Recovering Unrecoverable Data

The Need for Drive-Independent Data Recovery

A ChannelScience White Paper

Commissioned by
ActionFront
Data Recovery Labs, Inc.

Written by
Charles H. Sobey
April 14, 2004

ChannelScience
Detecting the Future of Data Storage SM

7300 Cody Court
Plano TX 75024-3837 USA
connect@ChannelScience.com

(972) 814-3441 Voice
(972) 208-9095 FAX

RECOVERING UNRECOVERABLE DATA

The Need for Drive-Independent Data Recovery

Table of Contents

1. Executive Summary....3
2. Introduction to Hard Disk Drive (HDD) Technology....4
   2.1 Areal Density and Price Trends....4
   2.2 What Happens to Data in a Hard Disk Drive?....5
      2.2.1 Organizing the Data....5
      2.2.2 Locating the Data....7
      2.2.3 Detecting the Data....8
      2.2.4 Decoding the Data....9
      2.2.5 Drive Burn-in and Optimization: "Hyper-Tuning"....11
3. Data Recovery Market....12
   3.1 Perception vs. Reality....13
   3.2 A Call for Transparency....13
4. Data Recovery Technology....14
   4.1 Traditional Hardware Replacement Methods....15
      4.1.1 Replace the PCB....15
      4.1.2 Replace the Firmware....16
      4.1.3 Replace the Head Stack....18
      4.1.4 Move the Disks to Another Drive....18
   4.2 "Magic Machines" and "Proprietary Processes"....18
      4.2.1 Spin-Stand Testers....19
      4.2.2 Magnetic Force Microscopes (MFM)....20
      4.2.3 The Spin-Stand MFM?....20
      4.2.4 Exotic Recovery....21
5. The Frontiers of What's Possible: What Makes Data Unrecoverable?....21
   5.1 When Firmware Replacement Fails....21
   5.2 When Head Stack Replacement Fails....22
   5.3 When Disk Remounting Fails....22
   5.4 When the DATA Fails....23
6. Future Success Depends upon Developing Drive-Independent Data Recovery Capabilities....24
7. The FIRST Public Demonstration of Drive-Independent Data Recovery: ActionFront's SignalTrace^{TM} Technology....25
8. Conclusions....28
9. References....29
About the Author....30

RECOVERING UNRECOVERABLE DATA

The Need for Drive-Independent Data Recovery

1. Executive Summary

When a hard disk drive containing valuable data no longer responds, the user's last hope is to send the drive to a data recovery company that specializes in drive hardware failures. There is a general perception that data recovery companies have "magic machines" for retrieving data in almost any situation. The reality is less glamorous. The most sophisticated, commercially successful recovery techniques involve careful part-replacement, in a cleanroom environment, of the heads, the spindle motor and base casting, the electronics board, and/or the drive's firmware and parameter tables. Part-replacement has historically been successful for data recovery about 40 to 60% of the time. *Claimed* data recovery success rates are much higher. While they may, in fact, approach 100% *for some drive models*, for other models and failure modes the success rate is near zero. Drive-independent data recovery methods are needed now to read these drives. Furthermore, as the data density of hard disk drives continues to increase the number of unrecoverable drives is expected to grow.

The reason for this lack of successful recovery can be traced to the methods drive manufacturers must employ to achieve both high data density and high production yields. Specifically, current drives are "hyper-tuned" in the factory to optimize the performance of *each section of each hard disk drive*. The data format, head, disk, electronics, and firmware parameters are all optimized together. This means that it is less likely that a head stack or electronics board or parameter tables from one drive – even of the same model – will work well when used as a replacement in a failed drive.

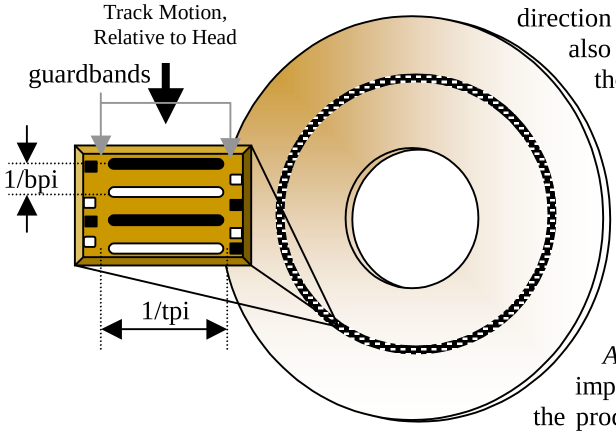
ActionFront Data Recovery Lab's SignalTrace™ technology is the only solution known to-date that demonstrates the capabilities needed for commercially viable recovery of user data that is otherwise unrecoverable using traditional part-replacement. SignalTrace™ technology replaces, instead, the exacting, optimized signal processing and positioning functions of the disk drive with custom hardware, software, and algorithms to precisely locate particular sectors of data and recover each bit individually – independent of the drive's specific hardware. Furthermore, its underlying design has the flexibility to provide this data recovery capability into the future as increasing data densities continue to require more hyper-tuning of disk drives in the factory.

