

AV Lite Proposal

To: T13 Technical committee
From: Jim Hatfield
 Seagate Technology
 389 Disc Drive
 Longmont, CO 80503
 Phone: 720-684-2120
 Fax: 720-684-2711
 Email: James.C.Hatfield@seagate.com
Date: January 14, 2003

- Dec. 18, 2002 - Approved by T13
- Jan. 13, 2003
 - TBD values assigned by T13
 - Minor clarifications added
 - Set DWE if Stream Error occurs during a write or flush command
- Jan 15, 2003
 - correct typos, misspellings

1 Introduction

The purpose of this proposal is to offer a simpler compromise to the tradeoff of performance vs. data quality. This has been made a priority among manufacturers and consumers of Audio/Visual devices.

The following chart summarizes some of the requirements.

In order to meet the cost needs of both Users, it would be useful to include the AV needs into the standard product.

Feature	Traditional User	AV User
Low Cost	Critical Importance	Critical Importance
Sequential Access Performance	Critical Importance	Critical Importance
Random Access Performance	Critical Importance	Nice to have
Data Integrity	Critical Importance	Important
Worst Case R/W Times	Less Important	Critical Importance

In the past, a number of approaches were tried:

Solution	Advantages	Disadvantages
SeaStream®	1. Controlled WorstCase R/W Time	1. Very Complex 2. Huge Code Size 3. Not Compatible with Standard DeskTop Product 4. Required Tight Coupling with Host Application
T13/D99128 Revision 8	1. Good Error Reporting 2. Driver Level Implementation 3. Mixed AV and Traditional Mode	1. New AT Commands 2. Difficult to Implement 3. Not Accepted As Standard
legacy drive	1. Simple Driver Level Development 2. Tradition AT command Set – Host Development can be largely done now. 3. Mixed AV and Traditional 4. Small Code Size and easy implementation.	1. Not a Standard
Proposed A/V Lite	1. Simple Driver Level Development 2. Traditional AT command Set – Host Development can be largely done now. 3. Standard AT Error Reporting (SMART Error Logging) 4. Mixed AV and Traditional 5. Small Code Size and easy implementation.	Not yet a Standard

2 Proposed Specification Changes:

The following AT/ATA-7 sections would be modified:

- Define and describe a new feature set, options, how to use it, etc.
- Flush Cache/Flush Cache Ext commands (5.13, 5.14):
- DMA Read/Writes (Ext) commands (5.26, 5.27, 5.57, 5.58):
- Set Features command (5.48):
- Identify/Identify Packet Device commands (5.16, 5.17):
- Read Log EXT command (5.30)

2.1 Feature Set Descriptions

2.1.1 (Add this section)

4.TBD AV Lite feature set

- Optional for devices not implementing the PACKET Command feature set
- Prohibited for devices implementing the PACKET Command feature set

The purpose of this feature is to define a mode of operation that balances performance with reliability.

Normally, reliably managing data is of paramount importance. There are, however, applications which prefer to sacrifice some data reliability in order to achieve higher levels of performance.

This mode applies only to these qualified commands: FLUSH CACHE, READ DMA, READ DMA EXT, WRITE DMA and WRITE DMA EXT.

The basic idea is for the host to define (to the device) a maximum time limit during which a group of commands is expected to complete. The device shall attempt to guarantee completion (of the group of commands) within the time limit.

The timer (in the device) has these mutually exclusive states: disabled, armed, running and expired. The timer does NOT apply to each individual command, but to the combined time required to execute a 'group' of commands.

If the device does not complete a qualified command before the timer expires, the device shall either abort the command or continue (possibly transferring incorrect data).

