# Improve CAM Power Efficiency Using Decoupled Match Line Scheme

Yen-Jen Chang Dept. of Computer Science National ChungHsing University Taichung, Taiwan ychang@cs.nchu.edu.tw Yuan-Hong Liao Dept. of Computer Science National ChungHsing University Taichung, Taiwan s9356042@cs.nchu.edu.tw Shanq-Jang Ruan Dept. of Electronic Engineering NTUST, Taipei, Taiwan sjruan@mail.ntust.edu.tw

# Abstract

Content addressable memory (CAM) is widely used in many applications that require fast table lookup. Due to the parallel comparison feature and high frequency of lookup, however, the power consumption of CAM is usually significant. In this paper we propose a decoupled match line scheme which combines the performance advantage of the traditional NOR-type CAM and the power efficiency of the traditional NAND-type CAM. In our design, a CAM word is divided into two segments, and then all the CAM cells are decoupled from the match line. By minimizing both the match line capacitances and switching activities, our design can largely reduce the CAM power dissipated in search operations. The results measured from the fabricated chip show that without any performance penalty our design can reduce the search energy consumption of the CAM by 89% compared to the traditional NOR-type CAM design.

# 1. Introduction

Content addressable memory (CAM) is a storage, which is addressed by the content (or data) rather than the memory address. Because CAM is a hardware table and can compare the search data with all the stored data in parallel, the speed of CAM lookup is much faster than the software lookup. Therefore, CAM is widely used in TLB, high associative cache, image processing, database and network routers, etc., all require fast table lookup. However, the power consumption of CAM is usually considerable due to the parallel comparison feature in which a large amount of transistors are active on each lookup. For example, in StrongARM [1] embedded processors, the fully-associative TLBs consume about 17% of the total chip power.

There are two traditional CAM designs; one is the NORtype CAM, and the other is the NAND-type CAM. The NORtype CAM provides the best performance in search operation, but it costs a large power consumption. In contrast, the NAND-type CAM trades the search performance for low power feature. As indicated in the previous researches, the main sources of power dissipation in CAM are the match lines and search lines (or bitlines). The power consumption of match lines can be reduced by reducing the voltage swing on the match lines [2][3], or by segmenting the match line [4][5]. Because the match lines are precharged conditionally in the segmentation techniques [4][5], the performance degradation is a vital and unavoidable disadvantage.

Combing the performance advantage of the traditional NOR-type CAM and power efficiency of the traditional NAND-type CAM, this paper presents a decoupled match line scheme, which can largely reduce the power consumption of CAM without any performance penalty. In the proposed CAM design, only the matched words can discharge the match line from 1 to 0. Similar to the match line segmentation methods, our design also divides a CAM word into two segments. The most distinct features of the proposed decoupled match line scheme are summarized as follows. (1) In our design, because all the pull-down paths are disconnected from the ground during the precharge phase, it is unnecessary to discharge all the bitlines to prevent the unexpected short-circuit power consumption. Thus, the power dissipated in bitline switching activities can be effectively reduced. (2) By using segmentation method, our design can largely reduce the match line switching activities, and thus the power consumption of match lines. (3) Because we decouple all CAM cells from the match line, the match line is lightweight that accelerates the discharge speed. This ensures our design has the same search performance as the traditional NOR-type CAM. (4) Because we provide a level restore path on the match line, our design has the immunity from the false match incurred by the possible race condition problem.

The proposed CAM design was fabricated with the TSMC 0.18 $\mu$ m technology. With the size of 128×32, the measurement results show that if a CAM word is divided into 4 and 28 bits, without any performance loss our design can improve the search energy efficiency up to roughly 89% compared to the traditional NOR-type CAM design, and the total area overhead is less than 6.1%.

The rest of this paper is organized as follows. Section 2 reviews the CAM organization and the previous work on CAM power reduction. Section 3 describes the circuitry developed for the *decoupled match line* scheme in detail. Besides the discussions on the importance issues, the comparison between our design and the related work is also

This work was supported by the National Science Council of Taiwan under grant No. NSC95-2221-E-005-049.



Fig. 1. A typical CAM cell. (a) XOR-type. (b) XNOR-type.

provided. Next, the measurement results are given in Section 4, and Section 5 offers some brief conclusions.

