Low Cost Power Failure Protection for MLC NAND Flash Storage Systems with PRAM/DRAM Hybrid Buffer

Jie Guo[1], Jun Yang[1], Youtao Zhang[2] and Yiran Chen[1]

[1]Department of Electrical and Computer Engineering, [2]Department of Computer Science

University of Pittsburgh, Pittsburgh, USA

[1]{jig26, juy9, yic52}@pitt.edu, [2]zhangyt@cs.pitt.edu

Abstract − **In the latest PRAM/DRAM hybrid MLC NAND flash storage systems (NFSS), DRAM is used to temporarily store file system data for system response time reduction. To ensure data integrity, super-capacitors are deployed to supply the backup power for moving the data from DRAM to NAND flash during power failures. However, the capacitance degradation of super-capacitor severely impairs system robustness. In this work, we proposed a low cost power failure protection scheme to reduce the energy consumption of power failure protection and increase the robustness of the NFSS with PRAM/DRAM hybrid buffer. Our scheme enables the adoption of the more reliable regular capacitor to replace the super capacitor as the backup power. The experimental result shows that our scheme can substantially reduce the capacitance budget of power failure protection circuitry by 75.1% with very marginal performance and energy overheads.**

I Introduction

The advantages in cost and power consumption enable wide applications of multi-level cell (MLC) NAND flash in consumer electronics such as smart phones. In conventional MLC NAND flash storage systems (NFSS), DRAM is used to buffer the file system data and the metadata of flash translation layer (FTL) to improve the system response time [1]. The volatility of DRAM imposes the potential risks of losing the data and metadata when unexpected power shut-off occurs, leading to the unrecoverable system errors [2]. The current solution is to move the DRAM data into the NAND flash arrays upon the power failure under the backup power, i.e., super-capacitors. However, besides expensive cost, super-capacitors generally suffer from severe aging effect [3], which are the major reliability concerns of NFSS.

Recently, some research was conducted on adopting the emerging phase change random access memory (PRAM) in NFSS design. Compared to NAND flash, PRAM offers many attractive advantages such as non-volatility, in-place update, longer life endurance ($10^7 \sim 10^9$) and lower dynamic power consumption [4]. By leveraging such features, PRAM/DRAM hybrid buffer can greatly improve the performance of NFSS and prolong its lifetime. However, the existing NFSS with PRAM/DRAM buffer still requires super-capacitors for the backup purpose.

In this work, we proposed low cost power failure protection schemes for MLC NFSS with PRAM/DRAM Hybrid Buffer to improve the reliability of NFSS. By reducing the energy consumption of power failure protection, we are able to use highly reliable regular capacitor to replace super-capacitors as backup power. In our scheme, PRAM is used to not only store the FTL metadata during the normal operation, but also backup the DRAM data during the power failure. By moving the data to PRAM instead of NAND flash during the power failures protection, the energy consumption is significantly reduced by eliminating the costly NAND flash operations. The unique contributions of our works are:

1. We quantitatively identified the backup power reliability issue in NFSS with PRAM/DRAM Hybrid buffer;
2. We proposed a scheme that substantially reduces the backup capacitance budget of power failure protection circuitry with negligible performance and energy overheads. Due to the reduction of power failure protection energy, we are able to replace the super capacitor with highly reliable regular capacitor to enhance system robustness. We also address the specific data corruption issues in MLC NAND NFSS.

II. Preliminary

A. MLC NFSS Power Failure

In a conventional MLC NFSS, file system data, including both file system metadata and user data, is buffered in a DRAM in front of NAND flash array [1]. FTL metadata, retaining the mapping info between the logic block address and the physic block address as well as the block information, is also stored in the DRAM [8]. During normal power-down process, the data in the DRAM must be moved to the NAND flash array. Power failure may cause data loss if the data migration does not complete in time.

In MLC NFSS, power failure may even damage the data already stored in NAND flash array. Take 4-level MLC NAND flash as an example, the two bits stored in a memory cell are assigned to two different data pages (called coupled pages or coupled page pair) and programed separately [7]. Whether a page is the 1st page (P1) or the 2nd page (P2) is determined by physical page number. P1 and P2 may not be always adjacent to each other. H.-W. Tseng et al. shows that the interrupted programming operations to P2 caused by power failures damage the P1 data as well [2]. Besides, power failures also cause corruptions of a data page if it is programmed right after an incomplete erase operation.

