WD AV-GP Large Capacity Hard Drives

by Kevin Calvert
Senior Engineering Program Manager

Introduction

This document provides important information to OEMs, integrators, and installers who want to deploy Audio/Video (AV)-class hard drives larger than 2.19 TB into high-capacity video recorders or other AV-based storage devices. The goal of this document is to provide an analysis of existing solutions and outline various integration options for large capacity drives (hard drives that exceed 2.19 TB).

Older systems using legacy operating systems and BIOSs that utilize the Master Boot Record (MBR) partition table scheme encounter a barrier at 2.19 TB (2,199,023,255,552 bytes). This is due to the ability to address only 2^{32} logical blocks with the most commonly used sector size of 512 bytes. Some operating systems, such as Windows® XP, only support booting from a Master Boot Record (MBR) partition formatted drive. Therefore, there are limitations when attempting to move to a higher capacity drive. The system's BIOS and operating system drivers need to agree on the capacity and geometry of a hard drive to boot and operate correctly. Agreement must take place across several software layers to successfully boot a system.

By working collectively with industry partners, system providers, and operating system vendors to ensure drive compatibility across multiple software layers, WD provides solutions for moving beyond the 2.19 TB capacity barrier.

Limitations of 512 byte Sector Size and Advanced Format (AF) Drives

To increase the amount of storage on a hard drive, manufacturers have been working with industry partners to implement larger native sector sizes. WD has transitioned to using sector sizes of 4096 bytes or a 4 KB physical sector size (Advanced Format or AF technology) on all large capacity hard drives. Although the drive may be using a sector size of 4096 bytes to store data on the media, it reports and emulates a disk using 512 bytes (512e) to avoid application incompatibilities.

Older Linux kernel releases along with earlier Windows releases were hard coded for 512 byte sector lengths along with the disk and other supporting utilities for SATA drives. Proper partition alignment is necessary to ensure the best possible performance.

Embedded Operating Systems Products

An embedded system is a computer system designed to do one or a few dedicated and/or specific functions, often with real-time computing constraints. It's embedded as part of a complete device often including hardware and mechanical parts. Embedded systems use operating systems customized to a specific hardware configuration that may not contain updates needed to handle a drive that has more than 2^{32} addressable blocks. Therefore, it's important that integrators check with their OS vendors to ensure all of the updates have been included in the version they are using.

Any software developed by the integrator must be coded to address hard drives greater than 2^{32} addressable blocks. A thorough review should be undertaken to evaluate if any code has been developed, and if it uses existing OS functions to address the hard drive or disk address directly.

AV Device Considerations

AV products are diverse in their applications and the use of storage products. Most products that use storage have some type of operating systems and system start up software that need to take into account the size (in addressable blocks) and the impact that size may have on the different layers of software contained in their products. This is coupled with ensuring that partitions are created on 4096 (4K) boundaries to yield maximum performance. Other considerations may include thermal characteristics and power requirements.

Startup Software

System startup software and other operating system layers and utilities need to agree on size and how they address the logical blocks to avoid system incompatibility issues. The following illustration depicts a typical system startup.

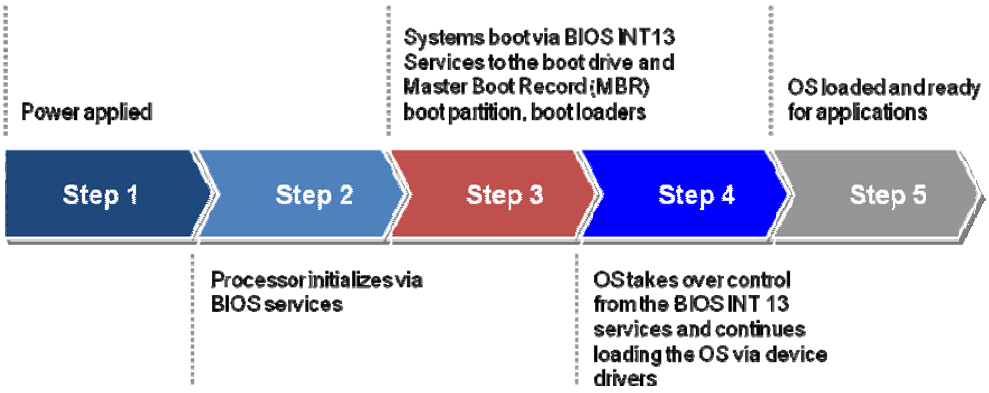

Figure 1: Typical System Startup

From step to step, agreement on drive recognition is critical to system startup and performance. It is therefore critical that all layers of software can handle storage capacities greater than 2^{32} blocks to avoid truncating partitions, or software components writing data on blocks of the disk that are not addressed.