# 2. Content Addressable Memory (CAM)

The core of content addressable memory is an array of CAM cells. As shown in Fig. 1, a typical CAM cell consists of store and compare units. The store unit is usually implemented as the traditional 6T SRAM cell. The compare unit needs two NMOS transistors to perform the comparison between the stored and search data. Besides the store and compare units, a pull-down transistor X, which is gate-controlled by the comparison result, is necessary to connect/disconnect the match line (ML) to/from the ground. Depending on the different applications, the compare unit can be implemented as XOR or XNOR functions, shown in Fig. 1(a) and Fig. 1(b), respectively. Note that both XOR and XNOR are implemented as pass-transistor logic (PTL) for minimizing the area cost. In the XOR-type CAM cell, if the stored data is equal to the search data, then the pull-down transistor X would be turned off to prevent the match line from being discharged to 0. In contrast, in the XNOR-type CAM cell if the store data is equal to the search data, then the pull-down transistor X would be turned on to discharge the match line to 0.

# 2.1 NOR-type CAM

Fig. 2 shows the traditional NOR-type CAM design, in which the CAM cell is XOR-type. All match lines are initially precharged to high. For a CAM word, because the pull-down transistors of each CAM cell are arranged in NOR type, the match line would be discharged if one or more cells are mismatched. Only when all cells are matched, i.e., the search data is identical to the stored data, the match line can retain logic high as in the precharge phase. Because the pull-down path is very short, in case of a mismatch the match line would be discharged to 0 quickly. Thus, the NOR-type CAM can provide the best performance in searching operation.

Note that the match line of the mismatched word has to be precharged to high before the next search. From Fig. 2, the pull-down transistors arranged in NOR type are beneficial for search performance, but they contribute a lot of drain capacitances to the match line. That results in more power dissipated in match line switching. Because in many applications most of the CAM words are mismatched, a large number of match line switches would consume a significant dynamic power. For example, in the CAM tag used in the TLB



Fig. 2. The traditional NOR-type CAM design.





or cache memory, at most one word is matched on each lookup, which implies that almost all the match lines would be discharged to 0, and then be charged to high. Consequently, the NOR-type CAM is power inefficient, although it can provide the best performance.

#### 2.2 NAND-type CAM

In contrast to the NOR-type CAM, an alternative NANDtype CAM is developed to reduce the power dissipated in search operation. Fig. 3 shows the schematic of the traditional NAND-type CAM, in which the CAM cell is implemented as XNOR-type instead of XOR-type. Compared to Fig. 2, besides the CAM cell, the pull-down transistors of each CAM cell are arranged in NAND type.

The match line would be initially precharged to high, and discharged to 0 only when all CAM cells are matched, i.e., the search data is identical to the stored data. Because the load capacitance of match line is small and only one match line is discharged to 0 during a search, the power consumption is minimal. However, the pull-down path is too long, such that the match line discharge is very slow in case of a match. Thus, the NAND-type CAM trades the performance degradation for a large power saving.

# 2.3 Related Work

There are many previous researches on CAM power reduction. Because our design would divide a CAM word into two segments, we only focus on the work related to the word



Fig. 4. The CAM word structure of the decoupled match line scheme.

segmentation. In [4], Zukowski et al. introduced a selective precharge technique to reduce the match line power consumption by breaking a CAM word into two stages. A small subset of CAM cells can be used to do a precalculation, and the results are used to do a conditional (selective) precharge. As indicated in [4], separating 7 out of 128 bits would reduce the CAM power consumption by roughly 85% with a modest delay penalty.

A similar CAM word structure was proposed in [5]. Besides the segmented match line, they also introduce a new CAM cell with single bitline. The single bitline design requires only one heavy loading bitline, and prevents the frequent switching. Therefore, their method can further reduce the power consumption of CAM, but the performance degradation is still inevitable.

An adaptive serial-parallel CAM [6] is another low power CAM structure, which can operate either in parallel or in serial mode. In serial mode the energy consumption is almost a quarter of the conventional parallel CAM, but the cycle time is 25% slower than the original CAM. In parallel mode, the energy consumption is still 33% better than the conventional parallel CAM without any performance penalty.

# 3. Low Power CAM Design Using Decoupled Match Line Scheme

The key idea behind our design is to combine the performance advantage of NOR-type CAM with the power efficiency of NAND-type CAM. As shown in Fig. 4, we divide a CAM word into two segments, i.e., SEG\_1 and SEG\_2, and the necessary control circuitry. In the SEG\_1, the CAM cell is implemented as XNOR-type and their pull-down transistors are arranged in the NAND type, denoted as NAND-type block in Fig. 4. The NAND-type block is connected to the ground only when all the CAM cells of SEG\_1 are matched. In contrast to SEG\_1, we use the XOR-type CAM cell to implement the SEG\_2, and their pull-down transistors

are placed in the NOR type, denoted as NOR-type block in Fig. 4. The NOR-type block is disconnected from the ground only when all the CAM cells of SEG\_2 are matched.