To prevent data loss and corruption, power failure protection circuitry is deployed in NFSS. Due to the large size of the backed-up data, super-capacitors are usually selected as the backup power for power failure protection [2][8]. However, the capacitance degradation of super-capacitors is a severe reliability issue for NFSS: according to the Arrhenius law lifetime model, the capacitance loss of super-capacitors under 60ºC can be as high as 30% within five years [3].

B. NFSS with PRAM buffer

The adoption of PRAM in NFSS has been actively studied in recent research. In [4], G. Sun et al. proposed a PRAM based data buffer to extend the lifetime of NFSS. In [5], J. Kim et al. designed a PRAM/DRAM hybrid buffer, named *hftl*, to enhance the performance of an embedded NFSS. File system metadata and user data are buffered in a DRAM first. Then the user data is moved into NAND flash arrays, while the file system metadata retained in a PRAM. FTL metadata is stored in the PRAM to reduce DRAM size. By leveraging in-place update characteristics of PRAM, *hftl* improves the random access performance and extended lifetime of the NFSS. In [6], D. Liu et al. proposed a wear-leveling scheme to mitigate wear-out of PRAM buffer. However, these works are not dedicated to address the power failure protection issues in the existing PRAM/DRAM buffer.

C. Motivations for Proposed Scheme

In the NFSS with DRAM buffer, data is flushed to NAND flash upon power failure. Power down energy E_{PD} can be expressed by:

$$E_{PD} = (P_{DRAM \to NAND} + P_{ctrl} + P_{Static}) \times t_{PD}. \qquad (1)$$

Here P_{ctrl} is the dynamic power consumption of the control circuitry which incorporates FTL. P_{Static} represents the static power of the circuitry. P_{DRAM_NAND} is the power consumption of the data migration from DRAM to NAND flash. t_{PD} is the power-down process time, which is determined by the data transferring bandwidth between DRAM and NAND flash as well as the possible garbage collection operations when the power failure occurs. Note that the garbage collection upon the power failure, which costs one erase cycle at most, must complete before the data migration starts. Thus, we have:

$$t_{PD} = \frac{S_{data}}{BW_{nand,write}} + t_{ber}. \qquad (2)$$

Here S_{data} is size of the DRAM data that has to be moved to the NAND flash in power down process. $BW_{nand,write}$ is the maximum write bandwidth of NAND flash arrays. t_{ber} is one erase cycle of NAND flash.

The maximum capacitance budget C_{sc} to sustain the power-down process can be calculated by [10]:

$$E_{PD} = \frac{1}{2}\left[C_{sc}V_{sc}^2 - C_{sc}V_{op}^2\right] \cdot k. \qquad (3)$$

Here V_{op} is the minimum capacitor voltage required by the power failure protection circuitry. V_{sc} is the charging voltage of the capacitor. k is the energy conversion efficiency, which is determined by the energy loss on the equivalent series resistance (ESR) and DC-DC converter. A typical value of k is ~0.9 [10]. Usually a 20% margin is added to C_{sc} to overcome the parametric variability of capacitors.

Eq. (1) – (3) can be used to calculate the capacitance budget of the power failure protection circuitry. For example, if we assume V_{op} equals the operating voltage of the NAND flash (3.3V) and V_{sc} is the power supply voltage (5V), the maximum capacitance budget of a 256GB NFSS with 8MB DRAM is 56.3mF based on the parameters shown in TABLE I. Only super-capacitor can provide such high capacitance.

In this work, we proposed low cost power failure protection

scheme to reduce the capacitance budget (C_{sc}) by leveraging the non-volatility of PRAM. As we shall show in the next section, redirecting the data backup destination from NAND flash to PRAM will not automatically guarantee the data integrity of a MLC NFSS. Thus, an enhancement on the algorithm of the data migration during the power-down process is required. The reduction of capacitance budget minimizes the power failure protection cost of the NFSS and enables the possibility of replacing the super-capacitors with the more reliable regular capacitors. Compared with super-capacitors, regular capacitors can offer a wider operating temperature range (125ºC vs. 70 ºC). Their MTTF is up to 80 years [11] whereas lifetime of super-capacitors heavily relies on temperature and voltage.