BIOS Requirements

Determining the necessary BIOS considerations is an important part of the decision-making process. One can either use UEFI or develop a solution where blocks larger than 2^{32} can be addressed and newer partitions types are recognized.

UEFI Support

The UEFI specification defines a new model for the interface between operating systems and platform firmware. UEFI is a community effort by many companies in the computer industry to modernize the booting process. The solution uses Globally Unique Identifier (GUID) Partition Tables, otherwise known as GPT, instead of MBR. This partitioning method provides for up to 18 Exebyte (2^{64}) of Logical Block Addressing. The interface consists of data tables that contain platform-related information, plus boot and runtime service calls that are available to the operating system and its loader. Together, these provide a standard environment for booting an operating system and running pre-boot applications.

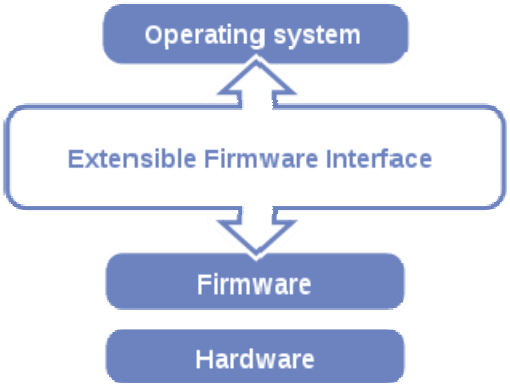

Figure 2: Unified Extensible Firmware Interface's (UEFI) Position in the Software Stack

Windows Support

Windows operating systems have different levels of support for large capacity drives. The table below summarizes this support.

	Windows xp 32-bit[2]	Windows xp 64-bit[2]	Windows Vista 32-bit	Windows Vista 64-bit	Windows 7 32-bit[1]	Windows 7 64-bit[1]
Boot	Not Available	Not Available	Not Available	Supported[3]	Not Available	Supported[3]
Secondary	Not Available	Not Verified	Supported	Supported	Supported	Supported

[1] To achieve full capacity and functionality with drives greater than 2.19 TB, WD requires the use of Microsoft Windows 7 Service Pack 1 (SP1) or later.

[2] Presently WD does not directly provide support for these applications. We have worked extensively with our partners and solutions may be available. HBA and RAID Controller vendors may have developed solutions for these applications.

[3] Boot support requires a system with UEFI Support and a 64-bit version of the OS.

How You Can Take Advantage of Large Capacity Drives with Windows

WD provides many options that you can use to take advantage of large capacity drives, even on current computer system configurations with today's technologies.

- Implementing a large capacity WD drive as the system startup drive requires a system with UEFI support and uses GPT partitions.
- Implementing a large capacity WD drive in legacy systems uses the large capacity drive as secondary storage and uses GPT partitions.
- Review the considerations for the Windows-based operating systems above to help determine what options are available for implementing large capacity drives on your current systems, and to help you plan future operating system and computer system purchasing plans.

Linux Support

Current commercial releases of major Linux Distributions support large capacity drives. However, distributions released before June of 2010 will not have full support built in with default installation tools. Please take caution and review the information details for the distribution that you will be using for correct operation and minimum requirements to be used with large capacity drives and drives using AF capabilities.

Kernels

Older Linux Kernel version (all 2.4 and pre-2.6.32) computing environments with a legacy BIOS and MBR partition table scheme encounter a barrier at 2.19 TB because they can address only up to 2^{32} logical blocks. Be sure to use kernels that contain support for drives greater than 2.19TB. The kernels released after April 2010 have support for large capacity drives using 4096 sector sizes.

Kernel 2.6.31 adds parameters which allow for non-native 512 byte sector drives - Advanced Format, 4K/512, etc. The parameters are presented in /sys/block/$DEVICE and are available for tools and applications. Newer partitioning tools use these parameters to optimize partition data but the older tools typical of most current Linux distributions ignore them.

- /sys/block/$DEVICE/alignment_offset is the offset in bytes of the lowest aligned logical block. Normally zero, if jumper 7-8 is used on a WD AF drive, this will report 3584 designating that logical sector 7 is the first aligned LBA to a 4K boundary.
- /sys/block/$DEVICE/queue/logical_block_size - block size used to address a location on the device - 512
- /sys/block/$DEVICE/queue/physical_block_size - smallest unit the device can operate on - 4096
- /sys/block/$DEVICE/queue/minimum_io_size - device's preferred minimum unit for random I/O - 4096
- /sys/block/$DEVICE/queue/optimal_io_size - device's preferred unit for streaming I/O - 0

There is a good overview of these parameters and more at *http://mkp.net/pubs/linux-advanced-storage.pdf*.

