September, 2009 Version 1.0

Wear Leveling Implementation on SQFlash

Author: Ethan Chen/ Tones Yen

E-mail: ethan.chen@advantech.com.tw; tones.yen@advantech.com.tw

Table of Contents

Introduction

LBA vs PBA

NAND is similar to a hard-disk drive. It is sector-based (page-based), and fitted for storing sequential data such as picture files. The host system performs its data transfer (reads and writes) to LBA (logical block address) only. For example, the FAT information would also be stored from LBA X to LBA Y. The assigned allocation area for the FAT does not move in LBA, whereas the actual FAT table for the PBA (Physical Block Address) moves randomly from time to time. The PBA is the physical address of the NAND Flash. A mapping table is also required for the controller to remap the LBA to the PBA of the NAND Flash.

Program Operations (Write)

Programming is necessary to change erased bits from 1s to 0s. Due to the nature of NAND Flash, the smallest erasable entity is a “block”, and when erased, all bits are set to 1, or all bytes to FFh. If the host issues a write operation to certain pages of partially occupied PBA, the controller needs to find a spare, i.e., previously erased, block, and merge the old information on the PBA with the new incoming information issued from the host. The typical programmable write-erase endurance of each block is 100k for SLC flash, and 10k for MLC flash. Therefore, it is essential to deploy an efficient wear leveling algorithm to ensure that rite and erase operations are distributed evenly over all the blocks in the SSD.

Wear Leveling

Wear leveling is an algorithm by which the controller in the storage device re-maps logical block addresses to different physical block addresses in the solid-state memory array, to allow an even spread of writing/erasing over all NAND blocks. The frequency of this re-map depends on the algorithm to find the "least worn" areas to which to write, and any data swapping capabilities are generally considered proprietary intellectual property of the controller vendor. They key for wear-leveling is to minimize extra NAND write/erase cycles and reduce performance impact.

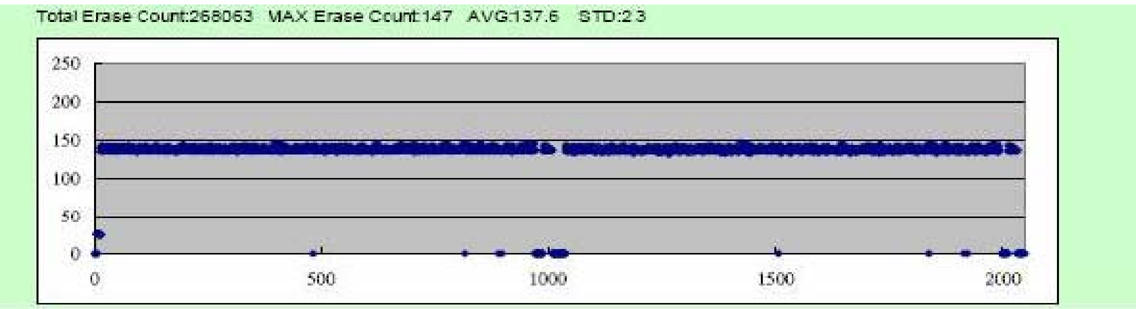

Figure 1. X-axis represents the PBA of the NAND Flash.

SQFlash Wear Leveling

SQFlash's Controller deploys two types of wear-leveling over the entire storage media: Dynamic and Static. Take the following scenarios for example:
A 32 GB device stores 4 GB of files with 28 GB capacity remaining free for wear leveling. The SQFlash controller will deploy dynamic wear-leveling on the Dynamic Data Areas (28 GB of free capacity) to allow the incoming programming requests to distribute evenly across the 28 GB of free space. As the dynamic wear-leveling is being executed for X amount of time, the controller will also implement the static wear leveling method on the 4 GBs of Static Data Areas. To implement this, the controller will swap between a free block and a static data area block. The controller seeks very "cold" static blocks and very "hot" free blocks for swapping by block vector.

Dynamic Wear Leveling

SQFlash controller reserves a total of 1~2% spare blocks (not including init defective blocks) for wear-leveling. Furthermore, one queue is created by the controller for queuing numbers of spare blocks.
Whenever a write command is received, the controller de-queues one spare block from the spare queue to start executing the write operation (Fig. 2). After that, when this block programming is finished, it en-queues the old mapped block into the spare queue and updates mapping between logical and physical flash blocks (Fig. 3).
Due to queuing the operation, this simple round-robin operation will keep using the oldest existing block from the queue. In order to prolong the life cycle of MLC NAND-Flash, all blocks will be amortized using this wear-leveling method.

