AN ADAPTIVE ECC SCHEME FOR RUNTIME WRITE FAILURE SUPPRESSION OF STT-RAM CACHE

by

Xue Wang

B. S. in Electronic Engineering, Tsinghua University, 2009

Submitted to the Graduate Faculty of

the Swanson School of Engineering in partial fulfillment

of the requirements for the degree of

Master of Science

University of Pittsburgh

2015

UNIVERSITY OF PITTSBURGH

SWANSON SCHOOL OF ENGINEERING

This thesis was presented

by

Xue Wang

It was defended on

November 10 2015

and approved by

Yiran Chen, PhD, Associate Professor, Department of Electrical and Computer Engineering

Hai Li, PhD, Associate Professor, Department of Electrical and Computer Engineering

Zhi-Hong Mao, PhD, Associate Professor, Department of Electrical and Computer Engineering

Thesis Advisor: Yiran Chen, PhD, Associate Professor, Department of Electrical and Computer Engineering

AN ADAPTIVE ECC SCHEME FOR RUNTIME WRITE FAILURE SUPPRESSION OF STT-RAM CACHE

Xue Wang, M.S.

University of Pittsburgh, 2015

Spin-transfer torque random access memory (STT-RAM) features many attractive characteristics, including near-zero standby power, nanosecond access time, small footprint, etc. These properties make STT-RAM perfectly suitable for the applications that are subject to limited power and area budgets, i.e., on-chip cache. Write reliability is one of the major challenges in design of STT-RAM caches. To ensure design quality, error correction code (ECC) scheme is usually adopted in STT-RAM caches. However, it incurs significant hardware overhead. In observance of the dynamic error correcting requirements, in this work, we propose an adaptive ECC scheme to suppress the runtime write failures of STT-RAM cache with minimized hardware cost, in which the cache is partitioned into regions protected by different ECCs. The error rate of a data is speculated on-the-fly and the data is allocated to a partition that provides the needed error correcting capability. Moreover, to accommodate the time-varying error correcting requirements of runtime data, the thresholds that determine data's destination cache partition will be adaptively adjusted. Our experimental results show that compared to conventional ECC schemes, our scheme can save up to 80.2% ECC bit overhead with slightly degraded write reliability of the STT-RAM cache. Moreover, the detailed analysis shows that through simultaneous optimization in cache access patterns and reducing STT cell programming workload, our method outperforms conventional ECC design in power and energy consumptions.

Keywords: NVM, STT-RAM, cache, write error, ECC.

TABLE OF CONTENTS

PREFACE . . . viii

1.0 INTRODUCTION . . . 1

1.1 Emerging Memory Technologies . . . 1

1.2 Reliability Challenges . . . 2

1.3 Related works . . . 2

1.4 Summary of proposed method . . . 3

2.0 PRELIMINARY . . . 4

2.1 STT-RAM basics . . . 4

2.2 STT-RAM Write Error Model . . . 6

3.0 TECHNICALMOTIVATIONS . . . 9

3.1 Observation and facts . . . 9

3.2 Spatial and Temporal Variability's of ECC Requirement . . . 10

4.0 METHOD . . . 12

4.1 Basic Concept of Sliding Basket . . . 12

4.2 ECC Requirement Driven Cache Partition . . . 16

4.3 Dynamic Sliding . . . 18

5.0 EXPERIMENT . . . 19

5.1 Simulation Platform . . . 19

5.2 Scheme configuration. . . . 20

6.0 RESULT . . . 21

6.1 ECC Overheads . . . 21

6.2 Reliability Enhancement . . . 22

6.3 Bit Matching Effect in Sliding Basket 23
6.4 Performance and Energy 25
6.5 Hamming Based replacement policy 26
6.5.1 Influence of FLIP reduction 27
6.5.2 Influence on missrate 28
6.5.3 Hamming First Replacement 29
7.0 CONCLUSIONS 31
BIBLIOGRAPHY 32

LIST OF TABLES

1 Bit overheads of different ECC schemes. 17
2 System configurations. 19
3 Scheme configurations. 20

LIST OF FIGURES

1 (a) STT-RAM basics; (b) The currents applied to write to STT-RAM. 4

2 The tradeoff between MTJ switching time and transistor size for a fixed block error rate of 3.64×10^{-12} at $22nm$. The block size is 512-bit. 5

3 Distribution of the FLIP of 512-bit cache blocks in some representative benchmarks. 7

4 BLER and bit overhead when applying different ECC schemes. 8

5 Spatial and temporal variabilities of ECC strength requirement in a L2 STT-RAM cache. 10

6 Cache set compositions of proposed scheme. 12

7 Handle blocks when write hits. 13

8 Flow chart of writing data into a block. 15

9 Comparison of Mean Time Between Failure. 22

10 Illustration of bit matching for reducing FLIP. 23

11 Comparison of average FLIP. 24

12 Comparison of Instruction Per Cycle (IPC). 24

13 Comparison of energy consumption. 25

14 Comparison of average FLIP. 27

15 Comparison of average missrate. 28

16 Comparison of average missrate. 30

PREFACE

Over the past two years of studying at University of Pittsburgh, I have learned a tremendous amount. Here, I want to thank my supervisor, Professor Yiran Chen and Professor Hai Li, for all their support and guidance during my Master Program. Also, I am very thankful to Professor Zhihong Mao for kindly consenting to be my committee. I appreciate all of the extremely helpful suggestion discussionS. I thank all the lab members in Dr. Chen and Dr. Li's research group as well as all my friends here in University of Pittsburgh.

1.0 INTRODUCTION

1.1 EMERGING MEMORY TECHNOLOGIES

To meet the exponentially increasing data processing capability in modern computer system, the emerging memory technologies (EMTs) Phase-change memory (PCM), Resistive random-access memory (RRAM), like spin-transfer torque random access memory (STT-RAM) are proposed to as candidates to replace conventional memories, such as SRAM and DRAM.

PCM leverages a phase change material such as Ge2Sb2Te5 (GST), which can switch states with significantly different resistance [1]. RRAM is a type of non-volatile memory, which stores information by changing the resistance across the dielectric solid-state material. Chua *et all.* proposed memristor, which is a 2-terminal non-volatile memory cell storing information by switching resistance [3]. STT-RAM stores information as the resistance states of magnetic tunneling junction (MTJ) devices. It offers nanosecond access time comparable to SRAM, high integration density close to DRAM, and zero standby power like Flash memory. STT-RAM is considered a promising candidate to replace SRAM, by providing solutions to the severe challenges like the prominent leakage power consumption and the significant degradation in device reliability.

Among these EMTs, STT-RAM features many attractive characteristics, including near-zero standby power, nanosecond access time, small footprint, *etc.* These properties make STT-RAM perfectly suitable for the applications that are subject to limited power and area budgets, like on-chip cache [4].

1.2 RELIABILITY CHALLENGES

A major challenge in STT-RAM design is the access reliability issue, *e.g.*, high write error rate. A write error in an STT-RAM cell occurs as the write pulse is removed before the completion of the resistance switching of its data storage device – *magnetic tunneling junction* (MTJ) [5]. The parametric variability of MOS transistor and MTJ [6] as well as the thermal-induced randomness in MTJ switching process [7] induce a large variation of MTJ switching property, making the write reliability control very difficult. Process variations, for example, induce deviations of the electrical characteristics of MOS transistors and MTJs from their nominal values, leading to read and write errors of memory [8, 9].

1.3 RELATED WORKS

Error correction code (ECC) has been widely adopted in STT-RAM to ensure the access reliability. The high error rate in STT-RAM designs (which indeed relies on the storage patterns) demands for a strong ECC scheme. However, a strong ECC usually implies large bit overhead and long data encoding and decoding time, which is against the requirement of the high-density and delay-sensitive on-chip cache applications. Popular ECC schemes, such as SEC-DED [10], BCH [11, 12, 13], i.e. Hi-ECC [14], ECP[15], etc., are designed by assuming the error rates of the stored data with different values are always identical. These ECC schemes, however, are not suitable for STT-RAM designs because they are generally designed for the worst-case that rarely happens and often neglect the unique structure and asymmetric characteristics of STT-RAM and lose in efficiency. For example,considering the asymmetric erroneous probabilities of programming 1 and 0, Wu *et al.* developed an asymmetric ECC scheme that strengthens the protection for '0→1' switchings [16]. Nonetheless, these ECC schemes were designed for the worst-case scenario that rarely happens in real applications, *e.g.*, ignoring the variability of data error rate across different memory blocks and program segments. Such pessimistic designs require to reserve considerable design margin, introducing significant hardware cost and performance overheads.

1.4 SUMMARY OF PROPOSED METHOD

In light of the dynamic error correcting requirement in the application of STT-RAM caches, in this work, we propose *Sliding Basket* – an adaptive ECC scheme to suppress runtime write failures. In Sliding Basket, the cache is partitioned into regions (*baskets*) protected by different ECCs. The error rate of a data is speculated on-the-fly and the data is allocated to a partition that provides the needed error correcting capability. Moreover, to accommodate the time-varying error correcting requirements of runtime data, the thresholds that determine data's destination cache partition will be adaptively adjusted. Our experimental results show that compared to conventional ECC scheme, Sliding Basket can save up to 80.2% ECC bit overhead with slightly degraded write reliability of the STT-RAM cache. Moreover, the detailed analysis shows that through simultaneous optimization in cache access patterns and reducing STT cell programming workload, Sliding Basket outperforms conventional ECC design in power and energy consumptions.

2.0 PRELIMINARY

2.1 STT-RAM BASICS

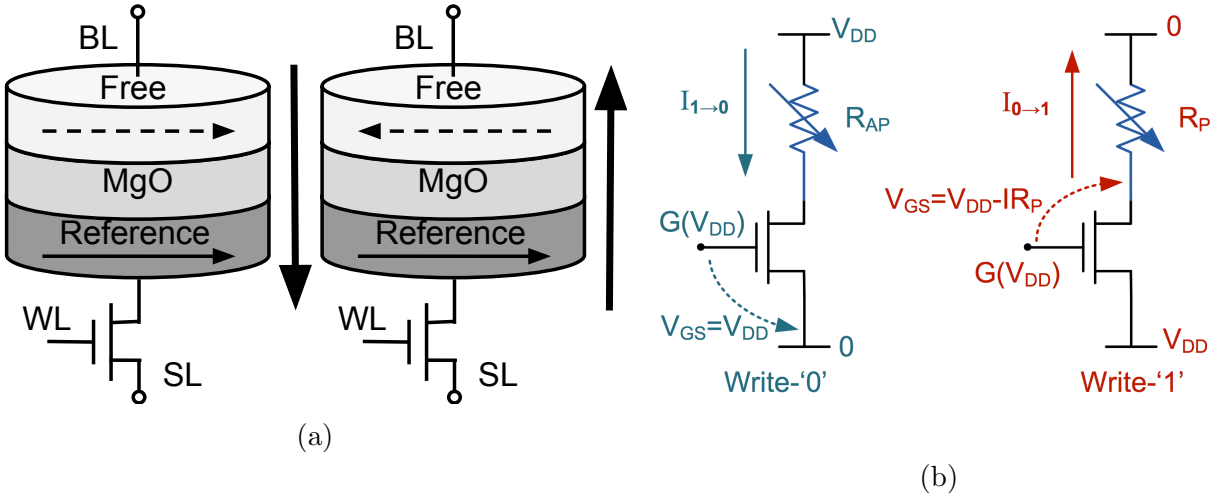

Figure 1: (a) STT-RAM basics; (b) The currents applied to write to STT-RAM.

