

Mitigating Wordline Crosstalk using Adaptive Trees of Counters

Sayed Mohammad Seyedzadeh[†], Alex K. Jones[§], Rami Melhem[†]
Computer Science Department[†], Electrical and Computer Engineering Department[§]
University of Pittsburgh
seyedzadeh@cs.pitt.edu, akjones@pitt.edu, melhem@cs.pitt.edu

Abstract—DRAM technology scaling has the undesirable side effect of degrading cell reliability. One such concern of deeply scaled DRAMs is the increased coupling between adjacent cells, commonly referred to as crosstalk. High access frequency of certain rows in the DRAM may cause data loss in cells of physically adjacent rows due to crosstalk. The malicious exploit of this crosstalk by repeatedly accessing a row to induce this effect is known as *row hammering*. Additionally, inadvertent row hammering may also occur due to the natural weighted nature of applications' access patterns.

In this paper, we analyze the efficiency of existing approaches for mitigating wordline crosstalk and demonstrate that they have been conservatively designed. Given the unbalanced nature of DRAM accesses, a small group of dynamically allocated counters in banks can deterministically detect “hot” rows and mitigate crosstalk. Based on our findings, we propose a Counter-based Adaptive Tree (CAT) approach to mitigate wordline crosstalk using adaptive trees of counters to guide appropriate refreshing of vulnerable rows. The key idea is to tune the distribution of the counters to the rows in a bank based on the memory reference patterns. In contrast to deterministic solutions, CAT utilizes fewer counters, making it practically feasible to be implemented on-chip. Compared to existing probabilistic approaches, CAT more precisely refreshes rows vulnerable to crosstalk based on their access frequency.

Experimental results on workloads from four benchmark suites show that CAT reduces the Crosstalk Mitigation Refresh Power Overhead in quad-core systems to 7%, which is an improvement over the 21% and 18% incurred in the leading deterministic and probabilistic approaches, respectively. Moreover, CAT incurs very low performance overhead ($\sim 0.5\%$). Hardware synthesis evaluation shows that CAT can be implemented on-chip with only a nominal area overhead.

Keywords—Row Hammering; DRAM; Technology Scaling;

I. INTRODUCTION

Dynamic random-access memory (DRAM) has been a widely used storage element in computer systems. Over time, process technology has scaled DRAM to achieve greater storage density by reducing the technology feature size [1–5]. However, shrinking process technology causes DRAM cells to become significantly less reliable [2, 6, 7]. As the chip density increases with technology scaling, the interaction between circuit components, such as transistors, capacitors, and wires increases, leading voltage fluctuations [8–10]. Specifically, when the cumulative interference to a DRAM wordline becomes strong enough, the state of nearby cells can change leading to memory errors [11, 12]. Vulnerability to wordline electromagnetic coupling (crosstalk) exists in recent sub 40nm commodity DRAM chips due to physical limitations of these technologies. Kim [13] showed that through frequently alternating the charge of specific memory locations, crosstalk can be used, intentionally, to affect the charge of adjacent cells [14]. However, in addition to intentional malicious attacks [15, 16], inadvertent row hammering is also possible

due to the unbalanced nature of some applications' access patterns that can lead to “hot” rows and induce crosstalk.

One simple solution to mitigate wordline crosstalk is to increase the refresh rate for all rows. Although, this approach is effective, it imposes an unnecessarily high power and performance overhead [6, 15, 17–23]. There are two main hardware approaches to mitigate wordline crosstalk in DRAM. The first relies on a random number generator [24, 25] in the memory controller to probabilistically refresh accessed rows [13, 26]. Although the idea behind this approach is simple, it results in an early refresh of rows that can be potentially affected by crosstalk [15], as well as unnecessary row refreshing in the absence of hot rows.

The second approach detects the most frequently accessed rows, or *aggressor rows*, and then refresh the rows that are adjacent to it, or *victim rows*. To protect new DDR4 modules against wordline crosstalk, DRAM architectures implement the Targeted Row Refresh (TRR) mechanism that allows victim rows to be refreshed on demand¹. Intel has announced the existence of pseudo-targeted row refresh in Xeon-class Ivybridge architectures to help in the crosstalk mitigation, but has not yet released the details of this mechanism. Several methods have been proposed to identify rows for TRR. A simple method to deterministically recognize aggressor rows, called Static Counter Assignment (SCA), is to dedicate a counter per row to keep track of the number of row activations. However, having one counter per row induces a significant area and power overhead to the memory system. To prevent this high cost, row activation counters can be stored in a reserved area of the main memory and a dedicated on-chip counter-cache can be used to minimize the performance penalty of retrieving counter values from main memory [26].

Due to row access locality in DRAM [28–30], many counters in SCA would be underutilized [31]. Thus, we propose a Counter-based Adaptive Tree (CAT) approach that dynamically assigns counters to frequently accessed aggressor rows. Hence, with a small number of on-chip counters it is possible to deterministically refresh victim rows. When CAT results in a highly unbalanced tree, it provides a significant advantage in refresh energy over a block-based uniform counter distribution with a similar number of counters, while CAT also converges to a balanced tree when accesses in memory are uniform. The dynamic counter tree of CAT is constructed by partitioning the memory rows into groups, each governed by a counter, based on the DRAM access pattern. To accomplish this, the **refresh threshold**, defined as the number of aggressor row accesses before crosstalk is

¹Note that Micron DDR4 documentation [27] suggests that the TRR module is an optional module, thus there may exist future DDR4 systems that do not incorporate TRR.

known to occur in victim rows, is subdivided into different split thresholds. These split thresholds identify points at which a group is likely to contain an aggressor row and increased precision is desired for more precise identification. When crossing a split threshold, CAT subdivides the block governed by a counter and provides a second counter to increase the precision. Thus, CAT builds a uniform tree for workloads with random row access frequency and a non-uniform tree for workloads with biased row access frequency. The key feature of CAT is that hot rows are instrumented using smaller groups, while rows with low access frequency (cold rows) are unlikely to induce crosstalk and are instrumented using larger groups. Thus, the selection of appropriate split thresholds is a key factor to the success of CAT.

CAT deterministically detects aggressor rows before they reach the refresh threshold, whether they result from intentional malicious attacks or applications' unbalanced access patterns. It refreshes victim rows to prevent crosstalk and guarantee reliability of the memory during normal operation as well as defend against row hammering attacks. We make the following contributions for wordline crosstalk mitigation:

- We demonstrate that, due to access locality in DRAM [28–30], instead of over-provisioning with one counter per row, a small number of counters can be implemented on-chip to refresh victim rows, while achieving low latency, low power consumption and satisfactory area overhead.
- We introduce a non-uniform counter assignment, CAT, to more precisely determine the aggressor rows and refresh vulnerable rows to improve the effectiveness of uniform counter assignment.
- We determine the suitable number and values of split thresholds for non-uniform counter assignment (*i.e.*, CAT) to best construct a, potentially unbalanced, tree for optimized alignment of access counters to increasingly small groups of rows that contain aggressor rows.
- We introduce a dynamically reconfigurable CAT scheme (DRCAT), that tracks and reacts to the temporal changes in memory access patterns resulting from either application context switching or different phases of a particular application in order to more precisely identify actual victim rows and reduce DRAM refresh energy.

II. BACKGROUND AND RELATED WORK

DRAM-based main memory is a multi-level hierarchy of structures. Each memory module is composed of a number of chips and is connected to the memory controller through a channel. Internally, each chip consists of multiple banks and each bank is organized as rows of DRAM cells. Each cell is composed of an access transistor and a capacitor, in which the data is stored. While accessing a row, all cells in the row are selected in parallel using a wordline. Due to capacitive coupling between cells on adjacent wordlines, if a wordline (aggressor row) is accessed frequently, voltage levels on neighboring wordlines (victim rows) can be affected leading to crosstalk. Mitigating wordline crosstalk is possible by refreshing the victim rows before the aggressor rows reach the refresh threshold.

A hardware approach to alleviate wordline crosstalk is for each DRAM access to refresh the victim rows adjacent to the accessed row based on a probability function [26]. In this probabilistic approach, called PRA (Probabilistic Row

Activation), the memory controller uses a Pseudo-Random Number Generator with a given probability (α) to determine when the memory controller should issue a refresh signal to refresh the two rows adjacent to the accessed row. When either the number of memory accesses or the probability α is high, this approach generates a significant number of refresh commands, thus exacerbating memory contention and increasing the energy cost [15].

As a hardware alternative to the probabilistic approach, a deterministic approach can be used to prevent aggressor rows from being accessed more than the refresh threshold before refreshing the victim rows. Maintaining a counter for each memory row is a significant overhead [32, 33]. To address this problem, an approach was proposed that stores the counters in a reserved area of DRAM and a set-associative counter cache was established in the memory controller to improve accessibility to frequently used counters [26]. Note that the primary idea in [26] is similar to that used for Counter-based caches [34], where threshold-based counters detect expired lines for proactive eviction. While, using counters allows for accurate counts of row accesses, caching the counters introduces the complexity of maintaining a cache (*e.g.*, tag matching, eviction policies) within the memory controller. Moreover, misses to the cache counter can be expensive.

Rewriting instructions, such as CLFLUSH [35], have been proposed as software countermeasures against wordline crosstalk and are now deployed in Google Native Client. Similarly, access to the Linux page map interface is now prohibited from userland [36]. These countermeasures have already been proven insufficient to mitigate malicious kernel attacks [37, 38]. In [15], a generic software mechanism, ANVIL, is proposed to detect aggressor rows by monitoring the last-level cache miss rate and row accesses with high temporal locality. A similar approach is proposed in [39] to monitor the number of last-level cache misses during a given refresh interval. Both approaches rely on software access to CPU performance counters.

III. MOTIVATION

In this section, we analyse the previously proposed hardware approaches and make key observations to motivate the dynamic counter assignment as a hardware solution that mitigates wordline crosstalk and combats row hammering.