#### 3.1 Search Operation

Similar to the traditional CAM, in our design there are two phases during a search. They are *precharge* and *match evaluation* phases, respectively. In the precharge phase, all the match lines are first precharged to high, and then in the match evaluation phase only the matched words would change the logic level of the corresponding match line from 1 to 0.

## Precharge Phase

In this phase, the control signal PRE is low. Thus, the match line (ML) is initially precharged to high. Because the pull-down path T1 and T2 are disconnected by NI and N2 transistors, respectively, both M1 and M2 nodes are precharged to high via PI and P2. Due to no paths to the ground, it is unnecessary to discharge all the bitlines to 0 to prevent the unexpected short-circuit during the precharge phase. Compared to the traditional CAM implementation, therefore, our design is more efficient in bitline power saving. In addition, in our design the match lines are precharged unconditionally. It is different from other segmentation techniques [4] [5], in which the match lines are precharged conditionally that would result in a performance penalty.

#### Match Evaluation Phase

After the precharge phase, the control signal PRE is asserted high and the search data have to be loaded on the bitlines to start the matching process. This phase is called *match evaluation phase*. Because we divide a CAM word into two segments, i.e., SEG\_1 and SEG\_2 as shown in Fig. 4, depending on the match results of each segment there are four possible cases in the match evaluation phase. It is a real match only when both the SEG\_1 and SEG\_2 are matched. These cases are described in detail as follows.



Fig. 5. The HSPICE waveform for each case.

#### Case 1: SEG\_1 is mismatched & SEG\_2 is mismatched/matched

Because SEG\_1 is a mismatch, in the NAND-type block at least one NMOS transistor is turned off that disconnects the pull-down path T1 from the ground. Therefore, node M1 retains high that turns off the tail transistor N2 to disconnect the pull-down path T2. This implies that no matter whether SEG\_2 is a match or mismatch, node M2 is still high to turn on N3. Because the path T1 is disconnected from the ground, the match line ML would maintain logic high as in the precharge phase. Fig. 5 shows the *HSPICE* waveform for each case, in which the lengths of SEG\_1 and SEG\_2 are assumed to be 4 and 28, respectively.

# Case 2: SEG\_1 is matched & SEG\_2 is mismatched

Because SEG\_1 is a match, in the NAND-type block all NMOS transistors are turned on that connects the path T1 to ground. Therefore, node M1 is discharged to 0 that turns on the tail transistor N2. If SEG\_2 is a mismatch, then in the NOR-type block at least one NMOS is turned on that connects the pull-down path T2 to the ground. Thus, node M2 is discharged to 0 to turn off transistor N3 to prevent the match line from discharging to 0 through path T1. From Fig. 5, it can be observed that our design still works well in this case.

# Case 3: SEG\_1 is matched & SEG\_2 is matched

Similar to the case 2, in this case the **M1** node is also discharged to 0 to turn on the tail transistor N2. Because SEG\_2 is a match, in the NOR-type block all NMOS transistors are turned off that disconnects the pull-down path **T2** from the ground. Thus, the **M2** node still retains logic high as in the precharge phase. That turns off *P4* and turns on *N3* to discharge the match line to 0 through the pull-down path **T1**. That indicates a real match. Fig. 5 shows the correct result.

# 3.2 Implementation Issues

Depending on the application, user can adjust the length of SEG\_1. If the length of a CAM word is n bits and the length of SEG\_1 is x bits, then the length of SEG\_2 would be n-x bits. In the SEG\_1, because all the pull-down transistors are arranged in serial mode (i.e., NAND-type block), and they are on the critical path to discharge the match line, the length of SEG\_1 is a powerful lever on the performance and power efficiency in our design.



Fig. 6. An example of the charge sharing problem incurred by large SEG 1.

#### SEG 1 Length vs. Race Condition

From Fig. 4, we note that the speed of **M1** discharge depends on the length of SEG\_1. This implies that there is a possible *race condition* problem in case 2, i.e., SEG\_1 is matched & SEG\_2 is mismatched. (a) If the **M1** discharge is fast enough, then the tail transistor N2 would be turned on quickly to discharge **M2**, such that N3 transistor is turned off quickly to prevent the match line from discharging. Therefore, the logic high level of match line can be retained correctly. (b) In the other case, if the **M1** discharge is too slow to prolong the on time of N3 transistor, then the match line would be discharged unexpectedly. It is a false match.