In an STT-RAM cell, data is represented by the resistance state of its MTJ device. As illustrated in Figure. 1(a), an MTJ is composed of two ferromagnetic layers (*i.e.*, reference and free layers) and an oxide insulator. The magnetization direction of the reference layer is fixed while that of the free layer is switchable under a polarized write current [5]. When the magnetization directions of the two layers are in parallel (anti-parallel), the MTJ is in low (high) resistance state, representing logic '0' ('1'). The write current through the MTJ is supplied by a NMOS transistor. Such an STT-RAM cell is denoted as '1T1J' structure.

As shown in Figure. 1(b), a current from V_{DD} to ground is programming the two ferromagnetic layers in the MTJ into parallel, resulting a logic '0' state. On the other hand, if the write current is applied from the opposite direction, the two ferromagnetic layers in the MTJ are programmed into anti-parallel, resulting a logic '1' state.

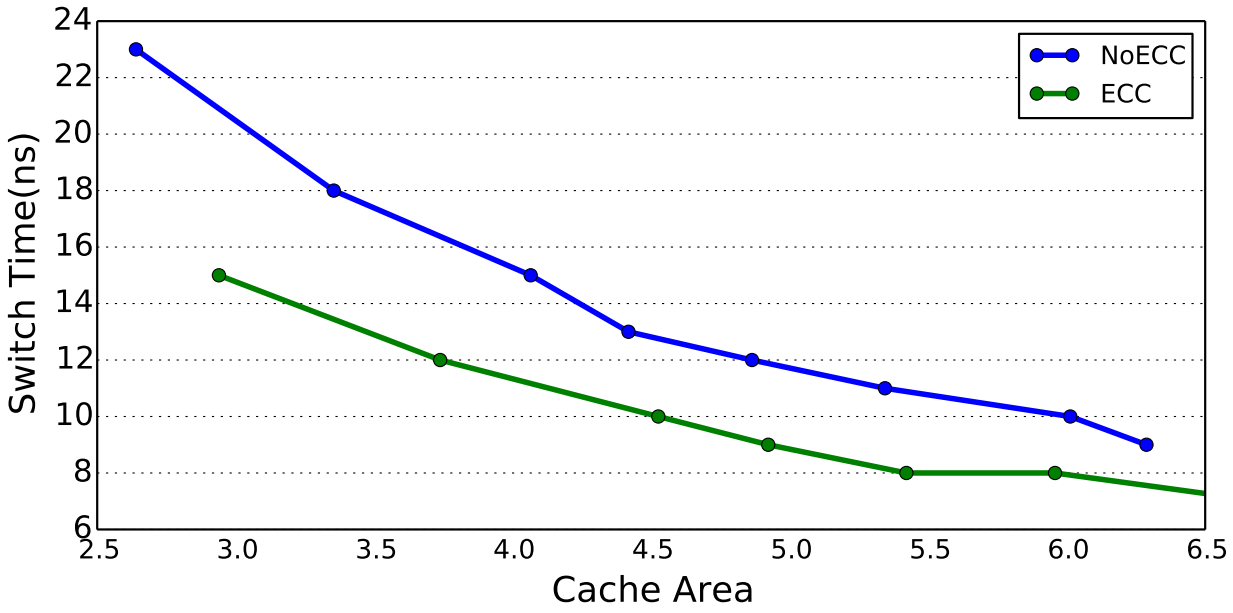

Figure 2: The tradeoff between MTJ switching time and transistor size for a fixed block error rate of 3.64×10^{-12} at $22nm$. The block size is 512-bit.

In an STT-RAM cell, a write error happens when the write current is removed before the MTJ resistance switching process completes [17][18]. Due to the relatively lower driving strength of the NMOS transistor and the higher switching current of the MTJ, the '0→1' switching demonstrates a much higher probability of write errors than the '1→0' switching [17].

We note that due to the intrinsic randomness in process variations and thermal fluctuations, write errors of STT-RAM cannot be completely eliminated. Raising the amplitude of write current (by increasing NMOS transistor size) and prolonging write pulse width (*i.e.*, MTJ switching time) can reduce the write error rate. Figure. 2 compares the STT-RAM switch time and cache area for 8MB L2 STT-RAM cache without ECC ("NoECC") and with (72, 64) single bit ECC ("ECC") while maintaining the same block write error rate of

3.64×10^{-12} for 8MB L2 STT-RAM cache. These results are simulated with NVSIM [19] with 22nm technology node. The "ECC" schemes reduced the area from 6.01mm to 4.52mm by 24.8% compared to the "NoECC" cache at the same switch time of 10ns.

Thus we conclude that a tradeoff exists between the applied write current amplitude and write pulse width under a fixed write error rate target [6]. However, the circuit-level solutions introduce extra area overhead and/or performance penalty, not even mentioning the increase in write energy consumption.

2.2 STT-RAM WRITE ERROR MODEL

As aforementioned, the write reliability of an STT-RAM cell is mainly determined by the failure rate of '0→1' switching rather than that of '1→0' switching. The write failure rate of an STT-RAM cache block, *i.e.*, the probability of having no more than t erroneous bits, can be approximated by [18]:

$$P(n_e \leq t) = \sum_{i=0}^{t} C_{\mathrm{FLIP}}^{i} \mathrm{BER}^{i} (1 - \mathrm{BER})^{\mathrm{FLIP}-i}, \tag{2.1}$$

where FLIP denotes the number of the bits flipping from 0 to 1, which includes the flipping in ECC bits if ECC is applied; n_e is the number of error bits; and BER is the write error rate of a single bit.

STT-RAM is often adopted in latency-sensitive scenarios, so the application of ECC is usually constrained to SECDED (*single error correction, double error detection*) schemes. A cache block can be divided into several data segments, each of which is separately protected by a set of ECC check bits. In such a case, the BLER (*BLock Error Rate*) of the cache block can be approximated by:

$$P_{err,blk}^{ecc} = 1 - \prod_{i=1}^{SegNum} P_i(n_e \leq 1). \tag{2.2}$$

Here, $SegNum$ represents the number of the divided segments in the cache block and BLER denotes the probability of having uncorrectable error(s) in the cache block.

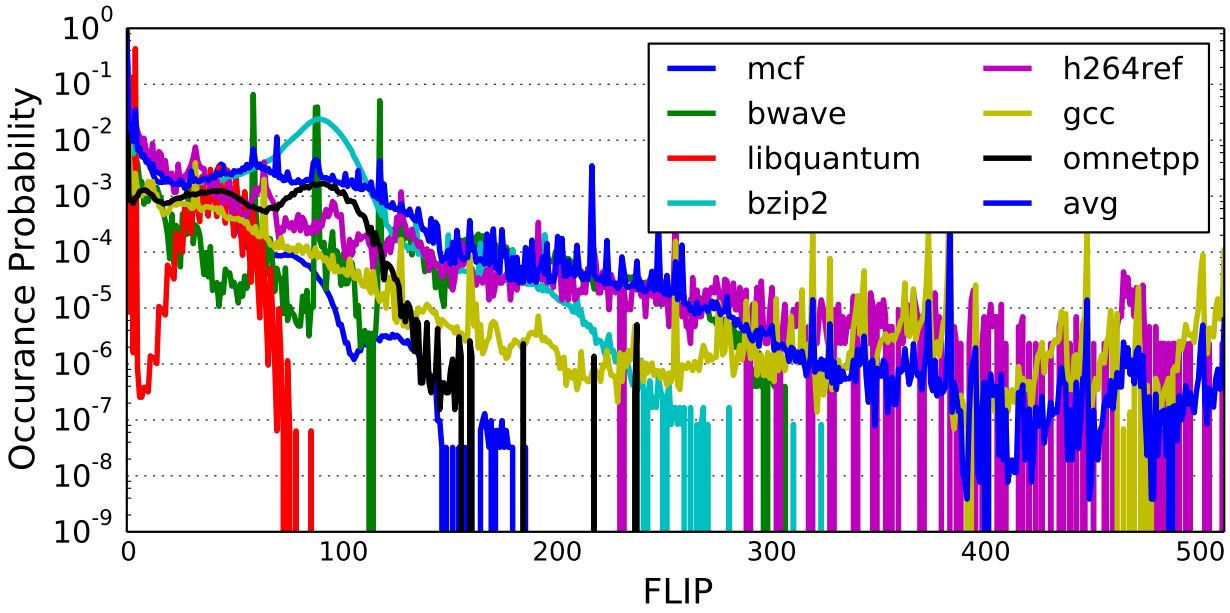

Figure 3: Distribution of the FLIP of 512-bit cache blocks in some representative benchmarks.

Fig. 3(a) summarizes the distribution of average FLIP of the 512-bit cache blocks in some representative SPEC2006 benchmarks. As the value of FLIP increases, the occurrence probability reduces rapidly. In other words, it is very rare that many bits are flipping from 0 to 1 simultaneously. Here we assume the read-before-write scheme [20] is applied, *i.e.*, write operation is performed only when the data being written is different from the one stored in the targeted memory cell.

Figure. 4 shows how the BLER varies with FLIP under several ECC schemes. Here we use (w, k) to denote an ECC with w codeword length and k ECC bits, with which the data bits are protected with 1 bit error correction and 2 bits error detection capability. Utilizing a strong ECC, *i.e.*, (72, 64), effectively suppresses the write failures but increases the ECC bit overhead. When the target failure rate, i.e. BLER, is set, the normal attempt in the application of STT-RAM caches is to select ECC based on the maximum possible FLIP. However this approach induces significant design pessimism in the case of smaller FLIPs.

