1 Performance Control Proposal

1.1.1 Performance Control Feature Set

The Performance Control feature set is mandatory for NAND flash based cards. It is dependent on implementing the General Purpose Log feature set and the 48-bit Addressing feature set. The purpose of this feature set is to enable the host to optimize the performance of the card, and for the card to guarantee a level of performance in streaming applications like A/V applications.

If the Performance Control feature set is supported, the following commands shall be supported:

1. Performance Assign Ext
2. Performance Management DMA Ext
3. Performance Release Ext

The following GPL feature set logs shall be implemented:

1. Performance Control Description Log

Support of the Performance Management feature set is indicated in IDENTIFY DEVICE data word XX bit X.

The Performance Control features may be used by the host to notify the card of its intended operations and of the LBAs of specific memory areas that may help the card improve performance. When a card conforms with a specific profile or profiles, the performance guarantee feature is supported, see section 3.1.

1.1.1.1 Performance Assign Ext – BBh

Register	Bit ->	7	6	5	4	3	2	1	0
Command (7)		BBh							
Device (6)		1	1	1	Drive	Reserved			
LBA High (5)	Current Previous	Performance Control record description word address in the page(7:0) X							
LBA Mid (4)	Current Previous	Performance Control record description log page number(7:0) Performance Control record description log page number (15:8)							
LBA Low (3)	Current Previous	X X							
Sec Cnt (2)	Current Previous	X X							
Feature (1)	Current Previous	02h – Write file, 03h – Read file X							

Figure 1: Performance Assign Ext

This command is a part of the Performance Control feature set. It causes the card to assign resources for a new performance controlled file stream, if available, and return the new 32 bit File Stream ID.

The assignment of File Stream IDs shall be volatile. After power-up, hardware rest, or soft reset the card shall have no assigned File Stream IDs.

Feature Register:
- Bit 0 shall be set to 0 when requesting a Write File Stream ID, and set to 1 when requesting a Read File Stream ID,
- Bit 1 shall be set to1,
- Bits 2-7 shall be set to 0.

LBA registers shall contain a reference to the specific Performance Control record intended to be used by this file stream. The records are documented using the GPL feature at log address 26h or TBD. Each log page may contain more than one record. The LBA high, current register shall contain the address of the record in the log page.

The command, if successful, returns the new File Stream ID in the LBA registers. See Table 1.

Table 1: Performance Assign result

Register	**Bit ->**	**7**	**6**	**5**	**4**	**3**	**2**	**1**	**0**
Status (7)		Status Register							
Device (6)		1	1	1	Drive	Reserved			
LBA High (5)	Current Previous	File Stream ID(23:16) X							
LBA Mid (4)	Current Previous	File Stream ID(15:8) X							
LBA Low (3)	Current Previous	File Stream ID(7:0) File Stream ID(31:24)							
Sec Cnt (2)	Current Previous	X X							
Error (1)	Current Previous	Error register X							

If the command is not successful, the LBA registers will contain zeros.

If the File Stream ID cannot be assigned, or the Performance Control record pointed to by the command is invalid, the card shall abort the command and generate an error. It shall return 1 in bit 4 (IDNF) of the Error Register, and return zeros in the LBA registers. The extended error information returned shall be code 14h – “ID Not Found”.

1.1.1.2 Performance Management DMA Ext – BBh

Register	Bit ->	7	6	5	4	3	2	1	0
Command (7)		BBh							
Device (6)		1	1	1	Drive	Reserved			
LBA High (5)	Current Previous	X X							
LBA Mid (4)	Current Previous	X X							
LBA Low (3)	Current Previous	X X							
Sec Cnt (2)	Current Previous	256 word block count(7:0) 256 word block count(15:8)							
Feature (1)	Current Previous	04h X							

Figure 2: Performance Management DMA Ext

The Performance Management DMA Ext command transfers a set of records detailing LBA ranges and their performance control related attributes. Each record is composed of 16 words. See Table 2. The command shall transfer the records using an integer number of 256 word blocks.

The Sector Count register shall contain the number of 256 word blocks to be transferred. The number of blocks shall be the integer portion of the number of records divided by sixteen. If the remainder of this operation is not zero, the host shall add one more block to the count. The minimum number of word blocks to be transferred shall be 1.