ACTION FRONT
DATA RECOVERY LABS

2. Introduction to Hard Disk Drive (HDD) Technology

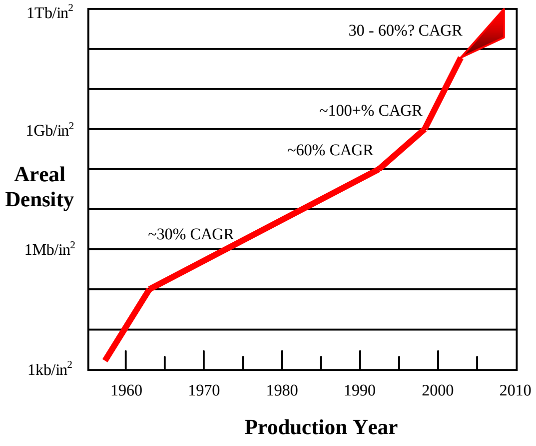
In 1956, IBM introduced the world’s first hard disk drive, the RAMAC (Random Access Method for Accounting and Control). It was approximately the size of two refrigerators placed side-by-side and stored about 5MB of data. It cost over $50,000 [1], or $10Million/GB! Currently, HDDs routinely provide over 100GB of storage in a 3 1/2” form-factor for less than $1/GB. This history of improvement out-strips the better-known development records in semiconductor density and telecommunications data rates. The constant research, high volumes, and low prices required by the disk drive industry have brought many great HDD-related companies into, and out of, existence. Even the HDD bell-weather, IBM, sold its drive business to Hitachi in 2002.

2.1 Areal Density and Price Trends



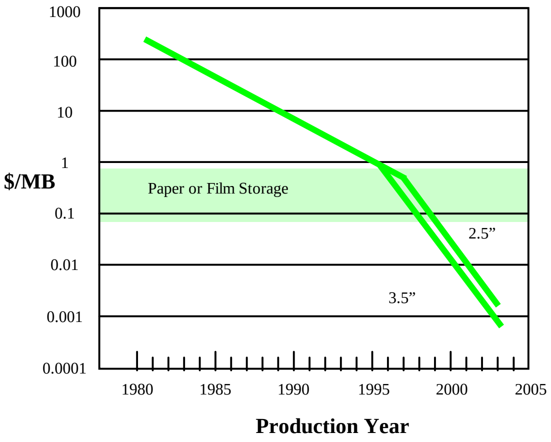
In hard disk drives, data is arranged in concentric circles, called *tracks*. To get more data on a track, the spacing between each bit in the down-track direction must decrease. The data density in this direction, also called the *linear density*, is measured in thousands of *bits per inch* (kbpi). Similarly, the track density across the disk is measured in thousands of *tracks per inch* (ktpi). The tpi metric not only reflects the width of the written track, but also the small *guardbands* that are needed between tracks to provide margin for head-to-track misalignment. These metrics are illustrated in the figure to the left.

Areal density is the metric used to quantify the impressive growth in HDD data storage capacity. It is the product of bpi and tpi, which reflects the amount of *user* data that can be stored reliably in one square unit of area on the disk surface. It is now measured in gigabits per square inch (Gb/in^2). Areal density has increased by almost 8 orders-of-magnitude since the introduction of the first disk drive. This trend is shown in the areal density plot to the left.



For decades, the compound annual growth rate (CAGR) of areal density was about 30%. With the introduction of MR (magneto-resistance) head technology around 1990-1991, the rate increased to 60%. When GMR (giant magnet-resistance) heads were introduced in the late 90’s, the CAGR temporarily increased to over 100%, during which time there was an increase in the number of companies that existed the industry or merged. The pace of areal density growth is now slowing and should settle somewhere between the historical rates of 30 and 60%.

Currently, an areal density of 100 Gb/in^2 might be achieved by a combination of 800 kbpi and 125 ktpi, for example. This provides a *bit aspect ratio* (BAR) of about 6:1 (bpi/tpi). The bit-to-bit spacing in this example is 1.25 microinches



(about 30 nanometers)[1]. The track-to-track spacing (*track pitch*) is 8 microinches (about 200 nanometers). The 5 – 10% guardbands between tracks are a fraction of a microinch (less than 20 nanometers). It is astonishing that drives routinely achieve this level of mechanical precision at a price per megabyte that has been falling at the rates shown in the graph to the left. For the past few years, it has been cheaper to store data on HDDs than on paper or film. Currently the price of HDD storage is about $1/*giga*byte.

When drives cost thousands of dollars, drive repair was a lower priced alternative to purchasing a new HDD. Today, the most economical option for dealing with a malfunctioning drive is to replace it with a new one. The new drive will likely be larger, cheaper, and faster. In fact, it is typically the data itself – even for the home user – that is much more valuable than the drive.