TABLE I NAND flash parameters

Capacity	Block size	Block number	Page size
	1MB	8192	8KB
Timing	*Max. Program Latency*	*Max. Read latency*	*Max. Erase Latency*
	2.5ms	50us	10ms
Power consumption	*Peak erase current (mA)*	*Peak program current (mA)*	*Peak read current (mA)*
	30	30	30

III. Low Cost Power Failure Protection Scheme

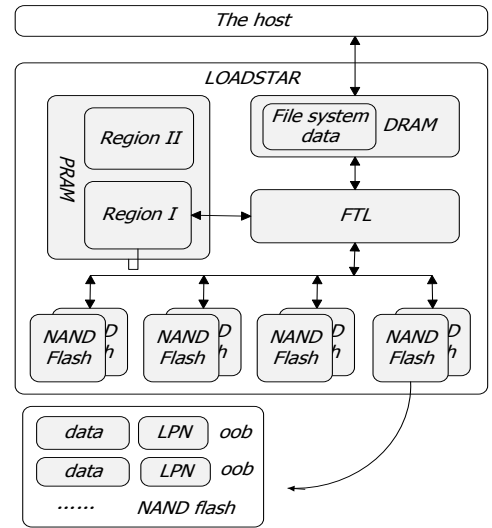### A. Overview of Proposed Scheme

Figure 1 architecture of proposed scheme

The overall architecture of low cost power failure protection scheme is shown in Figure 1. A DRAM is used to buffer the file system data before they are flushed into NAND flash array. The corresponding logic page numbers of the file system data are stored in out-of-band (OOB) area. A PRAM is partitioned into two regions. Region I stores FTL metadata (including page map table and bit map table) and Region II is reserved for the backup purpose. Page map table maintains the mapping relationships between the logical and physical page numbers in the NAND flash array. Bit map table records the validity of each NAND flash page. In

normal operations, only the data in the Region I is accessed by the FTL. When power failure occurs, the DRAM data which has not been flushed into the NAND flash array will be copied to the Region II instead of the NAND flash arrays. We note that in reality, region II can utilize the reserved garbage recycling space and does not necessarily incur the increase of PRAM capacity. In normal NFSS designs, not all NAND flash space is visible to file system because part of it needs to be reserved for garbage collection [5]. As an example, in a 256GB NFSS, 25% NAND flash space must be reserved to suppress the performance degradation induced by garbage recycling. A 256MB PRAM is needed to hold a page map table and a bitmap table with a capacity of 192MB and 8MB, respectively. Then the rest space (56MB) of the PRAM can be used for region II. A simple cache LRU (least-recently-used) based replacement policy is applied to the DRAM: when the size of the buffered data is larger than the pre-defined threshold T_{data}, the least-recently accessed data is evicted to the NAND flash array.

B. Details of Proposed Scheme

It is known that the programming of NAND flash is power consuming: according to [12], the peak operating current of a NAND flash array with 8 chips can be up to 0.5A (or 1.65W), which is around 6.6 times of the PRAM with the same write bandwidth. The energy consumption of the NFSS power-down process will be saved by moving the DRAM data into the Region II of the PRAM. All the operations to the NAND flash are terminated immediately. However, as we discussed in Section II-*A*, sudden interruptions of the write/erase operations will result in the data corruption in a MLC NAND flash. Therefore, we propose the following control scheme to guarantee the data integrity: during the normal operations of NFSS, the data of a P1 page is retained in the DRAM until the programming to its P2 page in the NAND flash array is completed. When power failure occurs, all the DRAM data that have not been copied into the NAND flash array will migrate to the PRAM, including the data that are being written into the NAND flash. Also, if a data being written into the NAND flash is written to a P2 page, the data of its corresponding P1 page must also be moved to the PRAM.

TABLE II Restoring registers

Register	function	Size
$N_{data,w}$	The number of backup pages in PRAM	2Bytes
OTR	Operation Type Register: It records the ongoing operation to a NAND flash chip	3bits
CPN	Current Page Number: it records the physical page number being operated.	4Bytes
RF	Rewrite Flag: it indicates whether there is any P1 page that needs to be reprogrammed during power-up process.	1bit

To retrieve the data during power-up process, we define four registers to record the status of the NAND flash at the time of power failures (listed in TABLE II). These registers are stored in the backup region of PRAM. Our scheme first copies all the data in the Region II of the PRAM into DRAM and then handles the follow cases in the different ways:

1. If the ongoing operation to the NAND flash chip during the power failure was a read access or a write access to a P1 page, or there was no any operation to the NAND flash chip, no further actions needs to be taken.