When this feature is supported and enabled:

The Host:

1. may terminate this mode by executing a SET FEATURES command to set a zero value for a WorstCaseTimer;
2. may enter this mode by executing a SET FEATURES command to set a **non-zero** value for a WorstCaseTimer;
3. may execute a SET FEATURES command to specify the action to be taken by the device if a command does not complete within the define time window (e.g. abort or continue with bad data: 'ReadContinuous / WriteContinuous' option)
4. shall 'arm' the timer for the first group of commands by executing a FLUSH CACHE command, knowing that the timer does not start until the host sends the first qualified command to the device;
5. shall execute each group of commands with this protocol:
 - a. execute a qualified command on the device. This starts the WorstCaseTimer.
 - b. execute any group of commands to the device. The number and type of commands in this group depends on the expected average completion time (for the specific commands on this specific device) and on the selected WorstCaseTimer value;
 - c. expect that the group of qualified commands shall complete before the time limit expires. Commands other than those listed above shall not be affected by the WorstCaseTimer.
 - d. complete the execution of a FLUSH CACHE command BEFORE the expiration of the time limit. This shall re-arm the time for the next group of commands. If the timer expires DURING a Flush Cache command, the buffered data is not guaranteed to still be in the buffer. It should be considered as 'lost' or 'destroyed'

The Device:

1. on receipt of a FLUSH CACHE command shall stop the WorstCaseTimer, re-arm it, but not allow it to run;
2. shall start the WorstCaseTimer on receipt of the first qualified command after a FLUSH CACHE command
3. shall not allow the WorstCaseTimer to affect the outcome of commands other than those listed above;
4. all qualified commands must complete before the WorstCaseTimer expires. If the command cannot complete because of retries (or any other reason), it shall be terminated (either aborted or finished with incorrect data).The data returned may include the bad sector.

When this feature is not supported or not enabled (e.g. Normal mode):

1. The WorstCase timer and the 'ReadContinuous / WriteContinuous' option are ignored (or not even implemented).
2. All configured retries shall be executed before an error is returned to the host.
3. If a Read command cannot complete because of retries, the data returned may include the bad sector.

2.2 IDENTIFY DEVICE changes

2.2.1 (Add the following to table 14)

Table 14 – IDENTIFY DEVICE information (continued)

Word	Content
84	Command sets supported. If words 82 and 83 = 0000h or FFFFh command set notification not supported. 11 1=A/V Lite Supported 12 1=ReadContinuous / WriteContinuous in A/V Lite Supported
87	Command set/feature enabled. If word 85, 86, and 87 = 0000h or FFFFh command set enabled notification is not supported. 11 1=A/V Lite Mode Enabled 12 1=ReadContinuous / WriteContinuous in A/V LITE (vs Abort on Error) Enabled
116	Number of msec for the WorstCaseTimer in A/V Lite

2.2.2 Modify the following sections**5.16.42 Words (84:82) Features/commands sets supported**

(Add the following text):

“If bit **11** of word 84 is set to one, the drive supports A/V Lite (i.e. a WorstCaseTimer can be set)”

If bit **12** of word 84 is set to one, the drive supports ReadContinuous/WriteContinuous on Errors in A/V Lite . “

5.16.43 Words (87:85) Features/command sets enabled

(Add the following text):

“If bit **11** of word 87 is set to one, A/V Lite has been enabled. If bit ~~6~~**11** of word 87 is cleared to zero, the drive is operating in normal PC mode.”

“If bit **12** of word 87 is set to one, the data sent to the host during Timeout conditions may not be accurate. If bit ~~7~~**12** of word 87 is cleared to zero, a A/V Lite Timeout shall abort the command. Any data sent to the host shall be accurate.”

2.2.3 (Add the following sections):**5.16.TBD Word 116: WorstCaseTimer for A/V Lite**

If A/V Lite is supported and enabled, this word equals the number of milliseconds (in 10-millisecond increments) that the WorstCase Timer has been initialized to.

e.g. a value of 70 indicates a 700 millisec WorstCaseTimer.

2.3 SET FEATURES Changes:

2.3.1 (Add the following to table 39)

Table 39 – SET FEATURES register definitions

Value (see note)	
E0h	Obsolete
20h	Set A/V Lite WorstCaseTimer in 10 milli <i>milli</i> sec increments
21h	Set A/V Lite Error Handling

2.3.2 (Add the following sections)

5.48.TBD Obsolete (E0h)

This command is included for legacy implementations.