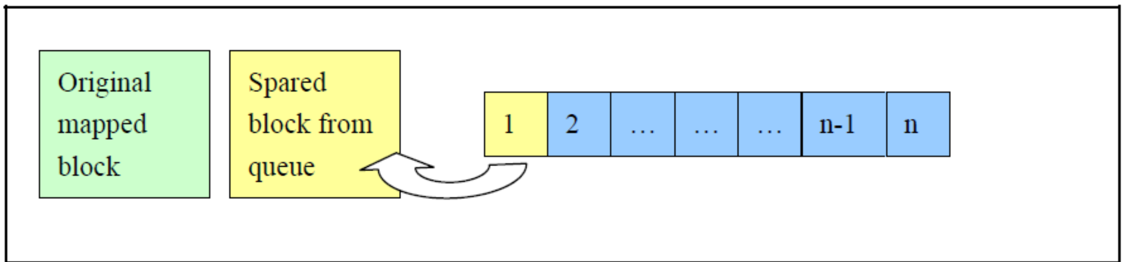

Figure 2. De-queue first block from spared queue

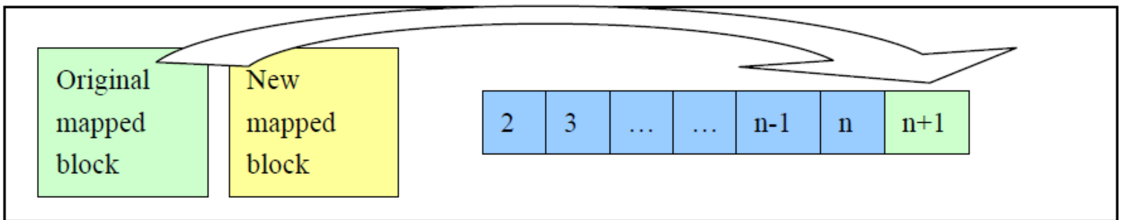

Figure 3. En-queue to last block at spared queue

Static Wear Leveling

BASE: Every Nth write command, change one logical block to a free block. (In this case N=1)

STEP 1: Assume we have one flash disk; the internal architecture will be like that shown in FIG-4.

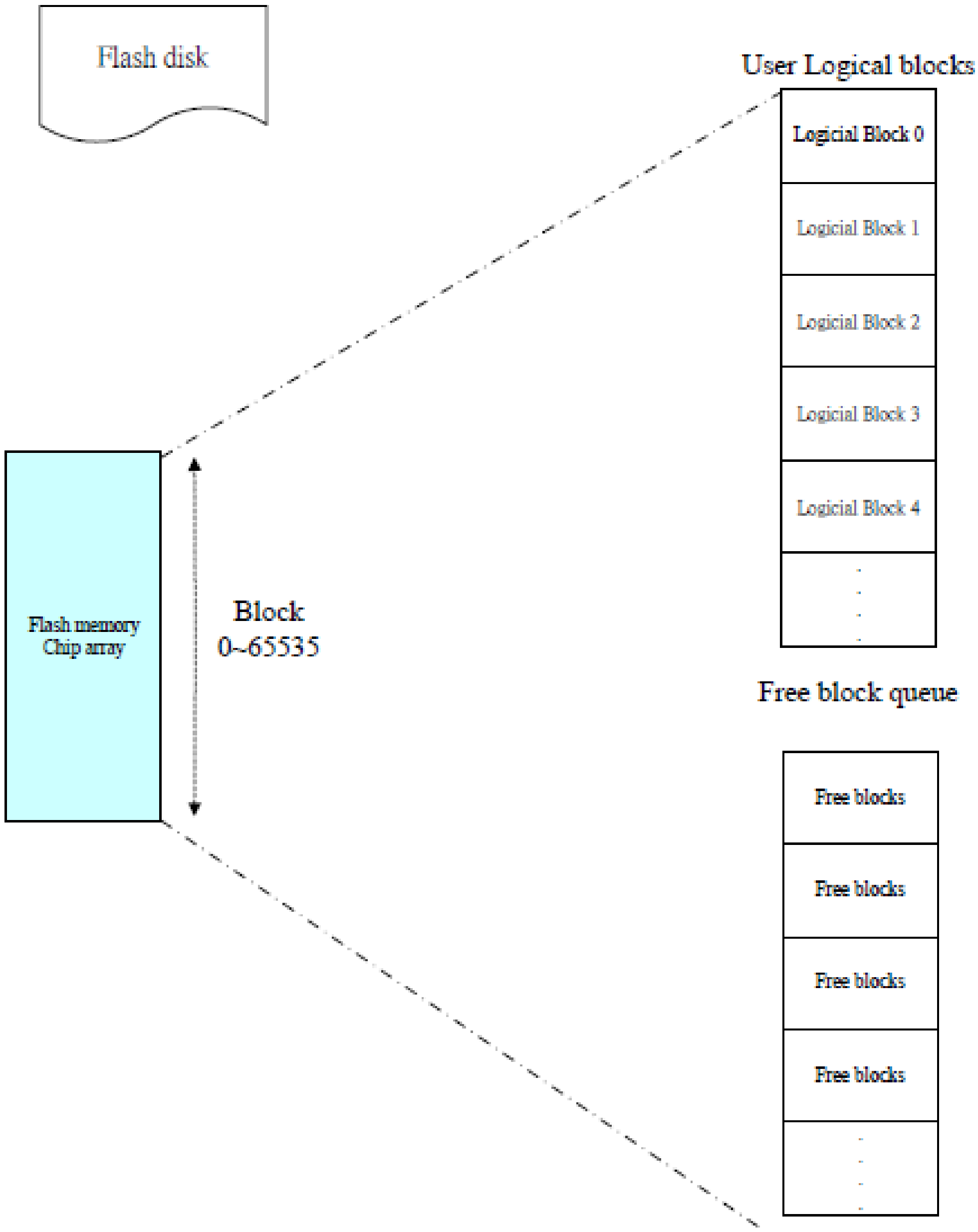

FIG-4

STEP 2: Logical and physical block mapping. (FIG-5)