A. Probabilistic Refresh Analysis

Using a probabilistic approach, such as PRA, to mitigate wordline crosstalk can protect against failure with a high probability, depending on the value of the refresh threshold, T , and the probability, α , of triggering a refresh. The probability of experiencing an error in Y years (defined as Y -years unsurvivability) for PRA is computed as:

$$\text{unsurvivability} = (1 - p)^T \times Q_0 Q_1 \quad (1)$$

where $p = \alpha$ is the probability of refreshing TWO victim rows on an access, Q_0 is the number of refresh threshold windows during a refresh interval, and Q_1 is the number of 64ms periods during Y years. The parameter T depends on the technology node. Specifically, scaling down DRAM increases voltage fluctuations in cells because of the interaction between circuit components. Therefore, the refresh threshold is projected to decrease for future memory technology [26].

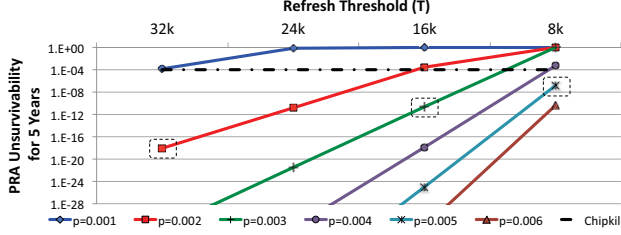


Figure 1: PRA unsurvivability for refresh thresholds 32k, 24k, 16k and 8k. PRA refreshes two victim rows with probability p .

Figure 1 compares the 5-years unsurvivability for different refresh thresholds when p ranges from 0.001 to 0.006. Assuming mild row accesses during refresh intervals, we set Q_0 to 10, 15, 20, and 40. Figure 1 shows that for $T=32K$ and $p > 0.001$, PRA's unsurvivability is lower than the Chipkill's unsurvivability of $1E-4$. Our key observation from this figure is that, for smaller values of T , larger values of p are needed to match the 5-years survivability of chipkill². In fact, PRA's failure probability increases exponentially when the refresh threshold scales down, as is expected in future technology nodes. This means that larger values of p (more frequent random refreshes) are needed to guarantee acceptable survivability.

Note that the reliability reported in Figure 1 assumes the use of a true pseudo random number generator, PRNG, such as the one proposed in [25]. This is important since the computed reliability is contingent on the randomness of the numbers generated by PRNG. Specifically, the unsurvivability in Eq. 1 will not apply if a simpler (less costly in terms of area and power) PRNG is used since the randomness of the generated numbers will not be independent enough. To study the effect of the randomness of the generated numbers, we conducted a Monte-Carlo simulation to estimate the unsurvivability of PRA when a LFSR-based PRNG [40, 41] is used. The results show that, using an LFSR-based PRNG largely reduces PRA's unsurvivability. For example, for $T=16K$ and $p=0.005$, PRA's unsurvivability reaches $1E-4$ after only 25 refresh intervals. To improve the reliability, a much larger value of p should be used with LFSR-based PRNGs which increases the refresh power and decreases the performance. A similar conclusion was reached in [16]. Hence, PRA requires true random number generators, which are known to be complex and to consume relatively large power [24, 25, 42–44], to achieve the probabilities shown in Figure 1.

B. Static Counter Assignment (SCA) Analysis

Using a deterministic approach for counting the number of accesses per row with on chip counters using SCA requires a large area and power overhead. One intuitive solution is to use fewer counters by partitioning the rows in each memory bank into fixed-size groups and assign one counter per group. To illustrate SCA, we assume that every bank in DRAM includes N rows and uses M counters. The refresh threshold, T , determines the size of every counter as $\log_2 T$ -bits. This approach, called SCA_M , divides the rows into M groups, each including $\frac{N}{M}$ rows. For every row activation, the row address maps to the appropriate counter. Then, the corresponding

counter counts the number of accesses. When the counter reaches the threshold, it is reset and a refresh signal is sent to the memory controller to refresh $\frac{N}{M} + 2$ rows; the $\frac{N}{M}$ rows in the group plus the two rows adjacent to the group, which guarantees the refresh of any row in or adjacent to the group subjected to the crosstalk.

The energy overhead in SCA originates from activating the counters when memory is accessed and refreshing $\frac{N}{M} + 2$ rows when a counter exceeds T . Figure 2 breaks down the energy overhead of SCA_M during a 64ms auto-refresh period when $N = 65536$ and the number of counters M ranges from 16 to 65536³. For a small number of counters, the energy resulting from refreshing victim rows (blue line) dominates the total energy of activating counters in SCA. In contrast, the total energy of activating counters in SCA is the dominant energy as the number of counters significantly increases (orange line).

Figure 2 shows that the total energy can be minimized at $M=128$. In this case, SCA_{128} not only reduces the energy overhead in comparison to SCA_{65536} , but also decreases the area overhead by two orders of magnitude (as will be explained in Section VII). In comparison, the counter cache approach [26], which stores counters in the reserved area of DRAM memory, reports data for a much larger counter storage cache, requiring capacity to store on the order of thousands of counters per bank. Ostensibly, this is to allow for enough flexibility to store the relevant counters to hot rows without a high miss rate due to capacity misses and/or thrashing. Thus, the energy overhead of counter storage cache will significantly exceed SCA_{128} due to the increased static power. For example, Figure 2 shows the optimistic energy (assuming no misses requiring accesses to the DRAM) of 2K and 8K per bank counter caches as horizontal lines. These lines intersect the SCA_{4096} – SCA_{16384} points, respectively, as they have the same amount of total counter storage⁴. Thus, the total energy consumption of SCA with $M \leq 4K$ counters is lower than that of the counter caches with different sizes. In particular, SCA_{128} can improve the total energy overhead and area overhead by 1.5 orders of magnitude in comparison to a 2K counter cache and nearly two orders of magnitude compared to an 8K counter cache [26].

Thus, our key observation of these deterministic approaches is that allocating one counter to each row in a DRAM bank with a cache counter can be effective but are somewhat conservative and leave room for improvement. Specifically, the analysis of row access frequency of DRAM banks on real workloads reveals that the row access frequency during the refresh interval is not uniform and mostly a small group of rows are activated in DRAM banks. For example, Figure 3 depicts the row access frequency of a given bank for two typical real workloads, *blackscholes* and *facesim*, within a time period of one refresh interval (64ms). Figure 3 clearly shows that a small group of rows dominate overall accesses. This motivates us to propose a dynamic counter assignment for wordline crosstalk mitigation.

³The refresh energy includes the average refresh energy of victim rows for 18 real workloads (Details in Section VI). We modified CACTI [45] to model the cache in the counter cache approach [26].

⁴The counter caches have additional storage to store the tag array. However, this storage is typically less than the data array making it inconsequential on a log plot.

²Similar analysis done in [26] shows that $PRA_{p=0.001}$ probability of failure is higher than $1E-4$.

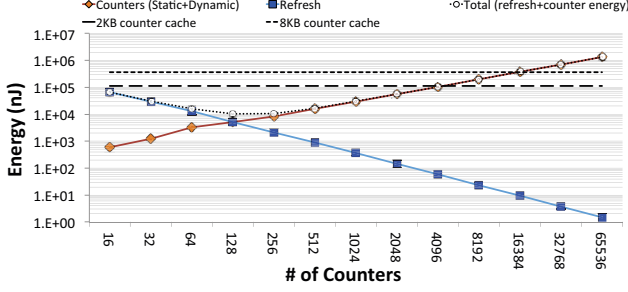


Figure 2: The energy overhead of SCA and counter caches [26] for different number of counters.

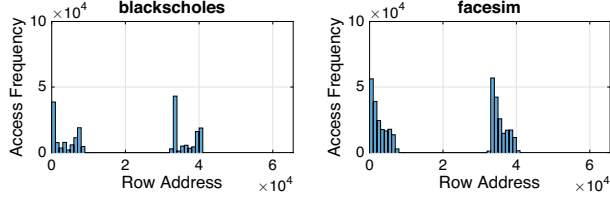


Figure 3: Row address frequency in a DRAM bank with 64K rows.

C. Our Goal

Our goal is to take advantage of a small number of counters per bank, but better target the aggressor rows to provide further benefits to overall energy consumption while mitigating crosstalk. Our novelty is the adaptive construction and dynamic reconfigurability of a "potentially unbalanced" tree of counters to match access patterns.

IV. COUNTER-BASED ADAPTIVE TREE

In order to better assign row partitions to access counters, the Counter-Based Adaptive Tree (CAT) is a new and practical dynamic row partitioning technique that considers access frequency of rows to more carefully assign counters to appropriately sized groups of rows in order to improve energy and area efficiency. To divide an initial group of rows (e.g., a bank or some other uniform coarse partition) into groups of suitable sizes, CAT defines different split thresholds that identify access frequency stages prior to reaching the refresh threshold. These split thresholds are used to build a non-uniform binary tree structure that maps hot rows to smaller groups, while cold rows, *i.e.* rows with relatively low access frequencies, are mapped to larger groups. This aligns access counters to small groups of rows that contain an aggressor row to more precisely identify actual victim rows.

A. A simple CAT Example

Figure 4 depicts two trees built by CAT, where a terminal node, \bullet , represents an active counter and intermediate node, \circ , represents an expired counter, which had been split into two counters. The level of a node is defined as its distance from the root, with the root being at level zero. The levels of the CAT are associated with unique split thresholds. Hence, when a node reaches the next threshold, it further subdivides the group, or *splits* the node, generating two children counters initialized to the current count value. This is accomplished by activating a second counter as a clone of the existing counter. The binary tree of counters continues to grow until all available counters are activated or a maximum allowed level, a parameter of the CAT algorithm, is reached.

