Data Set Management Commands Proposal for ATA8-ACS2

Dec. 12th, 2007
Revision 6

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

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

Phone: (425) 705-9157
E-Mail: natobr@microsoft.com

Table of Contents

Revision History		
Date	Revision	Description
2007-04-21	0	Initial Draft
2007-06-04	1	1. Changing the title of the proposal to reflect more advanced feature covered by this proposal. 2. Adding new dataset attributes for device optimization 3. Incorporating feedbacks on “Trim” attribute. • Non media specific command • Define the read response after “Trim” • Clarification of relation to Security Erase
2007-09-05	2	1. Moving normal return status of Data Set Management command from Status register to Count register. 2. Attributes Importance

Revision History		
Date	Revision	Description
		3. Attributes Clear bit
2007-10-05	3	1. Incorporated feedbacks from 9/16/07 T13 ad-hoc meeting 2. Moving “Trim” attribute ahead with space in protocol for extension of other attributes later.
2007-11-01	4	1. Name the attribute as “Trim”, instead “Deallocated” 2. Add definition of “Trim” for NV Cache 3. Use Table 142(as error output, instead Table 124
2007-11-14	5	As revised on Nov. 14th T13 ad-hoc meeting: 1. Refined the description of Trim attribute 2. Adding all required clauses for ATA-ASC spec 3. Using Table 110(General extended normal output) as normal output. 4. Using Table 129(Write log error) as error output
2007-12-12	6	As revised on Dec. 12th 07 Plenary: 1. Adding bit fielded Indentify word for future DSM attributes. 2. Removed the section 4.2 of term definition.

1. Introduction

The first version of this proposal was called “Trim” command proposal, the “trim” command carries the information related to deleted data blocks to device for optimization. But, it has showed that the information other than deleted data blocks provided by host can also be useful to device, one example would be the frequency of data write and read. A collected set of such information is called “data set attributes” in this proposal.

This proposal discusses the details necessary for defining and implementing Data Set Management commands on an ATA device. The advantage to Data Set Management commands is having methods to communicate an application’s I/O access behaviors within a host to ATA device. This new feature, when used with applications within a host that accurately provide I/O access behavior information about it’s data, allows an ATA device to be able take internal optimizations to provide the responsiveness needed by the host when later accessing that data.

This document provides a common understanding of the new concepts that Data Set Management introduces, provides a common language to discuss Data Set Management functionality, and proposes a new command that creates a Data Set Management abstraction.

(Note: Since rev. 3, this proposal has been addressing "Trim" attribute only with extensibility of other attributes proposed in previous version of this proposal. The goal is to allow "Trim" attribute be implemented first.)

1.1. The purpose of Trim Attribute

File deleting happens in a file system all time, but the information related to the deleted file only kept in OS, not to device. As a result, a device treats both valid data and invalid data(deleted file) in its storage media the same way, all necessary operations keeping data alive are applied. For example, background defect management and error recovery process are used for HDD. Merge, wear leveling and erase are applied to SSD. OS's process, such as format, Defragger, Recycle bin, temp file deleting and Paging can produce a considerable amount of deleted file. By telling device the deleted file information as invalid data, the device can reduce its internal operation on all invalid data as self-optimization.

2. Description of I/O Quality of Service

An application layer in a host that is a source for I/O operations may be able to know or predict the nature of the data it is storing on the ATA device. The nature of how this data will be accessed or used in the future is expressed by context attributes that identify how the data should to be stored. Context attributes are applied to LBA ranges in order to identify a **data set**. Together context attributes for a data set instructs the ATA device how that data set will likely be used in the future by the host.

If the host can provide accurate context attributes to the ATA device, then the ATA device can use this context to make informed decisions on how to best provide the quality of service, indicated by attributes. By providing context attributes, the host provides Data Set Management information to the ATA device.

3. Description of the Data Set Management Command

For the context attributes provided by the host to be meaningful, the context attributes defined in this document must allow the host reflect aspects of quality of service such as pattern, frequency, latency, and LBA ranges read and written by the host.

The primary piece of information provided in a Data Set Management command is the LBA ranges that represent the data set that the application stored on the ATA device. Often times the data set stored by the application is not contiguous and consequently it is important that the command be able to provide a series of discontinuous sectors of a flexible length. Since the command structure in ATA does not meet this requirement, the LBA ranges for a Data Set Management command must be passed to the ATA device through a data transfer just like the NVC feature set.