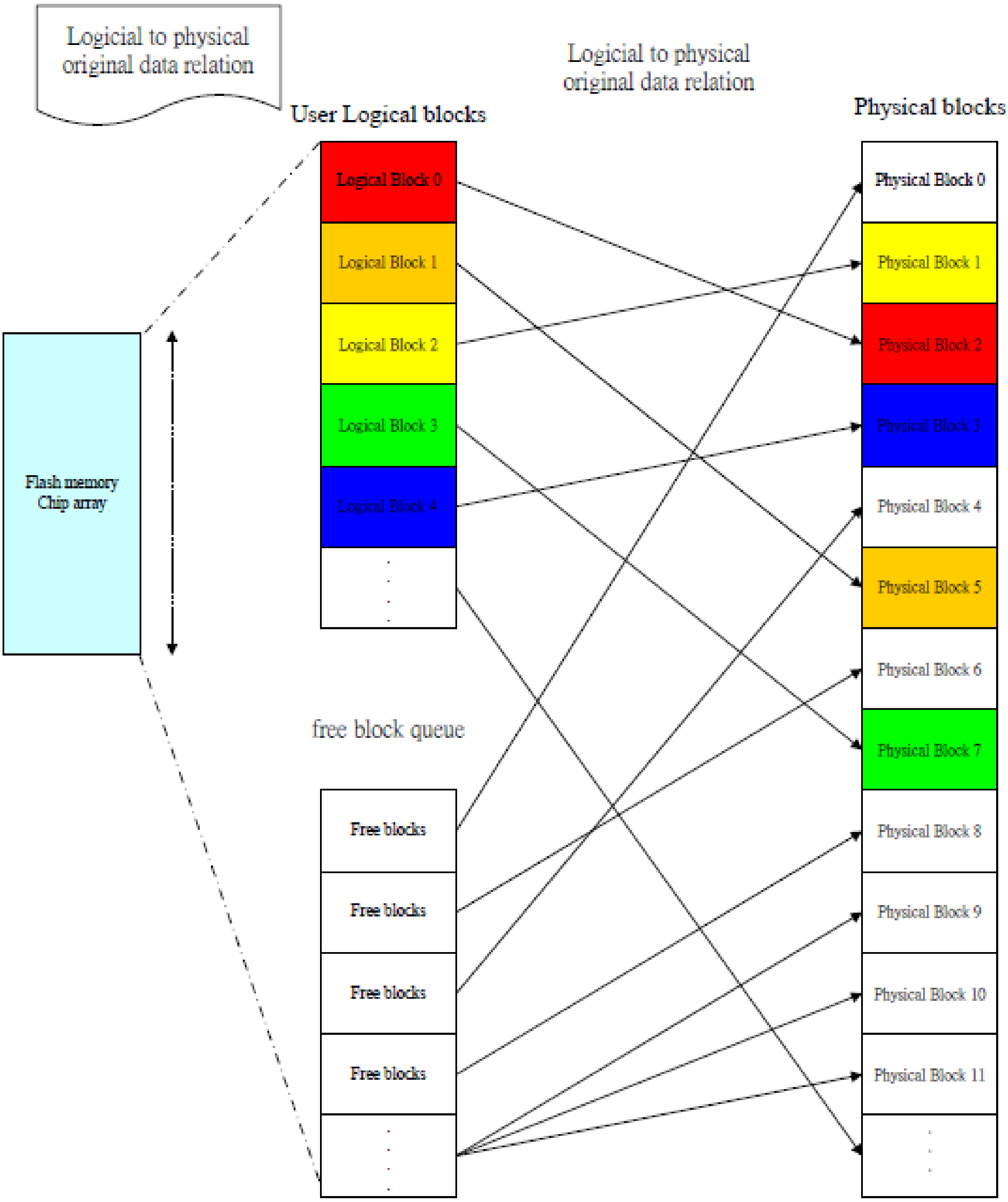

FIG-5

STEP 3: Write the logical block 2 N times (FIG-6)