2. If the operation to the NAND flash chip was a write access to a P2 page, we copy its corresponding P1 page from the PRAM to the NAND flash chip.

3. If the operation to the NAND flash chip was an erase operation, the erase operation must be repeated.

Algorithm I.1 Power-down process

Input: $N_{data,w}$
Output: DRAM data migrate into the PRAM
1: check whether $N_{data,w}$ is equal to 0
2:    **if** $N_{data,w}$ = 0 **then**
3:       store OTR and CPN of each flash chip into the PRAM
4:       **if** OTR is write operation and CPN is pointed to P2 **then** move the corresponding P1 page into the PRAM
5:       Set RF 1
6:       **if** there are data in the DRAM that have not stored in or being written to the NAND flash array **then**
7:         backup data and their logic page numbers into the PRAM
8:         record the number of backup data page in $N_{data,w}$
9:       **if** no data needs to be backed up in the DRAM **then**
10:         write 0 to $N_{data,w}$
11:    **if** $N_{data,w}$ ≠ 0 **then**
12:       keep PRAM register value of $N_{data,w}$ and data unchanged
13: **end** power failure protection

Algorithm I.2 Power-up process

Input: $N_{data,w}$, OTR, CPN and RF
Output: upload data in PRAM to DRAM and rewrite data into NAND flash array
1: **if** $N_{data,w}$ ≠ 0 **then**
2:    Load data in the PRAM into the DRAM
3:    Check whether OTR is write/erase operation
4:    **if** OTR is write operation **then**
5:       **if** RF = 1 **then**
6:         Write backed-up P1 into a clean P1 page in NAND flash
7:         Reset RF to 0
8:         Update CPN and OTR
9:    **if** OTR is erasing operation **then**
10:       Re-erase the block which CPN is pointed to
11:       Update CPN and OTR
12:    Reset $N_{data,w}$ to zero
13: **if** $N_{data,w}$ = 0 **then**
14:    no further operation
15: **end** power up process

The power-down and power-up processes of our proposed scheme are depicted in the Algorithm I.1 and Algorithm I.2. Note that our scheme is also resilient to the power failures occurring in the power-up process: When a power failure occurs during the rewrite or re-erase operations in the power-up process, the operations stop immediately. Since the rewrite can be conducted to only a P1 page, it will not lead to the data corruption of any other pages. The rewrite will restart during the next power-up process until it is successfully conducted. If the power failure happens when the PRAM data are being uploaded to the DRAM, the uploading stops immediately. No further actions need to be taken. The uploading will restart during the next power-up process as long as $N_{data,w}$ is not reset to 0. Here $N_{data,w}$ denotes the data page number backed up in the Region II of the PRAM and is retained in a separate register.

The DRAM is divided into three buffers to facilitate the power failure scheme− free buffer, data buffer and frozen buffer. Free buffer holds the data that comes from the host and has already been flushed to the NAND flash array. Data buffer holds the data that comes from the host but has not been flushed to the NAND Flash array, or the data that was loaded from the NAND Flash array. Frozen buffer holds the P1 pages whose P2 pages are not written or being written to the NAND Flash array. Frozen buffer does not directly receive the data from the host or the NAND flash array. Instead, if a page in the data buffer is written to a P1 page in the NAND flash array, it will be moved to the frozen buffer after the write operation completes. If the P2 page of a P1 page in the frozen buffer is successfully written into the NAND flash array, both pages will be moved to the free buffer.

Since part of the DRAM space is reserved for the frozen buffer, the total sizes of the data buffer and the free buffer is smaller than the total capacity of the DRAM. If the DRAM capacity is too small, the reduction of the data buffer size ($N_{data,w,max}$) will significantly degrade the DRAM hit rate. In our designs, the DRAM capacity is increased from the baseline design when it is smaller than a threshold, e.g., 32MB to maintain a reasonable data buffer size. If the DRAM capacity is larger than 32MB, we will only reduce $N_{data,w,max}$ to make the room for the frozen buffer. Since the original DRAM capacity and $N_{data,w,max}$ are large, only very marginal NFSS performance and energy penalties are incurred by the reduction of $N_{data,w,max}$.