5.48.TBD Set A/V Lite WorstCaseTimer in 10 ~~milli~~*milli*sec increments (20h)

This is only valid if the A/V Lite Feature set is supported.

Notes – (Only if A/V Lite is supported)

1. The host initializes the Sector Count Register with the desired value.
2. A value of zero in the Sector Count Register shall disable A/V Lite, and shall clear bit 11 to zero in word 87 of the IDENTIFY string.
3. The WorstCaseTimer shall be set to zero on any reset.
4. A non-zero value shall set bit 11 to one in word 87 of the IDENTIFY string. In addition, this will arm (but not start) the WorstCaseTimer.
5. A manufacturer may modify any timer set below it's minimum retry time to that drives minimum retry time. (e.g. if the user attempts to set this to 10 ~~milli~~*milli*sec and the manufacturer has a minimum retry time of 200 ~~milli~~*milli*secs, the drive's WorstCaseTimer shall be set to 200 ~~milli~~*milli*secs)
6. The WorstCaseTimer value shall be displayed in Word 116 of the Identify Device Sector.
7. This timer represents Worst Case time from the host perspective. The drive shall not exceed this time. Internally, the actual time that shall trigger a timeout shall be shorter than this value to ensure that this value is not exceeded.

e.g. Sector Count set to 70 indicates a 700 ~~milli~~*milli*sec WorstCaseTimer.

5.48.TBD Inputs

Register	7	6	5	4	3	2	1	0
Features	20h							
Sector Count	10 milli <i>milli</i> sec increments of the WorstCaseTimer DMA R/W cmds							
LBA Low								
LBA Mid								
LBA High								
Device	obs	LBA	obs	DEV				
Command	EFh							

5.48.TBD Set A/V Lite Error Handling (21h)

This is only valid if the A/V Lite Feature set is supported and enabled.

This enables or disables the “ReadContinuous / WriteContinuous in A/V Lite” feature.

A value of zero in the Sector Count Register shall cause aborts if the A/V Lite WorstCaseTimer expires on a R/W command, and shall clear bit 12 to zero of word 87 of the IDENTIFY result data.

A value of one in the Sector Count Register shall cause potentially incorrect data to be sent if the WorstCaseTimer expires, and shall set bit 12 to one in word 87 of the Identify Device Sector.

