



# Managed Flash Background Operations Series

## Part 3: Understanding Wear Leveling in NAND Flash Memory

NAND flash memory has had a significant global impact since its inception in 1987, enabling the mobility of content that businesses and consumers are now reliant upon every day. It features capacity and performance scaling that provides a very cost-effective way to store data as long as proper data management is applied. One important flash memory management function is wear leveling, as presented in this brief - the third in a series of Managed Flash Background Operations.

### Wear Leveling Overview

The concept of wear leveling has similarities to rotating tires on a vehicle. A reason that mechanics suggest tire rotations every 5,000 miles or so is to ensure that the tires wear evenly. The NAND flash memory implemented in devices also wears each time that cells are erased and programmed/written to. Wear leveling is a management task that balances cell use and wear.

NAND flash memory has a limited number of Write/Erase (W/E) cycles before it starts to fail and become unreliable. To address this, wear leveling algorithms have been developed that spread out the W/E cycles as evenly as possible to all available NAND flash memory blocks. When this process is implemented, all of the flash blocks in theory should wear out at the same time, thus maximizing the life of the flash device. In this scenario, every flash block should have similar erase counts until all of them reach end-of-life (EoL) simultaneously.

When performing wear leveling, it is important to determine an optimal time in order to balance performance and extend reliability. The wear leveling algorithm used may be set to a specific erase count (also known as the threshold) between one flash block and another. A specific group of flash blocks can be selected for use, and once those blocks reach the threshold number, they can no longer be selected for use, but not considered EoL. When this occurs, another group of flash blocks will be selected and used until they reach the threshold number, and then the cycle repeats. Figure 1 shows the maximum and minimum erase counts within any block of a managed flash device<sup>1</sup> at any given time with this implementation.

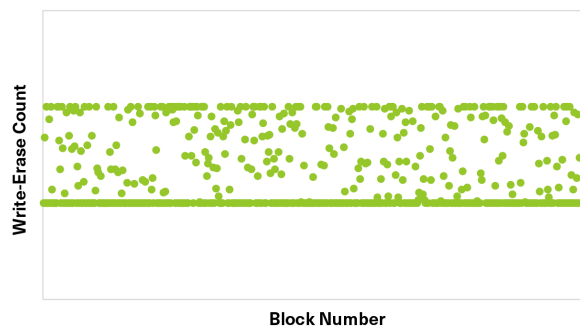


Figure 1: Data sample that shows the difference between the minimum and maximum Write/Erase counts kept at a certain level with wear leveling

In a use case when wear leveling is not enabled - when 10% of the flash device's total capacity is used - only 10% of the total flash blocks in the device would be programmed and the remaining 90% of the total flash blocks would never be programmed. As an example, if after one year the flash blocks that were programmed reach their W/E cycle limit and become unreliable, the flash device would be considered EoL despite the fact that 90% of the flash blocks were never used!

With wear leveling under the same circumstances, and using the same amount of writes, **ALL** of the flash blocks would be utilized. **All** of the flash blocks would reach 10% of its W/E cycle limit after one year, therefore, it would take ten years before all the flash blocks reached EoL. From this example, wear leveling can extend NAND flash memory life, **helping to enable its maximum lifetime potential to be achieved.**

## Wear Leveling Methods (Static and Dynamic)

There are generally two methods of wear leveling that most algorithms are created for: static wear leveling and dynamic wear leveling. In static wear leveling, all usable flash blocks, whether they contain data or not, would be made available for wear leveling. This includes reserved flash blocks as well as flash blocks used for internal system data. Static wear leveling tracks the W/E count of all flash blocks and makes sure that they are within a certain threshold of one another, and set by firmware.

In dynamic wear leveling, flash blocks that do not contain data (also known as free blocks) are selected by the lowest erase count to be used for the next write operation. The term dynamic means only incoming data being written to a flash block can be selected for wear leveling and flash blocks that contain data cannot be selected. Figure 2 shows differences between the wear leveling methods where dynamic wear leveling is easier to implement and static wear leveling is better for extending the life of NAND flash memory.