Table 2 – Range Record Structure

Word	Size	Name	Description
0 to 1	DWord	Q_T	Type of data range
2 to 3	DWord	Q_{ID}	File Stream ID when applicable, 0 otherwise.
4 to 7	QWord	Q_S	Start LBA of the data
8 to 11	QWord	Q_L	Length, in sectors.
12 to 15	QWord	Q_R	Reserved, set to 0.

The defined data range types are defined in Table 3. The card may use the information contained in the range records.

Hosts using the Performance Guarantee feature in compatible cards shall use data range types 1, 2, 3.

Table 3 – Data Range Types.

Type	Description	Card implementation and comments
00000000h	This record contains no information, and no more records are available.	Mandatory
00000001h	Fixed file system data. "File system data" is data used by the computer operating system to manage user files. Fixed file system data is file system data the location of which is fixed and unchangeable. Examples (in some file systems) include the Master Boot Record (MBR), File Allocation Table (FAT), UEFI GPT, exFAT tables, root directory. If this range contains information regarding a specific File Stream ID then the Q_{ID} field shall contain the File Stream ID.	Mandatory
00000002h	General file system data. This category includes all other file system data, for example – non-root directories and tables the location of which may shift over time or that are relevant to a small number of files only. If this range contains information regarding a specific File Stream ID then the Q_{ID} field shall contain the File Stream ID.	Mandatory
00000003h	File data. The Q_{ID} field shall contain the File Stream ID, or 0 if no one specific File Stream ID is associated with the range.	Mandatory
00000004h	Non File Stream ID related data that the host plans to access often. Example: OS virtual memory cache.	Mandatory
00000005h	Non File Stream ID related Read/Write user data.	Mandatory
00000006h	Non File Stream ID related Read Only user data.	Mandatory
00000007h to c33cf55eh	Reserved	
c33cf55fh	No data. This lets the card know that this LBA range data is discarded. When trying to read from a discarded LBA using any Read command the returned data is undefined.	Mandatory
c33cf560h to ffffffffh	Reserved	

The command shall transfer one or more 256 word blocks. The first occurrence of 0000h in the Q_T field shall terminate the record set. All unused words shall be set to 0.

If the sector count is zero the card shall abort the command.

If the card is unable to parse and recognize any record, the command shall abort with an error and return the position of the bad record in the LBA registers. See Figure 3.

LBA High (5)	Current Previous	Address of record in the block (0 to 7) X
LBA Mid (4)	Current Previous	256 word block number(15:8) X
LBA Low (3)	Current Previous	256 word block number(7:0) X

Figure 3: LBA registers contents when error encountered

All the records up to and not including the bad record shall be processed. The bad record and any record following it shall not be processed.

1.1.1.3 Performance Release Ext – BBh

Register	**Bit ->**	**7**	**6**	**5**	**4**	**3**	**2**	**1**	**0**
Command (7)		BBh							
Device (6)		1	1	1	Drive	Reserved			
LBA High (5)	Current Previous	File Stream ID(23:16) X							
LBA Mid (4)	Current Previous	File Stream ID(15:8) X							
LBA Low (3)	Current Previous	File Stream ID(7:0) File Stream ID(31:24)							
Sec Cnt (2)	Current Previous	X X							
Feature (1)	Current Previous	08h X							

Figure 4: Performance Release Ext

This command is a part of the Performance Control feature set. It causes the card to release the resources used for a previously assigned performance controlled file.

The Feature Register shall be set to 04h.

If the command is unsuccessful or the File Stream ID is not recognized the card shall abort the command and return 1 in bit 4 (IDNF) of the Error Register.

If the command fails because the device is unable to find the File Stream ID then the extended error information returned shall be code 14h – “ID Not Found” .

1.1.2 Performance Guarantee

Cards supporting the Performance Control feature set may support the Performance Guarantee feature.

Devices supporting this feature shall have one or more LBA ranges, each associated with at least one Performance Control record. See section2 , in particular Table 8.

The Performance Guarantee feature guarantees a certain card performance in megabytes per second when the host behaves in accordance with the profile the card implements and the rules defined in this section.