Increasingly, the home user’s drive is filled with often-priceless photos and movies. The time it takes to recover a failed drive can also be more costly than the drive itself – even when backups are available. (You do have backups, don’t you?) However, backups typically represent a snapshot of the data some time ago (last night, last week, last month). Therefore all recent work and transactions are still lost. Unfortunately, many companies that run backups diligently do not practice *restoring* data from backups. Sometimes the backups themselves are corrupted. Even in redundant systems, such as drive arrays, data loss due to multiple-drive failures is not uncommon.

For these reasons, no matter what precautions have been taken, a drive may need the services of a data recovery company. For criminal investigations requiring data forensic analysis, there is no substitute for the drive in question. It must yield its information even if it has been intentionally destroyed.

2.2 What Happens to Data in a Hard Disk Drive?

When you push the “Save” button, 💾, and write your data to the HDD, you expect it to be returned correctly when you open the file in the future. The actual specification for this expectation of *data integrity* is the *unrecoverable read error rate*. This is typically in the range of 1 bit in error for 10^{13} to 10^{15} bits read. Every part and function of the drive is essential for achieving this level of data integrity, however for the purposes of data recovery the topics discussed next are the most relevant. These include the logical-to-physical block translation system, the servo positioning system, the drive layout optimization routines, the data detection algorithms, and the data decoding.

The typical probability of an unrecoverable read error is less than 1 in a trillion bytes read.

2.2.1 Organizing the Data

Files, whether they represent text, a database, photo, song, movie, web page, executable program, or anything else, are stored as a series of *sectors*. A sector is a physical location on the disk that is designated to store (most commonly) 512 *user* bytes. Because of the encoding overhead and the requirements of the

---

[1] A microinch is one millionth of an inch; a nanometer is one billionth of a meter. There are about 25.4 nanometers in 1microinch.

detection algorithms (discussed below), about 600 bytes are actually stored in a sector.

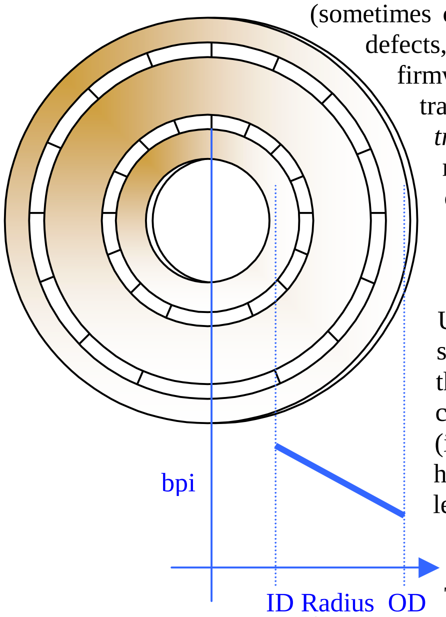
Sectors have traditionally been uniquely identified in a drive by *cylinder*, *head*, and *sector* (CHS) coordinates. The "head" number indicates on which surface the sector is located. The "cylinder" number identifies the specific concentric track on that surface where the sector can be found. And the "sector" number indicates which of the hundreds of sectors on the track contains the data that is sought.

The HDD does not know where files are located. It knows where sectors are located.

How does the drive know where your file is? It doesn't. That is the job of the operating system. The operating system keeps track of which *logical blocks* on which drive contain your file. For convenience, we will consider a logical block to be a data sector, although each block could also point to several consecutive sectors. The drive will request a logical block from the drive, for example block # 1,635,324. The HDD must map this logical block location into a physical block (CHS) location, for example cylinder 5,000 on head 1 at sector 452. There are fast algorithms for computing this, however the interesting complication is when the usual physical location for a logical block has a *defect* that precludes it from reliably storing data.

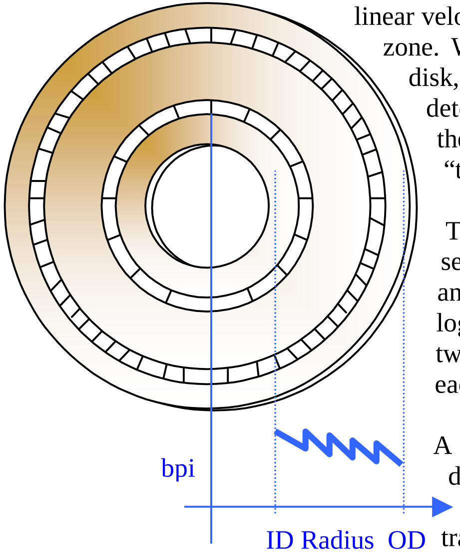
Such locations are found and mapped out during the manufacturing process. There are also provisions for doing this check and re-mapping when the drive is in use in the field. The drive has many spare sectors and even spare tracks to be used as replacements for defective sectors. This is transparent to the operating system under normal operation. The drive accepts the logical block address and performs the *logical-to-physical translation* itself. This varies from drive-to-drive, reflecting the mapping-out of defects found during the drive's surface scan self-test.