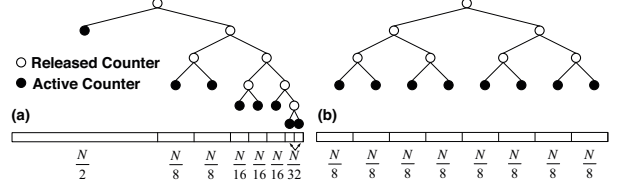


Figure 4: The adaptive tress of counters for the workload with (a) biased, (b) uniform row address frequency. The number of row addresses in the bank is N .

More precisely, assuming that we limit the number of levels in the tree to L , we define $L-1$ split thresholds T_0, \dots, T_{L-2} where $T_0 \leq \dots \leq T_{L-1}$ and $T_{L-1} = T$, recalling that T is the refresh threshold. Each of the M counters in a bank, C_0, \dots, C_{M-1} , has $\log_2 T$ bits and, initially, only C_0 is in active mode. When a counter at level l reaches the split threshold, T_l , it splits and two counters are activated at level $l+1$. This process continues until all the counters are activated or $l = L-1$. For example, Figure 4 shows two CATs for $L = 6$ and $M = 8$. The CAT in Figure 4(a) results from a non-uniform row access pattern, which causes more counters to be allocated to the hot row area (smaller blocks) and grows the tree through level 5. In contrast, when the row access frequency is uniform, counters are distributed uniformly throughout the bank addresses as shown in Figure 4(b). In this case, the CAT approach grows the tree only through level 3 and mimics SCA.

In CAT, N rows in one bank are initially treated as a group to which C_0 is allocated. As soon as C_0 reaches T_0 , CAT splits C_0 into C_0 and C_1 with the same starting value of T_0 . In this case, C_0 counts the number of accesses when the row address is between 0 to $\frac{N}{2} - 1$ and C_1 counts the number of accesses when the row address ranges from $\frac{N}{2}$ to $N - 1$. When C_1 reaches T_1 , CAT splits C_1 into C_1 and C_2 with the new initial value T_1 where C_1 and C_2 track row addresses in the ranges from $\frac{N}{2}$ to $\frac{3N}{2} - 1$ and from $\frac{3N}{2}$ to $(N - 1)$, respectively. CAT continues this process until it activates all counters and no group can be split into smaller sub-groups. At this point, the split thresholds of counters are set to T . The minimum number of rows in a given group depends on the number of defined split thresholds. With $L-1$ split thresholds (a CAT with at most L levels), the minimum number of rows per group is $\frac{N}{2^{L-1}}$.

B. Constructing the CAT

Algorithm 1 shows the process for refreshing rows under the CAT structure per memory bank. It has two main modules: the Counter Module (CM) that records the number of row accesses and the Reconfiguration Counter Module (RCM) that activates and initializes counters. Assuming M counters in a given bank, CAT requires an array of M counter modules that are implemented on-chip, and one RCM that can be implemented either on-chip or in software. Each counter module CM_i maintains two registers, L_i and U_i to store the lower and upper row addresses assigned to this counter, and a register l_i to store the index of the split threshold used for that counter. The RCM maintains a *last_activated* counter register.

Initially, at the start of each refresh interval, CAT is reset such that only the first counter module, CM_0 , is activated with $L_0 = 0$, $U_0 = N - 1$, $l_0 = 0$, and *last_activated* = 0. Each

time a row is accessed, its address is located in the range $L_i - U_i$ of some active C_i , and this counter is incremented (lines 5-7). When C_i reaches T_i , $flag_i$ is raised (lines 8-10), which triggers RCM to activate a new counter as long as the number of active counters is less than M and the counter level $l_i < L - 1$ (lines 15-16). When a new counter is activated, it is initialized by C_i (line 17) and the interval between L_i and U_i is split into two equal-size ranges where the lower bound of C_i remains unchanged and the upper bound of C_i is assigned to the upper bound of the new counter. Then, U_i shrinks to $U_i = \frac{U_i + L_i}{2}$ and the lower bound of the new counter is set to $U_i + 1$ (lines 18-20). The split thresholds of both counters are set to $l_i + 1$ (lines 21-22). For example, after initialization, when CM_0 reaches T_0 , CM_1 is introduced by subdividing CM_0 in half, such that $C_1 = C_0$, $L_0 = 0$, $U_0 = \frac{N}{2} - 1$, $L_1 = \frac{N}{2}$, and $U_1 = N - 1$ with both CM 's split thresholds being set to T_1 and $last_activated = 1$.

This process continues until some CM , CM_i , reaches the highest threshold $T_i = T$ (i.e., if $l_i = L - 1$, lines 10-12). In this case, C_i is reset and the signal R_i is raised to cause the memory controller to refresh all existing rows in the address range of $L_i - 1$ and $U_i + 1$. When all counters are activated, CAT will set the index of all split thresholds to $l_i = L - 1$ which causes $T_i = T$ (line 25).

C. Efficient CAT Management Using SRAM

To directly implement Algorithm 1, maintaining the range boundaries of row blocks requires more storage than the actual counters, themselves. Given that SRAM uses less area and static power than registers [46], we are motivated to design and optimize the CAT for SRAM. In this case, instead of storing row range boundaries, we use pointers to *store the structure* of the CAT as shown in Figure 5. During each access, the tree structure is traversed sequentially by chasing the pointers to find the counter assigned to a specific row address.

The CAT, shown visually in Figure 5(a), is composed of two types of nodes: leaf nodes (shown in light blue) that represent active counters and intermediate nodes (shown in white) that determine the tree's structure. Rather than store all the nodes in our data structure, shown in Figure 5(b), we store only the intermediate nodes. Thus, we use an array, I , of size $M - 1$ (the maximum number of intermediate nodes in a tree with M leaves) to store information about intermediate nodes. Separately, we use another array C , of size M to store the counters, shown in Figure 5(c). For each intermediate node, two pointers, L_ptr and R_ptr , point to information about its two successors. If the successor is another intermediate node, L/R_ptr contains the entry for that intermediate node. If the successor is a leaf, L/R_ptr contains the entry for the counter corresponding to that leaf. Two flags, L_leaf and R_leaf , indicate if the corresponding successor is an intermediate or leaf node. The length of each counter is $\log(T)$ bits and each pointer is $\log(M)$ bits. The root of the tree, I_0 , is deterministically stored in the first entry of the array I .

To determine whether to inspect the left or right entry, we examine the address A ($0 \leq A < N$). Starting at the root, the high order bit of A determines the successor that covers row A , thus accessing a leaf or an intermediate nodes as already described. More generally, when traversing an intermediate node at level l of the CAT, the l^{th} bit of A , counting from

Algorithm 1: CAT structure per memory bank

```

1 Parameters:  $N$ : # rows per bank;  $M$ : # counters per bank;
   $L$ : # thresholds; Input:  $row\_address$ ; Output:  $R_i$ : Refresh
  signal for refreshing all existing rows between  $L_i - 1$  and  $U_i + 1$ .
2 begin
3   Counter Module  $CM_i$  /*  $i = 0, \dots, M - 1$  */
4   if  $L_i \leq row\_address \leq U_i$  then
5     if  $C_i < T_i$  then
6        $C_i++$ ;
7     else
8       if  $l_i < L - 1$  then
9          $flag_i = 1$ ; /* Signal to trigger RCM */
10      else
11         $R_i = 1$ ; /* Signal to refresh corresponding rows */
12         $C_i = 0$ ;
13
14   Reconfiguration Counter Module (RCM) /* Activated when
15      $flag_i = 1$  */
16   if  $flag_i == 1$  for some  $i$  then
17     if  $last\_activated < M - 1$  &&  $l_i < L - 1$  then
18        $last\_activated++$ ; /* Increase # of active counters */
19        $C_{last\_activated} = C_i$ ;
20        $U_{last\_activated} = U_i$ ;
21        $U_i = \frac{U_i + L_i}{2}$ ;
22        $L_{last\_activated} = U_i + 1$ ;
23        $l_i++$ ;
24        $l_{last\_activated} = l_i$ ;
25   if  $last\_activated == M - 1$  then
26     for  $i = 0 : M - 1$  do
27        $l_i = L - 1$ ;

```

the most significant bit, is used to select the successor, which may be a leaf node or an intermediate node at level $l + 1$. Before the CAT is completely built, it is guaranteed that fewer than M counters and $M - 1$ intermediate nodes are utilized.

To illustrate the process of splitting a counter during the building of the CAT, we consider the example CAT articulated in Figure 5 by rolling back the last split operation. The last counter, C_7 was deployed by splitting C_6 into C_6 and C_7 and introducing I_6 . Let us assume that the current I_6 still points to a leaf node C_6 . At this point of the CAT construction, only 7 counters were deployed. This incomplete tree is represented by the same array I of Figure 5 with the fourth entry of the array being $[C_3, \mathbf{C_6}, 0, 0]$ (differences noted in bold) and the last entry being still undefined. To reach the state shown in Figure 5, C_6 reached split threshold; I_3 's R_ptr was replaced by a pointer to the next available entry in I , I_6 ; the last available counter, C_7 , is initialized to match C_6 ; and I_6 is set to $[C_6, C_7, 0, 0]$ as is shown in the figure.