A performance guarantee session starts when a Performance Assign Ext command is successfully executed, and ends when a Performance Release Ext command is successfully executed and the card has no more assigned File Stream IDs.

Writing or reading a stream is done in three stages.

1. File stream ID allocation,
2. File stream writing or reading,
3. File stream ID deallocation.

1.1.2.1.1 File Stream ID Assignment Process

During the file stream ID assignment process the host shall find the card capabilities and assign a file stream ID. The number of available file stream IDs for all performance control records may change as a result of this process. The host shall either use the process as defined in the profile referred to by the PR field in the performance control record, or make sure that only one file stream ID assignment process shall be in process at any given time.

1. The host shall issue Read Log (DMA) Ext commands to find the number of pages and to read the performance control records, see paragraph 2.3.

2. If a record's Type field is equal to 0 this is a write record, else this is a read record. Use a performance control record that matches the intended file stream operation: reading or writing. The host shall choose the record used for the file stream it is allocating.

3. For the record chosen, If S_{Tm} is greater than 0 and S_{Ta} is equal to 0 then there are no more available file stream IDs available to support this record, and the file stream ID assignment process should not proceed.

4. The host should use the performance, if specified, by the profile referred to in the PR field of the performance control record, see Table 8. Alternatively, the host may calculate the available bandwidth for this range by using parameter specified in the performance control record. It is $\mathbf{BW} = \frac{\mathbf{AU} * \mathbf{RU} * (\mathbf{Sector\ size\ in\ Bytes})}{\mathbf{T_{AU}} + \mathbf{T_F}} \left[\frac{\mathbf{Megabytes}}{\mathbf{Second}}\right]$.

5. If the host uses the full performance that the performance record allows, it shall not assign or use more file stream IDs.

6. If the host uses less than the full performance calculated earlier, the host may assign more file stream IDs based on the same or other records, as long as the sum of all the fractions of the required performance divided by the maximum performance used for each record does not exceed one.

7. The host shall issue a Performance Assign Ext command to assign the file stream ID.

1.1.2.2 File Stream Usage - Write

1. Choose the next AU to write the file stream into. Make sure it does not contain data that needs to be saved.

2. Issue the Performance Management DMA Ext command using the file stream ID and defining three record types: type 1 for the fixed file system area that will be affected. Type 2 for the file system area that is specific to the file being written, and the location of which is not constrained by the file system. Type 3 for the full AU to be written.

3. Issue any other required file system related commands in preparation for writing the AU. File system related read and write commands should have the LBA and sector count required for the operation, without regard to RU and AU sizes and boundaries.

4. Write the AU starting from the first RU in integer multiples of RU size. The shortest command used shall have a sector count equal to one RU, and the longest command used shall have a sector count equal to the number of sectors in the AU (RU*AU). See Table 8.

5. Issue any other required file system related commands required after writing the AU.

6. If an AU or a part of an AU is allocated to a file stream using the Performance Management DMA Ext command, and any data was written to the AU, then any part of the AU that was not written shall contain undefined data.

7. If the host expects a performance guarantee as specified by the profile, then the host shall follow the Performance Guarantee process as defined by this section and by the profile, and shall not issues any command that is not a part of this process.

1.1.2.3 File Stream Usage – Read

1. Choose the next AU to read the file stream from.

2. Issue the Performance Management DMA Ext command using the file stream ID and defining three record types: type 1 for the fixed file system area that will be read. Type 2 for the file system area that is specific to the file being read, and the location of which is not constrained by the file system. Type 3 for the data to be read.

3. Issue any other required file system related commands in preparation for reading the AU. File system related read and write commands should have the LBA and sector count required for the operation, without regard to RU and AU sizes and boundaries.

4. Each Read command shall use an LBA which is the first LBA in a RU (see Table 8) and have a sector count that is an integer multiple of the number of sectors in a RU. One Read command shall not read data from two different AUs, even if their logical block addresses are contiguous.

5. If the host expects a performance guarantee as specified by the profile, then the host shall follow the Performance Guarantee process as defined by this section and by the profile, and not issues commands that are not a part of this process.