In the field, the drive may acquire additional defects due to corrosion, handling, or other causes. These are typically identified in a table of exceptions (sometimes called the P-list and the G-list, for *primary* defects and *grown* defects, respectively). This table, the table of parameters, and the firmware are typically stored on the disk itself in the outermost tracks. These tracks are referred to as the *system area*, *maintenance tracks*, *diskware*, *negative cylinders*, etc. However, some drive models store the table in non-volatile memory on the printed circuit board. Clearly this table of exceptions is uniquely linked to the media in a particular drive. The table for one drive will not, in general, be the same for the media from another drive.



Up until the 1980's, drives typically had the same number of sectors on each track. However, the circumference of a track at the outer radius of the disk (called the OD, for outer diameter) is clearly much larger than the circumference of tracks at the ID (inner diameter). This means that the linear bit density (bpi) is highest only on the innermost track. All the other tracks contain less data than they have the potential to store. This is shown in the graphic to the left.

To maximize the amount of data that can be stored, each disk surface is divided into groups of adjacent tracks called *zones*. There are 8 to 32 (or more) zones per surface. From the ID to the OD, each zone is written with a



higher frequency to counteract the bit spacing growth caused by the higher linear velocities at the larger radii. The bpi still drops slightly across each zone. While zoning makes better use of the storage capacity of the disk, it also means that many unique optimization settings must be determined for each surface during manufacturing. The figure to the left shows the additional sectors in the OD zone and the bpi "taper" across the disk.

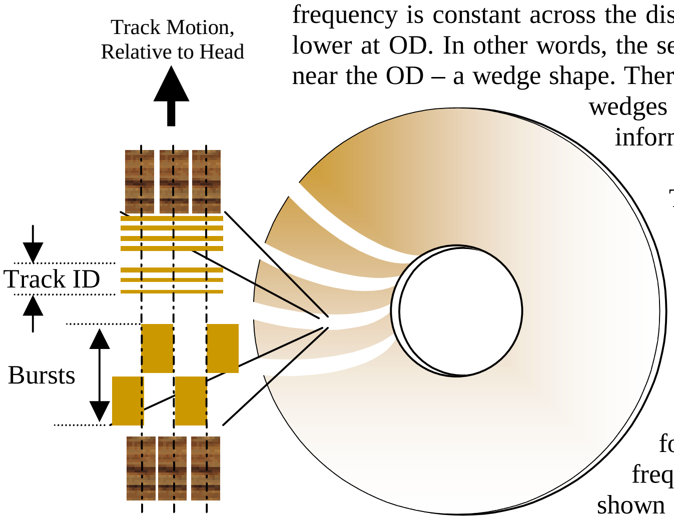
The user's file is likely to be stored across many sectors. These sectors may be spread across different tracks in different zones and even across different disk surfaces. Furthermore, the same logical blocks may be mapped into different physical sectors on two drives depending on the unique distribution of defects on each disk.

A track of data may be less than 10 microinches in width. The drive must find this track within a few thousandths of a second and follow the repeatable and random fluctuations of the track to less than one millionth of an inch. Most amazing is that this can be accomplished in a consumer product that sells for less than $100. The *servo positioning* system makes this possible by using a sophisticated feedback control algorithm that controls the fast seeking and precise track following.

2.2.2 Locating the Data

For the best performance, the servo system requires a very accurate measurement of the head's position relative to the track. Each HDD surface is divided into data sectors and *servo wedges*. The servo wedges are arc-shaped regions that extend from the ID to the OD. They contain a unique magnetic pattern that provides a reference to the center of the track.

The servo pattern is typically written at a much lower bpi than the data and its frequency is constant across the disk. It is not zoned. This means that the bpi is lower at OD. In other words, the servo pattern is shorter near the ID and longer near the OD – a wedge shape. There are typically 50 to 200 evenly spaced servo wedges per revolution. This *embedded servo* information is on each disk surface.



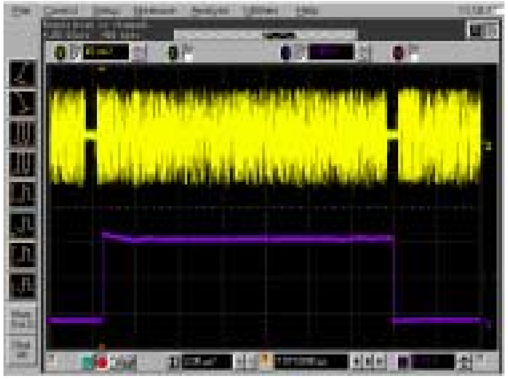
The figure to the left shows three data tracks (high bpi portions with guardbands in between) and an embedded servo field. The servo field begins with a single frequency pattern for establishing timing and amplitude references. A sync pattern indicates the beginning of the encoded cylinder number (or "track ID"). This is followed by three to six *bursts* of single-frequency magnetic transitions (only two are shown in the figure for clarity). These bursts provide accurate position information, relative to the track center.

