Formatting, Cloning and Duplicating Advanced Format Media

Technical Paper
February 9, 2010
Revision 0.9

IDEMA

Technical Editor:

Curtis E. Stevens
Phone: 949-672-7933
E-Mail: Curtis.Stevens@WDC.com

Table of Contents

List of Tables

List of Figures

1 Introduction

The disk drive industry is continuing to increase the capacity of its media. There are 3 basic methods for doing this:

1. increase the bit density;
2. increase the number of tracks per inch; and
3. increase the number of surfaces.

Increasing the bit density yields the best result. The benefit of increased bit density is that it is seen in every track and on every surface. Each track is composed of a series of sectors and each sector has gap, lead-in and error correction information. Advanced Format media increases the sector size. The increase in sector size reduces the number of gap and lead-in entries on each track thus providing more space for user data. Advanced Format media also provides an increase in data integrity by providing a more efficient Error correction scheme.

Many disk drive interface technologies already allow for increased sector sizes. However, the 512-byte sector has been the standard for over 30 years. As a result, many points in a computer system (e.g., laptop, server, DVR, PSP, Cell Phone) have become inflexible and only work using 512-byte sectors. Therefore, devices which employ Advanced Format media need to maintain a 512-byte sector at the drive interface. This creates an efficiency issue when writing on Advanced Format media. If the data being written does not start at the beginning of an Advanced Format sector, or end at the end of an Advanced Format sector, then the drive will take longer to write the data.

This paper provides both detection and usage requirements for Advanced Format drives. If the procedures in paper are properly implemented, then an Advanced Format media device will provide optimal performance.

2 Scope

This technical paper provides requirements for both detecting Advanced Format media and implementing file systems on Advanced Format media. These methods are described for SAS, USB Bulk-only Transport, USB Attached SCSI, and SATA.

2.1 Expected Specification Life

This technical paper provides information which is expected to be useful for approximately 10 years.

2.2 Revisioning

The revision numbers used in this document indicate the level of maturity and review level of the specification. The revisions are .6, .7, .8, .9, or 1.0, followed by a letter. The revision numbers can be interpreted as follows:

1. Rev 0.6 - Proposal, not reviewed by anybody except the author. A document advances from the .6 level to the .7 level when 1 or more people have read the document and provided input for a new revision.
2. Rev 0.7 - The specification in review by a closed group of people, usually known as a working group. The specification advances from .7 to .8 when the working group believes that the document is complete and is ready to be reviewed by a greater audience. Greater audiences include Customers, Suppliers, and other groups within WD.
3. Rev 0.8 - The specification is in review corporately and possibly by customers under NDA. At this level, the specification is content complete and some or all of the sections are usually in freeze. Sections that are not part of the critical path for implementation may still be under development. The specification advances to .9 or 1.0 (whichever is appropriate) when there are no changes required by the people reviewing the document.
4. Rev 0.9 - The specification is in review publicly. At this level, the document is content complete, and attempts are being made to freeze the document. Internal specifications do not usually take on a .9 revision. There is usually a set time-frame for public review, 30, 60, and 90 days are common. After the review period is over, comments are addressed and the document can advance to 1.0, or stay at .9 for another review cycle if there were substantial changes required.
5. Rev 1.0 - Release. A revision goes to 1.0 usually when implementation is complete. At this point it is common to remove the revision history and to publish the specification.

6. The revision number has a letter as a suffix. This letter advances as an indication of how many drafts have been reviewed by the target audience. If more than 25 drafts are created at any revision, the suffix goes from a-z and then aa, ab, etc.

2.3 References

Table 3 - Document References

Name	Location
ATA8-ACS	www.t13.org
SBC-3	www.t10.org

3 Overview

This disk drive industry is switching from a 512-byte physical sector to larger physical sector sizes. Drives with a physical sector size larger than 1,024 bytes are referred to as Advanced Format drive. Advanced Format drives have Advanced Format media. Advanced Format media has Advanced Format sectors. In 2009 disk drive manufacturers started deploying Advanced Format sectors with 4,096 bytes of user data. It is likely that the Advanced Format sector size will be increased again within the next 5 years, as a result disk drives should be checked for Advanced Format requirements from now until eternity.