C. Endurance of PRAM

In our scheme, two kinds of metadata – bitmap table and map page table are stored in the PRAM. Every write access to the bitmap table invalidates the previously mapped physical page of the current logic page and validates a newly mapped physical page. The flag bits corresponding to the two involved physical pages are inverted. Hence, the total programming cycles of this portion of the PRAM is exactly twice that of the NAND flash array. Since the endurance cycle of PRAM is considerably higher than that of NAND flash, the bitmap table region in the PRAM will wear out after the NAND flash array does. However, the map page table region in the PRAM may potentially face the risk of endurance because some write-intensive workloads or malign attacks may frequently write certain logic pages. The updating of the corresponding lines in the map table region will be significantly higher than any other lines. Since NAND flash array is usually protected by wear-leveling technique, it may be possible that the map page table region in the PRAM will wear out before the NAND flash array.

We propose to use a scheme similar to Delayed Write Policy [13] to minimize the write accesses to the PRAM. A small region in the DRAM is used to buffer the recently updated page map table entries. The entries of this region are evicted under a LRU replacement policy. When a power failure occurs, the buffered page map table entries in the DRAM will be flushed into the PRAM. The required capacity for buffering the page map table entries is very small, i.e., 64KB

for a 32-chip NAND Flash array. The incurred energy overhead for the power-down process is negligible.

We also propose a wear-leveling technique to enhance the PRAM endurance by leveraging the backup region in the PRAM based on the star-gap technique [13]. The page map table is divided into M blocks. A swap block with the same size of the block is also defined. A swap register is used to record the location of the swap block. For each block, a 4-byte counter is assigned to record the write access number. We use star-gap to conduct the wear-leveling within each block. For the wear-leveling between blocks, we compare the write access numbers of each block for every given time interval. If the write access number of a block A is β times greater than that of a block B, we copy the content of block B to the swap block and move the content of block A to block B. Then the block A becomes the swap block. The write operation to the mapped physical pages in the block A now are redirected to the mapped physical pages in the original swap block.

When a power failure occurs, the backup region available is composed of the swap block and all PRAM blocks that are not used for the bitmap table and the page map table. If a block swapping occurs at the power failure, the swapping stops immediately. The swap block may be re-appointed for the data backup. When the average write access numbers of every entry in a block is greater than a threshold a, the block (called semi-obsolete block) is used as the backup region in the PRAM only. Since the occurrence of power failures is infrequent in reality, using semi-obsolete blocks for data backup purpose should not generate severe concerns on the NFSS reliability. In our experiments, we set $M = 64$, $\beta = 50$ and $a = 5 \times 10^7$.

D. Cost and Overhead

Our scheme requires some PRAM space to store the coupled P1 pages of the incomplete programmed P2 as well as the operation registers used for power-up process. The consumed PRAM size can be calculated by:

$$S_{data} = N_{data\_wmax} \cdot (S_{page} + S_{LPN}) + S_{page} \cdot M + M \cdot (S_{CPN} + S_{OTR} + S_{RF}) + S_{Ndw} \quad . (4)$$

Here S_{page} is the data page size. S_{LPN} is the size of a logic page number. S_{Ndw} is the data size of $N_{data,w}$, which is between 0 and the maximum number of the data pages that the DRAM can store ($N_{data,w,max}$). S_{CSN}, S_{OTR} and S_{RF} are the sizes of every CPN, OTR and RF register, respectively. For an 8-chip NFSS with 4MB data buffer, S_{data} is only 4.07MB.

The power-down energy cost of the low cost power failure protection scheme E_{PD} can be calculated as:

$$E_{PD} = (P_{DRAM \rightarrow PRAM} + P_{ctrl}) \times t_{PD}, \quad (5)$$

Since there is no write or erase operation to NAND flash, t_{PD} can be expressed by

$$t_{PD} = \frac{S_{data}}{BW_{PRAM,write}}. \quad (6)$$

During power-up process, rewrite or re-erase operations are performed to NAND flash array and reloads the backup data from the PRAM to the DRAM. The maximum

power-up time t_{PU} is calculated by:

$$t_{PU} = t_{ber} + S_{data} / BW_{P \rightarrow D}. \qquad (7)$$

$BW_{P \rightarrow D}$ is the data transferring bandwidth from the PRAM to the DRAM. Increasing the parallelism of the PRAM design can also improve $BW_{P \rightarrow D}$, and reduces power up time by paying higher peak power. Our scheme may perform re-write or re-erase operations during power-up process, whose overhead is one erase cycle (t_{ber}) at most.