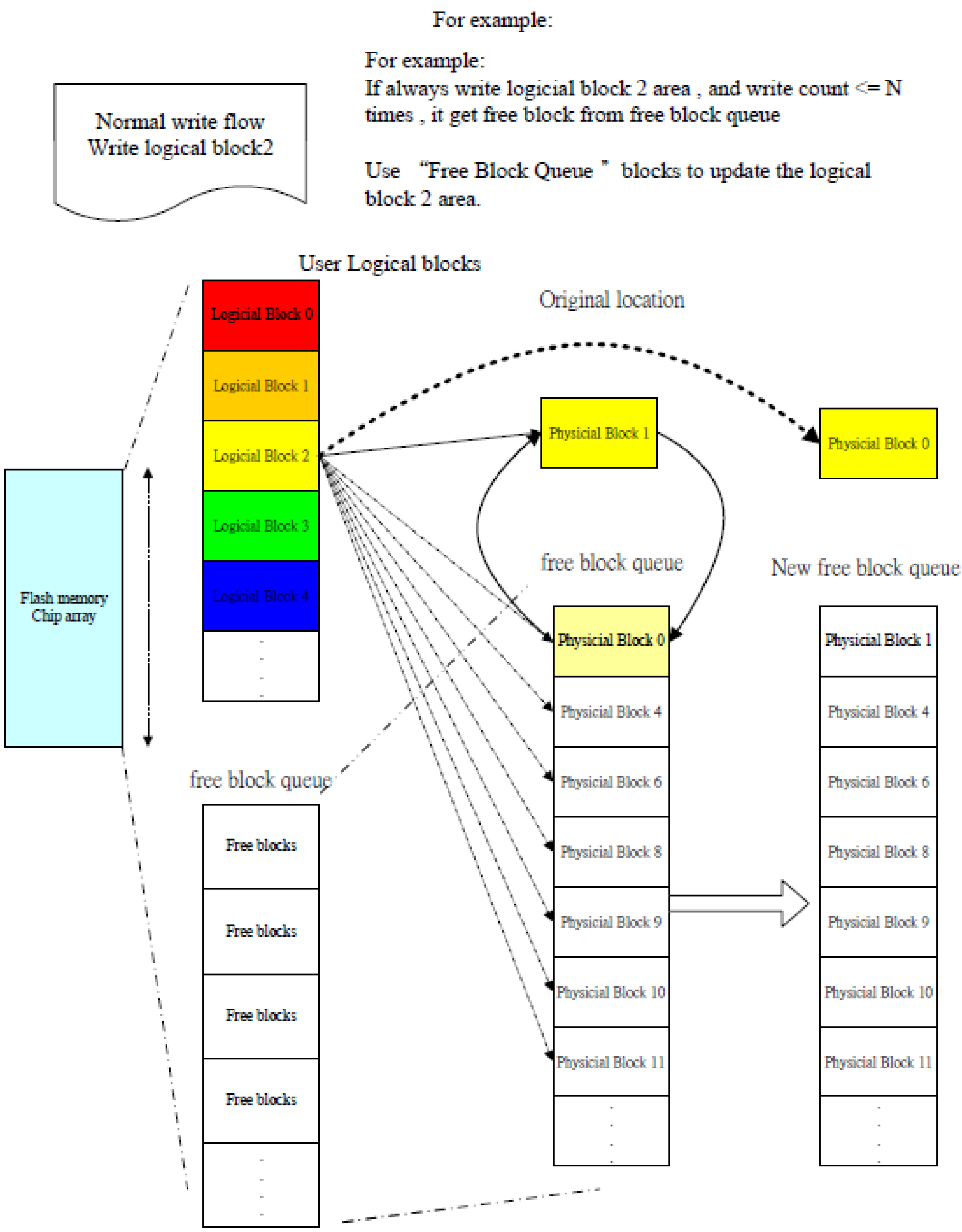

FIG-6

STEP 4: Change a seldom-accessed block from the logical area to the free block area. (FIG-7)