1.1.2.4 File Stream ID Deallocation Process

1. Issue the Performance Release Ext command. The card may release any internal resources used. Any Write activity into an AU that was allocated using a type 3 record (see Table 8) by issuing a Performance Management DMA Ext command shall cause the unwritten portion of the AU to contain undefined data.

2. If, after releasing the file stream ID, no more file stream IDs are assigned, the card may assert the busy line for up to 500 milliseconds before it is ready for the next operation.

2 Appendix B

(Normative)

2.1 Performance Control Log Definitions and Overview

A complete overview of the ATA General Purpose Logging (GPL) log definitions appears in *BSR INCITS 452-200x, AT Atachment - 8 ATA/ATAPI Command Set (ATA8-ACS), Annex A*.

This appendix describes the minimum requirements for the CFA Performance Control feature set related log implementation. This is the only log related feature required by the CFA specification. No other log is required for CFA implementations.

These logs are accessible via commands in the General Purpose Logging (GPL) feature set. Table 4 is a summary of these logs. There are 3 terms associated with logs: names, addresses and pages. The log name is a term that describes the data in the associated log. Each log name has an associated numeric value that is the log address. The log addresses are used by read and write log commands to access a specific log. Each log is composed of one or more pages.

Log pages shall be 512 bytes of data. Table 5 illustrates the conceptual layout of the CFA required logs. GPL feature set commands allow the host to specify the starting log page number as well as the number of log pages to be returned. Additional ATA features not supported by CFA are allowed, see the ATA standard.

Table 4 – CFA log address definitions

Log Address	Log Name	Feature Set	R/W
00h	Log directory	N/A	RO
26h or TBD	Performance Control Log	Performance Control	RO
80h-9Fh	Host specific logs		R/W or vendor specific
Key – RO – Log is read only R/W – Log is read or written			

Table 5 - CFA Conceptual Log Structure

Log Name	Log Address	Log Pages
Log Directory	00h	Page #0, required (the Log Directory)
Performance Control Log	26h or TBD	Page #0, required (first 512 byte log page)
		Page #1, optional (second 512 byte page)
		…
		Page #255, optional (last 512 byte page)

2.2 General Purpose Log Directory (GPL Log Address 00h)

Table 6 defines the 512 bytes that make up the General Purpose Log Directory.

Table 6 - CFA General Purpose Log Directory Structure (Log #0)

Word	Description
0	GPL Version = 0001h.
1	0000h or see ATA standard
2	0000h or see ATA standard
…	
26h or TBD	Number of Performance Control Log pages (0001h to ffffh). 0000h is not permitted.
…	
FFh	0000h or see ATA standard

The value of the General Purpose Logging Version word shall be 0001h. A value of 0000h indicates that no General Purpose Log Directory exists.

2.3 Performance Control Log (Log Address 26h or TBD)

The mandatory Performance Control Log shall contain at least one page, page 0. It may contain up to 65535 pages.

The Performance Control Log, starting at page 1, shall contain Performance Control records. The records are consecutive. Each record shall describe the performance of an LBA range for read type or write type operations. The LBA ranges shall not overlap between different records of the same type and profile. There is no guarantee that all LBAs will be included in the table.

Page 1 to the last page shall be contiguous. Each page, other than the last page, shall contain up to 8 Performance Control Description records starting at byte 0 of the page and contiguous with each other. See Table 8 for the Performance Control Description record structure.

The last page may contain at least 1 Performance Control Description records.

Page 0 shall contain an address map of the Performance Control Description log pages. See Table 7.

Table 7 - Structure of the Performance Control Description log directory (page #0)

Word	Description
0	Performance Control Description log version = 0001h.
1	Size of a Performance Control description record in words.
2,3 (DWord)	DWord - Number of Performance Control description records.
4 to FFh	Reserved

2.3.1.1 Performance Control log version

The value of the Performance Control log version shall be set to 0001h.

2.3.1.2 Size of a Performance Control description record

The size of each Performance Control description record shall contain a value of 20h (32 words – the size of the record defined in Table 8) or one of the following values: 20h (32d), 40h (64d), 80h (128d), 100h (256d). Any field in the record that is not defined in Table 8 shall be vendor specific.

2.3.1.3 Number of Performance Control description records

The number of Performance Control records shall be between 0 and 65535. A value of 0 in this field means that no Performance Control records exist, and the performance guarantee feature is not supported.