IV. Experimental Results

In this section, we present the simulation results about the effectiveness of our low cost power failure protection. We adopted Flashsim [14] in our FTL level simulations and modified it by adding PRAM/DRAM hybrid buffer and multi-channel NAND flash array access capability. The disk traces extracted under Windows 7 (Win 7) and Red Hat (RHEL 5) operating systems as well as the TPC-C [15] benchmark (TPC-C) for OLTP application are used as the benchmarks to represent various applications.

We simulated two NFSS' with 64GB and 256GB capacities, which respectively represent the lower bound and the upper bound of the data storage capacity for consumer electronic products like iPhone. The 64GB NFSS has a 4-channel-2-way NAND flash array and a PRAM of four 32MB chips. The 256GB NFSS has a 4-way-8-channel NAND flash array and a PRAM of four 32MB chips. The distance between P1 and P2 in a coupled page pair is 6 pages [2]. The sizes of DRAM buffer range from 2MB to 64MB. The performance and power consumption of DRAM buffers are estimated by CACTI 5.3 [16]. NAND flash device parameters are listed in TABLE I. PRAM parameters are derived from [4][17][18] and listed in TABLE III. The power consumption of FTL controller is set to 54mW [8]. When the data buffer size is below 32MB, we increase the DRAM capacity to make the room for frozen buffer and minimize the performance degradation; when the data buffer size is equal to or above 32MB, we keep the DRAM capacity unchanged but only adjust the maximum data buffer size to make the room for frozen buffer.

The total runtime and energy consumptions of different NFSS' at normal operations for all benchmarks are shown in Figure 2 and Figure 3, respectively. Following the increases of DRAM capacity, the total NFSS energy consumption decreases due to the improved DRAM hit rate. When the DRAM capacity is below 32MB, compared to the baseline design that backs up all the DRAM data into the NAND flash array, the extra energy overhead of our scheme due to the increased DRAM size is very marginal: at most 0.55% under all the benchmarks for both 64GB and 256GB NFSS'.

When the DRAM capacity reaches 32MB, the maximum data buffer size is adjusted in our scheme, leading to the potential degradation of DRAM hit rate and performance. The reduction of DRAM hit rate may cause the increases of the accesses to the NAND flash array and system energy consumption. However, for WIN-7, the decrease of the maximum data buffer size results in a negligible run time increase (< 0.08%) and no more than 0.13% energy increase

w.r.t. the corresponding baseline design with a 64MB DRAM. Interestingly, we observe the slight performance improvement when our scheme is applied for the 64GB and 256GB NFSS with 32MB DRAM in WIN7. We believe it is because the parallelism of the data access patterns to the NAND flash array under our scheme is somehow improved, allowing more simultaneous accesses. In RHEL-5, the runtime increases of the proposed NFSS' with a 32MB or 64MB DRAM w.r.t. the baseline are < 0.1% while the energy consumption increases are < 0.3%. In TPC-C, our scheme only causes slight run time increases (< 1.3% for all simulations) and energy overheads (< 1.4% for all simulations). Note that in RHEL-5, we found the NFSS performance degrades with the increase of DRAM buffer size. It is because that the increase of cache hit rate is very marginal but causes more data accesses being allocated to one or two channels. Then the garbage collection overhead is rapidly raised.

Table III 32MB PRAM chip parameters

Capacity	Bank number	Max read Bandwidth	Max write Bandwidth
256Mb	4	266MB/s	9MB/s
Standby power	Write power	Read power	Single word write
0.058mW	91mW	62mW	10μs

Our scheme also dramatically saves the power-down energy of NFSS, as shown in Figure 4. For a 2MB DRAM, The proposed scheme only consumes 29.4% of the power-down energy of the baseline for the 256GB NFSS. Although this benefit becomes less prominent when the DRAM capacity increases due to the relatively longer power-down time, the power-down energy consumption of our proposed scheme is still only about 32.3% of the one of the baseline even for the DRAM capacity of 64MB.