Given the above implementation, the maximum number of sequential SRAM accesses for traversing the CAT is equal to the maximum depth of the tree, L . That number of accesses may be reduced if instead of starting to build the CAT tree from its root, we start from a pre-set complete binary tree with λ levels for some $\lambda \leq \log M$. Consequently, to traverse the CAT, we can use the most significant λ bits of the row address, A , to directly access the appropriate intermediate node at level $\lambda - 1$, which reduces the maximum number of SRAM accesses to reach a leaf to $L - \lambda + 1$. For example, if we start from a uniform binary tree with $\lambda = \log M$ levels, the initial CAT will be a complete tree containing $M/2$ counters and $M/2 - 1$ intermediate nodes. The other $M/2$ counters can then be used to grow the CAT non-uniformly beyond λ levels and up to a maximum of L levels. Moreover, by pre-splitting the counters uniformly up to level $\lambda - 1$ (that is starting from a balanced CAT with λ levels), we can reduce

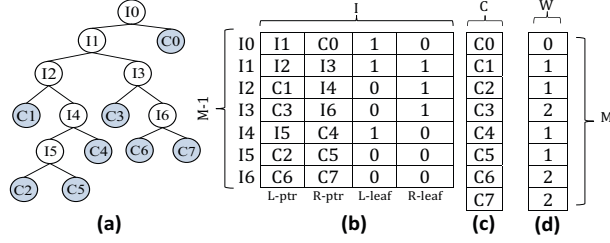


Figure 5: (a) A CAT using pointer chasing with $M = 8$ counters and 6 levels. (b,c) The data structure used to represent the CAT. (d) An array of weight registers used for reconfiguring the CAT (see section 5).

the size of the intermediate node array because we can avoid storing the intermediate nodes at levels smaller than λ .

D. Determining Split Threshold Values

The CAT adapts the distribution of the available counters to the rows in a bank depending on memory reference patterns. Specifically, the CAT is dynamically shaped to minimize the number of refreshed rows, and thus, the refresh power. Given a sequence of row references, the split thresholds determine the shape of the tree. In experimenting with the CAT technique, we found that its performance is sensitive to the values of the split thresholds. Given the combinatorial number of options for selecting the split thresholds, we present in this section a model to determine these thresholds in a way that minimizes the number of refreshed rows. We explain that model assuming that we start from a uniform CAT with $\lambda = m = \log M$ levels, and determine the split thresholds T_{m-1}, \dots, T_{L-2} used to grow the tree non-uniformly to a maximum of L levels.

For illustration, we consider a simple example in which 4 counters are used for the N rows in a bank. Specifically, assume that after a number of references, the CAT is represented by the tree structure shown in Figure 6(a). Depending on the reference pattern and the values of the split thresholds T_1 and T_2 , this structure can evolve to either the balanced tree structure of Figure 6(b) or the non-balanced tree structure of Figure 6(c). That is, whether counter C1 splits first (Figure 6(b)) or counter C3 splits first (Figure 6(c)) depends on the relative value of T_1 and T_2 . After one of the counters splits, the CAT reaches its final shape and the thresholds of all the counters are set to the refresh threshold T . Hence, it is crucial to choose the split thresholds so that the CAT assumes the form of the tree that minimizes the number of refreshed rows.

Continuing with the 4-counter example, we note that a counter at level 0, 1, 2 and 3 will be assigned $4w$, $2w$, w and $w/2$ rows, respectively, where $w = N/4$. Now, assume that the bank receives R row references (accesses) during the regular refresh interval. If the references are uniformly distributed across rows, then each counter in Figure 6(b) will receive $R/4$ references, and if the refresh threshold is T , then each counter will reach this threshold $R/4T$ times. Each time a counter reaches T , the w rows assigned to it are refreshed. Hence the total number of refreshes is

$$Cost_{SCA} = w \times R/T \quad (2)$$

An unbalanced CAT is expected to reduce the number of refreshed rows if the R references are sufficiently biased towards a small group of rows. To determine the “amount” of bias that will favor the CAT of Figure 6(c) over the uniform

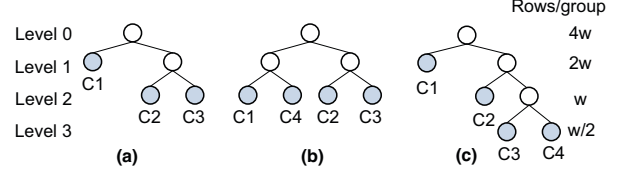


Figure 6: Two possible evolutions of the CAT of (a) to a balanced tree structure (b) or an unbalanced structure (c).

tree of Figure 6(b), we define this bias using a variable x such that the group of rows assigned to counter C4 receives x more references than the other rows. That is the ratio of references caught by counters C1, C2, C3 and C4 is $2w : w : w/2 : x + w/2$. This means that each of the four counters will receive $r1 = 2w\alpha$, $r2 = w\alpha$, $r3 = 0.5w\alpha$ and $r4 = (x + 0.5w)\alpha$ references, respectively, where $\alpha = R/(x + 4w)$. Consequently, the refresh threshold T will be reached in counter C1 $r1/T$ times, and in each time the $2w$ rows assigned to it will be refreshed. Similarly, T in C2 will be reached $r2/T$ times, and in each time the w rows assigned to it will be refreshed. Finally, T in C3 and C4 will be reached $r3/T$ and $r4/T$ times, respectively, and in each time the corresponding $w/2$ rows will be refreshed. Hence, the total number of refreshed rows is

$$Cost_{CAT} = ((2w)^2 + w^2 + (\frac{w}{2})^2 + (x + \frac{w}{2})\frac{w}{2})\frac{\alpha}{T} \quad (3)$$

From (2) and (3), we conclude that $Cost_{CAT} < Cost_{SCA}$ when

$$x > 3w. \quad (4)$$

After determining the critical bias that causes the CAT to outperform the uniform tree, we proceed to find the thresholds T_1 and T_2 that will force, after a short sequence of accesses, say R_s , the tree of Figure 6(a) to evolve to the tree of Figure 6(c) if $x > 3w$ and to the uniform tree otherwise. For this, we note that if the reference bias is $x = 3w$, then after R_s references, the counters C1, C2 and C3 in Figure 6(a) will record $2w\beta$, $w\beta$ and $4w\beta$ accesses, respectively, where $\beta = R_s/7w$. Hence, if T_2 is set to be $2T_1$, then C3 will reach T_2 before C1 reaches T_1 when $x > 3w$, thus converging to the CAT of Figure 6(c). On the other hand, if $x < 3w$, then C1 will reach T_1 before C3 reaches T_2 , thus leading to the uniform tree of Figure 6(b). To completely specify the split thresholds, we chose $T_2 = T/2$, which guarantees that the CAT converges before any counter reaches the threshold T . Consequently, $T_1 = T/4$.

The same reasoning used in the 4-counter example can be generalized to the case of $M = 2^m$ counters. Due to space limitation, this extension is not presented in this paper and is provided in the technical report. The generalized model is used to determine the split thresholds for all the experiments presented in this paper. For example, when applied to the tree with $M = 64$ counters and $L = 10$ levels, the values of the thresholds computed by the model are: $T_5 = 5155$, $T_6 = 10309$, $T_7 = 12886$, $T_8 = 16384$, and $T_9 = T = 32768$.

V. RECONFIGURING THE CAT TO TRACK CHANGES IN ACCESS PATTERNS

The CAT assigns the available counters to the rows of a bank according to the pattern of row accesses. However, the row access pattern changes with time, which necessitates

a mechanism for the reconfiguration of the CAT to track these changes. In the next two sections, we propose two such mechanisms. The first, PRCAT, periodically reconstructs the CAT and the second, DRCAT, dynamically reconfigures the CAT by reassigning counters from cold to hotter regions of the bank.

A. Periodically Reset CAT (PRCAT)

In this scheme, the CAT tree is rebuilt at epochs equal to the auto-refresh interval (64ms for several DRAM generations [47]). For LPDDRx devices that support burst refresh [48], this simple scheme tracks the number of row accesses exactly. It can also be applied to modern DDRx devices that support distributed refresh at the expense of some inaccuracy in tracking the number of accesses. Specifically, because row refreshes are out of sync with the resetting of the CAT, recent information about row accesses are lost when the CAT is reset. Moreover, PRCAT resets the CAT periodically, even when the row access patterns do not change, potentially incurring the overhead of reconstructing the CAT unnecessarily. In the next section, we describe a CAT reconfiguration scheme which avoids these two shortcomings at a small cost for keeping additional information about the usage of the counters in a CAT.

B. Dynamically Reconfigured CAT (DRCAT)

The DRCAT allocates weights to counters to track the number of times each counter reaches the refresh threshold. After the CAT is completely built, the DRCAT identifies the counters allocated to regions that become cold and reallocate them to regions that become hot. A 2-bit weight register is used to record the weight of each counter. As described in the last section, when a counter reaches the refresh threshold, its corresponding rows are refreshed and its value is reset to zero. However, to keep track of the hotness of row regions, the weight corresponding to that counter is incremented (with an upper bound of 3) and the weights corresponding to all other counters are decremented (with a lower bound of 0). If the weight of the incremented counter reaches its maximum limit, two counters having zero weights (cold regions) are merged and the released counter is used to split the hot counter.

To illustrate the scheme, consider the CAT example shown in Figure 5, where all counters have been activated and the weights of the counters are kept in the register W depicted in Figure 5(d). Assume that at a given time during operation, the values of the weight registers are $[0,1,1,2,1,1,2,2]$ and counter $C6$ reaches its limit (we used 2-bit counters). After the rows corresponding to $C6$ are refreshed, the values of the weights are updated to $[0,0,0,1,0,0,3,1]$ and the following steps are taken to reconfigure the CAT:

(1) From the table shown in Figure 5(b) an intermediate node in the CAT which has two counters as children ($L_leaf = R_leaf = 0$) with the weight of both being zero is selected. If such a node is found, the two counters are merged, one counter is freed and we go to step 2. In our example, $C2$ and $C5$ are leaves and both weight registers are zero. Hence, $C5$ is promoted to its parent node and the fifth row of the table is updated to $I4=[C5,C4,0,0]$ as shown in Figure 7(b). Furthermore, $C2$ and the sixth row, $I5$, of the table are released.