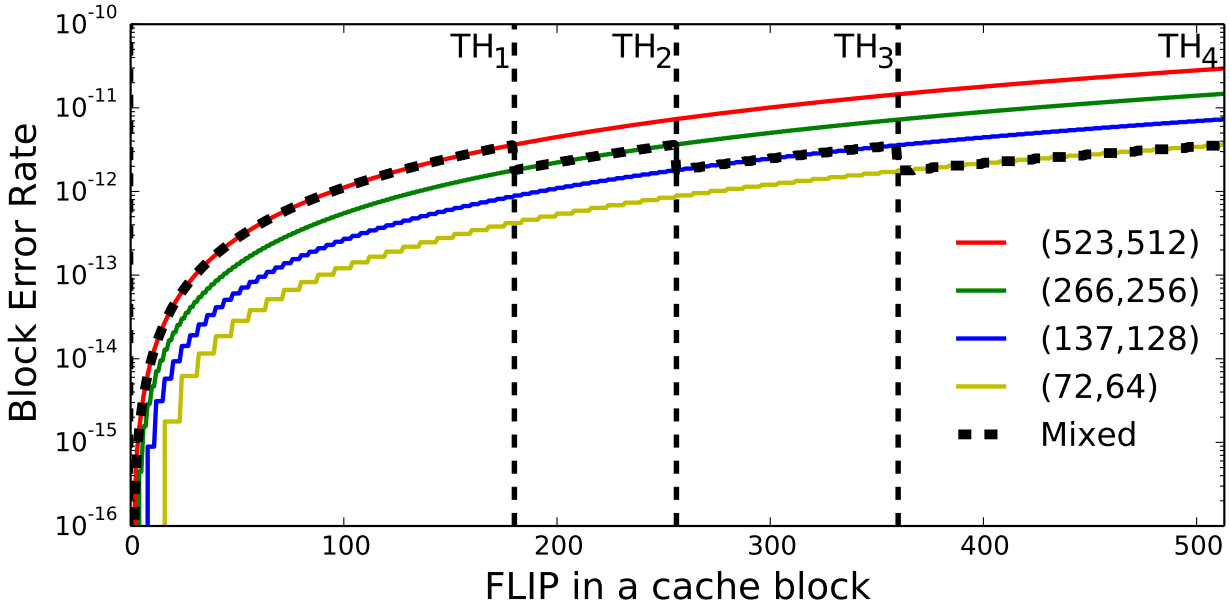

Figure 4: BLER and bit overhead when applying different ECC schemes.

3.0 TECHNICALMOTIVATIONS

3.1 OBSERVATION AND FACTS

Applying ECC can enhance the write reliability of STT-RAM, which can be further translated to memory cell arFigure.ea reduction and/or write performance improvement as shown in Figure. 2. In conventional designs, all the cache blocks are equipped with the same ECC scheme. The STT-RAM write reliability can be roughly measured by the worst BLER, which corresponds to the maximum FLIP under the protection of the strongest ECC available to the design. As shown in Figure. 2 the maximum BLER is 3.64×10^{-12}, which occurs when FLIP = 512 under the protection of (72, 64). If the FLIP of the cache block is smaller than the maximum FLIP, we may apply a weaker ECC to protect the block while still ensure the BLER not higher than the maximum BLER. In general, such a scheme can be represented as a pair of $[ECC_1, ..., ECC_n]$ and $[TH_1, ..., TH_n]$, where $ECC_i, i = 1, ..., n$, are different ECCs from the weakest to the strongest; and $TH_j, j = 1, ..., n$, are the maximum FLIP that can be protected by ECC_j. Here $TH_n = S$, which is the cache block size. Here we assume the error correcting strength of ECC_i increases monotonically when i increases. For illustration purpose, Figure. 4 compares the failure rates when applying a mixed ECC scheme of [(523, 512), (266, 256), (137, 128), (72, 64)] and [180, 256, 360, 512]. The failure rates corresponding to different FLIPs are bounded by the maximum BLER at FLIP = 512. The ECC bit overhead of every single ECC scheme is also presented in the figure for comparison purpose.

The above observation motivates us to propose *Sliding Basket* scheme that aims at reducing the ECC bit overhead by fully utilizing the error correcting strength of ECC schemes to satisfy the ever-changing needs of the cache block data.

3.2 SPATIAL AND TEMPORAL VARIABILITY'S OF ECC REQUIREMENT

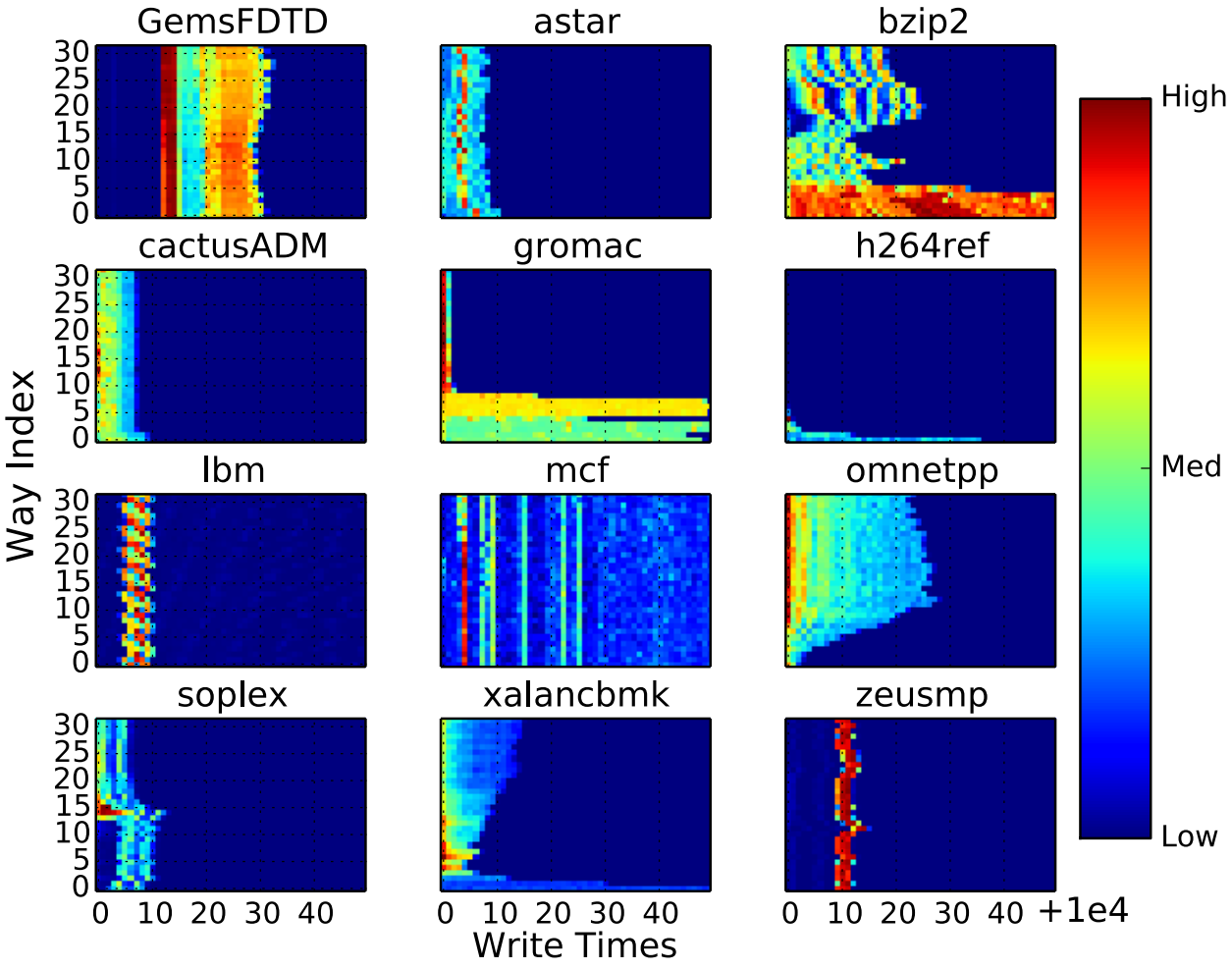

Figure 5: Spatial and temporal variabilities of ECC strength requirement in a L2 STT-RAM cache.

Figure. 5 shows the required ECC strength of cache block that is mainly decided by the FLIP of data, varies across the 32 ways in a L2 STT-RAM cache. This property is referred to as the *spatial variability* of the ECC requirement. Also, the required ECC strength of every cache block changes over time, representing the *temporal variability*. In the following three subsections, we will show: 1) how to leverage the spatial variability of the ECC requirement to reduce the ECC bit overhead by configuring the ECC schemes of cache blocks at way level

and guiding data to be stored at the corresponding cache block with proper ECC protection; and 2) managing the temporal variability of the ECC requirement through dynamically adjusting the thresholds to guide the data allocation during program execution.

4.0 METHOD

4.1 BASIC CONCEPT OF SLIDING BASKET

For simplicity, we assume two ECC schemes with different error correcting strengths available to our design, *i.e.*, strong and weak. All the cache blocks belonging to the same set are partitioned into two groups: Group H and Group L, which belong to the different cache ways. The cache blocks in Group H and Group L are protected by the strong ECC and the weak ECC, respectively.

Figure. 6 illustrates the overhead of ECC bits for Group H and Group L, in which the gray area denotes the ECC coding bits for every block respectively.