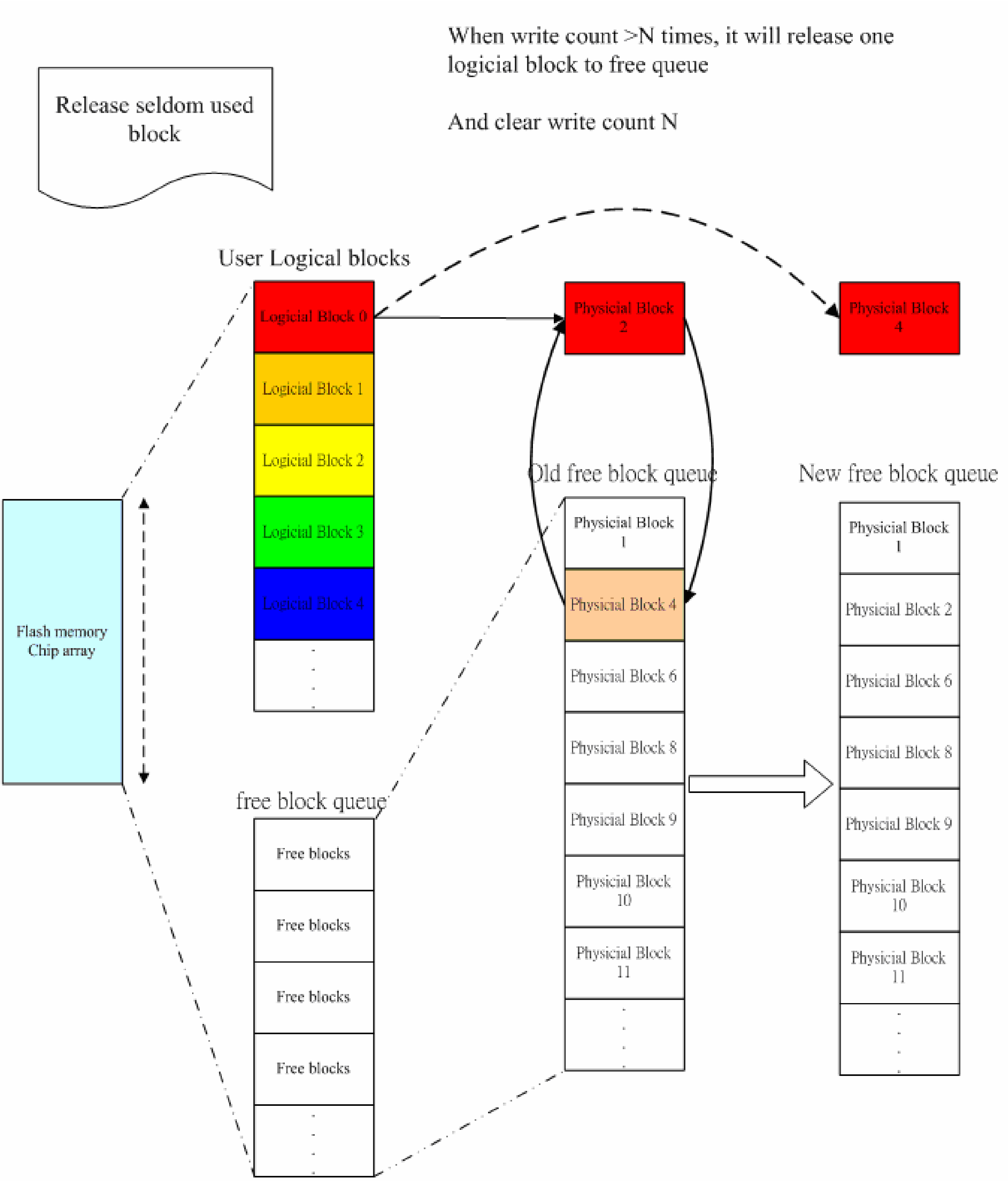

FIG-7

STEP 5: Mapping between the logical block and physical block (FIG-8)