The decrease in the power-down energy consumption also substantially reduces the backup capacitance budget of the proposed scheme: for example, compared to the baseline scheme, with the 64MB DRAM, the capacitance budget reduction is respectively 77.1% and 75.1% for a 64GB NFSS and a 256GB NFSS, respectively. It becomes possible to use on-board Tantalum capacitor array as the backup capacitors becomes possible.

VI. Conclusion

Super capacitor is often used as backup power in MLC NFSS. However, its unreliable performance imposes severe threats to system robustness. In this work, we propose a low cost power failure protection scheme for MLC NFSS with PRAM/DRAM hybrid buffer to back up the data in the PRAM instead of NAND flash during the power failure. The significant energy reduction of power failure protection makes it possible for us to replace super capacitors with more reliable regular capacitor to enhance system robustness. The DRAM and PRAM control schemes are also modified to maintain the data integrity of the MLC NAND flash array. Our simulation results show that our scheme can reduce the power-down energy dissipation and the backup capacitance
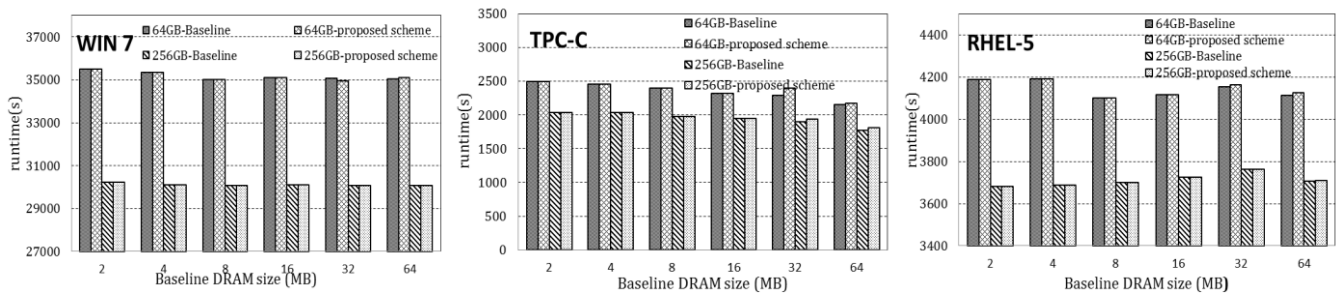
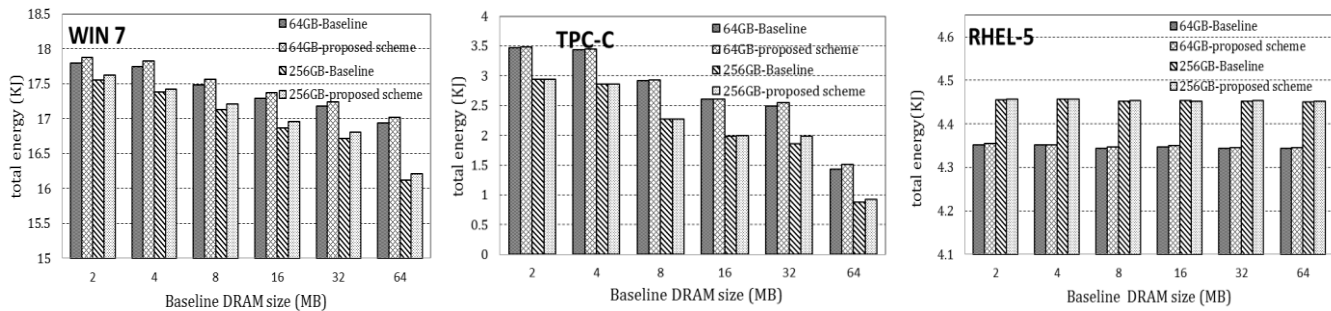Figure 2 Total Runtimes for WIN 7, RHEL 5, and TPC-C.