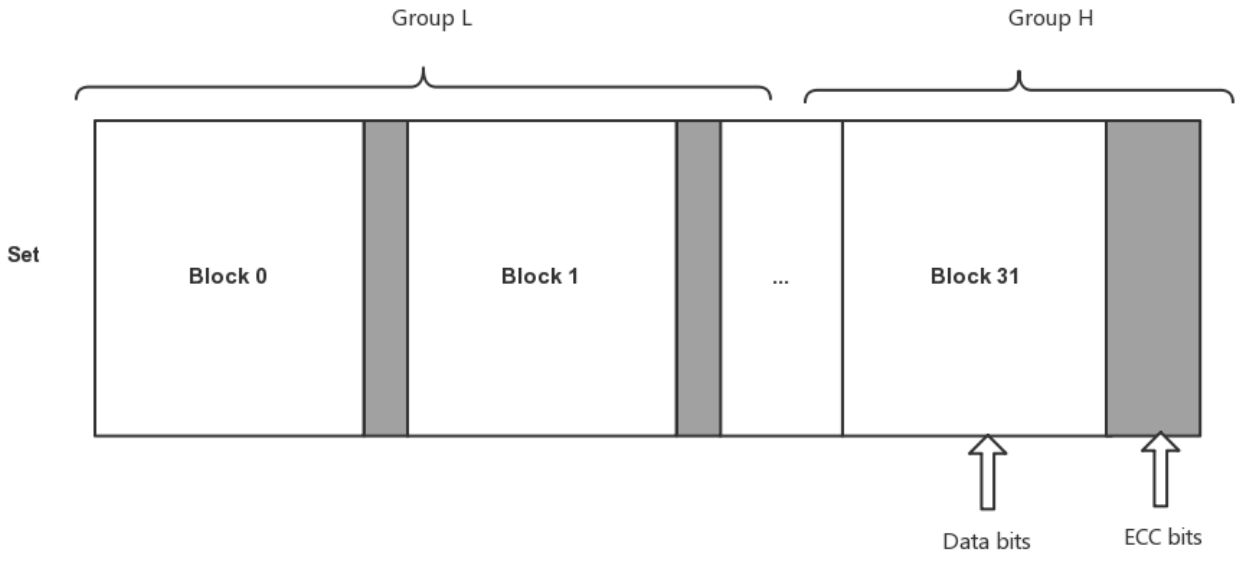

Figure 6: Cache set compositions of proposed scheme.

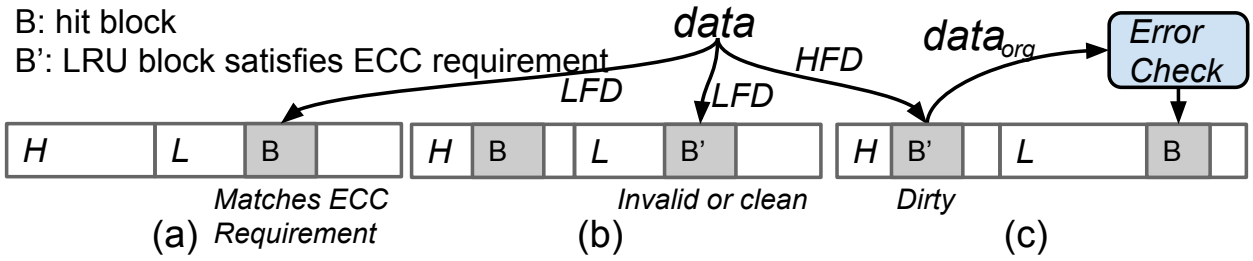

Figure 7: Handle blocks when write hits.

When a data is scheduled to be written into the STT-RAM cache, its required ECC strength can be first estimated based on its FLIP, as discussed in Section 2.2. Because only two ECC schemes are available in the presented example, only one FLIP threshold is needed to categorize the data into two types, say, *High Flip Data* (HFD) and *Low Flip Data* (LFD). Once the estimated required ECC strength is obtained, the following data allocation procedure will be applied to guarantee the reliability of the data:

In a write miss, the HFD and the LFD are placed into Group H and Group L, respectively, by following LRU policy within each group. In a write hit, if the hit Block B belongs to a group which has the ECC exactly matching the need of the data, B is updated with the new data directly, as shown in Figure. 7(a). Otherwise, a block in the group with the ECC exactly matching the need of the data, say B', will be selected to store the data based on LRU policy. If Block B' is invalid or clean or belongs to the group with an ECC strength weaker than that of the group containing Block B (*e.g.*, in Group L as shown in Figure. 7(b)), the original data in Block B' is evicted and replaced with the new data. If Block B' belongs to a group with an ECC strength stronger (*e.g.*, Group H in Figure. 7(c)) than that of the group containing Block B, the data in Block B' is first moved to Block B and then the new data is written to Block B'. An error check and correction will be performed on the original data of Block B' before writing it to Block B. The whole procedure is depicted in Figure. 8.

procedure DATA ALLOCATION FLOW OF SLIDING BASKET(1)

Select Group based on ECC requirement

$data \leftarrow$ *incoming data*

$B \leftarrow$ *hit block if hit*

$B' \leftarrow$ *lru block from selected group*

$data_{org} \leftarrow$ *data original in B'*

Write miss:

Find LRU block *B'* in selected group

Write *data* to *B'*

Write hit:

if *B matches ECC requirement* **then**

Update *data* to *B*, see Figure. 4(a)

else

Find LRU Block *B'* in selected group

if *B' is clean or* $\in$ *weaker ECC group* **then**

Evict *data* in *B*, see Figure. 4(b)

else

Perform error check for $data_{org}$ in *B'*

Move $data_{org}$ to *B*, see Figure. 4(c)

Write *data* to *B'*

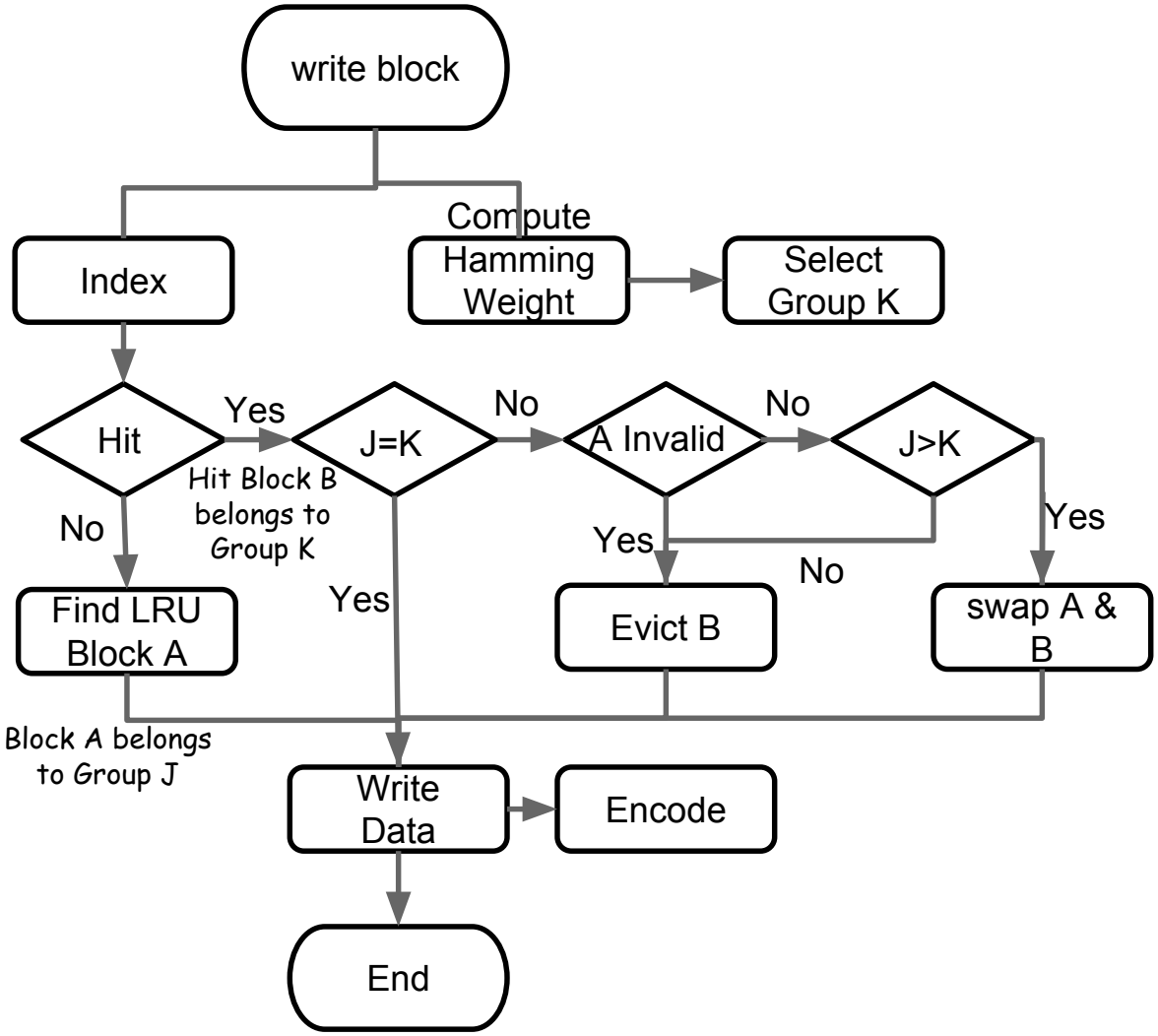

Figure 8: Flow chart of writing data into a block.

The rationales of the above flow are the follows: as the majority of the data have low FLIPs during program execution, the capacity of the groups with weak ECC, *i.e.*, Group L, shall be larger than that of the groups with strong ECC, *i.e.*, Group H. Hence, if the block to be replaced (*i.e.*, B', which is the least recently used block in the group) is in Group L, replacing Block B' with the new data imposes minimum impact on system performance due to the large capacity of Group L (see Figure. 7(b)). However, if the block to be replaced (*i.e.*, B') is in Group H, directly overwriting the original $data_{org}$ in Block B' may greatly affect system performance, because miss rate may be caused by frequent access to Group H due

to its limited capacity and small associativity. Thus for both consideration in performance and reliability, we need to move the original data in B' to the originally hit Block B (see Figure. 7(c)).

There exists a small probability that the weak ECC associated with Block B in Group L may not be able to provide sufficient protection to the data moved from Block B'. Adding an error check and correction step in moving the data from Block B' into Block B can effectively reduce the possibility of writing an erroneous data. In fact, as Block B' is identified in Group H based on LRU policy, the possibility of continuing to use the data moving from Block B' is already very low.

4.2 ECC REQUIREMENT DRIVEN CACHE PARTITION

By assuming the data can be perfectly allocated to the cache group with exactly matched ECC strength, at a specific time, an ideal partition of the Sliding Basket cache is:

$$N_i = \begin{cases} \lceil \sum_{j=TH_{i-1}+1}^{S} P_j \times A \rceil, i = n \\ \lceil \sum_{j=TH_{i-1}+1}^{S} P_j \times A \rceil - \sum_{j=i+1}^{n} N_j, i < n \end{cases} . \tag{4.1}$$

Here S denotes the size of cache block; A is the associativity of the cache; P_j is the occurrence probability when FLIP $= j$; N_i is the number of ways protected by the ECC_i; $TH_0 = 0$. The case of $i = n$, in which the strongest ECC is needed, is handled separately to make sure to cover the worst-case. The above "ideal" partition guarantees to provide the least necessary ECC protection to the data with minimum ECC bit overheads, which can be calculated by:

$$O_{ECC}^{verhead} = \sum_{i=1}^{n} N_i \times O_{ECC_i}^{verhead}. \tag{4.2}$$

Here $O_{ECC_i}^{verhead}$ denotes the bit overhead of ECC_i to cover one block in the corresponding data group.