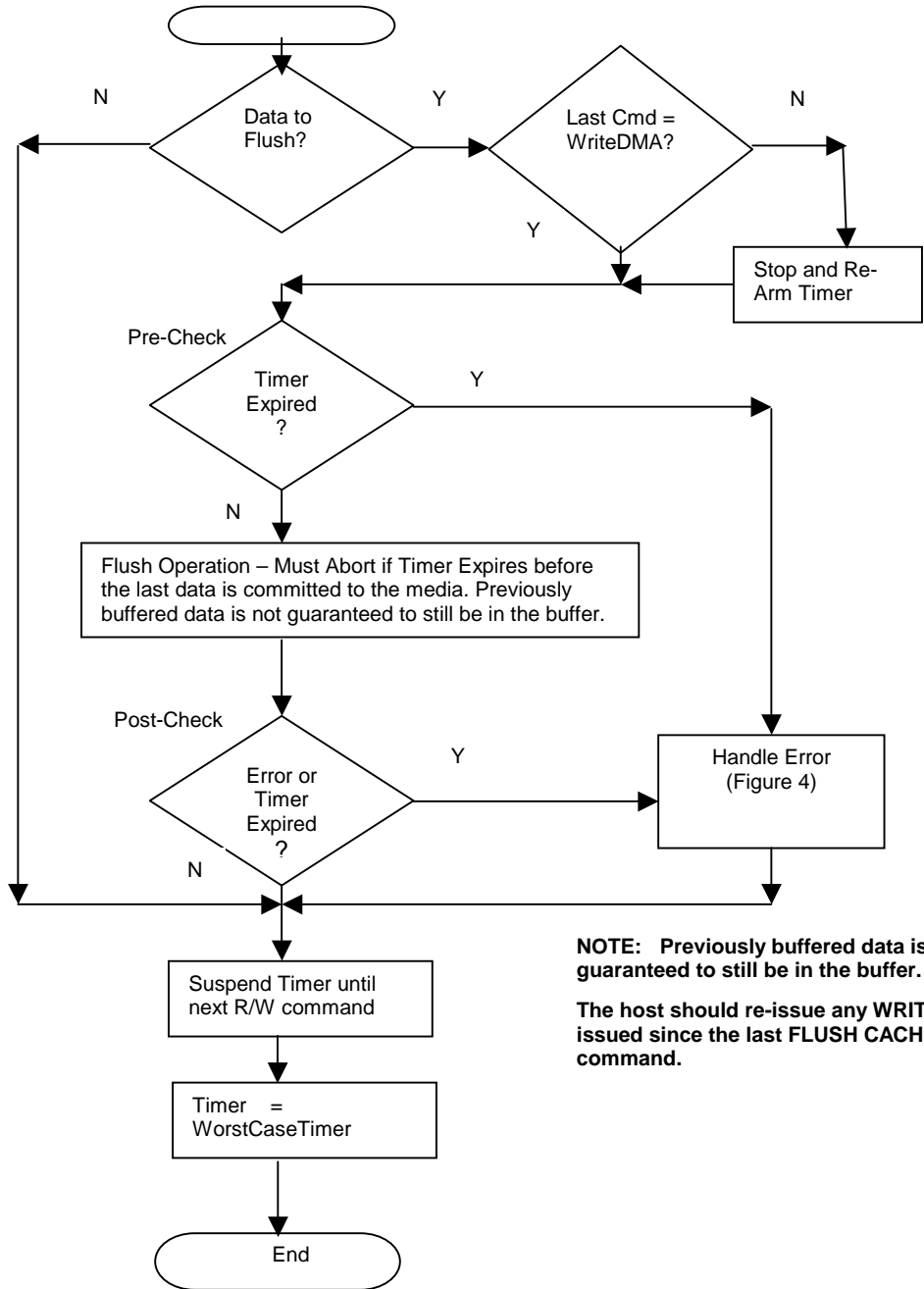
5.48.TBD Inputs

Register	7	6	5	4	3	2	1	0
Features	21h							
Sector Count	0 = Abort; 1 = ReadContinuous / WriteContinuous							
LBA Low								
LBA Mid								
LBA High								
Device	obs	LBA	obs	DEV				
Command	EFh							

2.4 FLUSH CACHE Changes

If A/V Lite is supported and enabled, the following algorithm shall be used:

Figure 1



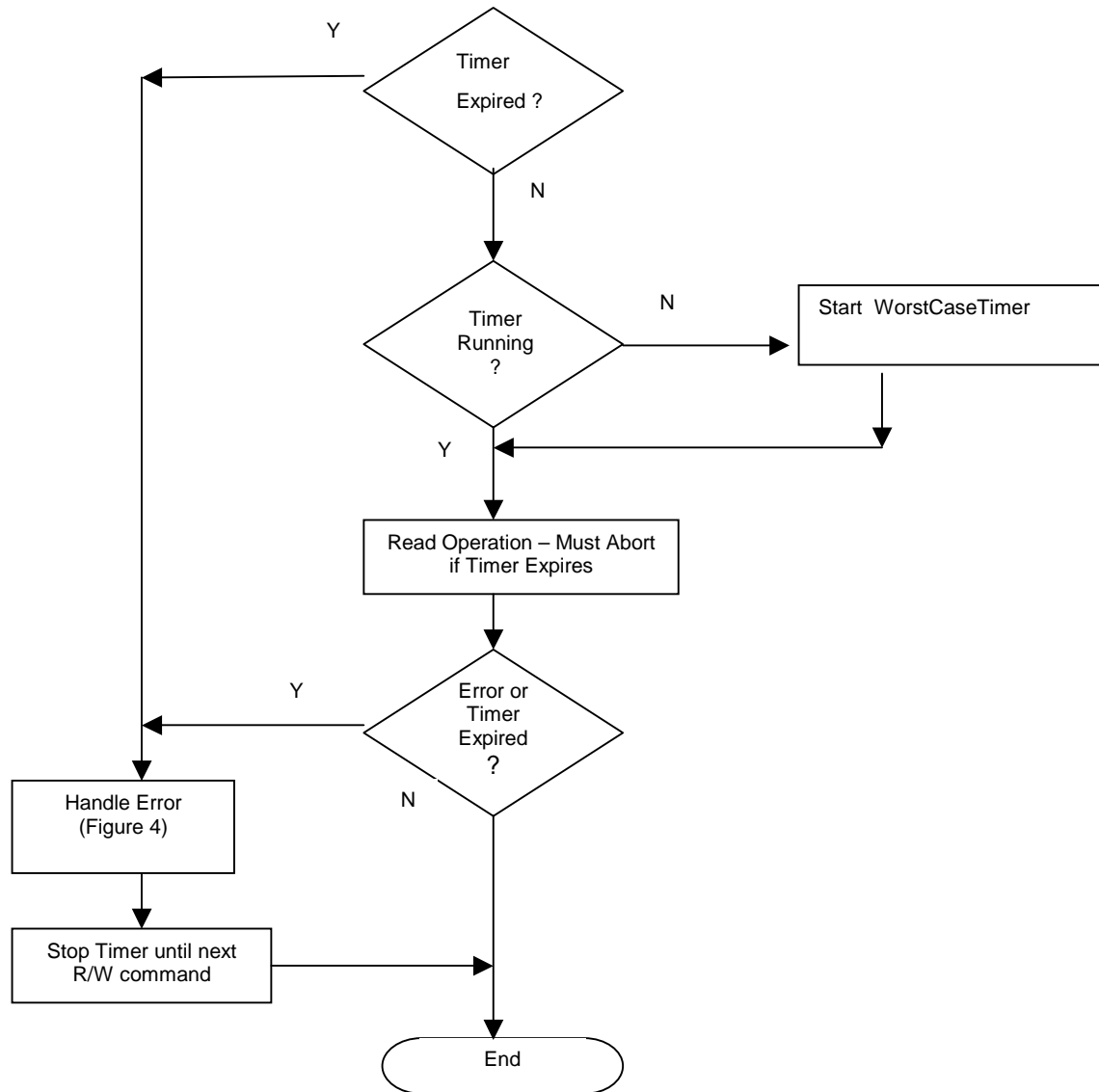
NOTE: Previously buffered data is not guaranteed to still be in the buffer.

The host should re-issue any WRITES issued since the last FLUSH CACHE command.

2.5 READ DMA and READ DMA EXT command changes

If A/V Lite is supported and enabled, the following algorithm shall be used:

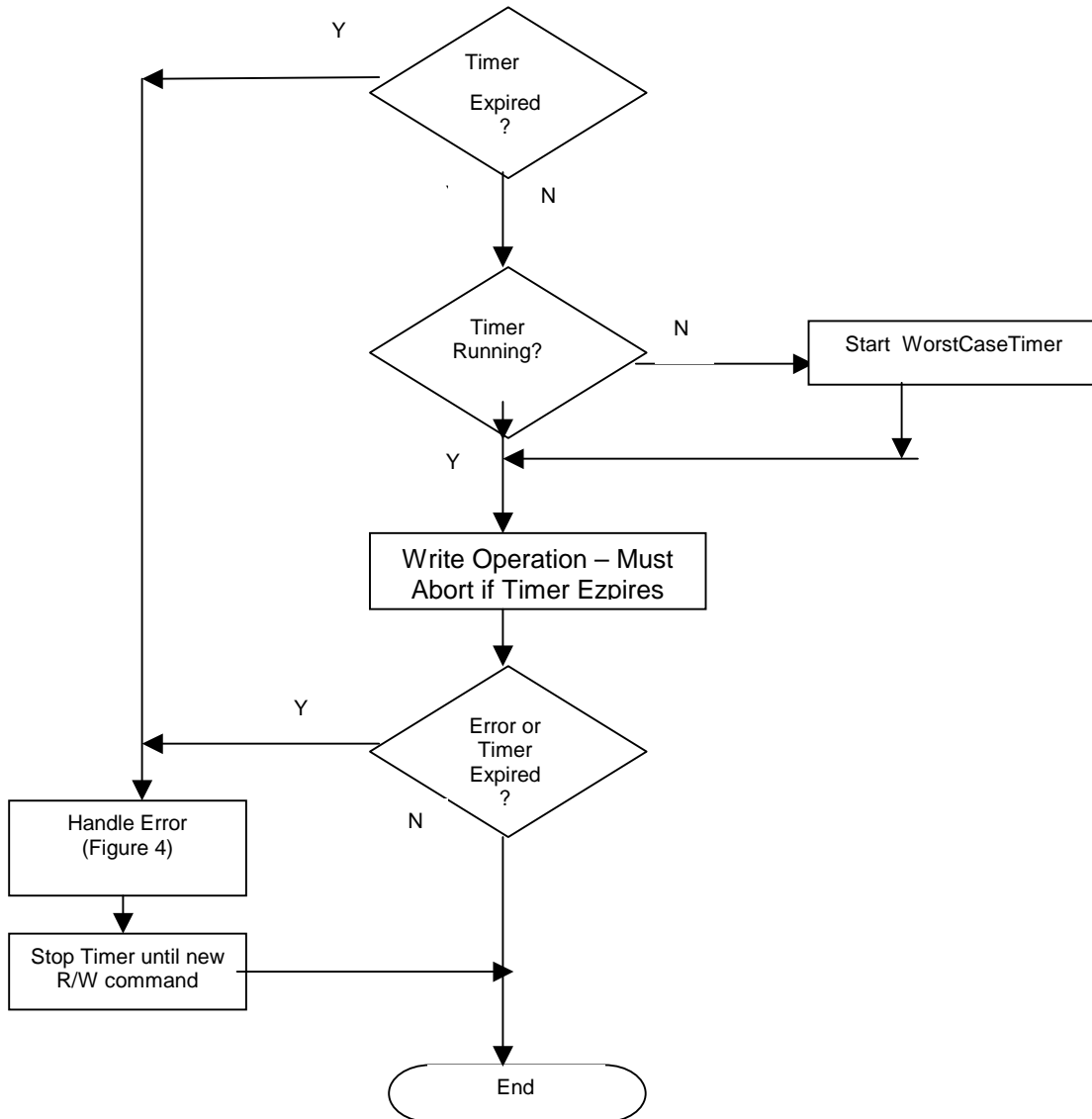
Figure 2



2.6 WRITE DMA and WRITE DMA EXT command changes

If A/V Lite is supported and enabled, the following algorithm shall be used:

Figure 3

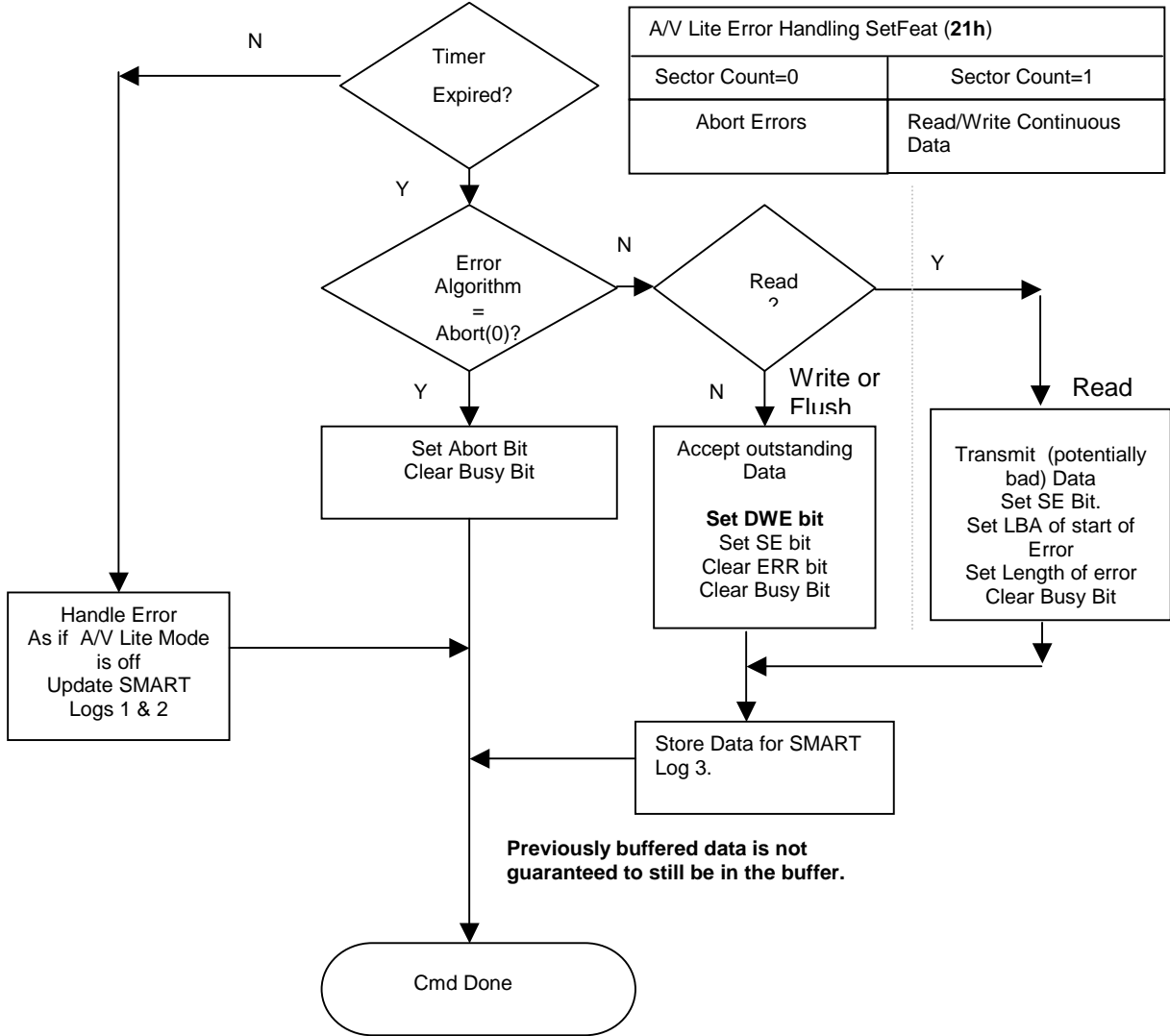


2.7 ERROR HANDLING

Two error handling algorithms are supported:
 Abort the Command and report the LBA of the Aborted Command; or
 Transfer “garbage” data if data is not currently available.

The algorithms are selected via the Set Features Command –
 Features Register = **20h** and **21h**

Figure 4



5.26.6 Error Outputs (READ DMA)
5.57.6 Error Outputs (WRITE DMA)

Register		7	6	5	4	3	2	1	0
Error		ICRC	UNC	MC	IDNF	MCR		NM	obs
Sector Count		Length of Error (7:0)							
LBA Low		LBA (7:0)							
LBA Mid		LBA (15:8)							
LBA High		LBA (23:16)							
Device		obs	na	obs	DEV	Reserved			
Status		BSY	DRDY	SE	SRV	DRQ	DWE	na	ERR

Sector Count register-
 bits (7:0) number of contiguous sectors containing potentially bad data, beginning with the LBA of the first sector with an uncorrectable error.