(2) We split the region tracked by the hot counter using the counter freed in step 1. In our example, we show splitting

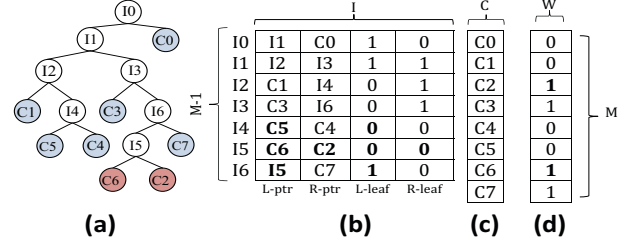


Figure 7: The CAT of Figure 5 after reconfiguration.

$C6$ by replacing the L_ptr in its parent node (entry $I6$) by the index of the released row ($I5$) and set its corresponding flag to 1 to indicate that $I5$ will represent an intermediate node. Finally, we update $I5=[C6,C2,0,0]$ to point to $C6$ and $C2$ and reset the corresponding flags.

(3) We update weight of the newly split counters to 1 to ensure they remain split for a reasonable period of time while preventing them from being quickly split in succession.

The DRCAT adds a negligible area overhead to the PRCAT design. For example, PRCAT uses 2 bytes per counter for $T=14K$ and in this case, it occupies area overhead similar to DRCAT. The reason is that DRCAT uses the first 14 bits for the counter and the two last bits for the weight register. With respect to latency, the DRCAT traverses the tree to find the cold counters and their parent intermediate node. Since the reconfiguration of the tree happens infrequently and traversing the tree is not on the critical path, system's performance is not affected by the reconfiguration.

Note that, in addition to tracking the change in the hot spot of memory accesses, the reconfiguration of the CAT according to the weights of the counters has the flexibility of adapting to multiple hot spots in the access patterns.

VI. EXPERIMENTAL METHODOLOGY

To evaluate the proposed technique, we performed simulations using the memory system simulator USIMM modeling 55nm DRAM [47]. Unless stated otherwise (in Section VIII), the default simulation environment was set to model memory traffic from a dual core CPU. The total memory capacity is 16 GB with a total of 16 banks divided into two ranks, with 64K rows per bank. The Last level cache size is 512KB per core in our simulation. Detailed simulation parameters for USIMM are listed in Table I. The DRAM timing constraints follow a Micron DDR3 SDRAM data sheet [46, 49]. Verilog implementations of the control logic for the different wordline crosstalk mitigation schemes were created to provide an area and energy overhead comparison. These Verilog codes were synthesized using Synopsys Design Compiler and evaluated for power using Synopsys PrimeTime, targeting a 45nm FreePDK standard cell library [50–55]⁵. We have changed the number of counters per bank in the designs between 32 and 512 and, for CAT, allowed the trees to grow up to 14 levels to study the trade-off between performance, crosstalk mitigation refresh power, and hardware overhead. For the evaluation of PRA, we accounted for the energy to generate a random number every row access.

To provide realistic workloads for evaluating the wordline crosstalk mitigation schemes, we used workloads from the

⁵ It is commonplace for DRAM to trail CMOS by a technology generation. Systems with 45nm CPUs were concurrent with 55nm DRAM.

Memory Scheduling Championship [56]. These workloads cover a variety of benchmarks including commercial applications and selected benchmarks from the PARSEC, SPEC, and Biobench suites. Furthermore, we use 12 kernel attacks to mimic malicious codes in Section VIII-D.

One metric used to compare different crosstalk mitigation schemes is the crosstalk mitigation refresh power overhead (CMRPO). The CMRPO is the average power consumed for deciding which rows to be refreshed in order to mitigate crosstalk. It is computed relative to the regular refresh power in the absence of any crosstalk mitigation (2.5mW to refresh 64K rows during a 64 ms refresh interval [17, 49]).

While rows are refreshed in a bank to mitigate crosstalk, that bank cannot be accessed, possibly delaying subsequent memory requests to that bank. To estimate this delay, we define the execution time overhead (ETO) as the delay in execution time due to memory requests to banks being refreshed (to mitigate crosstalk) relative to the execution time when no provisions are made to mitigate crosstalk.

Table I: System Configuration

Processor	Two 3.2GHz cores, Memory bus speed: 800 MHz 128-entry ROB, Fetch width: 4, Retire width: 2 Pipeline depth: 10
Memory controller	Bus freq.: 800 MHz, Write queue capacity: 64 Address mapping: rw:rk:bk:ch:col:offset Management policy: closed-page with FRFCFS
DRAM	2 channels(each 8GB DIMM), 1 rank/channel 8 banks/rank, 64K rows/bank, 64B cache line

VII. EVALUATION

We compare crosstalk mitigation schemes: DRCAT, PRCAT, SCA (implemented with SRAM) and PRA (refreshes two victim rows but not the aggressor row). In this section, we conduct experiments on a dual-core system using refresh thresholds of $T=32K$ and $T=16K$ and a maximum of $L=11$ levels for DRCAT and PRCAT. In Section VIII-A, we will study the effect of the maximum number of CAT levels and the value of the refresh thresholds on power and performance. Moreover, we will report results for quad-core systems. We assume that either the memory controller knows which rows are physically adjacent to each other [57] or the DRAM chip is responsible for refreshing the row and its neighbors [58].

A. Hardware Overhead

Table II shows the hardware cost for managing and maintaining the counters for SCA, DRCAT and PRCAT with $L=11$ levels and $T = 32K$ as the number of counters per bank ranges from 32 to 512. We separately report the dominant sources of hardware energy overhead. These sources include: (1) the dynamic energy per access of the designed circuits plus the SRAM storage, and (2) the static energy during a 64 ms refresh interval of circuits plus the SRAM storage. The SRAM energy is extracted from CACTI [45] and the circuit energy (combinational and io-pad) is derived from Synopsys PrimeTime. Note that for DRCAT and PRCAT, the dynamic energy per memory access accounts for multiple accesses to SRAM (from 2 to $L - \log(M/4)$) while for SCA, SRAM is accessed only twice to read and write the counters. A

modified version of Table II is used for DRCAT and PRCAT when the maximum tree depth changes in the experimental tests.

The results show that the dynamic energy per access of PRCAT is roughly twice that of SCA for the same number of counters. With respect to area overhead and static energy, Table II clearly shows that PRCAT and SCA occupy equal area and consume similar static power when the number of counters of SCA is twice that of PRCAT. For example, $PRCAT_{64}$ and SCA_{128} occupy iso-area. Moreover, this area is one order of magnitude smaller than the area needed by the leading counter-based approach that stores in memory one counter per row and uses a 32KB on-chip counter cache [26] (equivalent storage to 2,048 counters per bank). Thus, implementing 64 or even 256 counters per bank is feasible. Our implementation shows that the average latency for PRCAT is 3.6ns (circuit latency plus repeated access to SRAM) which is much lower than the row activation latency in the DRAM memory [59].

In comparison to PRCAT for $T=32K$, DRCAT uses a 2-bit weight register per counter to reconfigure the structure of CAT. The results in Table II show that the circuit design and SRAM storage of DRCAT, on average, augments 4.2% area overhead to the system compared to PRCAT. Also, DRCAT increases the dynamic energy per row access by 5% over PRCAT. Furthermore, it incurs 4ns latency. When DRCAT reconfigures counters, its latency is about 7.5ns. The main reason for the extra latency is the traversal of the tree as explained in Section V-B. However, updating the DRCAT and accessing the memory can be done in parallel.

Table II also shows the specification of a PRNG [25] for PRA in 45nm technology⁶. We select one PRNG for PRA that is applied for all banks during row accesses. The energy per bit (the efficiency) for PRNG is computed as $Power/Throughput$. For $p = 0.002$ and $p = 0.003$, PRNG generates 9 bits ($\sim \log(1/0.003)$ or $\log(1/0.002)$) so that PRA can decide if victim rows should be refreshed when a row is accessed. The energy for generating 9 random bits is denoted by eng_PRNG . A similar conclusion was reached in [16].

B. CMRPO

We use the results shown in Table II to compute CMRPO for a benchmark during its execution by adding the following components needed to mitigate crosstalk: (1) The dynamic power (product of dynamic energy per memory access and the total number of memory accesses during execution divided by the execution time), (2) the static power (static energy during a refresh interval divided by the refresh interval), and (3) the refresh power (product of the average number of rows refreshed to prevent crosstalk with the energy to refresh one row (1nJ per row [60]) divided by the execution time).

Figure 8 shows the CMRPO for different approaches when $T = 32K$. It reveals that both $DRCAT_{64}$ and $PRCAT_{64}$ with $L=11$ achieve a CMRPO of 4%, which is an improvement over the 11% in the cases of PRA and SCA. Note that the CMRPO for PRA includes refreshing an average of two victim rows every 500 accesses and generating 9 PRNG bits every access, with the PRNG generation being dominant.

⁶An PRNG design with low static power is reported in [24]. However, this design is much slower than the design in [25] which leads to a larger Energy/bit consumption.

Table II: Hardware energy (per bank) and area of DRCAT, PRCAT and SCA for different number of counters, M , and the specification of the PRNG used for PRA [25]. The reported energy for PRNG (eng_PRNG) is for generating 9-bits per row access.