The number of performance control records per log page can be found by dividing 256 by the contents of Word 1 of Page 0. The legal values are 1, 2, 4, and 8.

The log page number from where a specific Performance Control record can be read can be calculated by subtracting 1 from the number of Performance Control description records, then dividing the result by the number of performance control records per page calculated earlier. The log page number being sought is the integer portion of the result of that division operation plus one.

Table 8 - Performance Control Description Record Structure

Byte	Size	Name	Description
0 to 3	DWord	Type	0 – Write performance record 1 – Read performance record
4 to 7	Dword	S_{Tm}	The maximum number of concurrent streams. 0 if the Performance Control functionality is not supported. ffffffffh if there is no limit to the number of streams that can be simultaneously allocated.
8 to 11	Dword	S_{Ta}	Number of available unallocated streams. This value must be 0 or a positive number not greater than S_{Tm}.
12 to 15	Dword	RU	Size of a Recording Unit (RU) in 512 bytes blocks. For example, if the RU is 128KB this number will be 256.
16 to 19	Dword	AU	Number of RUs in an Allocation Unit (AU). Continuing the previous example, if the size of the RU is 128KB, and the size of the AU is 8MB, then this number will be 64 (64*256*512 bytes = 8MB, or 8,388,608 Bytes)
20 to 27	QWord	OFS	The LBA of the first sector in the first RU in the first AU this record is associated with.
28 to 31	DWord	PR	The profile code for a specific performance level and specific file system usage. Profile codes are assigned in a separate document. See 0.
32 to 35	Dword	T_F	Maximum time for reading the file system and / or writing standard file system updates as defined for PR (the Profile code) required for processing one AU, in microseconds. This includes all the updates, for example: FAT, directory, and other tables. If the value of T_F, for example, is 120 milliseconds, then this field will contain the value 120,000. If PR is 0 then this field may contain 0. A value of 0 means T_F is undefined.
36 to 39	Dword	Reserved	Set to 0
40 to 43	Dword	T_{AU}	Maximum time to read or write one AU, one RU at a time, using the method defined by the profile as specified by the PR field, in microseconds. For example, if writing an 8MB AU lasts 280 milliseconds at most, this field will contain a value of 280,000. If PR is 0 then this field may contain 0. A value of 0 means T_{AU} is undefined.
44 to 47	Dword	N_{AU}	Number of consecutive AUs in this range. The range size is N_{AU}*AU*RU*(LBA size).
48 to 51	Dword	R_{MAX}	Maximum number of ranges that may be defined in one PEFORMANCE MANAGEMENT DMA EXT command. 00000000h means no support for the PERFORMANCE MANAGEMENT DMA EXT command. ffffffffh means an unlimited number of ranges may be defined in one command.
52 to 59		Reserved	Set to 0
60 to 63	DWord	-LAST	Set to 0 if this is the last record, 1 if there is at least one more record.