TABLE 1 summarizes the bit overheads of the four concerned ECC schemes and their FLIP thresholds for the 512-bit cache block. The average occurrence probability of the FLIPs

between two adjacent thresholds over all the benchmarks are also depicted in the 5^{th} line of the table. The corresponding "ideal" cache partition of this FLIP distribution is shown in the bottom line of the table. Based on Eq. (4.1), the data with the required ECCs of (266, 256) and (137, 128) are covered by (72, 64). Thus, no cache ways are assigned to these two ECC schemes in the cache partition in TABLE 1. This partition incurs a 2.47% ECC bit overhead as calculated by Eq. (4.2).

To accommodate the temporal variability of the ECC requirement, we introduce some margins to the cache partition by raising the number of cache ways assigned to the strong ECC from the average case. More details on the design tradeoff between the cache partition and system performance/reliability will be discussed in Section 6.

Table 1: Bit overheads of different ECC schemes.

Data bits	512	256	128	64
ECC bits	11	10	9	8
Overhead	2.15%	3.91%	7.03%	12.5%
TH_i	180	256	360	512
$\sum_{j=TH_{i-1}+1}^{TH_i} P_j$	99.16%	0.74%	0.06%	0.04%
N_i^{avg}	31	0	0	1

Obtaining FLIP of a cache block needs to compare both the incoming data and the data originally stored in the cache block. In the Sliding Basket flow presented in Algorithm **??**, however, deciding the destination cache block also requires to know FLIP. To solve this "chicken-and-egg" problem, we propose using HW (*Hamming Weight*), which denotes the number of '1's in the data, to approximate the FLIP of the destination cache block to guide the operation of Sliding Basket. In [6], Bi *et al.* proved that there is a strong correlation between HW and FLIP of the data. Indeed, HW is a good pessimistic approximation (upper bound) of FLIP. Hence, the basic Sliding Basket scheme presented in Section 4.1 can be safely modified by using HW (instead of FLIP) as the threshold to allocate the data into different cache groups.

4.3 DYNAMIC SLIDING

Obviously a fixed cache partition (even with margin to the group with strong ECC) cannot perfectly accommodate the temporal variability of the ECC requirement. Besides the possible write failures caused by temporarily inadequate capacity of the groups with the needed ECC strength, the inflexibility of the fixed cache partition could also harm system performance by over- or under-utilizing some particular cache groups. To solve this issue, we propose to dynamically adjust the thresholds that are used to allocate the data to overcome the shortcoming of the fixed cache partition: a miss rate counter is added to monitor the usage pattern of each group. The TH of a group will be reduced (raised) when a significant miss rate increase (decrease) is detected.

5.0 EXPERIMENT

5.1 SIMULATION PLATFORM

TABLE 2 summarizes the baseline system configuration used in our experiments. The timing and energy parameters of the STT-RAM cache are extracted from NVSim [19] at $22nm$ technology. Here the baseline access latency of the STT-RAM cache has already taken into account the ECC encoding and decoding latency of (72, 64), which is set to be 1 clock cycle according to [21]. To be conservative, we assume that the ECC encoding and decoding of (523, 512) take one more cycle to finish than that of (72, 64). All the simulations are performed on GEM5 simulator [22] with 64-bit Alpha instruction set.

Table 2: System configurations.

Core	4GHz, OOO 8-width, 64-energy LSQ, 64-entry instruction queue, 192-enrty ROB
Caches	32KB L1I, 2-way, 2-cycle, 64B line
	32 KB L1D, 4-way, 2-cycle, 64B line, write back
	8MB STT-RAM L2 cache, single-core, 32-way, 16 banks, 1 port, 64B line, write back, 20 MSHRs Read energy 0.241nJ, Read delay 12 Cycles, Write energy 0.882nJ, Write delay 30 Cycles
Main memory	DDR3-1600, 2-channel, open-page, FR-FCFS [23]

Table 3: Scheme configurations.

Scheme Configuration	ECC bits*	Overhead	Cache Area
No ECC	0	0	6.02 mm^2
SECDED(72, 64)	64	12.5%	4.53 mm^2
SECDED(523, 512)	11	2.15%	4.14 mm^2
Opt 1: 26(523, 512)+6(72, 64)	20.93	4.09%	4.22 mm^2
Opt 2: 27(523, 512)+5(72, 64)	19.28	3.76%	4.20 mm^2
Opt 3: 28(523, 512)+4(72, 64)	17.62	3.44%	4.19 mm^2
Opt 4: 31(523, 512)+1(72, 64)	12.65	2.47%	4.16 mm^2

*ECC bits denote the average ECC bits number for a 64B line in a 32-way cache set.

5.2 SCHEME CONFIGURATION.

We select 15 representative benchmarks from SPEC CPU 2006 suite [24] in our evaluations. SimPoints [25] is used to extract a single simulation point of benchmarks. Each simulation point contains 500 million instructions. The caches and memory system are warmed up with 100 million instructions before jumping into a simulation point.

To evaluate the impact of cache partitioning on the system performance and reliability, we include not only the cache partition of the average case ("Opt 4" in TABLE 3) but also some partitions with larger Group H ("Opt 1-3") in our simulations. Since the occurrence probability of the data with high FLIP/HW is very low (refer TABLE. 1), only two ECC schemes – (523, 512) and (72, 64), are adopted in Sliding Basket scheme. A threshold TH_1 is used to guide the data allocation between the two groups and set to 180 at beginning of execution. The miss rate counter is checked every one million cycles and TH_1 changes by 10 if the variation of the miss rate is larger than 5%. This optimal setup is selected based on the exhausted simulations, which are found to be the optimal configurations based on our experimental observation.

6.0 RESULT

6.1 ECC OVERHEADS

For comparison purpose, the bit overheads of SECDED (523, 512) and (72, 64) are also included in TABLE 3, representing the best and the worst design overheads of conventional ECC schemes, respectively. Also, the areas of the L2 caches with and without ECC protections extracted from NVSim [19] are listed in TABLE 3 including different cache partitions of Sliding Basket. The BER of STT-RAM cells is set to 1.5×10^{-8}, which leads to a BLER of 3.64×10^{-12} when (72, 64) is applied. The write pulse width is set to $10ns$. To achieve the same BLER without applying ECC, the size of NMOS transistor in the STT-RAM cell must be increased to supply a larger write current, resulting in 32.9% more cache area compared to that of (72, 64) (*i.e.*, $6.02mm^2$ vs. $4.53mm^2$).

The ECC bit overheads for Sliding Basket configurations "Opt 1-4" vary between 2.47% and 4.41%, which are only 17.8% to 32.7% of that (12.5%) of (72, 64). Note that the area estimations have included the contribution of ECC logic and other support circuits of these schemes.

6.2 RELIABILITY ENHANCEMENT

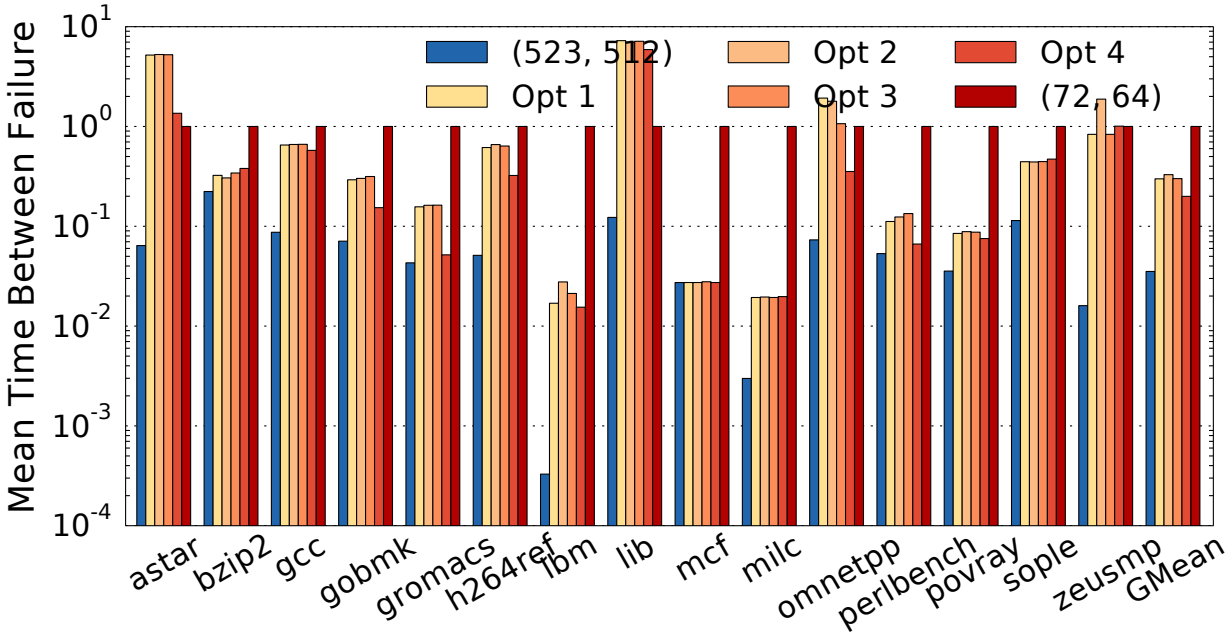

Figure 9: Comparison of Mean Time Between Failure.

Fig. 9 compares the *Mean Time Between Failure* (MTBF) of all the ECC schemes and cache configurations. A larger MTBF implies a better system reliability. On average, the MTBF achieved by Sliding Basket is 20.0% ("Opt 4") $\sim$ 32.9% ("Opt 2") of the one of (72, 64), that is, 5.67$\times$ $\sim$ 9.35$\times$ of that of (523, 512). All the data are normalized to the one of (72, 64). The results show that raising the size of Group H from the cache partition based on the average case ("Opt 4") can effectively enhance the system reliability, reaching the highest MTBF at "Opt 2". Continuing to increase the size of Group H, however, does not further enhance the system reliability.

Interestingly, we found that in some benchmarks, such as `astar`, `omnetpp`, `lib`, and `zeusmp`, *the highest MTBFs achieved by Sliding Basket are even* 1.13$\times$ $\sim$ 6.81$\times$ *better than that of the strongest ECC scheme (72, 64)!* A detailed analysis on this interesting observation shall be given in the next subsection.