The first two bursts, typically called the "A burst" and the "B burst," are shown written off center. When the head is exactly on track center, it will get a certain amount of signal from the A burst and then an equal amount from the B burst. The relative amount (amplitude or energy) of each burst signal provides a precise measure of the head's position relative to the track center. Because the servo information is written before any tracks of data, the servo bursts actually *define* the center of the data tracks. The track ID indicates *which* track center.

The servo system also identifies each sector. It does this by maintaining synchronization with the "first" servo wedge in a revolution and timing from there to indicate the beginning and end of each data sector on the track. This timing relationship changes from zone-to-zone, but the wedge-to-wedge servo timing remains constant.

2.2.3 Detecting the Data

Every data sector is a sequence of binary 1s and 0s, stored as a pattern of magnetic *transitions*. A magnetic transition is a change from a north facing-magnet to a south-facing magnet or vice versa. These are sometimes called "north-north transitions" and "south-south transitions," which stresses their "polarity" differences. The GMR head and its amplifier respond with a voltage pulse for each transition that is read. The polarity of the pulse indicates the transition's polarity.



An oscilloscope screen shot of a typical data sector readback waveform (top trace) is shown in the figure to the left. The trace at the bottom is the *read gate*. This is generated based on timing offsets from the rotational synchronization generated by the servo system. As stated above, the timing offsets vary from zone-to-zone. Detection of a sector begins with the read gate's assertion and ends with its de-assertion.

The detection of data is equivalent to the detection of the presence or absence of the pulses, and their polarity. However, detection must take place in a noisy environment, so mistakes can be made. Furthermore, the readback signal can be distorted in many ways, including due to slightly off-track placement of the head. At high bpi the pulses overlap, which causes pulse position shifting known as *intersymbol interference* (ISI). This makes identifying the data sequence especially difficult. Drives today use variations and extensions of partial-response maximum-likelihood (PRML) sequence detection [2, 3] in order to correctly detect data in such environments. In the future even more sophisticated techniques, such as iterative detection, will likely be employed.

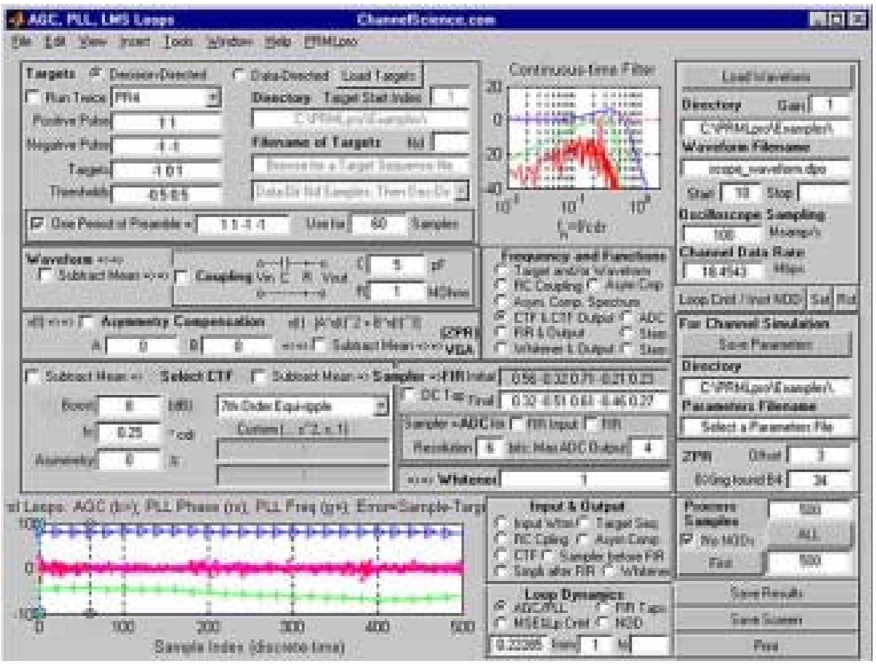
For good error rate performance, it is necessary to establish the proper gain for each sector and lock the detection process to the precise frequency and phase of the readback waveform. This places three specific requirements on the stored data.

1) *Every* data sector must start with a single-frequency sequence of transitions. This is usually called the *preamble* and is about 10 to 15 bytes long. The preamble makes it much easier to establish the proper gain and timing synchronization for the sector. Every servo field also starts with a single frequency preamble for the same reason.

