary content addressable memory circuits," Ph.D. dissertation, Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, Ontario, Canada, 2006.

[11] T. Ogura, M. Nakanishi, T. Baba, Y.Nakabayshi, and R. Kasai, "A 336-kb content addressable memory for highly parallel image processing," in Proceedings of the IEEE Custom Integrated Circuits Conference (CICC), May 1996, pp. 273-276.

[12] F. Yu, R. H. Katz, and T. V. Lakshman, "Gigabit rate packet pattern matching using TCAM," in Proceedings of the IEEE International Conference on Network Protocols (ICNP), Oct 2004, pp. 174-183.

[13] S. Choi, K. Sohn, J. Kim, J. Yoo, and H.-J. Yoo, "A TCAM-based periodic event generator for multi-node management in the body sensor network," in Proceedings of the IEEE Asian Solid-State Circuits Conference (ASSCC), 2006, pp. 307–310.

[14] http://www.alldatasheet.com

[15] H. Kadota, J. Miyake, Y. Nishimichi, H. Kudoh, and K. Kagawa, "An 8-kbit content-addressable and reentrant memory," IEEE Journal of Solid-State Circuits, vol. 20, no. 5, pp. 951–957, Oct 1985.