6.3 BIT MATCHING EFFECT IN SLIDING BASKET

The principle of Sliding Basket is allocating the data with similar ECC requirement (indexed by Hamming Weight) to the group with the matched ECC strength. Besides, an important byproduct of Sliding Basket – *the potential reduction of* FLIP that naturally enhances the write reliability of STT-RAM cache – can be observed.

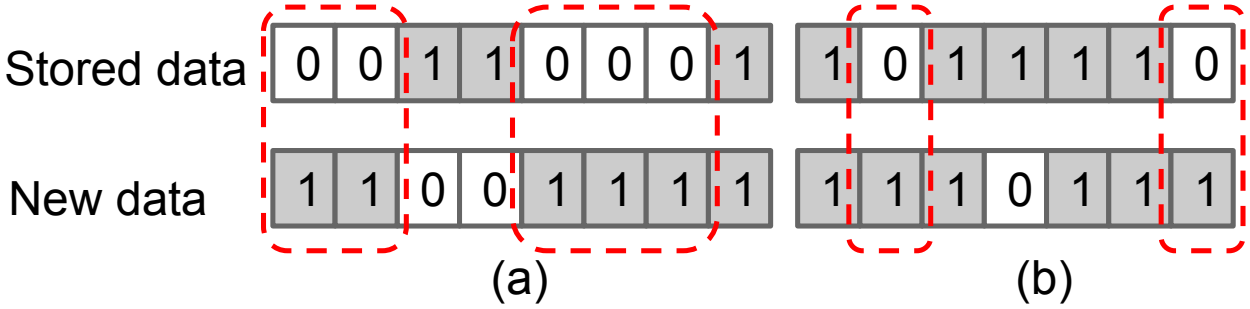

Figure 10: Illustration of bit matching for reducing FLIP.

Fig. 10 explains the reason for the FLIP reduction. In Sliding Basket, the data with similar HW (*i.e.*, the number of '1's) are allocated to the same group. Considering the locality of the cache data, such an operation potentially increases the probability for these '1's to show up at the same location of the new data and the stored data. Since read-before-write scheme is applied, the overlapped bits with the same value will not be updated during the write operation. In fact, in Group H, since the data are all with very high HW (more '1's), the overlapping rate of '1's between the new data and the stored data is even more prominent. Thus, compared to the data allocation in conventional cache design, a smaller average FLIP can be expected in Sliding Basket and a better write reliability may be achieved.

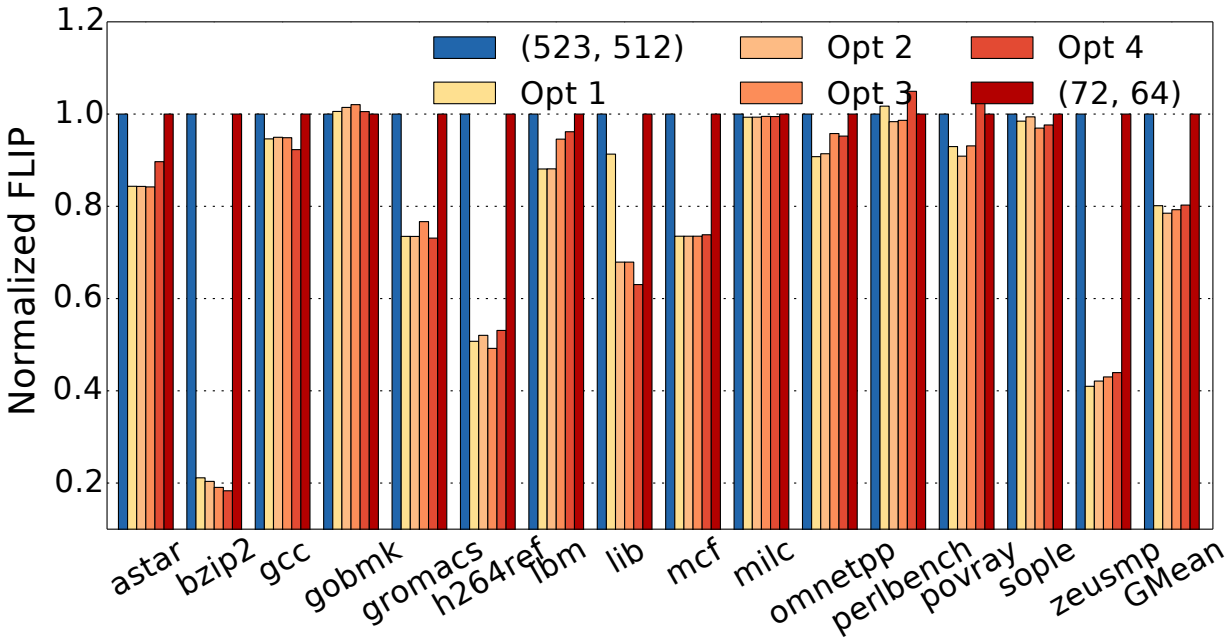

Figure 11: Comparison of average FLIP.

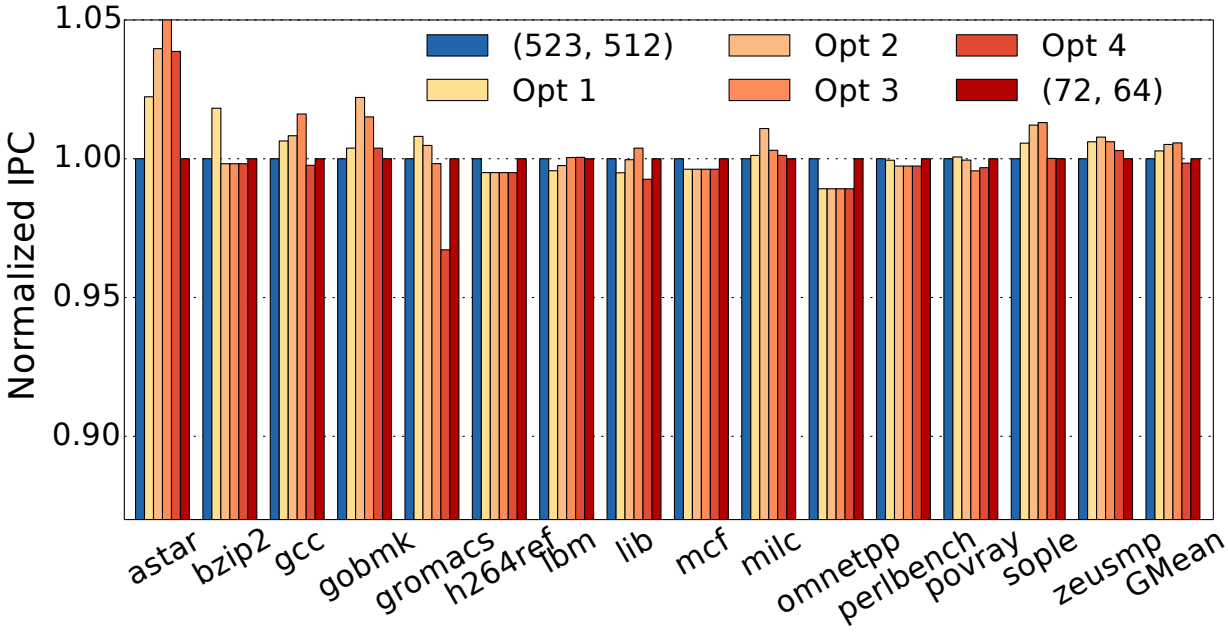

Figure 12: Comparison of Instruction Per Cycle (IPC).

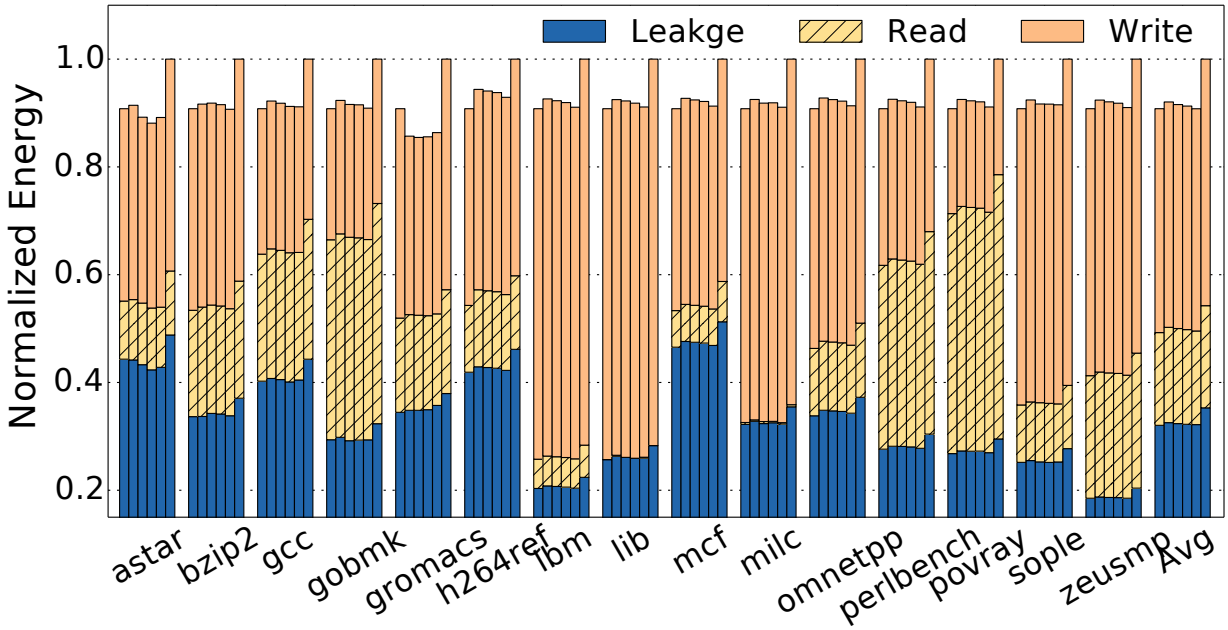

Figure 13: Comparison of energy consumption.

To verify our hypothesis, we compared the average FLIP of the cache data (not including the ECC bits) under different ECC schemes and cache partitions, as shown in Fig. 11. Compared to the level of (72, 64), Sliding Basket can reduce the FLIP up to 21.5% ("Opt 2") across all 15 benchmarks. The highest FLIP reduction (80.2%) is observed at `bzip2`.