OS Utilities

Booting Linux requires boot loader support, kernel compiled support for GPT, and utilities support that support large capacity drive disk utilities like "format" or "parted." Boot loaders like "grub" released after June 2010 have support for sector sizes of 4096 bytes as well as large capacity drives.

fdisk

Avoid using fdisk. It is an old design and even the main page recommends using cfdisk or parted. There are some notes on using older fdisk versions to align drives (including SSD) described on various websites. The method is to force CHS parameters which results in boundaries that match 4K - 56 sectors per track, for example, instead of 63.

util-linux-ng version 2.17 *http://karelzak.blogspot.com/2010/01/util-linux-ng-217.html* released on January 8, 2010 includes fdisk 2.17 which allows alignment of AF drives according to the kernel parameters listed above. The default use is "DOS compatible mode" which sets the first partition to start on sector 63. Press 'c' to toggle the DOS flag - turn it off and the first partition will start on sector 8 as a default.

parted

parted in its common form (typically 1.8.8) allows partitions to align at the sector level so we can set up partitions on 4K boundaries (i.e., start partition 1 at sector 64 rather than the typical sector 63). This is a very manual process as each partition boundary has to be aligned to 4K using this tool (parted) prior to OS installation or creating new partitions. Devising an algorithm to run on a spreadsheet or doing a tech note with some sample valid parameters for various drives is possible, but not preferred.

New parted 2.1 includes alignment features as long as kernel >= 2.6.31 is used and libblkid >= 2.17 which is part of util-linux-ng v.2.17. This is plausible for the hardcore Linux users who are customizing systems for embedded, etc., but not for the typical desktop user of Linux (Ubuntu, Debian, Fedora etc.). Parted is now at 2.2. See page on parted 2.x for details of configuration.

Linux Parted for Advanced Format Drives

Both parted 2.x and libblkid 2.17 are very new and not yet in standard distributions so we will have a period of pain before they are integrated. Also, the graphical version (gparted) does not seem to be near to supporting alignment, so typical OS installation scripts such as Ubuntu's can't be used at the moment. The Ubuntu team intends to include parted 2.2 in the upcoming 10.04 release.

gparted

gparted is the GUI partitioning tool for Linux, currently at version 0.5.2 [13].gparted began working for an aligned solution between versions 0.4.5 and 0.5.1.

Deselect the "Round to Cylinders" check box when creating partitions. Version 0.5.1+ aligns to MB boundaries. Earlier versions only align the start of the first partition. Allow at least 1 MB at start of the first partition for boot code and at least 1 MB at start of each logical partition in extended partitions.

If "Round to Cylinders" is selected, creating the file system may fail. The failure is due to the mkfs.XXXX stage which now prompts the user with a misalignment warning. Don't rely on this as a check! File system creation will succeed in some case such as NTFS, swap etc. Also, if underlying tools are old and do not report AF, gparted can create misaligned partitions. For example, gparted live CD/USB tools [14] have gparted 0.5.2 but util-lnux-ng 2.16 which does not report AF parameters up to partitioners.

disk utility

Used in recent Ubuntu distributions. Default in Ubuntu 10.04 - gparted not installed automatically any more. Version 2.29.90 will warn on misaligned partitions but does not offer a remedy. Partitions can be created misaligned as it uses new "-F" option on mkfs.XXXX to suppress misalignment warning.

e2fsprogs

Release 1.41.10 from February 2010 includes mke2fs which has alignment check. This is called from mkfs.XXXX and adds a user question to continue (y/n) even though partition alignment is wrong - "very poor performance" is suggested if the user continues.

How You Can Take Advantage of Large Capacity Drives in Linux

- Use fdisk from util-linux-ng >= 2.17.2 or parted/gparted
- Use +size {M, G} convention to specify "Last sector" (e.g. +5G to create 5GiB partition) then fdiskl aligns the size to physical block boundary
- Remember that fdisk(8) always follows your wishes -- it means that if you explicitly define first/last sector number then the partition could be misaligned.
- Start the extended partition at sector 64 (the default is 63), and end it at sector (*total amount of sectors on the drive – 1*)

For service and literature:

http://support.wdc.com
www.westerndigital.com

800.ASK.4WDC	North America
800.832.4778	Latin America (Spanish)
+800.6008.6008	Asia Pacific
+800.ASK.4WDEU	Europe (toll free where available)
+31.88.0062100	Europe, Middle East, Africa

Western Digital, WD, and the WD logo are registered trademarks in the U.S. and other countries. Other marks may be mentioned herein that belong to other companies. Product specifications subject to change without notice.

© 2011 Western Digital Technologies, Inc.
All rights reserved.

Western Digital
3355 Michelson Drive, Suite 100
Irvine, California 92612
U.S.A.

2579-771660-A00 May 2011