M	Energy:dynamic (nJ per row access) and static (nJ per refresh interval)						Area (mm ²)				
	DRCAT		PRCAT		SCA		DRCAT	PRCAT	SCA		
	dynamic	static	dynamic	static	dynamic	static					
32	3.05E-04	5.77E+03	2.91E-04	5.55E+03	1.41E-04	3.16E+03	3.16E-02	3.04E-02	1.86E-02	Area	4.004E-3
64	4.30E-04	1.39E+04	4.09E-04	1.32E+04	1.92E-04	8.81E+03	6.12E-02	5.86E-02	4.04E-02	Throughput(Gbps)	2.4
128	5.83E-04	2.77E+04	5.50E-04	2.63E+04	2.22E-04	1.44E+04	1.16E-01	1.11E-01	6.04E-02	Power(mW)	7
256	8.72E-04	5.44E+04	8.25E-04	5.13E+04	3.12E-04	2.39E+04	2.23E-01	2.11E-01	1.00E-01	Eff.(nj/b)	2.90E-3
512	1.17E-03	1.06E+05	1.10E-03	1.02E+05	4.25E-04	4.52E+04	3.93E-01	3.75E-01	1.72E-01	eng_PRNG(nj)	2.625E-2

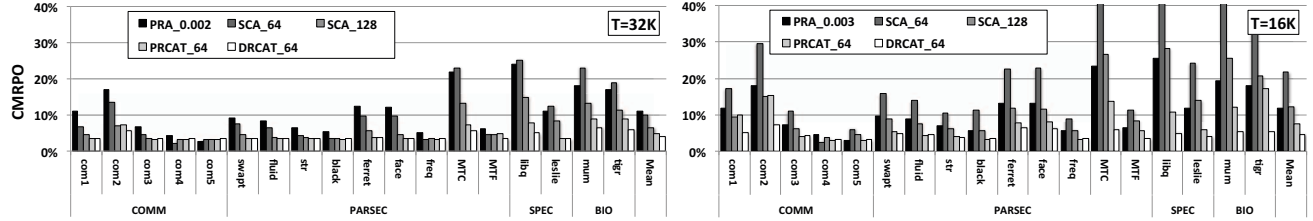


Figure 8: The CMRPO (as a percent of the regular refresh power). DRCAT and PRCAT use 64 counters and up to 11 levels.

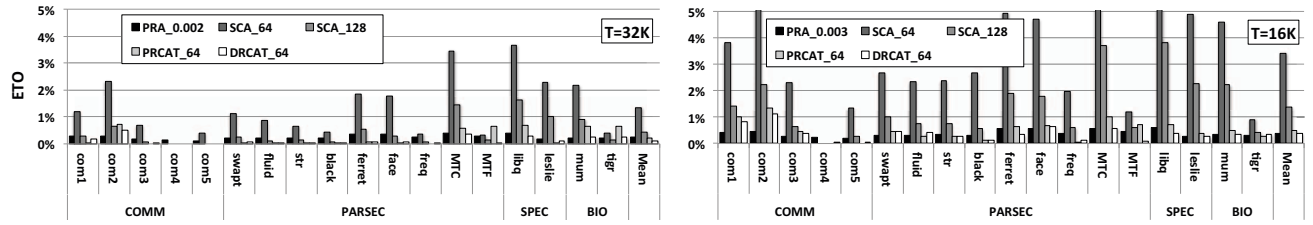


Figure 9: ETO resulting from refreshing vulnerable rows. DRCAT and PRCAT use 64 counters and up to 11 levels.

According to Table II, on average, for every 50 row accesses, PRA consumes energy equal to that of refreshing one row in DRAM.

For $T=16K$, we use $PRA_{0.003}$, rather than $PRA_{0.002}$ since the probability of failure for $PRA_{0.002}$ is greater than $1E-4$ (Chipkill reliability) according to Figure 1. Figure 8 shows that CMRPO for $DRCAT_{64}$ in dual-core systems is 4.5%, which is an improvement over the 12% and 22% incurred in $PRA_{0.003}$ and SCA_{64} , respectively. Also, considering iso-area, $DRCAT_{64}$ achieves a CMRPO improvement over the 13% incurred in SCA_{128} . Figure 8 indicates that reducing T from 32K to 16K will increase considerably CMRPO for SCA while slightly increasing CMRPO for PRCAT and DRCAT.

C. Execution Time Overhead

To evaluate performance, we report the execution time overhead (ETO) resulting from refreshing victim rows. When rows vulnerable to crosstalk are refreshed, any read or write request to the bank containing the refreshed rows is stalled, which leads to the execution time overhead.

Figure 9 shows the ETO for different workloads. For $T = 32K$, $PRA_{0.002}$, SCA_{64} , SCA_{128} , $PRCAT_{64}$ and $DRCAT_{64}$ incur low ETO of 0.26%, 1.32%, 0.43%, 0.23%, and 0.16% respectively. For $T = 16K$, the ETOs of $PRA_{0.003}$, SCA_{64} , SCA_{128} , $PRCAT_{64}$ and $DRCAT_{64}$ are 0.39%, 3.42%, 1.38%, 0.49% and 0.35% respectively. Note that ETO for $PRA_{0.003}$ when $T = 16K$ is roughly 1.5 times larger than ETO for $PRA_{0.002}$ when $T = 32K$ because it probabilistically refreshes 50% more rows. On the other hand, ETO for SCA_{128} when

$T = 16K$ is higher than ETO for SCA_{64} when $T = 32K$. This shows that when the refresh threshold is reduced, doubling the number of counters statically does not reduce the number of refreshed rows, which results in less accurate row tracking and thus larger refresh energy.

VIII. SENSITIVITY STUDY

A. Sensitivity to the Number of Counters and the Maximum CAT depth

Figure 10 shows CMRPO for DRCAT when the number of counters changes from 32 to 512 and the number of levels changes from 6 to 14, and compares results with those of SCA. From the figure, we note that increasing the number of CAT levels does not significantly impact CMRPO when the number of counters is relatively large. This is because, in this case, the static power consumed by the counters dominates the CMRPO, and hence, any improvement in the number of refreshed rows has minimal effect. Conversely, with a small number of counters, the energy for refreshing vulnerable rows is a large component of the CMRPO. Thus, having more levels in the tree saves refresh energy by targeting vulnerable row.

Due to the trade-off between static power and the power consumed to refresh vulnerable rows, the minimum CMRPO happens when DRCAT employs 64 counters and when SCA employs 128 counters for $T=32K$. Note that the refresh power of $DRCAT_{64}$ with L7 is close to SCA_{64} since it only increases row resolution one more level beyond SCA_{64} . However, $DRCAT_{64}$ incurs more static and dynamic power than SCA_{64} ; hence, its CMRPO is larger. The same argument applies to

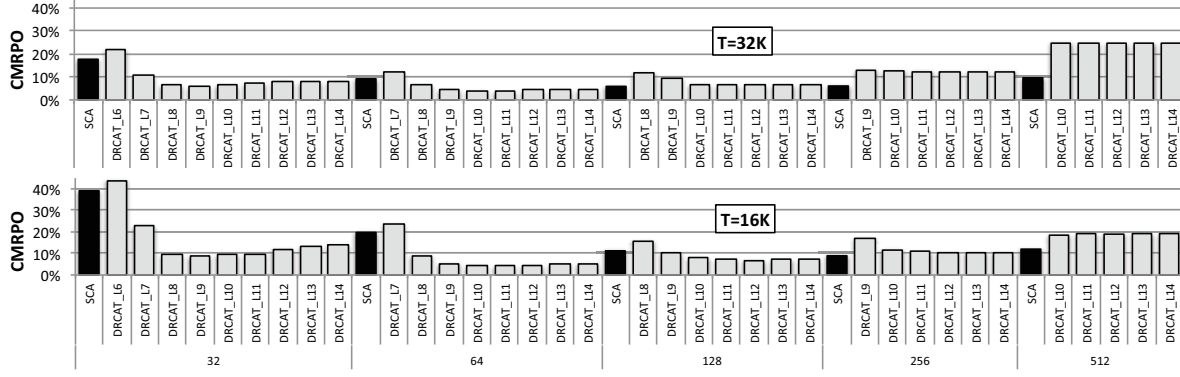


Figure 10: Crosstalk mitigation power overhead per bank for DRCAT using from 32 to 512 counters and different maximum CAT levels (6 to 14).

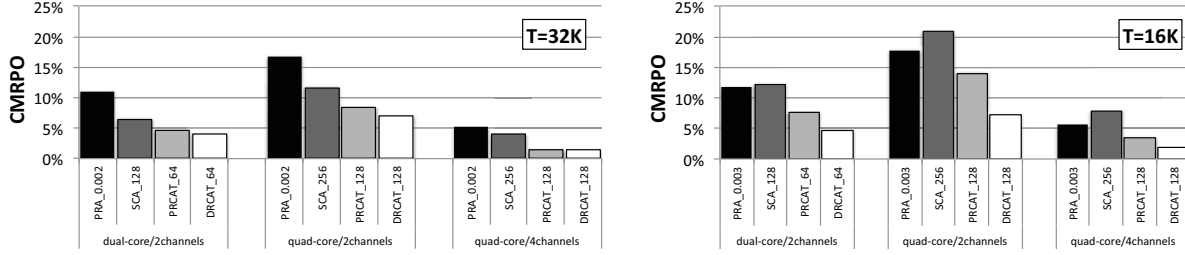


Figure 11: Effect of different mapping policies and number of cores on CMRPO (per bank). The banks in dual core and quad core systems include 64K and 128K rows, respectively.

explain why for fewer counters, CMRPO of SCA_{32} is smaller than that of $DRCAT_{32}$. When the threshold decreases from 32K to 16K, SCA will refresh victim rows more frequently and its CMRPO grows by 12% while the minimum CMRPO of $DRCAT_{64}$ changes very little.

We studied the sensitivity of ETO to the number of counters and the tree depth (the results are not shown in this paper). The key observation is that, for both refresh thresholds, when using at least 64 counters and $L \geq 9$, DRCAT incurs an ETO $< 1\%$. Results also show that ETO is inversely correlated to the refresh threshold. Another observation is that for a given fixed number of counters, increasing the tree depth does not necessarily reduce the number of refreshed victim rows; with a deeper tree, the number of rows associated with a certain counter will be reduced, but the number of rows associated with other counters will increase. In other words, trying to be precise in one area of the memory may lead to a gross imprecision in another area of the memory, which creates a trade-off that leads to an optimum value for the maximum tree depth.