6.4 PERFORMANCE AND ENERGY

Fig. 12 compares the performance of all the simulated ECC schemes and cache partitions. Our results show that Sliding Basket achieves a system performance generally comparable to the conventional ECC scheme across the 15 benchmarks. The largest performance degradation of Sliding Basket (merely 0.16%) occurs at `gromacs` with the configuration of "Opt 4" due to the increased miss rate in Group H. In fact, except for "Opt 4", the IPCs (*Instructions Per Cycle*) of "Opt 1-3" are all improved from the (72, 64) baseline by 0.28%, 0.51% and 0.57%, respectively, even the 1-cycle extra ECC encoding/decodeing latency has been

included in the simulations. Our detailed analysis shows that in these configurations, their L2 cache miss rates all decrease from the baseline. A possible reason is that the HW based data allocation flow improves the data eviction effectiveness by the enhancing the correlation of the new data and the data stored in the cache block.

The results in Fig. 12 also show that raising the capacity of Group H effectively improves the system performance and reaches the highest performance at "Opt 3". Continuing to increase the capacity of Group H, however, does not offer a better performance.

An energy consumption lower than the one of conventional ECC schemes is achieved in almost every Sliding Basket configurations in the 15 benchmarks, as shown in Fig. 13. Among all cache partitions, the largest average energy saving is obtained by "Opt 4" because of its largest capacity of (523, 512) among all partitions. For "Opt 4", the largest energy saving is achieved at `gromacs`, or 13.6% lower than that of (72, 64). Besides the slightly improved IPC, such a wide energy saving achieved by Sliding Basket is mainly due to the reduction of flipping bits caused by: 1) the bit matching effect discussed in Section 6.3; and 2) the smaller number of bits that actually needs to be written by applying weak ECC (and hence, less ECC bits) to the majority of the cache blocks.

6.5 HAMMING BASED REPLACEMENT POLICY

To further study the influence of hamming based replacement policy, the following experiments are performed. In the experiments, instead of considering temporal locality measured by LRU information, the correlation of data, measured by the hamming distance of old data and new data in a cache block, is also considered when replacing a block on a cache miss.

6.5.1 Influence of FLIP reduction

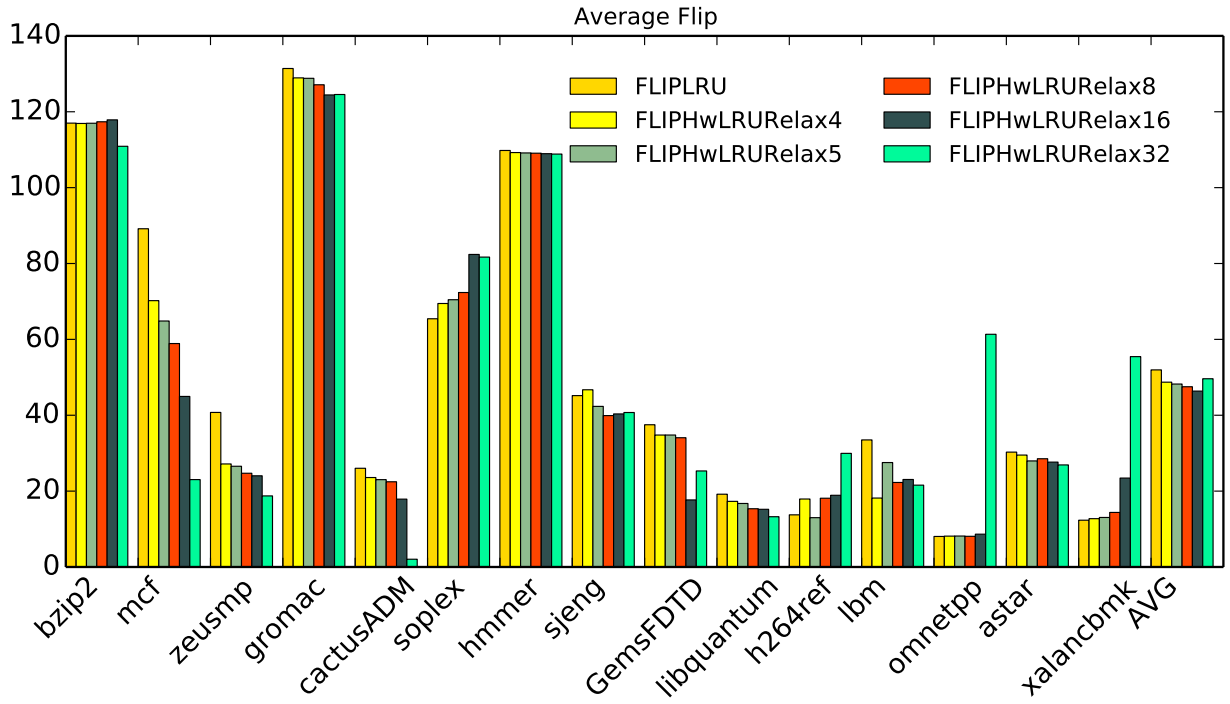

Figure 14: Comparison of average FLIP.

Figure. 14 shows the FLIP reduction when integrating hamming weight based replacement with LRU replacement policy. "FLIPLRU" indicates the average FLIP of LRU replacement, "FLIPHwLRURelax4" indicates the average FLIP when replacing the cache block with the minimal hamming distance for the first 4 least recent used blocks, assuming the LRU range is 4.

As shown in Figure. 14, FLIP can be reduced by replacing blocks with minimal hamming weighted or the least recent used blocks. By relaxing LRU from 4 to 32, that is, replacing the block with the minimal hamming weight among the first 4 to 32 LRU blocks, the FLIP can be reduced monotonically for most benchmarks, i.e. mcf, zeusmp, etc. Except for soplex, h264ref, FLIP increases when relaxing the LRU ranges.

6.5.2 Influence on missrate

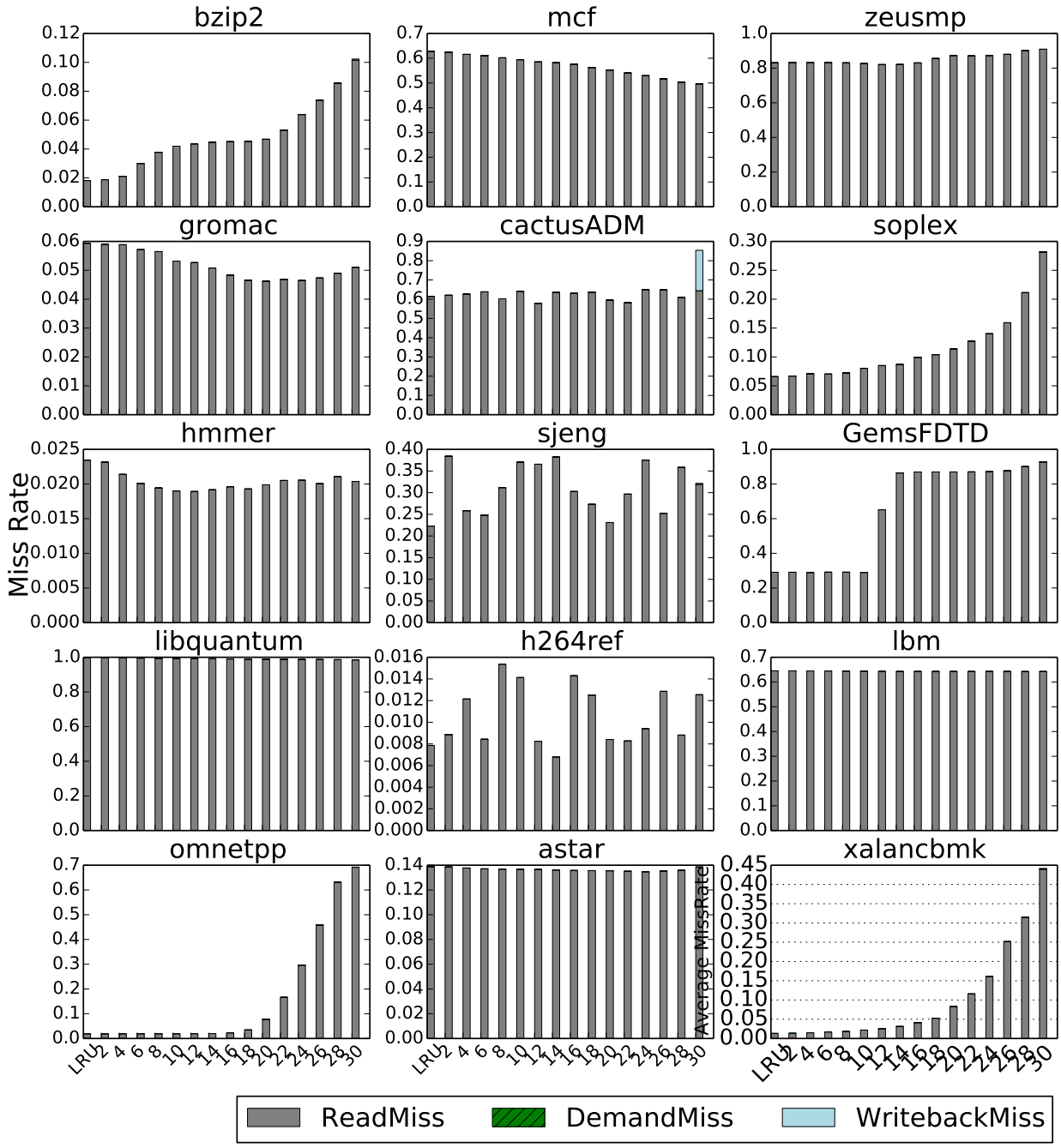

Figure 15: Comparison of average missrate.

To study the influence on performance, Figure. 15 plots how the miss rate varies with Hamming Weight and LRU mixed replace policy and difference LRU ranges. Figure. 15 compares

the miss rate of LRU replacement policy (the first bar) and LRU mixed with Hamming weight replacement, with LRU relax range from 2 to 32, (from the second bar to the last bar). We can see that relaxing the LRU to the first few least recent used block with minimal hamming weight does not bring too much performance loss. More interestingly, relaxing the LRU range can reduce the miss rage for some benchmarks like mcf, astar, libquantum, etc.

Such a results reveals the data correlation represented by the hamming weight of the old data and new data written to the same cache location(block), indicating that the LRU might not be the ideal replacement policy for some benchmarks and to maximize the temporal locality of cached data. Previous works by Belady, Sanchez, et all[26, 27] also proved that there exists replacement policy better than the widely accepted LRU replacement policy in terms of the temporal and spacial locality.

6.5.3 Hamming First Replacement

Figure. 16 illustrates the miss rate of another design of mixing LRU and Hamming weight based replacement. In this design, we sort the blocks within a set with its hamming weight with the new coming data. The replacement of a block is arbitrary by the sum of hamming weight ranking and LRU rankings.