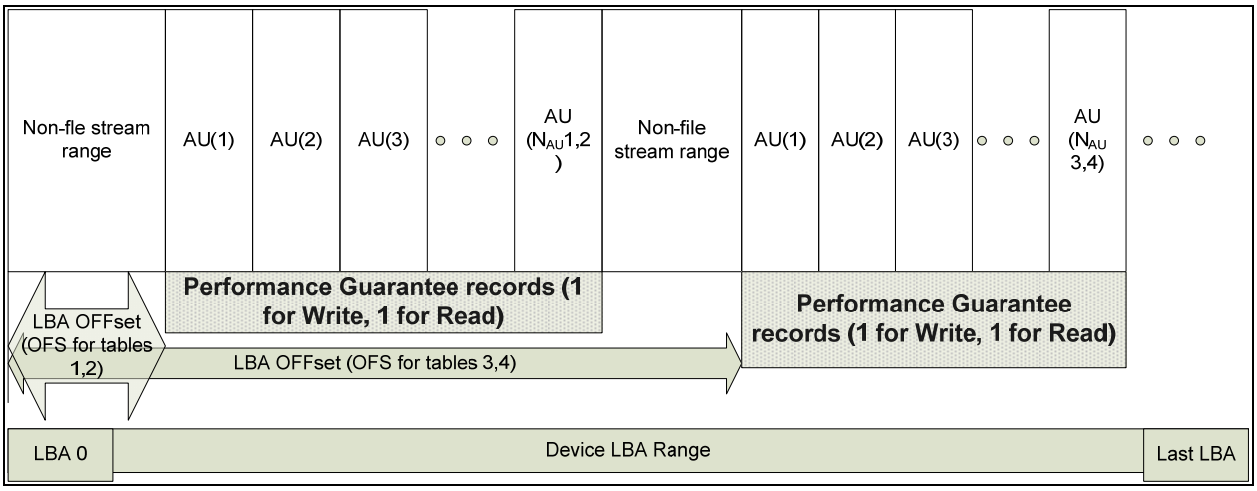

Figure 5: Illustration of a device division into Performance Control records and associated AUs

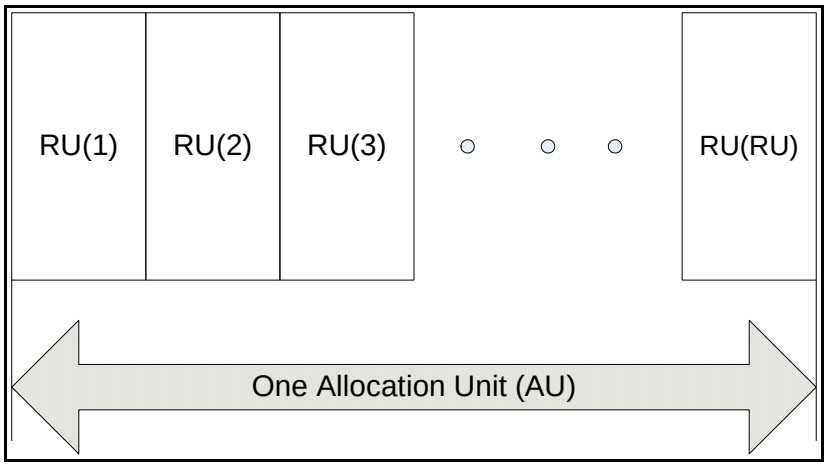

Figure 6: Division of an AU into RUs

3 Appendix C

(normative)

3.1 Profiles

Each Performance Control record defined in Table 8 shall contain in the PR field a profile number. Any profile number other than 0 shall be defined in a CFA Profile Specification.

A profile has to define at least the following:

- The target direction and speed of the card (e.g. Write to the card at a sustained rate of 20MB/sec (20,971,520 Bytes per second).
- Interface type to be used for testing the cards for compliance with the profile, e.g. PIO-4, UDMA-5, …
- File system used. Examples:
    - FAT32,
    - exFAT,
    - A combination: Cards up to 32GB to be formatted using the FAT32 file system, higher capacities to be formatted using the exFAT file system.
    - Ext4
- The specific command sequence used to read and update the file system and the frequency at which this may happen. Example:
    - Once per AU, or once per 1000 milliseconds, whichever occurs first.
- Maximum allowed T_F. This number will specify the maximum overhead time allowed to write one AU, including file system operations and internal card operations. The minimum host buffer size can be calculated based on this number.
- Any other assumption, directions, recommendations, etc. that will allow host and card manufacturers to design their systems to meet the performance requirement.

3.2 Performance Control

Performance Control is a feature set that coordinates writing to and reading from the card between the host and the card to increase card performance and endurance.

Optional functionality provided by the Performance Control feature is for the host to write to the card and read from the card in alignment with various physical pages defined in the card memory. If the card implements at least one performance control record, and its RU and AU fields contain a number other than zero, then the host can try and improve card performance and endurance.

3.2.1 Using the Card

Following any of the rules below shall not degrade card performance or endurance. Following any of these rules should increase card performance and / or endurance.

1. Data is written in multiples of one RU at a time, and the first sector of each write command is the first sector of an RU.

2. Data is written sequentially into one AU without writing any other data (including file system information and files in other locations) while the AU is not fully written or the file is closed.

3. Data is written into one AU from the beginning of the AU to the end of the AU sequentially. The host does not expect any previous data in the same AU to be later available.

Figure 7 – File stream usage example – flow diagram