



Understanding hard drive media defects white paper - Servers

Product Documentation

Abstract

This IBM Server White Paper provides a brief description of Hard Disk Drive (HDD) design and media manufacturing. The objective is to sensitize readers to the fact that today's technology cannot create perfect media, but that through correct, thorough testing and defect management, the problem is often transparent to end users.

Content

Summary

This IBM Server White Paper provides a brief description of Hard Disk Drive (HDD) design and media manufacturing. The objective is to sensitize readers to the fact that today's technology cannot create perfect media, but that through correct, thorough testing and defect management, the problem is often transparent to end users.

IBM Server White Paper

Date: December 1997

Note: This document is intended for the general audience.

One of the most critical components of IBM Servers is the Storage Subsystem which includes the RAID controller and the Hard Disk Drives (HDD). Typical HDD capacities found in today's IBM Server Systems range from 2.25GB to 9.1GB (1Giga Byte = 1 Billion bytes). RAID 5 and RAID 1 architecture technology provides the ability to continue operation in the event of HDD failures and provides the ability to rebuild lost data. Although HDD reliability has greatly improved in the past five years, the areal densities (defined below) have grown at a staggering 60% compound growth rate during that same time.

Due to the complexity of HDD's and the nature of the technology, media defects are a fact of life in ALL HDD's manufactured today. However, HDD's employ effective error correction techniques and data threshold analysis and reassignments to help prevent data loss. This paper will explain the reason why media defects occur initially and why they may occur throughout the life of the drive.

- Brief description of HDDs
- Media manufacturing
- HDD manufacturing defect mapping
- HDD defect management
- Summary

Brief description of hard disk drives (HDDs)

An HDD is a very complex electromechanical device which employs many technologies. It is comprised of the HDA (Head Disk Assembly) and the PCB (Printed Circuit Board). For the purpose of this paper, we will concentrate on the HDA.

The HDA consists of a spindle motor, disks (the media) , Read\Write heads, an actuator to move the head assembly (Head Stack) to the target data block, all contained within a sealed enclosure. The data is written onto rotating disks which are magnetically treated. The rotational speed for Server HDD's ranges from 5400RPM to 10000RPM with the majority of drives used today being at 7200RPM.

While the disk rotates, the Read\Write heads fly above the disk. The fly height is typically 1.8 to 2 micro inches (and getting lower as technology progresses). As a comparison, a human hair is 3000 micro inches (a micro inch is one millionth of an inch). The actuator moves the head stack assembly (up to 20 heads may be installed in a head stack assembly) onto the desired location (track) to write or read the data. The disk is segmented radially into tracks and each track is made up of sectors. The sector is the least addressable unit on the disk drive - this is where the data reside. A sector is 512 bytes long. So a 4.5GB drive will contain 8,789,062 sectors. In reality, disk drives have many more sectors than the stated capacity.

This is because perfect media (the disks) are not possible with today's technology. Thus, in the manufacturing process, some sectors are found to be unusable (deficient magnetic coating, pits, etc.), and are reassigned to spare sectors somewhere else on the drive. The drive will have the advertised capacity when leaving the factory as well as throughout the life of the drive.

The amount of data that can be stored on a disk is a measure of it's areal density - defined as the number of data bits stored on the disk per square inch. Thus a typical 4.5GB drives has an areal density of .8Gbits to 1Gb per square inch. In the next 5 years drives will reach areal densities of greater than 5Gb/sq. inch (IBM Almaden research has already demonstrated 5Gb/sq inch in a laboratory environment). In order to achieve these densities, improvements in head and media technologies have to occur. The advent of MR (Magneto Resistive) and GMR (Giant MR) heads as well as improvements in media have paced the advances in areal densities. In order to understand why media defects can occur a brief overview of media manufacturing is presented below.