To prevent the incorrect match incurred by the race condition, we add a PMOS transistor, P4, to provide the level-restore capability. Once the M2 node is discharged to 0, regardless of discharge speed, P4 transistor would be turned on to supply the lost charge. Consequently, our design provides the immunity from the potential race condition problem. Its effect can be realized from Fig. 5, where there is a small pulse in case 2. The lost charge would be supplied quickly.

#### **SEG** 1 Length vs. Charge Sharing

If the length of SEG\_1 is too long, the *charge sharing* problem would possibly occur when SEG\_1 is mismatched and SEG\_2 is matched. As shown in Fig. 6, the worst case is that all the pull-down transistors are turned on but the most left one. In this case, the charge of **M1** node would be shared among the intermediate nodes,  $i_0 \sim i_5$ , such that the voltage level of **M1** node is decreased. Because SEG\_2 is matched, *N3* is turned on to discharge the match line. If the voltage level of match line is too low, then it results in a false match.

Fig. 7 shows the voltage level of match line under the consideration to the worst charge sharing problem for various SEG\_1 lengths, in which the load capacitance of match line is assumed to be 4fF. In this simulation, because the threshold voltage of PMOS transistor is -0.438V in the TSMC 0.18 $\mu$ m model, the charge sharing problem would result in a false match when the voltage level of match line is lower than 1.8V-0.438V=1.362V, as the dash line shown in Fig. 7. From this figure we conclude that if the length of SEG\_1 is larger than 4, our design has a possible false match. Therefore, the



Fig. 7. The voltage level of match line under the consideration to the worst charge sharing problem for various SEG\_1 lengths.



Fig. 8. The probability of M2 discharge for various SEG\_1 lengths.

length of SEG\_1 is constrained within 4 bits throughout this paper.

#### ■ SEG\_1 Length vs. Power Saving

As described above, short SEG\_1 can prevent the charge sharing problem, but it increases the probability of **M1** discharge. Suppose, for example, that the length of SEG\_1 is one bit. For a random pattern, the probability of **M1** discharge would be 50% on average, i.e., the probability of tail transistor *N2* turned on is also 50%. Because there are *n*-1 pull-down transistors in the NOR-type block, the probability of **T2** path connected to the ground would increase largely. It results in a significant power dissipated in the discharge of the **M2** node with large drain capacitances. Ideally, the probability of **T1** path conducting, i.e., p(T1 conducting), and the probability of **T2** path conducting, i.e., p(T2 conducting), as shown in the following equation:

```
p(\mathbf{T1 conducting}) \times p(\mathbf{T2 conducting}) = (1/2)^{x} \times (1 - (1/2)^{n - x}) = (1/2)^{x} - (1/2)^{n}
```

, where *n* and *x* are the lengths of the entire word and SEG\_1, respectively. In this equation, we assume that the match probability is 1/2 for each CAM cell. In the SEG\_2, because all the pull-down transistors are arranged in the NOR type, the **T2** path is disconnected only when they are all turned off. Thus, p(T2 conducting) is equal to  $(1-(1/2)^{n\cdot x})$ . Fig. 8 shows the probability of **M2** discharge for various SEG\_1 lengths, in which *n*=32 is assumed. Clearly, the probability of **M2** discharge decreases sharply as the length of SEG\_1 is



Fig. 9. The match delay for various SEG\_1 lengths.

increased. This implies that the search operation would consume more power when we decrease the length of SEG\_1.

# 4. Results

For a solid result, in this paper we use TSMC  $0.18\mu$ m CMOS technology to implement the proposed CAM with decoupled match line scheme, and a conventional NOR-type CAM used for comparison. Both they are with size of  $128\times32$ , i.e., 128 words by 32 bits. The core was broken into four blocks for both the performance and power efficiency.

# Performance

In this paper, the metric used to evaluate the CAM performance is the match delay, which is defined as the elapsed time from signal PRE is asserted high to the match line discharged to 0 in case of a match. Fig. 9 shows the match delay for SEG 1 length from 1 to 6 bits. Due to no segmentation, the match delay of the conventional CAM design is fixed at 0.641ns, as the dash line shown in this figure. As revealed in Section 3.2, if the length of SEG 1 is larger than 4, then there is a possible false match due to charge sharing problem. From this figure, we summarize the most important aspects as follows. (1) In our design the match delay increases with the length of SEG 1. It is expected, because the match line discharge relies on the M1 discharge that connects T1 and T2 paths to the ground, and M1 discharge speed depends on the number of transistors in the NAND-type block. Therefore, the length of SEG 1 is critical to the match performance of our design.

