

Notification of Deleted Data Proposal for ATA8-ACS2

April 21th, 2007
Revision 0

Author: Frank Shu
Microsoft Corporation
One Microsoft Way
Redmond, WA. 98052-6399
USA

Phone: (425) 707-1779
E-Mail: frankshu@microsoft.com

Table of Contents

1	INTRODUCTION	3
2	DESCRIPTION OF NOTIFICATION OF DELETED DATA PROPOSAL	3
3	THE DEFINITION OF NEW “TRIM” COMMAND FOR NOTIFICATION OF DELETED DATA.....	3

Revision History		
Date	Revision	Description
2007-04-21	0	Initial Draft

1 Introduction

Solid State Drive(SSD), made by flash memory, is unlike Hard Disk Drive(HDD) in which stored data can be updated at the same physical location of storage medium. For SSD, new updated data is actually merged with the existing data of the same memory block into a new pre-erased memory block, this merge operation involves multiple read and write, it has become a major performance factor for SSD. On the other hand, since only the host computer knows which file has been deleted and the associated data has become invalid, SSD device treats all data in its storage as valid data, and carries them over merge operation unnecessarily.

To inform SSD device with deleted data information by host computer, makes it possible that SSD device can then reduce its merge overhead by not processing those invalid data, and further make pre-erased area for next write without merge operation.

Such optimization is also applicable for SSD ware leveling process. NAND flash memory has a limited erase cycles; about 100K for SLC(Single Layer Cell) flash and 10K for MLC flash, all SSD devices employ some kind of ware leveling to prevent bad memory cell over the use of SSD. The ware leveling basically moves data block from one location to another in SSD device over time; it can be looked as a different type of merge operation, and will even more benefit from this new technique.

Reducing merge operation also means reduce erase cycle, which increases the reliability of SSD device. While the deleted data information is focused on SSD device, it may also provide usefulness for HDD device for its background maintenance process as well.

2 Description of Notification of Deleted Data Proposal

To notify device with deleted data information by host computer, a new command and associated data need to be defined. The rest of this proposal will describe the new command definition in detail.

It's true today for both SSD and HDD devices that host computer's data is addressed by LBA and presented with the contents at the LBA. Also, host computer is capable storing the data for a file in a scatter-gather manner, so multiple data segments are possible. As for deleted data information, all segments of data need to be notified to device by host computer when a file is removed. As device will further remap LBA to its internal page and block for SSD, and sector, track and cylinder for HDD, this remapping information will be not necessary to dealt with and is out of the scope of this proposal.

The new proposed command is named as "Trim" command with Op code XXh. Its task register file is defined in the next section of this proposal.

3 The Definition of New "Trim" Command for Notification of Deleted Data

The Trim command definition for notification of deleted data.

ATA Command Format for Trim Command

Word	Name	Description
00h	Feature	XXh
01h	Count	Number of 256 word-blocks of LBA Range Entry to be transferred. 0000h specifies that 65,536 blocks are to be transferred
02-04h	LBA	Reserved
05h	Command	XXh

LBA Range Entry:

An individual LBA range, called an LBA Range Entry, is represented by 8 bytes in a data block during the data transfer of Trim command. The LBA is expressed by the LBA Range Entry's first 6 bytes, and the range length is a zero based number (i.e. 0=0, 1=1, etc.) represented by the next 2 bytes. If the 2 byte range length is 0 then the LBA Range Entry is not valid and should be discarded. With 512 bytes per LBA, the maxima invalid data size can be notified by a Trim command is 2GB. The large than 2GB invalid data will requires more than one Trim command to be sent to device from host computer.

Upon receive of notification of deleted data information from host computer via Trim command, device should mark all transferred LBA(s) as invalid data. Therefore, those data can be disregard during following merge and ware leveling operations for better performance and endurance of SSD device. In the case of deleted data occupies over erase block(s) of SSD, SSD device can apply pre-erase to the block(s) for better write performance.

(end)