As shown in Figure. 16, the miss rate degraded significantly when replacing with this scheme. This may be because the overweighted of hamming weight in the replacement harms the data temporal locality.

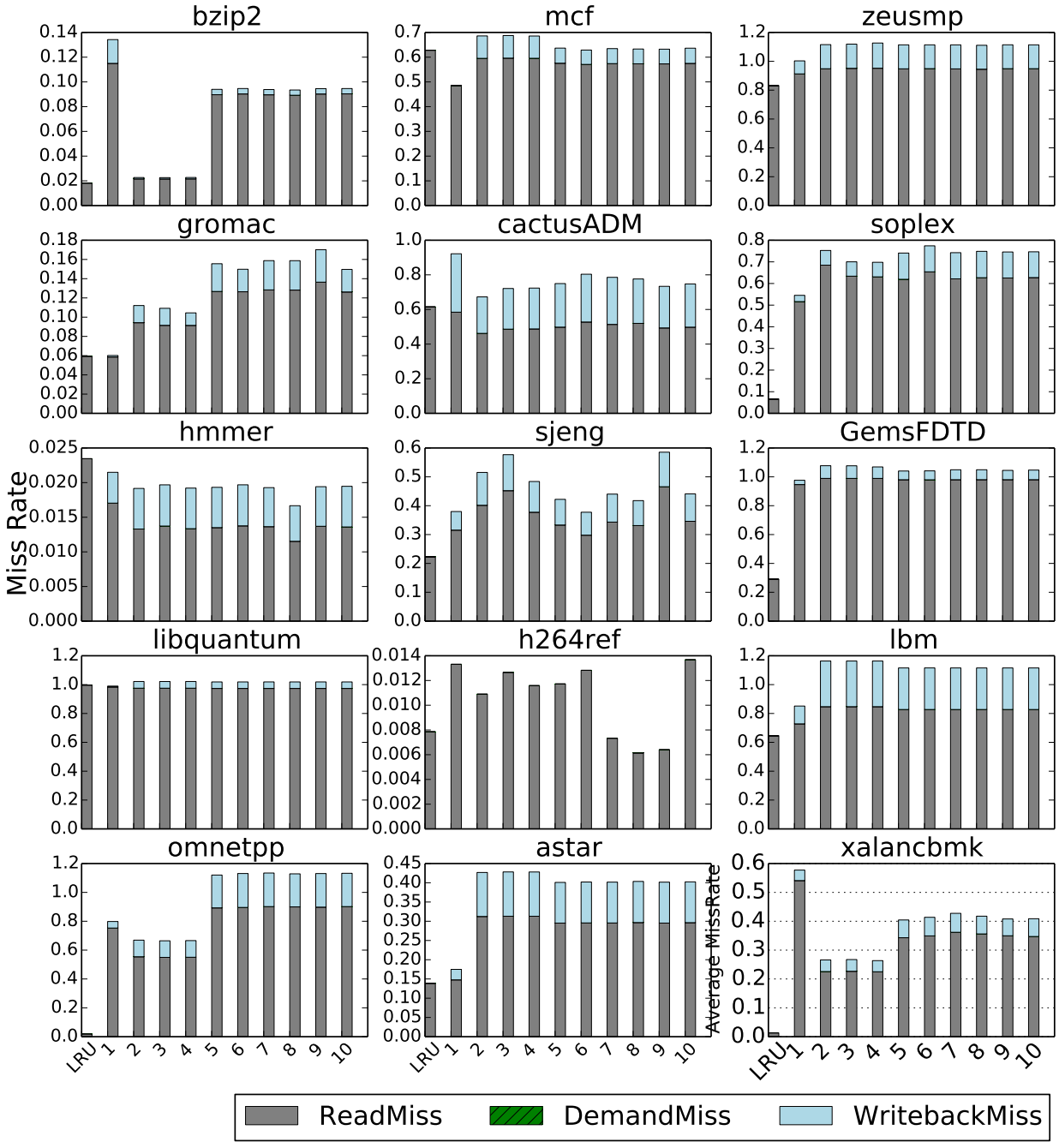

Figure 16: Comparison of average missrate.

7.0 CONCLUSIONS

In this work, we examine the dynamic needs of STT-RAM cache for ECC protections across different data blocks and program segments and propose Sliding Basket. It is an adaptive ECC scheme that can allocate every data to a cache group with the just needed ECC strength. As such, the associated hardware cost can be minimized. Our simulations show that Sliding Basket can save up to 80.2% ECC bit overheads with slightly degraded runtime reliability, compared to conventional ECC schemes. System performance and cache energy efficiency are also improved, benefiting from the enhanced data eviction effectiveness and the reduced bit flipping rate.

BIBLIOGRAPHY

[1] Y. Chen, C. Rettner, S. Raoux, G. Burr, S. Chen, R. Shelby, M. Salinga, W. Risk, T. Happ, G. McClelland, *et al.*, "Ultra-thin phase-change bridge memory device using gesb," in *International Electron Devices Meeting*, pp. 777–780, Citeseer, 2006.

[2] Y. N. Joglekar and S. J. Wolf, "The elusive memristor: properties of basic electrical circuits," *European Journal of Physics*, vol. 30, no. 4, p. 661, 2009.

[3] L. O. Chua, "Memristor-the missing circuit element," *Circuit Theory, IEEE Transactions on*, vol. 18, no. 5, pp. 507–519, 1971.

[4] A. Jog, A. K. Mishra, C. Xu, Y. Xie, V. Narayanan, R. Iyer, and C. R. Das, "Cache revive: architecting volatile stt-ram caches for enhanced performance in cmps," in *DAC*, 2012.

[5] M. Hosomi, H. Yamagishi, T. Yamamoto, K. Bessho, Y. Higo, K. Yamane, H. Yamada, M. Shoji, H. Hachino, C. Fukumoto, *et al.*, "A novel nonvolatile memory with spin torque transfer magnetization switching: Spin-ram," in *IEDM Technical Digest*, 2005.

[6] X. Bi, Z. Sun, H. Li, and W. Wu, "Probabilistic design methodology to improve run-time stability and performance of stt-ram caches," in *ICCAD*, 2012.

[7] X. Bi, H. Li, and J.-J. Kim, "Analysis and optimization of thermal effect on stt-ram based 3-d stacked cache design," in *VLSI*, 2012.

[8] Y. Zhang, L. Zhang, W. Wen, G. Sun, and Y. Chen, "Multi-level cell stt-ram: Is it realistic or just a dream?," in *Proceedings of the International Conference on Computer-Aided Design*, pp. 526–532, ACM, 2012.

[9] Y. Zhang, W. Wen, and Y. Chen, "The prospect of stt-ram scaling from readability perspective," *Magnetics, IEEE Transactions on*, vol. 48, no. 11, pp. 3035–3038, 2012.

[10] M.-Y. Hsiao, "A class of optimal minimum odd-weight-column sec-ded codes," *IBM Journal of Research and Development*, vol. 14, no. 4, pp. 395–401, 1970.

[11] C.-L. Chen and M. Hsiao, "Error-correcting codes for semiconductor memory applications: A state-of-the-art review," *IBM Journal of Research and Development*, vol. 28, no. 2, pp. 124–134, 1984.

[12] W. Xu, Y. Chen, X. Wang, and T. Zhang, "Improving stt mram storage density through smaller-than-worst-case transistor sizing," in *DAC*, 2009.

[13] T. K. Moon, "Error correction coding," *Mathematical Methods and Algorithms. Jhon Wiley and Son*, 2005.

[14] C. Wilkerson, A. R. Alameldeen, Z. Chishti, W. Wu, D. Somasekhar, and S.-l. Lu, "Reducing cache power with low-cost, multi-bit error-correcting codes," *ACM SIGARCH Computer Architecture News*, vol. 38, no. 3, pp. 83–93, 2010.

[15] S. Schechter, G. H. Loh, K. Straus, and D. Burger, "Use ecp, not ecc, for hard failures in resistive memories," in *SIGARCH*, 2010.

[16] W. Wen, M. Mao, X. Zhu, S. H. Kang, D. Wang, and Y. Chen, "Cd-ecc: content-dependent error correction codes for combating asymmetric nonvolatile memory operation errors," in *ICCAD*, 2013.

[17] Y. Zhang, X. Wang, Y. Li, A. K. Jones, and Y. Chen, "Asymmetry of mtj switching and its implication to stt-ram designs," in *DATE*, 2012.

[18] W. Wen, Y. Zhang, Y. Chen, Y. Wang, and Y. Xie, "Ps3-ram: a fast portable and scalable statistical stt-ram reliability analysis method," in *DAC*, 2012.

[19] X. Dong, C. Xu, Y. Xie, and N. P. Jouppi, "Nvsim: A circuit-level performance, energy, and area model for emerging nonvolatile memory," *TCAD*, 2012.

[20] K.-J. Lee, B.-H. Cho, W.-Y. Cho, S. Kang, B.-G. Choi, H.-R. Oh, C.-S. Lee, H.-J. Kim, J.-M. Park, Q. Wang, *et al.*, "A 90 nm 1.8 v 512 mb diode-switch pram with 266 mb/s read throughput," *IEEE Journal of Solid-State Circuits*, 2008.

[21] B. Zimmer and M. Zimmer, "Maximizing energy-efficiency through joint optimization of l1 write policy, sram design, and error protection," 2012.

[22] N. Binkert, B. Beckmann, G. Black, S. K. Reinhardt, A. Saidi, A. Basu, J. Hestness, D. R. Hower, T. Krishna, S. Sardashti, *et al.*, "The gem5 simulator," *SIGARCH*, 2011.

[23] S. Rixner, W. J. Dally, U. J. Kapasi, P. Mattson, and J. D. Owens, *Memory access scheduling*. 2000.

[24] J. L. Henning, "Spec cpu2006 benchmark descriptions," *SIGARCH*, 2006.

[25] E. Perelman, G. Hamerly, M. Van Biesbrouck, T. Sherwood, and B. Calder, "Using simpoint for accurate and efficient simulation," in *SIGMETRICS*, 2003.

[26] L. A. Belady, "A study of replacement algorithms for a virtual-storage computer," *IBM Systems journal*, vol. 5, no. 2, pp. 78–101, 1966.

[27] D. Sanchez and C. Kozyrakis, "The zcache: Decoupling ways and associativity," in *Microarchitecture (MICRO), 2010 43rd Annual IEEE/ACM International Symposium on*, pp. 187–198, IEEE, 2010.