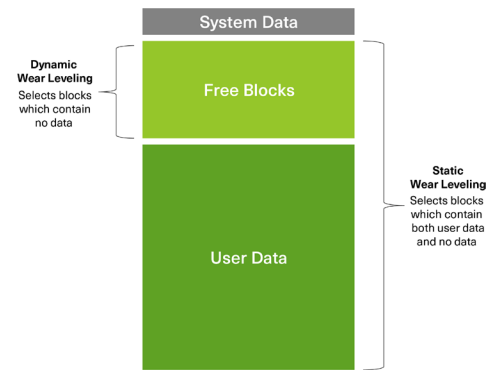


Figure 2: Static and dynamic wear leveling

## Wear Leveling Groups

Wear leveling algorithms may include further implementations to segment data where each group (also known as a zone) can be wear leveled only within their respective zones. For example, one group may contain data stored in the boot area while another group may contain system data, and yet another group may contain user applications. Since the data in each of the three groups is similar, the flash blocks within those groups can be wear leveled together. However, for memory devices such as e-MMC<sup>2</sup>, not all of them may use wear leveling groups and each vendor may have different implementations altogether.

As it relates to e-MMC and UFS<sup>3</sup> devices, a specific partition may be designated for each type of data. For example, critical system code may be stored in the Enhanced User Data Area, which is partitioned in Single Level Cell<sup>4</sup> (SLC) mode to offer higher reliability. Replay Protected Memory Block (RPMB) and the managed flash system data would be considered similar data, and belong within the same group.

Non-critical data, such as applications, music, or other data that is stored elsewhere (such as in the cloud), and can be downloaded again even if it is corrupted in flash memory, would reside in a separate Non-Enhanced User Area group which is partitioned in Triple Level Cell<sup>5</sup> (TLC) mode.

In this example, the SLC partition would be wear leveled together as one group and the TLC partition would be wear leveled together as another group. Figure 3 shows this example of a general e-MMC partition with different wear leveling groups.

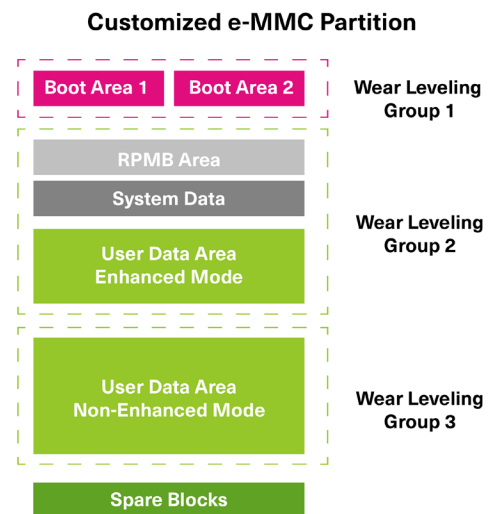


Figure 3: Example of wear leveling groups

## Wear Leveling in Managed Flash Devices

One advantage of managed NAND flash memory is that wear leveling has been implemented in the firmware as part of the overall flash management. The user does not need to design a wear leveling algorithm. This ensures that the e-MMC or UFS device will reach their maximum endurance specified by the manufacturer.

In e-MMC devices, the lifetime of the device can be read from the DEVICE\_LIFE\_TIME\_EST registers in the Extended CSD Register<sup>6</sup> (slice 268) as depicted in Table 1.

**Extended CSD Register:**

CSD- slice	Size	Name	Value	Description
268	1	DEVICE_LIFE_TIME_EST_TYP_A	0x00	Not defined
			0x01	0% - 10% device lifetime used
			0x02	10% - 20% device lifetime used
			0x03	20% - 30% device lifetime used
			0x04	30% - 40% device lifetime used
			0x05	40% - 50% device lifetime used
			0x06	50% - 60% device lifetime used
			0x07	60% - 70% device lifetime used
			0x08	70% - 80% device lifetime used
			0x09	80% - 90% device lifetime used
			0x0A	90% - 100% device lifetime used
			0x0B	Exceeded its maximum estimated device lifetime
Others	Reserved			

*Table 1: Example of estimated lifetime in an e-MMC managed flash device from the Extended CSD Register  
(This table is reproduced, with permission, from JEDEC specification JESD84-B51)*

Included within UFS devices is a Device Health Descriptor<sup>7</sup> that provides indications about the device's life expectancy, and is similar to e-MMC devices (Table 1). The lifetime of UFS devices can be read from the 'bDeviceLifeTimeEstA attribute offset 0x03' as shown in Table 2.