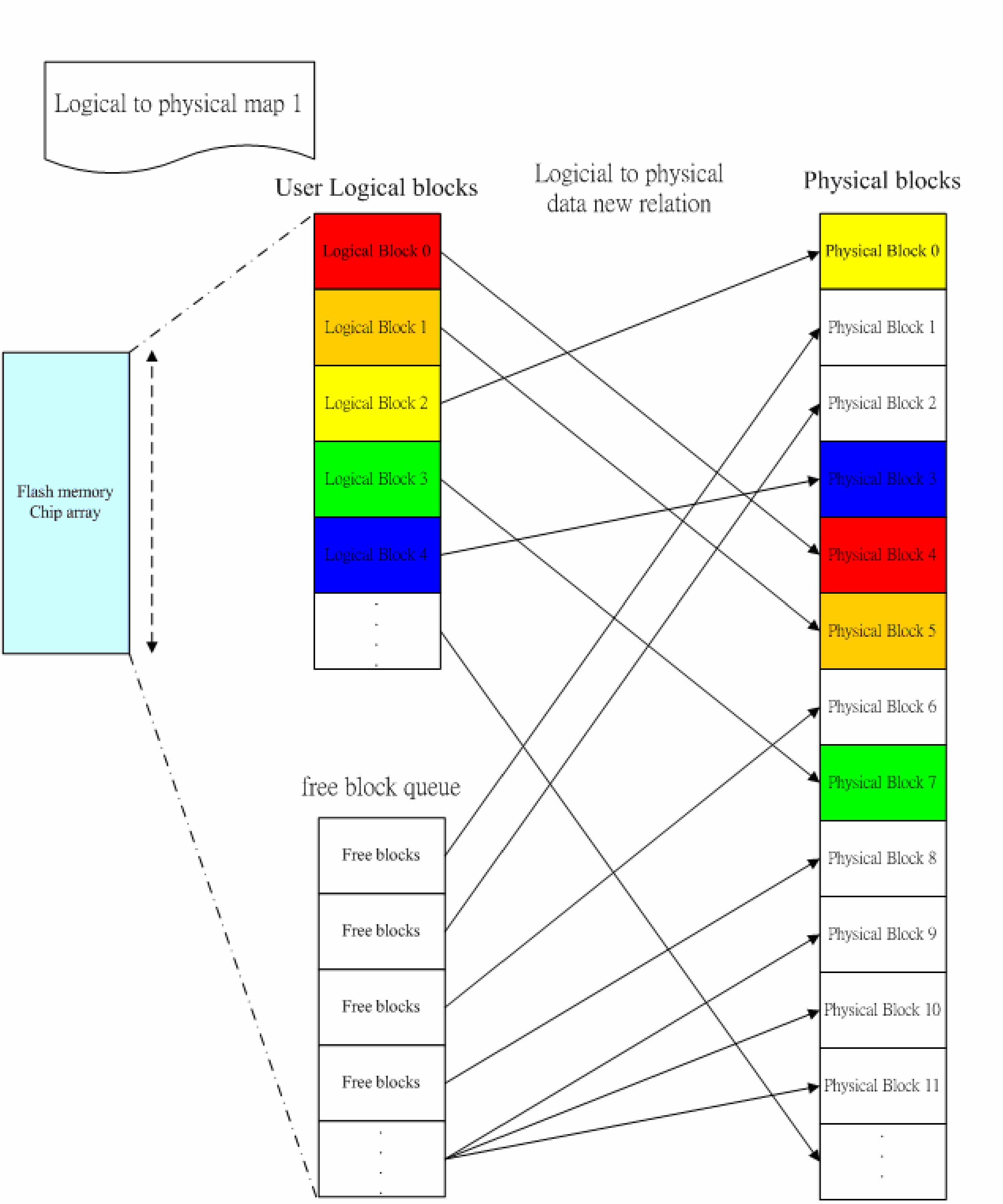

FIG-8

STEP 6: Write logical block 2 again. (FIG-9)

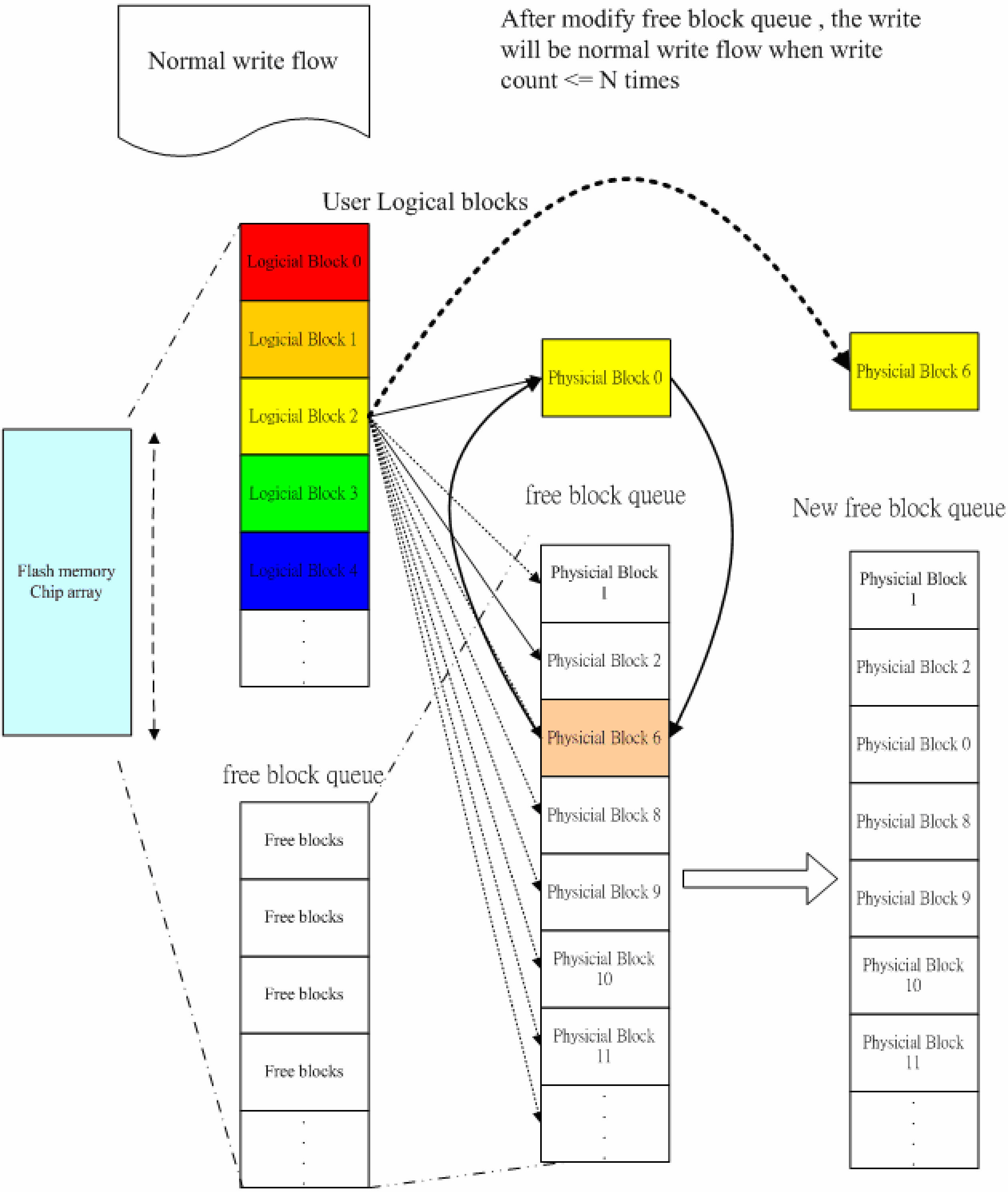

FIG-9

STEP 7: Swap the seldom used block again. (FIG-10)