We conclude that for DRCAT, the optimal number of counters and the maximum CAT depth affect both the CMRPO and the ETO. For $T = 32K$ and $T = 16K$ and using between 32 and 128 counters, a maximum of $L = 11$ levels minimizes CMRPO and results in a low ETO. For CAT with more counters, the maximum CAT depth is inconsequential for CMRPO. In fact, using DRCAT leads to larger CMRPO than using SCA. We did the same analysis for PRCAT and our results show that CMRPO for PRCAT is about 4% and 7% for $T=32K$ and $T=16K$ with 10 and 11 CAT levels, respectively. Also it incurs very low performance overhead ($< 0.5\%$) for both thresholds.

B. Sensitivity to Mapping Policy and Number of Cores

To analyze the effect of address interleaving, we experiment with dual-core systems using two standard mapping policies of USIMM [47]: (1) the 2-channel mapping policy (used in the experiments so far) and (2) a 4-channel mapping policy that maximizes memory access parallelism. Note that when keeping the size of each memory bank fixed, the 4-channel policy in USIMM quadruples the number of banks in the system. We also experiment with a quad-core system using the 2-channel and 4-channel mapping policies. The CMRPO of DRCAT, PRCAT and SCA are reported in Figure 11 for iso-area storage. Figure 11 shows that, when using the 2-channel mapping policy, the CMRPO for quad-core systems is larger than the dual-core systems. This is because having more cores reduces the spatial locality in the L2 cache, thus generating more memory traffic and forcing more refreshes. SCA is affected more than the other schemes by the increased traffic because of the inability to accurately track the row accesses due to the uniform distribution of counters to rows. This effect is amplified when $T = 16K$ resulting in the CMRPO for SCA exceeding that of PRA for the quad-core system. In this case, DRCAT reduces the CMRPO in quad-core systems to 7%, which is an improvement over the 21% and 18% incurred in SCA and PRA, respectively. Figure 11 shows that for quad-core systems, the 4-channel policy reduces CMRPO versus the 2-channel policy for all schemes. This is expected since in the 4-channel policy, the number of banks increases from 16 to 64, thus decreasing the number of refreshed rows.

Although we do not show the results for ETO in this section, we should note that ETO remains low for all schemes irrespective of the mapping policy or the number of cores.

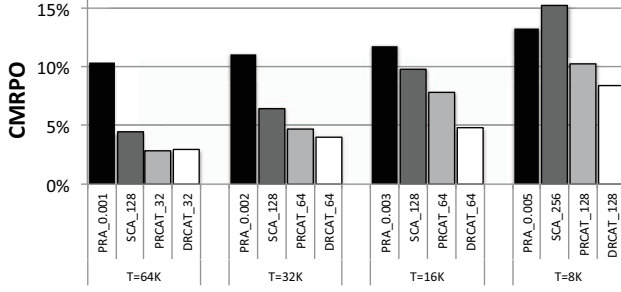


Figure 12: CMRPO for refresh thresholds $T = 64K/32K/16K/8K$.

The largest ETO is incurred when the 2-channel policy is used with quad-cores and $T = 16K$. Specifically, in this case ETOs for PRA_{0.003}, SCA, PRCAT and DRCAT are 0.47%, 1.45%, 0.6%, 0.38% respectively. The relatively high ETO for SCA is due to the fact that the number of refreshed rows is relatively high.

C. Sensitivity to Refresh Thresholds

Scaling down DRAM technology exacerbates the crosstalk problem leading to a decrease in the refresh threshold [26]. This motivates the sensitivity analysis on different refresh thresholds presented in Figure 12, which shows the CMRPO for four refresh thresholds on a dual-core system with the 2-channel mapping policy. We used PRA_{0.001}, PRA_{0.002}, PRA_{0.003} and PRA_{0.005} for $T = 64K, 32K, 16K$ and $8K$, respectively to guarantee that the unsurvivability is better than $1.0E-4$. The figure shows that, for thresholds 64K to 16K and dual core systems, DRCAT incurs CMRPO less than 5% which is an improvement over PRA's 12%. Also, it improves the CMRPO over PRCAT because the CAT is dynamically reconfigured rather than being periodically reset. Note that for $T=8K$, DRCAT and PRCAT need to double the number of counters to mitigate crosstalk, but still incur less than 10% CMRPO. With respect to ETO, all approaches incur very low overhead. Specifically, for $T = 8K$, the ETOs for PRA, SCA, PRCAT, DRCAT are 0.58%, 1.44%, 0.8%, and 0.48%, respectively. We conclude that CAT improves CMRPO relative to the other schemes for both current and future technologies.

D. Performance Under Malicious attacks

To evaluate the performance of the counter-based approaches under malicious attacks, we use 12 kernel attacks [16] that randomly select few target rows (4 rows per bank and a total of 64 target rows for 16 banks with dual-core/2-channels configuration) and access the target rows more frequently than other rows in DRAM. We integrate the kernel attacks with regular access rows of memory-intensive workloads (which we call benign workloads). We select three attack modes *Heavy* (75% target rows + 25% benign access rows), *Medium* (50% target rows + 50% benign access rows) and *Light* (25% target rows + 75% benign access rows). Note that the distribution of target rows in the kernel attacks follows the Gaussian distribution. Figure 13 shows the average execution time overhead for the benign workloads for three refresh thresholds. As expected, more intensive attacks leads to higher ETO since it causes more refreshes. While the ETO for PRCAT and DRCAT is less than 0.9% and 0.6% for different attacks and refresh thresholds, the ETO of SCA

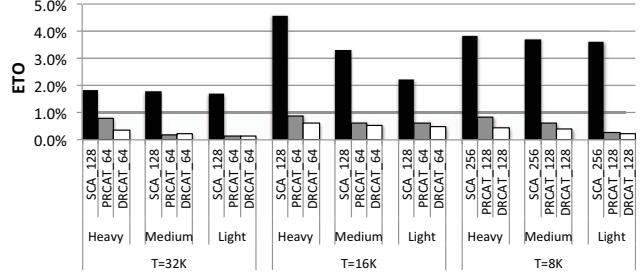


Figure 13: ETO for three kernel attack modes: Heavy (75% target rows + 25% benign access rows), Medium (50% target rows + 50% benign access rows) and Light (25% target rows + 75% benign access rows).

grows to 4.5% for $T=16K$ under heavy attacks. ETO for $T = 8K$ is lower than for $T = 16K$ because the number of counters is doubled.

We conclude that when malicious attacks target specific rows in DRAM, CAT-based approaches are more efficient than SCA approaches at mitigating the attacks since they confine attacked rows to smaller groups of rows to be refreshed.

IX. CONCLUSION

We introduce the notion of a tree-based non-uniform row partitioning for detecting rows vulnerable to crosstalk in memory banks. We develop a low-cost implementation of this notion with three key ideas: (1) we propose a low-cost implementation to maintain and access Counter-based Adaptive Trees that assign counters to rows non-uniformly and detects more precisely rows vulnerable to crosstalk. (2) We introduce a scheme to compute the split thresholds that cause the trees to dynamically evolve and match the row access patterns. (3) We introduce a scheme, DRCAT, for dynamically reconfiguring the CAT to track the temporal changes in memory access patterns resulting from either changing the running applications or changing the phases of a running application.

Our results show that DRCAT outperforms the leading approaches for wordline crosstalk mitigation. Specifically, for quad-core systems and refresh threshold of $T = 16K$, DRCAT reduces the CMRPO to 7%, which is an improvement over the 21% and 18% incurred in deterministic and probabilistic approaches, respectively. Moreover, DRCAT incurs very low performance overhead ($< 0.5\%$). Hence, we conclude that dynamic row partitioning is an effective solution to detect rows vulnerable to crosstalk in DRAM. Clearly, this hardware solution avoids wordline crosstalk during normal execution and protects against malicious attacks that explore vulnerability to wordline crosstalk.

X. ACKNOWLEDGEMENTS

We thank the anonymous reviewers for their feedback. This work is supported by CS50 merit pre-doctoral fellowship award from the university of Pittsburgh.

References

- [1] Stavros Volos, Djordje Jevdjic, Babak Falsafi, and Boris Grot. An effective dram cache architecture for scale-out servers. Technical report.
- [2] W Mueller and et al. Challenges for the dram cell scaling to 40nm. In *IEDM 2005*.
- [3] Xianwei Zhang, Youtao Zhang, Bruce R Childers, and Jun Yang. Restore truncation for performance improvement in future dram systems. In *HPCA 2016*.