The ATA8-ACS and SBC-3 standards have provisions for a disk drive to report Advanced Format sector sizes and other performance optimization information. These standards are used for SATA, SAS, USB, and IEEE 1394 based interface technologies. The physical, link, and transport layers of these interface technologies have already deployed Advanced Format requirements, there is no modification required to these system layers for optimal performance.

All command layer functions remain the same with Advanced Format drives. However, Advanced Format drives have better performance if the starting sector address and the length of each write command meet the requirements provided by the device (referred to as alignment requirements). These alignment requirements create a need to change to the layout of file systems for optimal performance on an Advanced Format drive.

3.1 Operating System Support

Windows Vista SP1 and all versions of Windows Seven create partitions and file systems optimized for performance on Serial ATA devices.

Windows XP, Windows Vista RTM, Linux, Apple Tiger, Apple Leopard, and Apple Snow Leopard do not partition and format Advanced Format SATA devices using the information provided by the device.

Windows XP, Windows Vista, Windows Seven, Linux, Apple Tiger, Apple Leopard, and Apple Snow Leopard do not partition and format Advanced Format USB devices using the information provided by the device.

If the Advanced Format device is used with an operating system that does not partition and format using the information provided by the device, then external optimization may be necessary. Duplicators, cloners, and reformatters following the recommendations in this white paper create a fully optimized file system on an Advanced Format device.

Due to fundamental changes in the reporting structure of the drive, hard drive suppliers may elect to not report device physical sector size and alignment in IDENITFY DEVICE words 106 or 209 for purposes of backward compatibility with legacy device drivers. After initial debug, all hard drive suppliers shall report device physical sector size and alignment in IDENITFY DEVICE words 106 or 209.

4 Advanced Format Detection

4.1 SATA

Table 4 shows the location of Advanced Formatting requirements in the ATA IDENTIFY DEVICE data (see ATA8-ACS).

Table 4 - SATA Advanced Format Requirements

Location	Description
Word 106 bits 15:14	If bit 15=0 and bit 14=1, then Advanced Format requirements are present
Word 106 bit 13	If bit 13=1, then the device has alignment requirements
word 106 bits 3:0	If bit 13=1, then word 106 bits 3:0 specify the number of logical sectors per physical sector (LSPS). The number is specified as a 2^x value where x is the value from bits 3:0. For example, an Advanced Format drive with a physical sector size of 4,096 bytes would place 0011b in bits 3:0. This would indicate that the physical sector size was 2^3 sectors or eight 512-byte logical sectors per physical sector.
Word 209 bits 15:14	If bit 15=0 and bit 14=1, then Advanced Format requirements are present and the Advanced Format drive is reporting specific alignment requirements. If bit 15=0, then the Advanced Format drive requires alignment 0.
Word 209 bits 13:0	Indicates the alignment required by the device for optimal performance. This value represents the number that the drive will add to each logical sector address (LSA) specified in read and write commands.

If an Advanced Format drive has 4,096 byte physical sectors it reports the following in IDENTIFY DEVICE data:

1. word 106: 6003h
2. word 209: 4001h (for adding one to each LBA) or 4000h (device adds nothing to each LBA)

Figure 1 shows the formatting of an Advanced Format drive when the values list above are indicated by the device.