Media manufacturing

An ideal storage disk will have no imperfections and store data for an indefinite amount of time. The latter is theoretically achievable with today's magnetic materials assuming the environment within the HDA does not

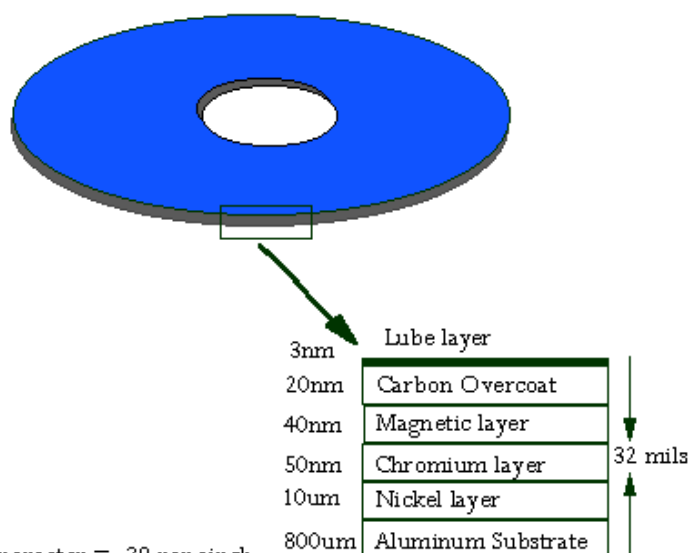
Your privacy choices  mination or malfunction of components within the disk enclosure. The former is not possible today but is often manageable.

The data is recorded onto the disk from signals emanating from the head transducer. A recorded region is called a bit cell. For an areal density of 1Gb/sq inch, the size of a bit cell is 1 billionth of a square inch. A 3.5" 1Gb/sq. inch disk can have in excess of 4 billion bit cells per surface. In order to achieve these staggering numbers, strict manufacturing processes and advances in magnetic materials are required.

A disk is defined as a thin film medium. That is multiple thin film layers of various materials are deposited onto the disk through a sputtering process. The general structure of a disk is shown below. The substrate on which the magnetic and other materials are deposited is always aluminum for 3.5" HDD's (glass substrates are used for 2.5" drives to improve shock characteristics).

Prior to material deposition which is performed in a clean room environment, the aluminum substrate is machined sized and ground to an acceptable finish. As can be seen in the diagram multiple layers of various materials are required to manufacture disks.

Thin Film Disk components



Note: nm= nanometer = 39 nanoinch
um= micrometer = 39 microinch

Each step of this process can introduce imperfections in the media. An imperfection the size of a bit cell will therefore prevent the cell from having the appropriate magnetic properties; thus a media defect is created. It is possible that the imperfection may encompass more than one bit cell and cause additional media defects.

Another source of media defects is through the HDD manufacturing process. Although extreme precautions are taken in the handling of disks, some microscopic scratches can occur as disks are mounted onto the motor spinners and the disk assemblies are merged onto the media. In certain cases contamination during the drive build process can also cause particles to be deposited onto the media and cause latent defects.

Although defects are inevitable, HDD test processes can map them out prior to leaving the factory. However, some latent defects may not be detectable initially and may translate into inaccessible sectors in the field.

HDD manufacturing defect mapping

Defect mapping is performed during the manufacture of the HDD. A Surface Analysis Test is performed whereas the disk is written repeatedly with stressed data patterns and subsequently read back. Any sectors that cannot be read back successfully are removed from the sector map. A list of bad sectors is kept within the drive (called the P-list). After the SAT, the drives are subjected to a final test which may uncover additional media defects. Those sectors are also mapped out and the P list adjusted. Additional testing is performed during System Manufacturing to detect any defects which may have not been mapped out by the HDD manufacturer (unlikely) or which may have "grown" due to latent media defects. (possible). Defect discovery at System Manufacturing is rare but is nevertheless performed to ensure that the drives are defect free when shipped to our customers. As drives are used in the field, grown defects can continue to occur due to many circumstances: latent imperfection on the disk, media damage due to mishandling of the drive and harsh environments. However, these defects can be reallocated by the drive to available spare sectors in the drive.