2) It is possible that the beginning of the user's data might look just like the repetitive pattern of the preamble. To precisely indicate the end of the preamble a unique, easily identifiable transition sequence called the *sync mark*, or *frame sync*, is written in between the preamble and the user's data. The sync mark is typically 2 to 6 bytes long and may be written in two locations in case the first sync mark is missed or damaged.
3) After the sync mark is found, gain and timing lock must be maintained throughout the user's data that follows. In order to ensure this, it must contain pulses at least every two to three bytes so that gain and timing locks can be adjusted. For example, if the user stored an all-zeros pattern there would be no transitions to generate pulses to use to maintain synchronization. For this reason, the users data is *run-length limited* (RLL) encoded before being written to the disk. This can expand the amount of data that must be written by about one percent to as much as 12.5%, depending on the RLL code used.

If the few bytes of sync mark are missed or damaged, the entire data sector may be unreadable.

The PRML detection techniques require a *target* for the expected pulse shape and for how pulses interfere with each other. To ensure that the waveform is close to this target, a combination of fixed and adaptive filtering is applied to the readback signal. For best performance, all of these channel parameters must be optimized ("tuned") for each zone of each head in each drive. ChannelScience's read channel simulation software package, PRML*pro*$^{\mathrm{TM}}$ (shown in the figure to the left), models most of the signal processing used for detecting the sequences of 1s and 0s from captured readback waveforms from magnetic disk, tape, and optical drives [4].



Even with all of these steps, the post-detection *raw error rate* is only about 10^{-5} to 10^{-8}. In order to achieve the specified unrecoverable error rates of 10^{-13} to 10^{-15}, error correction coding must also be used.

2.2.4 Decoding the Data

Inside a modern HDD, the users data is encoded about 5 times before being written to the disk. This is done to 1) Ensure no incorrect data is provided to the user, 2) Correct as many errors that may occur in detection as possible, and 3) Improve the quality of detection by improving timing recovery and by mitigating the effects of certain error-prone patterns. Because of these levels of encoding, the user's data itself is not written to the disk. Instead it is the encoded user data that is stored. Even if a tool such as PRML*pro*™ is used to recover the "data," it is actually detecting the *encoded* data. To yield useful information that can be reassembled into files, the various encoding steps must be *de*coded.

The user's data is never stored on the disk. Instead it is the *encoded* user data that is stored.

One "encoding" step is actually a data randomizer, also called a *scrambler*. The scrambler may be thought of as a circuit that pseudo-randomly flips various bits from 1 to 0 or vice versa. Surprisingly, this serves a few useful purposes. 1) Repetitive patterns are broken up. That is, it is less likely that a common pattern, (*e.g.*, a control character, space, carriage return) that might be a difficult pattern to detect will appear over and over, thereby degrading the bit error rate. 2) Electromagnetic interference (EMI) that might be generated by the electronics, in response to a repetitive pattern at a certain frequency, can be reduced. 3) A common pattern with a lot of zeros may be "scrambled" into a pattern with more ones. This can help the gain and timing control loops remain locked to the waveform. 4) It is also possible to scramble adjacent tracks differently. This can provide some decorrelation between tracks, which might improve detection when the head is slightly mispositioned off track center.

Because the bits are flipped *pseudo*-randomly, the flipping sequence can be regenerated during readback so that the data is exactly *un*scrambled. Precise location of the sync mark is necessary for this to succeed. Notice that the scrambler does not *prohibit* any pattern. For example, it is possible for the user to store a bit sequence that is "scrambled" into an all-zeros pattern. For this reason, it is still necessary to apply an RLL code to the scrambled user data.

A common RLL code for PRML channels maps 16 scrambled data bits into 17 code bits. This is a coding overhead of about 6% (17/16). This type of code ensures that there are no more than a certain number of zeros (maybe 10 to 15) in between ones. This causes pulses to be present in the readback waveform often enough for gain and timing to be tracked. There are other RLL codes that have much higher rates than 16/17. There are also RLL codes that are designed to eliminate certain patterns that are more error-prone. It is possible that different RLL codes are used in different zones of a single disk surface.

Currently, most drives combine RLL codes with a parity check code. This typically adds one or two bits to the RLL code overhead. For example, a 64/65-rate code (64 user bits are encoded into 65 RLL code bits) would become a 64/66-rate code when a single parity-check bit is added. The benefit of adding this small amount of parity is that the dominant errors made by the detector can be identified and corrected with a small increase in circuitry and code overhead.