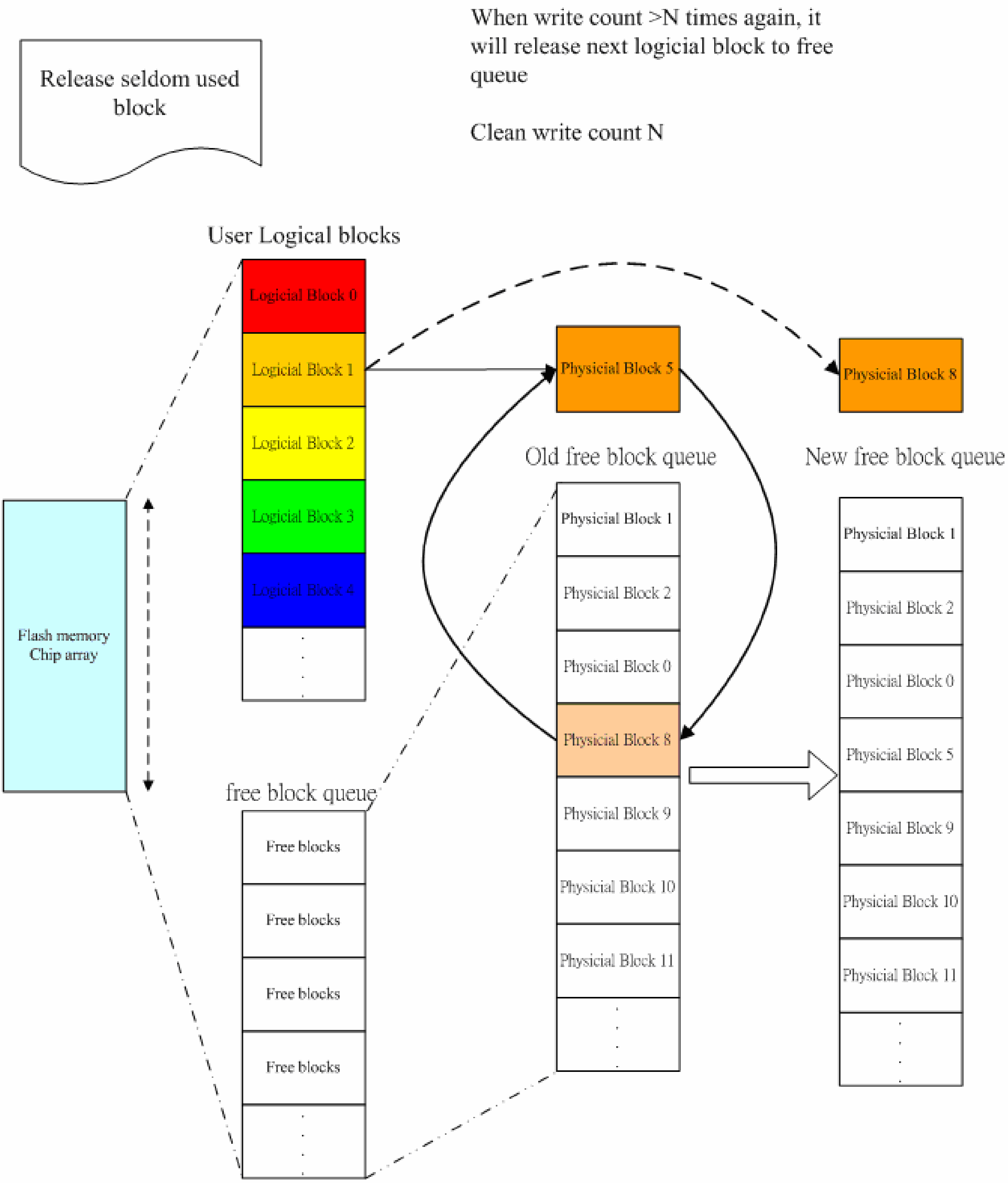

FIG-10

STEP 8: Mapping relationship of logical blocks and physical areas (FIG-11)