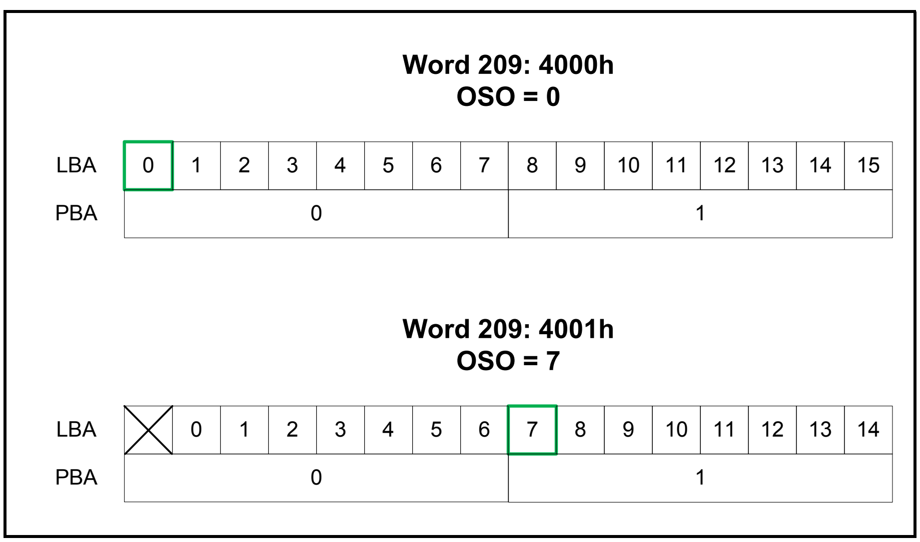

Figure 1 - SATA Alignment

The exponent reported in word 106 bits 3:0 indicates the number of bits in the logical block address that are checked to determine proper alignment. In Advanced Format sectors with 4,096 bytes, address bits 2:0 are all that need be tested for proper alignment (Logical Sector Address (LSA) is logically anded with 7h). For sectors with 4,096 bytes, the Operating System Offset (OSO), used later in this paper, is calculated using the following formula: OSO = (8 -

Alignment) logically anded with 07h. For example, when IDENTIFY DEVICE data word 209 indicates 4000h, the OSO=0, when word 209 indicates 4001h, the OSO=7. The OSO is the first properly aligned sector on the Advanced Format drive.

The alignment of any logical sector in a SATA Advanced Format drive reporting 2^3 LSPS is determined using the following formula:

Alignment = If (LSA and 7h) = 0 then 0 otherwise 8 - (LSA and 7h)

LSA = Logical Sector Address.

All major file system structures need to start with Alignment = OSO for optimal performance in an Advanced Format drive.

4.2 SAS, USB Bulk Only Transport, USB Attached SCSI, and IEEE 1394

Fields have been added to the READ CAPACITY (16) command in SBC-3 to report Advance Format device requirements. These fields are: Logical Blocks Per Physical Block Exponent (LSPS) and Lowest Aligned Logical Block Address (OSO). The physical sector size of an Advanced Format drive is specified as a 2^x value where x is the LSPS. For example, an Advanced Format drive with a physical sector size of 4,096 bytes would set the LSPS to 3. This would indicate that the physical sector size was 2^3 sectors or eight 512-byte logical sectors per physical sector.

The OSO is the first aligned logical sector address (LSA) on an Advanced Format drive. Figure 2 shows the formatting of an Advanced Format drive when the values list above are indicated by the device for both OSO=0 and OSO=1.