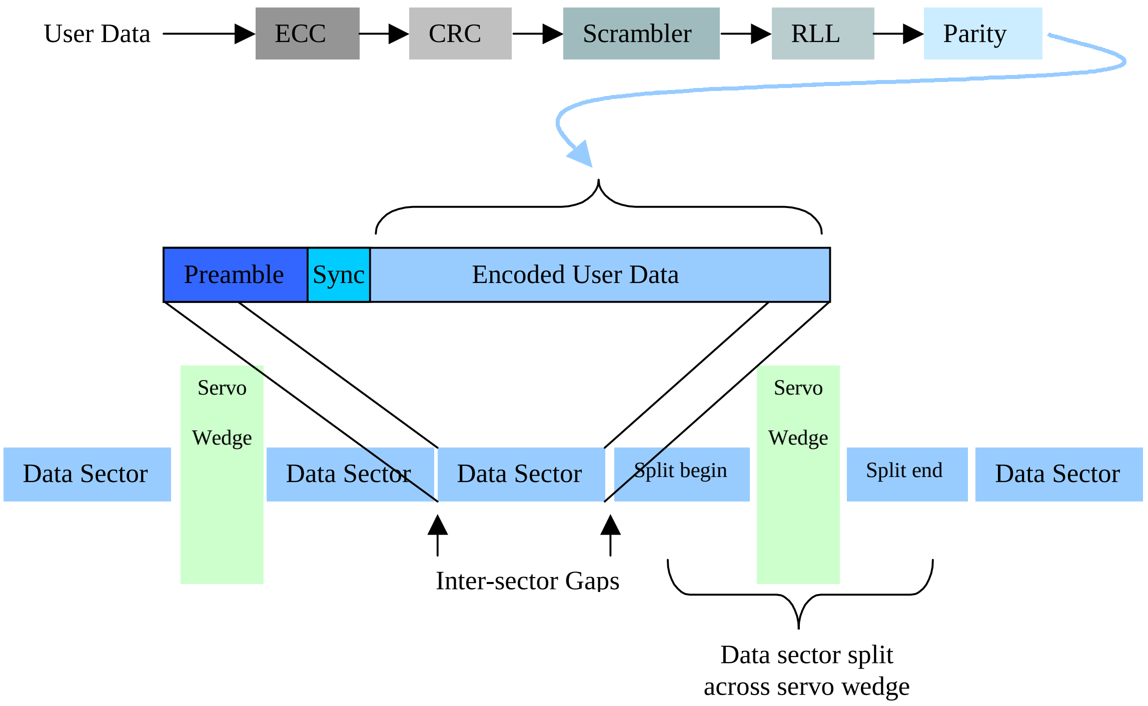
There is a very, very small chance that the user's data will be returned *incorrectly*.

However, all of these encoding methods combined still do not achieve the unrecoverable read error rate goal of better than 10^{-13}. This is possible only with error correction coding (ECC). ECC calculates *parity bytes* for the users data, which provide structured redundancy that can be used during decoding to detect and correct errors. The ECC encoded user data is what is scrambled and RLL encoded. Typically, Reed-Solomon encoding is used because of its good burst error correction capability and the economy of its implementation. Bursts of errors occur because a scratch or other small mark corrupts a group of consecutive bits. It is not uncommon to have the ECC capability to correct over 200 bit errors in a sector.

The ECC can fail in two ways. One way is that there are too many errors in a sector to correct. This is an unrecoverable read error. However, the drive will typically try several *heroic recovery* methods, such as re-reads, off-track reads, and even some reoptimization, to try to detect the data successfully before reporting an unrecoverable read error (also called a *hard error*). The other way ECC fails is much more dangerous.

If there are a few more errors in a sector than the ECC can correct, and they occur in a certain way, it is possible that the ECC decoding *miscorrects* the data. This is disasterous in financial transactions, for example. The probability of miscorrection, also called the probability of *data corruption*, is not commonly specified on drive data sheets. Ideally the probability is much less than 10^{-20}. To ensure that it is very unlikely that data will be miscorrected, the ECC encoded data is often "wrapped" with a CRC (cyclic redundancy check) code. This has a very strong capability to *detect* errors, but is not used for correction. This provides the final check that the data is correct as delivered back to the computer over the interface.

The figure below shows the encoding sequence and the organization of sectors on a track. Notice that to get the most benefit from zoning, sometimes data sectors are "split" across servo wedges. The second part of a split sector must also start with a preamble and a sync mark. The detected data sequences from both portions are concatenated and the decoding and descrambling proceed as usual.



2.2.5 Drive Burn-in and Optimization: "Hyper-Tuning"

After assembling the components, drive manufacturers *burn-in* every HDD. Depending on the quality of the drive and the demands of its intended market, the burn-in procedure may take about an hour or more than a day. The drives may go through testing for seek performance, power consumption, data handling, interface compliance, shock and vibration performance, temperature and power extremes, surface scanning for defects, noise measurements, etc.

However, it is during this time that the drive's parameters for detection, data organization, and positioning are determined. These optimized parameters are typically saved to a table that is stored on the extreme outer tracks ("negative cylinders") of the disk drive. Some manufacturers store duplicate copies of this table on each surface.