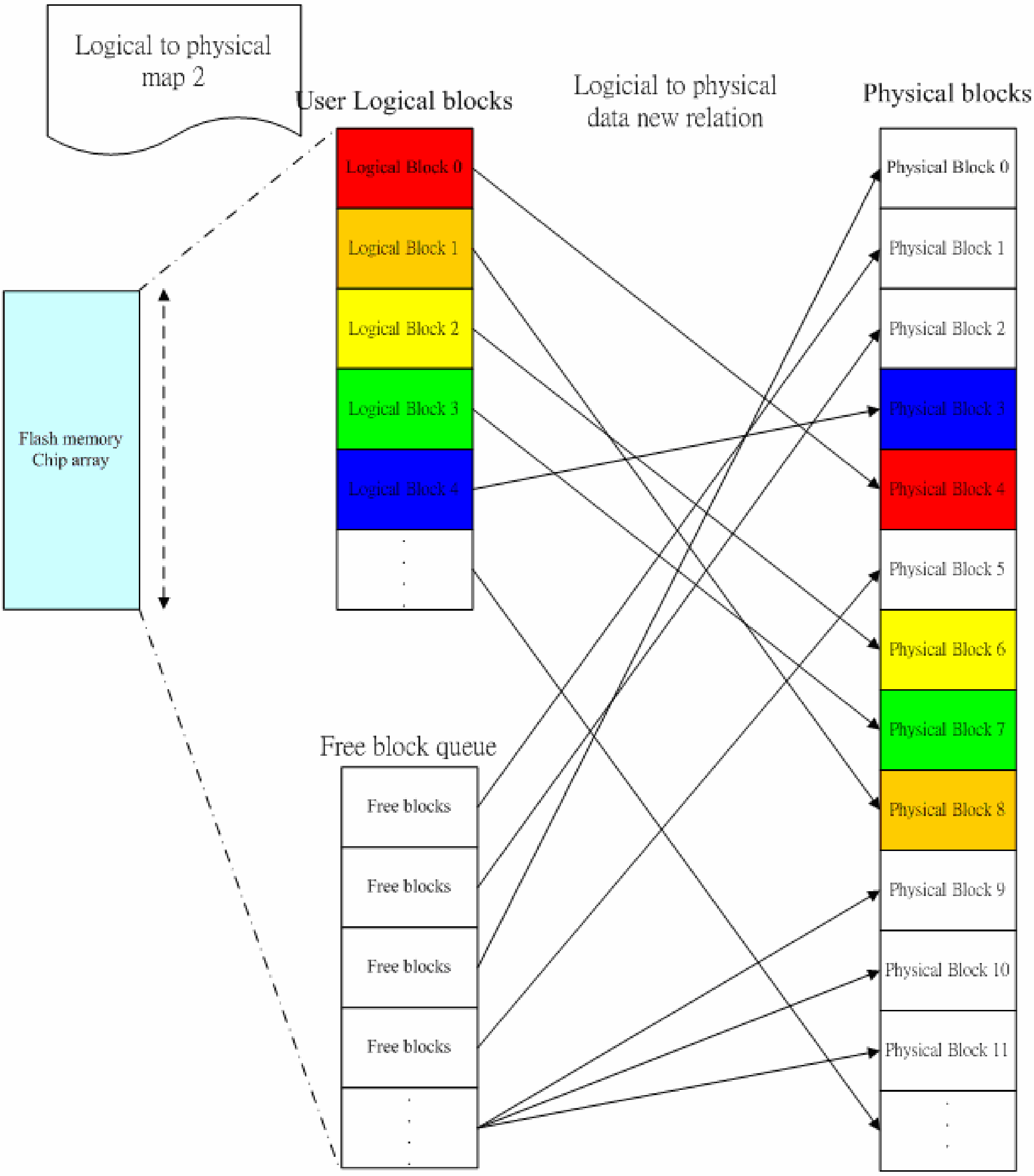

FIG-11

Advantech Co., Ltd. –Founded in 1983, Advantech delivers visionary and trustworthy industrial computing solutions that empower businesses. We cooperate closely with solution partners to provide complete solutions for a wide array of applications in diverse industries, offering products and solutions in three business categories: Embedded ePlatform, eServices & Applied Computing, and Industrial Automation groups. With more than 3,400 dedicated employees, Advantech operates an extensive support, sales and marketing network in 18 countries and 39 major cities to deliver fast time-to-market services to our worldwide customers. Advantech is a Premier Member of the Intel® Embedded and Communications Alliance, a community of embedded and communications developers and solution providers. (Corporate Website: www.advantech.com).

About Embedded Core –Advantech Embedded Core Services provide integrated solutions, from boards to modules to software. Through cross collaboration, our design-in services can fulfill electronic engineering demands at the design-in phase and promise to bring benefits which can shorten the design and integration cycle, thereby minimizing uncertainty and risk. (Emb'Core Website: www.advantech.com).

Advantech Co., Ltd. –Founded in 1983, Advantech delivers visionary and trustworthy industrial computing solutions that empower businesses. We cooperate closely with solution partners to provide complete solutions for a wide array of applications in diverse industries, offering products and solutions in three business categories: Embedded ePlatform, eServices & Applied Computing, and Industrial Automation groups. With more than 3,400 dedicated employees, Advantech operates an extensive support, sales and marketing network in 18 countries and 39 major cities to deliver fast time-to-market services to our worldwide customers. Advantech is a Premier Member of the Intel® Embedded and Communications Alliance, a community of embedded and communications developers and solution providers. (Corporate Website: www.advantech.com).