**Device Health Descriptor:**

Offset	Size	Name	Value	Description
0x03	1	bDeviceLifeTimeEstA	0x00	Information not available
			0x01	0% - 10% device lifetime used
			0x02	10% - 20% device lifetime used
			0x03	20% - 30% device lifetime used
			0x04	30% - 40% device lifetime used
			0x05	40% - 50% device lifetime used
			0x06	50% - 60% device lifetime used
			0x07	60% - 70% device lifetime used
			0x08	70% - 80% device lifetime used
			0x09	80% - 90% device lifetime used
			0x0A	90% - 100% device lifetime used
			0x0B	Exceeded its maximum estimated device lifetime
Others	Reserved			

*Table 2: Example of estimated lifetime in a UFS managed flash device from the Device Health Descriptor  
(This table is reproduced, with permission, from JEDEC specification JESD220E)*

One way that these registers can be used is to calculate the ratio of the current NAND flash memory block W/E cycle count and the maximum endurance specified by the manufacturer. Wear leveling is executed as part of managed flash background operations in e-MMC/UFS devices and performed seamlessly and hidden from the user. The results are reported back to these registers in real-time as a value that is equivalent to a percentage range of the device's lifetime used to date, depicting the most accurate estimation of the device's lifetime used. Without wear leveling, the average W/E cycle number could have a much lower lifetime percentage range than what the actual used lifetime of the device is.

## Summary

Wear leveling is used to extend the life of devices that implement NAND flash memory to ensure that all of the flash blocks reach their maximum endurance limit specified by the manufacturer. A wear leveling algorithm can evenly distribute the W/E cycles throughout all of the flash blocks in managed devices within a certain threshold level. There are different wear leveling methods, such as static and dynamic. There are also wear leveling groups that may be assigned to the algorithm based on the type of data stored that enables the group to be wear leveled together. Wear leveling is implemented in managed flash memory as the algorithm performs automatically and without user intervention. The lifetime estimation that can be read from e-MMC and UFS is only accurate if the wear leveling algorithm is implemented.

The next brief in the Managed Flash Background Operations Series covers Garbage Collection, which is an important task for cleaning up the managed flash blocks when valid and invalid data is spread across many blocks.

General information for KIOXIA memory products is available [here](#).

### FOOTNOTES:

<sup>1</sup> A managed flash device combines raw NAND flash memory and an intelligent controller in one integrated package, enabling memory management to be performed internally.

<sup>2</sup> Embedded MultiMediaCard (e-MMC) is a specification developed by JEDEC for mobile applications. The current release is v5.1, published in February 2015.

<sup>3</sup> Universal Flash Storage (UFS) devices are based on the UFS specification, of which, the v4.0 specification is the current release issued by JEDEC and published in August 2022.

<sup>4</sup> Single Level Cell (SLC) is a flash memory cell that contains one bit per cell for data storage.

<sup>5</sup> Triple Level Cell (TLC) is a flash memory cell that contains three bits per cell for data storage.

<sup>6</sup> The Extended CSD Register is in accordance to JEDEC specification JESD84-B51.

<sup>7</sup> The Device Health Descriptor is in accordance to JEDEC specification JESD220E.

### TRADEMARKS:

JEDEC is a registered trademark of the JEDEC Solid State Technology Association. All other company names, product names and service names may be trademarks or registered trademarks of their respective companies.

### DISCLAIMERS:

© 2023 KIOXIA Corporation. All rights reserved. Information in this tech brief, including product specifications, tested content, and assessments are current and believed to be accurate as of the date that the document was published, but is subject to change without prior notice. Technical and application information contained here is subject to the most recent applicable KIOXIA product specifications.

JEDEC standards and publications are copyrighted by the JEDEC Solid State Technology Association. All rights reserved.