The secondary piece of information provided in a Data Set Management command is the context attributes that identify the quality of service requested by the host. The context attributes that the host provides is also valuable for any command that provides a LBA range, such as read or write commands.

4. Data Set Management Commands Proposal

The primary focus of the proposed Data Set Management command is to enable the host to share quality of service context attributes for LBA ranges with the ATA device, which is a new concept to the ACS standard. Consequently many of the concepts that will be used in discussion and in the remainder of this document are also new. This section establishes the goals and requirements for the Data Set Management proposal and defines terminology for new Data Set Management concepts.

4.1. Data Set Management Commands Requirements

The purpose of Data Set Management command is to create a mechanism by which the host may share quality of service information with the ATA device. Requirements for the Data Set Management proposal break down into the mechanism's ability to deliver the quality of service context information and the abstraction that defines what form the context information takes.

The Data Set Management mechanism must:

- specify a mechanism to identify a flexible number of possibly discontinuous LBA ranges as a data set
- limit LBA ranges to physical sector aligned and physical sector multiple sizes
- provide a mechanism for identifying which quality of service context attributes the ATA device supports for future expansion.

The quality of service context information must:

- reflect the pattern, frequency, latency for one or more data sets
- be able to be contained in a standalone command that specifies a LBA range as well as an extension to a read command and a write command.

5. Proposed Changes to ATA8-ACS2

5.1. Change the clause 3.*1*

Data Set : A set of **LBA Ranges** to be treated by device as single group

5.2. Change to Security actions table

DATA SET MANAGEMENT Command aborted Executable Executable

Long Logical Sector Feature Set

DATA SET MANAGEMENT 256

5.3. Identify

One reserved bit of the ATA-8 ACS IDENTIFY DEVICE words 119 is used by the device during device enumeration to indicate support. The chosen IDENTIFY DEVICE word would be augmented with the following description:

[TBD1] the DATA SET MANAGEMENT command is supported

Bit [0] when set to one indicates that the device supports the Trim bit of Data Set Management command.

Bit 15:1 = Reserved.

5.4. DCO

Likewise, one reserved bit of "Command set/feature set supported part 2" DCO word is used by device to indicate Data Set Management support. The chosen DCO word would be augmented with the following description:

1= Reporting support for the DATA SET MANAGEMENT command is changeable

Word x[TBD] bit y[TBD] of DCO Identify set to one indicates that support for the DATA SET MANAGEMENT command is changeable.

1= Reporting support for the DATA SET MANAGEMENT command is allowed

Word x[TBD] bit y[TBD] of DCO Set sets to one then the device may support the DATA SET MANAGEMENT command.

6. Proposed New Commands for ATA8-ACS

Data Set Management Ext – xxh, DMA

6.1. Feature Set

This 48-bit command is optional for devices implementing the general feature set.

6.2. Description

The purpose of DATA SET MANAGEMENT command is for host to provide file system information for device optimization.

ATA Command Format for Data Set Management Command

Word	Name	Description	
00h	Feature	**Bit**	**Description**

		15:1	Reserved
		0	Trim
01h	Count	Number of 512-byte blocks of LBA Range Entry to be transferred, 0000h specifies that 65,536 blocks are to be transferred.	
02 – 04h	LBA	**Bit**	**Description**
		47:0	Reserved
05h	Command	XXh(TBD3)	

Trim

When trim is set to one the data in the logical blocks specified by the DATA SET MANAGEMENT command's output data becomes indeterminate.

Once a trimmed LBA has been written(e.g., a write command or a SECURITY ERASE UNIT command), the data in that logical block becomes determinate.(i.e., the logical block contains the written data.)

Trim shall not add or remove LBAs from the NV Cache Pinned Set.

6.3. Normal Outputs

See Table 110, Generic extended normal output.

6.4. Error Outputs

See Table 129, Write Log Error.

(*editor's note: Count and LBA fields should be reserved.)

6.5.

Output Data

See 7.20.3.6 for the format of the output data.

6.6. Changes to the Command Codes Table

DATA SET MANAGEMENT Optional for ATA device

6.7. Changes to the Historical Command Assignments Table

Yes.