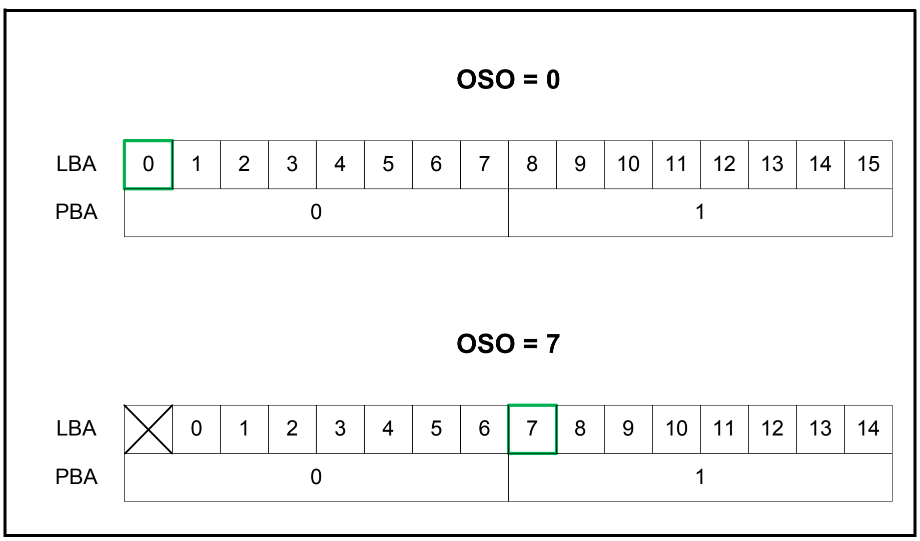

Figure 2 - SCSI Alignment

The alignment of any logical sector in a SCSI Advanced Format drive reporting 2^3 logical sectors in a physical sector is determined using the following formula:

Alignment = LSA and 7h

LSA = Logical Sector Address.

All major file system structures need to start with Alignment = OSO for optimal performance in an Advanced Format drive.

5 Advanced Format Partitioning

Advanced Format drives with OSO=0 (see 4.1 and 4.2 for how to calculate the OSO) are designed to provide best performance with operating systems that place the first partition at logical sector address 2048 (e.g., Windows Seven).

Advanced Format drives with OSO=7 are designed to provide best performance with operating systems that place the first partition at logical sector address 63 (e.g. Windows XP).

It is important to first determine the OSO of the drive before creating partitions. If the Advanced Format device OSO=7, then place the first partition at logical sector 63 and additional partitions on OSO boundaries. If the Advanced Format device OSO=0, then place the first partition at logical sector 2048 and additional partitions on OSO boundaries.

Partitions should be sized to meet the physical sector size requirements of the Advanced Format device. If the Advanced Format sector size is 4,096 bytes, then size the partition so the low order 3 bits of the size are zero (i.e., make the partition size a multiple of 4,096 bytes).

Windows XP, Windows Vista, and Windows Seven have maximum performance when the recommendations in this clause are followed. These operating systems format optimal performing file systems when the partitions are created following these recommendations. See clause 6 for formatting an NTFS file system on an Advanced Format drive that:

1. does not follow this recommendations in this clause;
2. requires a different file system (e.g., FAT-16, FAT-32, exFAT, HFS, ext4, etc.).

6 Advanced Format File System Formatting

Advanced Format drives require that all data structures start at the OSO of the device (see 4.1 and 4.2 for how to calculate the OSO). This means that for optimal performance in a file system, the following need to start at the OSO and have a length which is a multiple of the physical block size:

1. Partition Logical Start Address (LSA)
2. Metadata LSA (e.g., FAT, $MFT, $BITMAP, iNodes, etc.)
3. Clusters

In many file systems, there is no metadata between clusters. In these file systems, if the first cluster starts on the OSO, then all the clusters begin on the OSO. Figure 3 shows an example of a generic file system using a legacy start address for the first partition of logical sector 63.