HDD defect management

HDD's employ sophisticated defect management techniques to prevent data loss and promote data integrity. Earlier it was stated that there are many more sectors available in the drive beyond the drive advertised capacity. Typically each track has an additional sector beyond the required number of sectors and a drive may have thousands of spare sectors available. Those sectors are used in the event that a data sector becomes defective. In the case of defective sectors, the data is recovered (if possible) and rewritten onto the spare sector. The new sector is now part of the drive sector map and no loss of capacity or data has occurred. Data errors can be classified as soft errors and hard errors. Soft errors are recoverable. That is if the data was not read properly initially, the Error Recovery Procedures (ERP) of the drive can recover the data. ERP algorithms are very sophisticated and involve hardware correction (ECC on the fly - Error Correction Code), multiple reread of the data, track offset reading and application of firmware ECC. During the ERP process, the drive will determine if the sector requires reassignment to a spare sector. If so, the spare sector is identified and the data moved to that sector. If the data is unrecoverable it is a Hard error and the sector is no longer accessible. In this case data recovery and rebuild is done through the RAID subsystem.

Summary

This paper provided a brief description of HDD design and media manufacturing. The objective is to sensitize readers to the fact that today's technology cannot create perfect media, but that through correct, thorough testing and defect management, the problem is often transparent to end users. Latent media errors may not be detected in seldom-used files or in not-yet-used sectors and will only be identified if data is written/read to/from the drive. Your privacy choices  accomplishes this task in the background while allowing concurrent user disk activity. Data Scrubbing is recommended by IBM and described in "Using IBM RAID Adapters to Avoid Data

Loss," another IBM Server White Paper referenced in the ADDITIONAL INFORMATION section under White Papers.

Additional Information

Web Sites

IBM maintains extensive and timely information on the world wide web. Visit the following sites for more information on IBM PC Servers and other IBM products. These sources contain product information, performance data, and technical literature.

IBM Home Page <http://www.ibm.com>

White papers

The following White Papers pertain to RAID and hardfile technologies. These provide procedures for ensuring the highest protection and availability of customer data and are viewable on-line in PDF format at :

<http://www.ibm.com/systems/support> (<http://www.ibm.com/systems/support>)

From this site select "Other intel processor based servers" and then select "Online publications".

1. Using IBM RAID Adapters to Avoid Data Loss.
2. High Availability Using the IBM ServeRAID Adapter.
3. High Availability of Your RAID Subsystem with
 - IBM SCSI-2 Fast/Wide PCI-Bus RAID Adapter
 - IBM F/W Streaming RAID Adapter/A

Notice

International Business Machines Corporation 1997. All rights reserved.

References in this publication to IBM products, programs or services do not imply that IBM intends to make these available in all countries in which IBM operates. Any reference to an IBM product, program, or service is not intended to state or imply that only IBM's product, program, or service may be used. Any functional equivalent program that does not infringe any of IBM's intellectual property rights may be used instead of the IBM product, program or service. Information in this paper was developed in conjunction with use of the equipment specified, and is limited in application to those specific hardware and software products and levels. IBM may have patents or pending patent applications covering subject matter in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to the IBM Your privacy choices  IBM Corporation, 500 Columbus Avenue, Thornwood, NY 10594 USA. The information contained in this document has not been submitted to any formal IBM test and is distributed AS IS

WITHOUT WARRANTIES OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. The information about non-IBM (VENDOR) products in this manual has been supplied by the vendor and IBM assumes no responsibility for its accuracy or completeness. The use of this information or the implementation of any of these techniques is a customer responsibility and depends on the customer's ability to evaluate and integrate them into the customer's operational environment. While each item may have been reviewed by IBM for accuracy in a specific situation, there is no guarantee that the same or similar results will be obtained elsewhere. Customers attempting to adapt these techniques to their own environments do so at their own risk. This publication could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time.

The following terms are trademarks or registered trademarks of the International Business Machines Corporation in the United States and/or other countries.

OS/2 NetFinity

Microsoft, Windows, Windows NT, and the Windows logo are registered trademarks of Microsoft Corporation. UNIX is a registered trademark in the United States and other countries licensed exclusively through X/Open Company Limited.

Other company, product, and service names may be trademarks or service marks of others.

Applicable countries and regions

- Worldwide

Document Information

More support for:

Older System x->Netfinity 7000

Document number:

796648

Modified date:

23 January 2019

Your privacy choices 