Invention Disclosure: Re-Writing Schemes for Solid-State Storage Devices

Yuval Cassuto, Luoming Zhang and Amin Shokrollahi

I. Background and need of invention

The quality of solid-state storage devices (e.g. NAND Flash, NOR Flash, Phase-Change Memory) is determined by the following features

1) Read/write access speeds
2) Device lifetime and endurance

These features are impeded by inherent limitations of the device physics and architecture. For example, in NAND flash, memory *pages* (the basic access unit of the device, at order of 2KBytes) cannot be individually re-written. Rather, a full *block* (order of 128KBytes) must be erased before a page can be re-written. This property limits the write-access speed, and increases the wear of cells due to the need to frequently erase large portions of the memory.

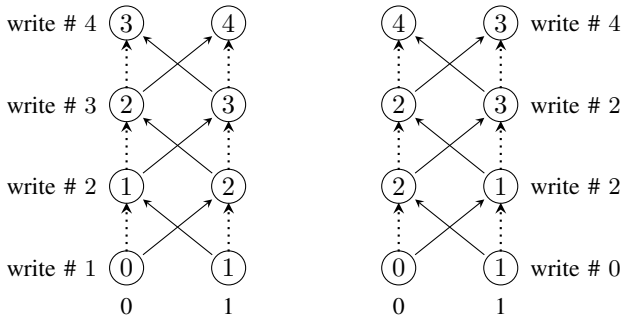
To reduce the frequency of block-erase operations, the solid-state drive employs a software module to optimize the placement of logical page addresses (seen by the host) into physical page addresses (used by the device). In NAND flash, this software module is called Flash Translation Layer (FTL). Not only does the FTL have to reduce the *total* number of block erases, but it also needs to guarantee a *leveled* wear of cells throughout the device. Writes from the host are highly non-uniform: *hot* logical page addresses are re-written at high frequency (e.g. file-system logs) and *cold* logical page addresses may be written only once in the lifetime of the device (e.g. archived photos or documents). Therefore, a *wear leveling* operation is crucial for extending the lifetime of the device. Wear-leveling is a costly operation both in terms of excess capacity needed for page relocations, as well as in terms of increased average wear due to copying of static data.

Solid-state storage devices often utilize a *multi-level cell* (MLC) technology, programming each memory cell to one of $q > 2$ discrete levels. The multi-level cell is an extension of the original *single-level cell* (SLC) technology, which allows a cell to be in one of only $q = 2$ levels. The MLC concept is commonly used to increase the storage capacity by storing $\log_2 q$ bits in each cell (hence q is restricted to integral powers of two: $4, 8, 16,$ etc.). MLC devices typically offer slower access speeds, due to longer write/read operations needed to attain/resolve the precise stored level.

II. Summary of invention

For MLC storage devices, the invention allows multiple unrestricted re-writing of pages, without need of block-erase operations. Each proposed scheme stores k user bits in n physical cells of q levels, and can guarantee a certain number t of *unrestricted* writes before an erase operation is mandated. In this application of the MLC technology, the number of levels q can be **an arbitrary integer**, and not limited to an integral power of 2 as in the common usage of MLC. **Meanwhile, the variations of the cells' level in each write become smaller which can make the write faster.**In addition to lowering the erase frequency at write time, the schemes allow for efficient read operations by constraining the cells to be in one of only 2^k states after every write operation. Small read windows can provide shorter read times and/or simplified lower-cost read circuitry.

A simple example for such a scheme is given in the figure below for storing $k = 1$ bit in $n = 1$ cell. The guaranteed number of writes for this example is $t = q - 1$. The nodes in this figure are numbered with physical-cell levels. Nodes on the left correspond to a $'0'$ stored value, and nodes on the right correspond to a $'1'$ stored value. Each vertical layer represents a write generation, and arrows represent level transitions between write generations (solid lines when the data bit changes, and dotted lines when it stays the same). Given that after i writes the cell is at level at most i, we get $t = q - 1$ writes without erasures.



At any write generation, the cell can be in only one of two possible levels. This number is optimal for $k = 1$.

The invention also proposes schemes to jointly store *hot* and *cold* pages, such that hot pages can be re-written multiple times, while cold pages are assumed to be written only once between erase operations. Schemes for hot-cold data bits allow increasing the number of writes, and additionally provide "automatic" wear leveling between hot and cold logical pages without need to relocate static data.

III. High level merit and potential

The new schemes offer read/write access and lifetime boosting capabilities close to the physical layer of the device. This may allow memory-chip manufacturers to implement a

"smart" memory cell, which can be utilized by any device-controller vendor to complement and enhance access and lifetime optimization modules at higher levels.