- [4] Saurabh Sinha, Greg Yeric, Vikas Chandra, Brian Cline, and Yu Cao. Exploring sub-20nm finfet design with predictive technology models. In *DAC 2012*.
- [5] Qingyuan Deng, David Meisner, Luiz Ramos, Thomas F Wenisch, and Ricardo Bianchini. Memscale: active low-power modes for main memory. In *ACM SIGPLAN Notices 2011*.
- [6] Janani Mukundan, Hillery Hunter, Kyu-hyoun Kim, Jeffrey Stuecheli, and José F Martínez. Understanding and mitigating refresh overheads in high-density ddr4 dram systems. *ACM SIGARCH CAN 2013*.
- [7] Kinam Kim et al. Technology for sub-50 nm dram and nand flash manufacturing. *IEDM Tech. Dig 2005*.
- [8] SY Cha. Dram and future commodity memories. *VLSI Technology Short Course 2011*.
- [9] Jack A Mandelman, Robert H Dennard, Gary B Bronner, John K DeBrosse, Rama Divakaruni, Yujun Li, and Carl J Radens. Challenges and future directions for the scaling of dynamic random-access memory (dram). *IBM Journal of Research and Development 2002*.
- [10] Yasuhiro Konishi, Masaki Kumanoya, Hiroyuki Yamasaki, Katsumi Dosaka, and Tsutomu Yoshihara. Analysis of coupling noise between adjacent bit lines in megabit drams. *IEEE Journal of Solid-State Circuits 1989*.
- [11] Kyungbae Park, Chulseung Lim, Donghyuk Yun, and Sanghyeon Baeg. Experiments and root cause analysis for active-precharge hammering fault in ddr3 sdrum under $3 \times$ nm technology. *Microelectronics reliability 2016*.
- [12] Kyungbae Park, Sanghyeon Baeg, ShiJie Wen, and Richard Wong. Active-precharge hammering on a row induced failure in ddr3 sdrums under $3 \times$ nm technology. In *IIRW 2014*.
- [13] Yoongu Kim, Ross Daly, Jeremie Kim, Chris Fallin, Ji Hye Lee, Donghyuk Lee, Chris Wilkerson, Konrad Lai, and Onur Mutlu. Flipping bits in memory without accessing them: An experimental study of dram disturbance errors. In *ACM SIGARCH CAN 2014*.
- [14] Daniel Gruss, Clémentine Maurice, and Stefan Mangard. Rowhammer.js: A remote software-induced fault attack in javascript. *arXiv preprint arXiv:1507.06955*, 2015.
- [15] Zelalem Birhanu Aweke, Salessawi Ferede Yitbarek, Rui Qiao, Reetuparna Das, Matthew Hicks, Yossi Oren, and Todd Austin. Anvil: Software-based protection against next-generation rowhammer attacks. In *ASPLOS 2016*.
- [16] Mohsen Ghasempour, Mikel Lujan, and Jim Gar-side. Armor: A run-time memory hot-row detector., <http://apt.cs.manchester.ac.uk/projects/ARMOR/RowHammer/index.html>. Accessed: 2015-08-11.
- [17] Amir Rahmati, Matthew Hicks, Daniel Holcomb, and Kevin Fu. Refreshing thoughts on dram: Power saving vs. data integrity. In *WACAS 2014*.
- [18] Jagadish B Kotra, Narges Shahidi, Zeshan A Chishti, and Mahmut T Kandemir. Hardware-software co-design to mitigate dram refresh overheads: A case for refresh-aware process scheduling. In *ASPLOS 2017*.
- [19] Joohee Kim and Marios C Papaefthymiou. Block-based multiperiod dynamic memory design for low data-retention power. *VLSI 2003*.
- [20] Prashant Nair, Chia-Chen Chou, and Moinuddin K Qureshi. A case for refresh pausing in dram memory systems. In *HPCA 2013*.
- [21] Taku Ohsawa, Koji Kai, and Kazuaki Murakami. Optimizing the dram refresh count for merged dram/logic lsis. In *ISLPED 1998*.
- [22] Mohammad Arjomand, Mahmut T Kandemir, Anand Sivasubramaniam, and Chita R Das. Boosting access parallelism to pcm-based main memory. In *ISCA 2016*.
- [23] Song Liu, Karthik Pattabiraman, Thomas Moscibroda, and Benjamin G Zorn. Flicker: saving dram refresh-power through critical data partitioning. *ACM SIGPLAN Notices 2012*.
- [24] Kaiyuan Yang, David Fick, Michael B Henry, Yoonmyung Lee, David Blaauw, and Dennis Sylvester. 16.3 a 23mb/s 23pj/b fully synthesized true-random-number generator in 28nm and 65nm cmos. In *ISSCC 2014*.
- [25] Suresh Srinivasan and et al. 2.4 ghz 7mw all-digital pvt-variation tolerant true random number generator in 45nm cmos. In *VLSI 2010*.
- [26] Dae-Hyun Kim, Prashant J Nair, and Moinuddin K Qureshi. Architectural support for mitigating row hammering in dram memories. In *CAL 2015*.
- [27] Micron Technology. Micron inc. DDR4 SDRAM MT40A2G4, MT40A1G8, MT40A512M16 data sheet, 2015.
- [28] Min Kyu Jeong, Doe Hyun Yoon, Dam Sunwoo, Mike Sullivan, Ikhwon Lee, and Mattan Erez. Balancing dram locality and parallelism in shared memory cmp systems. In *HPCA 2012*.
- [29] Satoshi Imamura, Yuichiro Yasui, Koji Inoue, Takatsugu Ono, Hiroshi Sasaki, and Katsuki Fujisawa. Power-efficient breadth-first search with dram row buffer locality-aware address mapping. In *GDDMM 2016*.
- [30] Aditya Agrawal, Amin Ansari, and Josep Torrellas. Mosaic: Exploiting the spatial locality of process variation to reduce refresh energy in on-chip edram modules. In *HPCA 2014*.
- [31] Seyed Mohammad Seyedzadeh, Alex K Jones, and Rami Melhem. Counter-based tree structure for row hammering mitigation in dram. In *CAL 2017*.
- [32] Kuljit S Bains and John B Halbert. Distributed row hammer tracking, March 29 2016. US Patent 9,299,400.
- [33] Zvika Greenfield, Kuljit S Bains, Theodore Z Schoenborn, Christopher P Mozak, and John B Halbert. Row hammer condition monitoring, January 20 2015. US Patent 8,938,573.
- [34] Mazen Kharbutli and Yan Solihin. Counter-based cache replacement algorithms. In *ICCD 2005*.
- [35] Mark Seaborn and Thomas Dullien. Exploiting the dram rowhammer bug to gain kernel privileges. *Black Hat*, 2015.
- [36] Kirill A Shutemov. Pagemap: Do not leak physical addresses to non-privileged userspace. In *Retrieved on November 2015*.
- [37] Erik Bosman, Kaveh Razavi, Herbert Bos, and Cristiano Giuffrida. Dedup est machina: Memory deduplication as an advanced exploitation vector. In *SP 2016*.
- [38] Daniel Gruss, Clémentine Maurice, and Stefan Mangard. Rowhammer.js: A remote software-induced fault attack in javascript. In *DIMVA 2016*.
- [39] N Herath and A Fogh. These are not your grand daddy's cpu performance counters: Cpu hardware performance counters for security. *Black Hat*, 2015.
- [40] <https://users.ece.cmu.edu/~koopman/lfsr/>.
- [41] http://se.mathworks.com/matlabcentral/mlc/downloads/downloads/submissions/22716/versions/3/previews/coding_1_02/lfsr.m/index.html.
- [42] Sheng-hua Zhou, Wancheng Zhang, and Nan-Jian Wu. An ultra-low power cmos random number generator. *Solid-State Electronics 2008*.
- [43] Marco Bucci, Lucia Germani, Raimondo Luzzi, Pasquale Tommasino, Alessandro Trifiletti, and Mario Varanonoovo. A high-speed ic random-number source for smartcard microcontrollers. *IEEE Transactions on Circuits and Systems I 2003*.
- [44] Marco Bucci, Lucia Germani, Raimondo Luzzi, Alessandro Trifiletti, and Mario Varanonoovo. A high-speed oscillator-based truly random number source for cryptographic applications on a smart card ic. In *TC 2003*.
- [45] Naveen Muralimanohar, Rajeev Balasubramonian, and Norman P Jouppi. Cacti 6.0: A tool to understand large caches. *University of Utah and Hewlett Packard Laboratories, Tech. Rep*, 2009.
- [46] Bruce Jacob, Spencer Ng, and David Wang. *Memory systems: cache, DRAM, disk*. Morgan Kaufmann, 2010.
- [47] Niladrish Chatterjee, Rajeev Balasubramonian, Manjunath Shevgoor, S Pugsley, A Udipti, Ali Shafiee, Kshitij Sudan, Manu Awasthi, and Zeshan Chishti. Usimm: the utah simulated memory module. *University of Utah, Tech. Rep*, 2012.
- [48] Ishwar Bhati, Mu-Tien Chang, Zeshan Chishti, Shih-Lien Lu, and Bruce Jacob. Dram refresh mechanisms, penalties, and trade-offs. In *TC 2016*.
- [49] 4Gb DDR3 SDRAM - MT41J512M8, Micron Technology. <https://www.micron.com/support/tools-and-utilities/power-calc>. 2011.
- [50] FreePDK45. <http://www.eda.ncsu.edu/wiki/>.
- [51] SeyedMohammad Seyedzadeh, Alex Jones, and Rami Melhem. Enabling fine-grain restricted coset coding through word-level compression for pcm. In *HPCA 2018*.
- [52] Seyed Mohammad Seyedzadeh, Donald Kline Jr, Alex K Jones, and Rami Melhem. Mitigating bitline crosstalk noise in dram memories. In *MEMSYS 2017*.
- [53] Seyed Mohammad Seyedzadeh, Rakan Maddah, Alex Jones, and Rami Melhem. Leveraging ecc to mitigate read disturbance, false reads and write faults in stt-ram. In *DSN 2016*.
- [54] Seyed Mohammad Seyedzadeh, Rakan Maddah, Donald Kline, Alex K Jones, and Rami Melhem. Improving bit flip reduction for biased and random data. In *TC 2016*.
- [55] Seyed Mohammad Seyedzadeh, Rakan Maddah, Alex Jones, and Rami Melhem. Pres: Pseudo-random encoding scheme to increase the bit flip reduction in the memory. In *DAC 2015*.
- [56] Memory Scheduling Championship (MSC). <http://www.cs.utah.edu/rajeev/jwac12/>.
- [57] Ad J van de Goor and Ivo Schanstra. Address and data scrambling: Causes and impact on memory tests. In *DELTA 2002*.
- [58] Kuljit S Bains, John B Halbert, Christopher P Mozak, Theodore Z Schoenborn, and Zvika Greenfield. Row hammer refresh command, January 12 2016. US Patent 9,236,110.
- [59] Wongyu Shin, Jungwhan Choi, Jaemin Jang, Jinwoong Suh, Youngsuk Moon, Yongkee Kwon, and Lee-Sup Kim. DRAM latency optimization inspired by relationship between row-access time and refresh timing. In *TC 2016*.
- [60] Mrinmoy Ghosh and Hsien-Hsin S Lee. Smart refresh: An enhanced memory controller design for reducing energy in conventional and 3d die-stacked drams. In *MICRO 2007*.