Figure 3 Energy Overheads for WIN 7, RHEL 5, and TPC-C.

budget by more than 77.1% and 75.1% for 64GB and 256GB NFSS', respectively. The maximum penalties of the runtime and energy consumption for the normal operations of the NFSS are only 1.3% and 1.4% respectively.
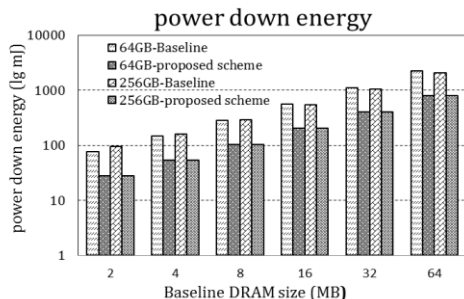
IV. Acknowledgement

Figure 4 Power down energy of the proposed scheme.

References

[1] F. Chen, R. Lee, X. Zhang, "Essential Roles of Exploiting Internal Parallelism of Flash Memory based Solid State Drives in High-Speed Data Processing", 17th High Performance Computer Architecture, Feb. 2011, pp 266-277.

[2] H.-W. Tseng, L. M. Grupp, R. E. Spada, S. Swanson, "Understanding the impact of Power Loss on Flash Memory," 48th Design Automation Conf, June. 2011, pp.1-6.

[3] G. Alcicek, H. Gualous, P. Venet, R. Gallay and A. Miraoui, "Experimental study of temperature effect on ultracapacitor ageing", Conf. on Power Electronics and Applications, 2007, Sept. 2007, pp. 1-7.

[4] G. Sun, Y. Joo, Y. Chen, D. Niu, Y. Xie, Y. Chen and H. Li, "A Hybrid Solid-State Storage Architecture for the Performance, Energy Consumption, and Lifetime Improvement," 16th Int'l Symp. on High Performance Computer Architecture, Jan. 2010, pp. 1-12.

[5] J. K. Kim, H. G. Lee, S. Choi, and K. I. Bahng. "A PRAM and NAND flash hybrid architecture for high-performance embedded storage subsystems", 8th ACM Int'l Conf. on Embedded software, 2008. pp 31-40.

[6] D. Liu, T. Wang, Y. Wang, Z. Qin, and Z. Shao, "Pcm-ftl: A write activity-aware nand flash memory management scheme for pcm-based embedded systems," in RTSS'11, 2011, pp. 357–366.

[7] K. Takeuchi, T. Tanaka, and T. Tanzawa, "A multipage cell architecture for high-speed pro-gramming multilevel NAND flash memories," Symp. VLSI Circuits, 1997, pp. 67-68.

[8] Y.-C. Chen, "Applying Super Capacitors to avoid the Power Cycling Issue of Solid State Drives" Thesis for Master of Science Dept. of Electrical Eng. Tatung University, 2009

[9] Y. Liu, C. Zhou and X. Chen. "Hybrid SSD with PCM", 11th Non-Volatile Memory Tech. Symp., Nov. 2011, pp. 1-5.

[10] W. Choi, P. Enjeti, and J.W. Howze, "Fuel Cell Powered UPS Systems: Design consideration," 34th IEEE Power Electronics Specialist Conf., 2003, pp. 57-64.

[11] J. Huang, L. Mei, and C. Gao, "Life Prediction of Tantalum Capacitor Based on Gray Theory Optimization Model", IEEE Int'l Conf. on Quality and Reliability, Sept. 2011, pp. 166-171.

[12] G. Hong, "Analysis of peak current consumption for large-scale, parallel flash memory", Workshop for Operating System Support for Non-Volatile RAM, April 2011

[13] M. K. Qureshi, J. Karidis, M. Franceschini, V. Srinivasan, L. Lastras, and B. Abali, "Enhancing Lifetime and Security of PCM-Based Main Memory with Start-Gap Wear Leveling," 42nd IEEE/ACM Int'l Symp. on Microarchitecture, 2009, pp. 14-23.

[14] A Simulator for Various FTL scheme, http://csl.cse.psu.edu/?q=node/322

[15] http://www.tpc.org

[16] http://www.hpl.hp.com/research/cacti/

[17] B.C. Lee, E. Ipek, O. Mutlu and D. Burger, "Architecting Phase Change Memory as a Scalable DRAM Alternative," 36th Int'l Symp. on computer architecture, 2009, pp.2-13

[18] C. Villa, D. Mills, G. Barkley, H. Giduturi, S. Schippers and D. Vimercati, "A 45nm 1Gb 1.8V Phase-Chang Memory", Int'l Solid-State Circuit Conf. 2010, Feb. 2010, pp. 270-271