LBA
 Indicates the LBA of the first sector with an uncorrectable error.

Status register -
SE (Stream Error) shall be set to one if an error has occurred during the execution of the command and Read /Write Continuous Mode is set. In this case the LBA returned in the Sector Number registers shall be the address of the first sector potentially in error, and the Sector Count registers shall contain the number of consecutive sectors that may contain errors.

DWE (Deferred Write Error) shall be set if an error was detected in a deferred write to the media. This error is from a previously issued command. If DWE is set to one and SE is cleared to zero, the values in Sector Count, Sector Number, Cylinder High, and Cylinder Low are undefined.

ERR shall be set to one if an Error register bit is set to one. If Read /Write Continuous Mode is set, the error bit shall be cleared to zero.

5.27.6 Error Outputs (READ DMA EXT)
5.58.6 Error Outputs (WRITE DMA EXT)

Register		7	6	5	4	3	2	1	0
Error		ICRC	UNC	MC	IDNF	MCR		NM	obs
Sector Count	HOB=0	Length of Error (7:0)							
	HOB=1	Length of Error (15:8)							
LBA Low	HOB=0	LBA (7:0)							
	HOB=1	LBA (31:24)							
LBA Mid	HOB=0	LBA (15:8)							
	HOB=1	LBA (39:32)							
LBA High	HOB=0	LBA (23:16)							
	HOB=1	LBA (47:40)							
Device		obs	na	obs	DEV	Reserved			
Status		BSY	DRDY	SE	SRV	DRQ	DWE	na	ERR

Sector Count register-

bits (7:0) number of contiguous sectors containing potentially bad data, beginning with the LBA of the first sector with an uncorrectable error (read with Device Control register HOB bit cleared to zero)

bits (15:8) number of contiguous sectors containing potentially bad data, beginning with the LBA of the first sector with an uncorrectable error (read with Device Control register HOB bit set to one)

LBA

Indicates the LBA of the first sector with an uncorrectable error

Status register -

SE (Stream Error) shall be set to one if an error has occurred during the execution of the command and Read /Write Continuous Mode is set. In this case the LBA returned in the Sector Number registers shall be the address of the first sector potentially in error, and the Sector Count registers shall contain the number of consecutive sectors that may contain errors.

DWE (Deferred Write Error) shall be set if an error was detected in a deferred write to the media. This error is from a previously issued command. If DWE is set to one and SE is cleared to zero, the values in Sector Count, Sector Number, Cylinder High, and Cylinder Low are undefined.

ERR shall be set to one if an Error register bit is set to one. **If Read /Write Continuous Mode is set, the error bit shall be cleared to zero.**

2.8 READ LOG EXT Changes

2.8.1 A/V Lite Errors (new section)

Write stream errors shall be placed in log 21h. Read stream errors shall be placed in log 22h.

These logs are formatted identically to SMART Error Log 1 and SMART Error Log 2. It shall contain a manufacturer-defined number of sectors (recommended 20). This number is defined in READ LOG EXT Log 0.

Data for this log is temporarily saved whenever a timeout occurs. As workload permits, this data shall be written to the READ LOG EXT Log **21h or 22h**. This logging of errors is not to affect performance. Performance during A/V Lite Operation is more critical than Error Logging. If this means that errors fail to log, then some errors shall not be logged. This being stated, if the Error Log can be updated for future reporting without affecting performance, it should be done.

The purpose of this log is to give the host a list of errors on the drive that can be reallocated or not accessed in the future. The host shall be able to reallocate these sectors by writing to them in normal (not A/V Lite) mode. The host may also choose simply not to access these LBA's via its own algorithms.