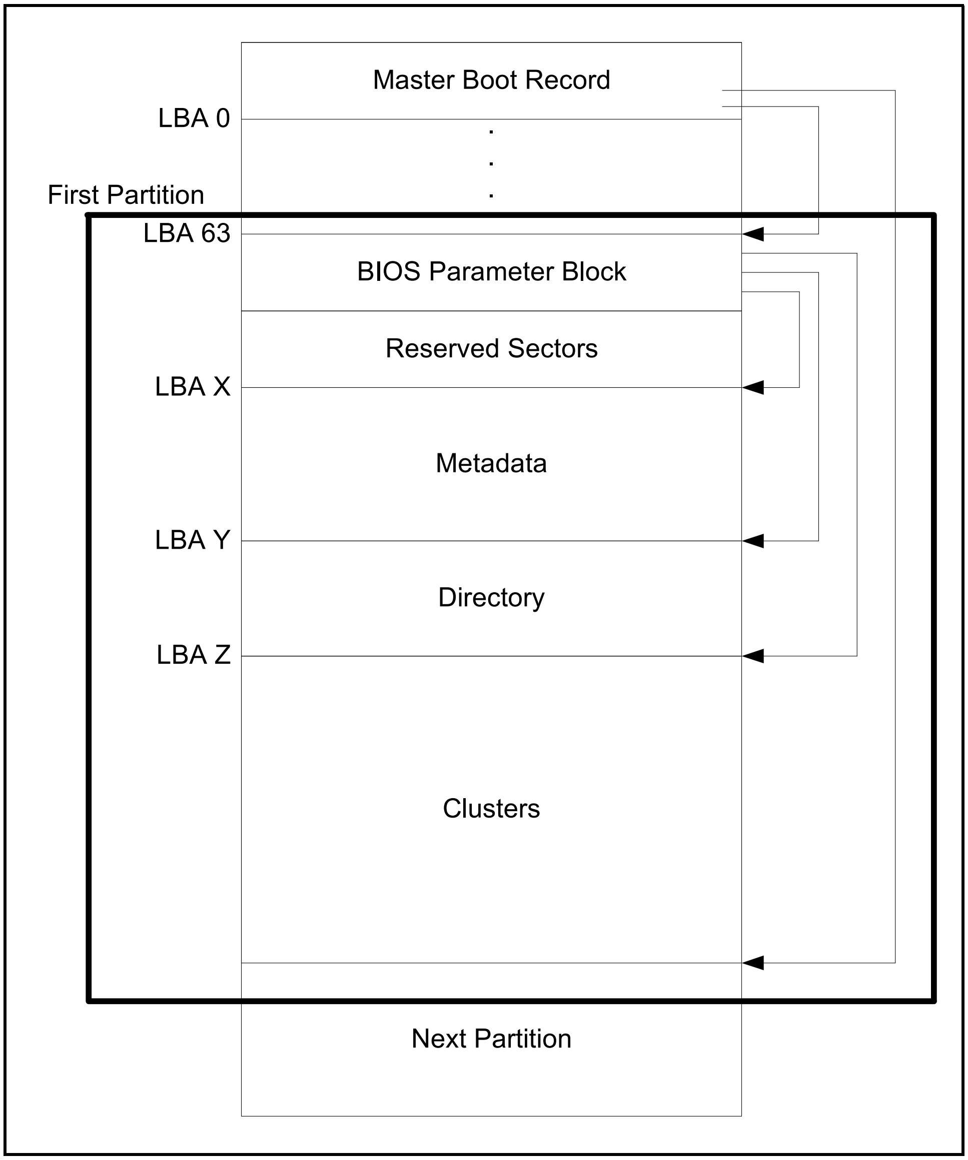

Figure 3 - Example file system

In this example, if the Advanced Format drive reports an OSO of 7 and an LSPS of 3 (i.e., 4,096 byte physical sectors), then the following needs to be true for a fully optimized user experience:

1. LBA X and 7h = OSO;
2. LBA Y and 7h = OSO; and
3. LBA Z and 7h = OSO.

If figure 3 were an NTFS file system and OSO=7, then the file system would be properly formatted by Windows XP. If OSO=0, then the file system needs to be changed so that LBA X, LBA Y, and LBA Z meet the OSO requirements listed above.

If figure 3 were an NTFS file system formatted by Windows Seven, then the first partition start LBA would be LBA 2048 instead of LBA 63. In this case, if OSO=0, then the file system would be properly formatted by Windows Seven. If OSO=7, then the file system needs to be changed so that LBA X, LBA Y, and LBA Z meet the OSO requirements listed above.

During the formatting process, if the file system does not meet the OSO requirements of the Advanced Format drive, then one way to create an aligned file system is to insert 7 reserved sectors after the BPB. This causes an NTFS file system to change from OSO=0 to OSO=7.

If there are multiple partitions on an Advanced Format drive, then all the partitions have to be modified to meet the OSO requirement of the Advanced Format drive.

Format operations on an Advanced Format drive should not create data structures smaller than the physical sector size of the Advanced Format drive. For example, if the Advanced Format drive reports a physical sector size of 4,096 bytes, then the cluster size of the file system should be a multiple of 4,096 bytes and the metadata should all be addressed in 4,096 byte units.

