To: INCITS Technical Committee T10
From: Fred Knight, Network Appliance
Email: knight@netapp.com
Date: Sept 9, 2008
Subject: SBC-3 Thin Provisioning Management Commands

1) Revision history

Revision 0 (Sept 9, 2008) First revision (r0)

2) Related documents

spc4r16 – SCSI Primary Commands – 4
sbc3r15 – SCSI Block Commands – 3
ssc3r04a – SCSI Sequential Commands – 3
08-149r1 – Thin Provisioning Commands
T13/e07154r6 – ATA8-ACS2 accepted TRIM proposal

3) Overview

Traditional storage devices pre-allocate physical storage for every possible logical block. There is a fixed one-to-one relationship between the physical storage and the logical storage (every logical block is permanently mapped to a physical block). Generally speaking, the physical capacity of the device is always the same as the logical capacity of the device (plus spares if any). The READ CAPACITY command reports the usable number of logical blocks to the application client. Historically, this has been referred to simply as the capacity of the device. These devices are fully provisioned.

Thinly provisioned devices also report the capacity in the READ CAPACITY command, but they do not allocate (or map) their physical storage in the same way that fully provisioned devices do. Thinly provisioned devices do not necessarily have a permanent one-to-one relationship between the physical storage and the logical storage. Thinly provisioned devices may report a different capacity (in the READ CAPACITY command), than their actual physical capacity. These devices often report a larger capacity than the actual physical capacity for storing user data.

One typical use of storage is a creation and deletion process. Files are created, possibly modified, and saved as new files (with the old one being deleted). Databases are created, where records are added, updated, and deleted.

Fully provisioned storage must allocate space to retain all possible data represented by every block described by the logical capacity (their physical capacity must be the same (or greater) than their reported logical capacity). These devices are always capable of receiving write data into the pre-defined and pre-allocated space.

Thinly provisioned devices may or may not pre-allocate space to retain write data. When a write is received, physical storage may be allocated from a pool of available storage to retain the write data and a mapping established between the location of that physical storage and the appropriate location within the logical capacity (a physical to LBA mapping). As long as this allocation and mapping process is successful, the write operates in the same way that it does on a fully provisioned storage device. However, if all the available physical capacity has been used, and no space can be allocated to retain the write data, the write operation must fail. This failure must have a new unique ASCQ.

In addition, there are several management functions related to thin provisioning. One example is the need to determine which blocks contain persistent user data (and protection information) vs. blocks which do not contain persistent user data. This is important in data copy operations or mirror operations where the desire is to copy only the persistent user data, but to not copy those blocks which do not contain persistent user data.

This is a first draft to begin the discussion of Capacity Management Operations (possibly to include determining and setting the soft threshold point, and other operations).

Existing text is shown in **BLACK**, new text is shown in **RED**, and comments (not to be included) are shown in **BLUE**.

Proposal:

5. x Data Set Management (DSM) IN command

5. x. 1 DSM IN command overview

The DSM IN command shall be implemented by device servers supporting Thin Provisioning (see 4.4.1.2). The DSM command (see table x.1) requests that the device server send provisioning status information to the application client.

Table x.1 – DSM Command

Bit / Byte	7	6	5	4	3	2	1	0
0	OPERATION CODE (9Eh)							

1	Reserved	SERVICE ACTION (nnh- 01h)
2	(MSB) LOGICAL BLOCK ADDRESS	
9	(LSB)	
10	(MSB) ALLOCATION LENGTH	
13	(LSB)	
14	RESERVED	
15	Control	

The OPERATION CODE field is defined in SPC-4 shall be set to the value defined in table x.1.

The Service Action field is defined in SPC-4 and shall be set to the value defined in table x.4.

The LOGICAL BLOCK ADDRESS field specifies the LBA of the first logical block accessed by this command. If the specified LBA exceeds the capacity of the medium, then the device server shall terminate the command with CHECK CONDITION status with the sense key set to ILLEGAL REQUEST and the additional sense code set to LOGICAL BLOCK ADDRESS OUT OF RANGE.

The ALLOCATION LENGTH field specifies the maximum number of bytes that the application client has allocated for returned parameter data. An allocation length of zero indicates that no data shall be transferred. This condition shall not be considered an error. The application client should specify an ALLOCATION LENGTH that is a (multiple of 16) + 8. The device server shall terminate transfers to the data-in buffer when:

a) the number of bytes specified by the ALLOCATION LENGTH field has been transferred,
b) when data representing all logical blocks higher than the specified logical block address has been transferred, or
c) the device server has transferred at least 1 LBA descriptor.

The contents of the parameter data shall not be altered to reflect the truncation, if any, that results from an insufficient allocation length.

The contents of the CONTROL byte are defined in SAM-4.

5. x. 2 DSM IN parameter list

5. x. 2.1 DSM IN parameter data

The DSM IN PARAMETER DATA (see table x.5) contains an eight-byte header followed by one or more LBA DESCRIPTORS. The LBA supplied in the CDB shall be the LOGICAL BLOCK ADDRESS field in the first LBA DESCRIPTOR. The LBA DESCRIPTORS shall be in ascending LBA order. The LBA in descriptor n+1 shall be the LBA of descriptor n + the extent length of descriptor n.

Table x.5 DSM parameter data

Bit / Byte	7	6	5	4	3	2	1	0
0	DSM IN PARAMETER LIST LENGTH (n-3)							
3								
4	Reserved							1DESC
5	Reserved							
7								

	LBA DESCRIPTOR LIST
8-23	LBA DESCRIPTOR 1
24-39	LBA DESCRIPTOR 2
…	…
n	LBA DESCRIPTOR m

The DSM IN PRARMETER LIST LENGTH field indicates the length in bytes of the LBA DESCRIPTOR LIST. The relationship between the DSM IN PARAMETER LIST LENGTH field and the CDB ALLOCATION LENGTH field is defined in SPC-4.

A 1DESC bit set to one indicates that the device server will only report one LBA DESCRIPTOR.

A 1DESC bit set to zero indicates that the device server may report more than one LBA DESCRIPTOR.

5. x. 4.1.1 LBA DESCRIPTOR

The LBA DESCRIPTOR (see table x.6) contains LBA provisioning status information.

Table x.6 LBA STATUS DESCRIPTOR

Bit / Byte	7	6	5	4	3	2	1	0
0	(MSB)			LOGICAL BLOCK ADDRESS				
7								(LSB)
8	(MSB)			EXTENT LENGTH				
11								(LSB)
12	RESERVED							MAP
13	RESERVED							
14	REESRVED							
15	RESERVED							

A MAP bit set to one indicates that the specified LBA range is mapped to a physical block range and contains persistent user data and protection information if enabled.

A MAP bit set to zero indicates that the specified LBA range is not mapped to a physical block range and does not contain persistent user data or protection information if enabled.

The EXTENT LENGTH field specifies the number of logical blocks represented by this descriptor.