(2) One interesting observation from this result is that when the length of SEG\_1 is less than 4 bits, the match delay of our design is even shorter than that of the conventional CAM. This is because our design decouples all CAM cells from the match line, such that the match line is lightweight. Once the path **T1** is conducting, it can discharge the lightweight match line quickly. Although the small NANDtype block would degrade the match performance slightly, the lightweight match line can compensate for the loss of match performance. From Fig. 9, when SEG\_1 is 4 bits, both designs have almost the same match delay.

# Power and Energy

Fig. 10 shows the power consumption during a search for various SEG\_1 lengths. As analyzed in Section 3.2, clearly, the search power consumption can be reduced sharply as we increase the length of SEG\_1. When the SEG\_1 length is 6



Fig. 10. The search power consumption for various SEG\_1 lengths.

bits, the search power consumption is only about 0.28mW. Compared to the conventional NOR-type CAM, whose search power consumption is fixed at 3.04mW, our design can reduce roughly 90% of the search power consumption. However, increasing SEG\_1 length would result in the performance degradation. It is a tradeoff between power and performance.

For a fair comparison, energy is a suitable metric, which is the product of the match delay (performance) and search power (power). Table 1 lists all the detailed measurement results. From Table 1, our design can improve energy efficiency up to 89% as the SEG\_1 length is 6 bits. The improvement difference between SEG\_1=6 and SEG\_1=4 is marginal. However, if the SEG\_1 length is larger than 4 bits, due to charge sharing a possible false match does exist in our design. For a reliable system, when the SEG\_1 length is 4 bits, our design can improve the energy efficiency by roughly 88.7% compared to the conventional NOR-type CAM.

#### Area Cost

Compared to the conventional NOR-type CAM, our design costs 8 additional transistors which all come from the control circuitry. For a CAM word with 32 bits, the layout size of the conventional NOR-type and our design are 7.54 $\mu$ m×131.21 $\mu$ m and 7.54 $\mu$ m×139.21 $\mu$ m. Note that the height of the proposed CAM is purposely retained the same as the height of the conventional NOR-type CAM, i.e., 7.54 $\mu$ m, such that both designs have the same power dissipated in the bitline switching. The area overhead is roughly 6.1%. Because the CAM words are part of the entire CAM system, the total CAM area overhead is less than 6.1%.

Table 1. The detailed measurement results.

SEG_1 Length	Match Delay (ns)	Search Power (mW)	Energy (pJ)	Energy Improvement
1	0.411	1.322	0.543	72.1%
2	0.492	0.652	0.321	83.5%
3	0.572	0.423	0.242	87.6%
4	0.648	0.340	0.221	88.7%
5	0.710	0.298	0.212	89.1%
6	0.768	0.279	0.214	89.0%
Conv.	0.641	3.040	1.949	

# 5. Conclusions

In this paper, we propose a *decoupled match line* scheme to reduce the power consumption of CAM. Compared to the conventional NOR-type CAM, the contribution of this paper is that we not only decouple all the CAM cells from match line to reduce search power consumption but also maintain the high search performance.

#### References

- [1] T. Juan, T. Lang and J. Navarro, "Reducing TLB Power Requirements," in Proc. of International Symposium on Low Power Electronics and Design, 1997, pp. 196-201.
- [2] H. Miyatake, M. Tanaka and Y. Mori, "A Design for High-Speed Low-Power CMOS Fully Parallel Content-Addressable Memory Macros," IEEE Journal of Solid-State Circuits, Vol. 36, June 2001, pp. 956-968.
- [3] I. Arsovski and A. Sheikholeslami, "A Mismatch-Dependent Power Allocation Technique for Match-Line Sensing in Content-Addressable Memories," IEEE Journal of Solid-State Circuits, Vol. 38, Nov. 2003, pp. 1958-1966.
- [4] C. A. Zukowski and S. Y. Wang, "Use of Selective Precharge for Low-Power Content-Addressable Memories," in Proc. of International Symposium on Circuits and Systems, 1997, pp. 1788-1791.
- [5] K. H. Cheng, C. H. Wei and S. Y. Jiang, "Static Divided Word Matching Line for Low-Power Content Addressable Memory Design," in Proc. of International Symposium on Circuits and Systems, 2004, pp. 629-632.
- [6] A. Effhymiou and J. D. Garside, "An Adaptive Serial-Parallel CAM Architecture for Low-Power Cache Block," in Proc. of International Symposium on Low Power Electronics and Design, 2002, pp. 136-141.