In a modern drive, self-servowriting may be employed. This means the drive's servo pattern is written during burn-in, by the drive itself. This gives incredible flexibility for determining a bpi/tpi combination that provides the desired capacity at the most robust performance point for a particular head/media pair. This is sometimes referred to as *adaptive formatting*. It is also necessary to measure various physical parameters of the head including the offset between the head's write element and the read element, resistance, temperature, pulse asymmetry, etc. Various head-parameters must be optimized, such as the current for writing, the current for read biasing, the *write precompensation* that is needed to partially linearize the readback signal, etc.

After the bpi/tpi, zoning, writing parameters, and reading parameters are determined, the detection parameters are optimized. These must be determined for every zone of every surface of every drive. A 6-surface drive with 16 zones requires 96 groups of channel optimization settings to be stored in the parameters table. These channel settings include equalization and noise-whitening filter coefficients; gain, timing, and adaptation parameters; detection target; RLL code selection; etc. Similar settings must also be stored for detecting the servo wedge information.

With almost every new generation, a drive parameter that was fixed becomes variable. This new variable must then be optimized, which leads to the "hyper-tuning" that occurs routinely in modern disk drives.

3. Data Recovery Market

At the time of this writing, a Google® search on "data recovery" returns over 1.5 million hits. Many data recovery companies and do-it-yourself techniques are listed. The search results also show that some refer to simply restoring lost data from backups as "data recovery." For the purposes of this white paper, *professional data recovery* is defined as "*the process of obtaining usable data from downed computers and backups and corrupted or deleted file sets*." The process is labor intensive, highly technical and usually performed in a controlled lab environment [5]. This is in contrast to the term "disaster recovery," which usually means *restoring* lost data from good backups.

Disaster Recovery **is the process of restoring data from good backups.**

Professional Data Recovery **is the process of obtaining usable data from downed computers and backups and corrupted or deleted file sets.**

Because of the vast number of data recovery options and the lack of any industry trade organization or standard-setting body, reliable statistics about success rates and "typical" recoveries are difficult to obtain. However, a leading data recovery company, ActionFront Data Recovery Labs, Inc., estimates that about three hundred thousand drives are sent to data recovery companies worldwide each year. About one half to three-quarters have some issue with reading. This might be sporadic reading, a high number of errors, excessive retries, etc. About one third to one half are *completely* unreadable. That is, they aren't recognized by the host, do not spin up, do not send back any data, etc. For reference, it is estimated that about 260 million new drives were shipped in 2003 [6].

3.1 Perception vs. Reality

On the web, data recovery companies often cite success rates in excess of 90%. No documentation or independent verification of these claims is provided. In reality, it is likely that these claims reflect success with a particular model of drive – or only the drives that they accept for recovery – and not across all drives. The success rates for other drive models may be close to zero.

It is also common to showcase extreme examples of data recovery, such as from drives damaged by fires, floods, smashed cases – even bullets and explosions! These are interesting (and sometimes surprisingly simple) recoveries, but they hide the true nature of the vast majority of hardware failures: The drives just stop working because of mundane reasons.

These include:
- Failure of solder traces, electronic components, or connectors on the printed circuit board (PCB)
- Exceeding a S.M.A.R.T. (Self-Monitoring, Analysis, and Reporting Technology) threshold
- Damaged or corrupted firmware
- Uncorrected bug in factory firmware
- Damage to “system areas” of the disk that are used for calibration, testing, storing firmware and parameters tables
- Spindle or voice-coil motor failure (*e.g.*, short circuit, open circuit)
- Seized bearings
- Breakdown of bearing grease
- Disk shift or mis-alignment
- Head damage
- Overheating
- And many others

3.2 A Call for Transparency

There are many reputable data recovery companies. But it is difficult for the end-user, whose drive containing his tax return failed just before the filing deadline, to determine which company to trust. Furthermore, he does not know if his failed drive is one of the models with which they have a 90% recovery rate or a near-zero rate.

A reputable data recovery company will tell you whether or not they have a good success rate with your drive. However, they might not have received a particular model yet for recovery and will not know for certain how likely a recovery is. Furthermore, even if the success rate is good for a drive, your drive may be damaged in such a way that recovery is not possible. There is always a chance that the data cannot be recovered. When your critical data is on the line, you want to be sure that the data is unrecoverable because of the drive and not because of the lack of skill at the data recovery company that you chose.

An independent data recovery *trade organization* **would benefit the end-user and the industry.**

In the past, some drive manufacturers have had qualification programs in which they identified approved data recovery companies. Such programs appear to have been dropped, perhaps because of the dangers of being implicated in lawsuits if the recovery fails or even makes things worse. On the websites of several major drive manufacturers, their “help” for data recovery now is to provide a Google® search link for “data recovery” and warn the user that they are making no recommendations.

It would be helpful to the end-user, and to the most reputable and capable companies in the industry, for an independent data recovery trade association to be formed. This association could provide a certification program for data recovery specialists and it might also gather and publish statistics on success rates for recoveries on different models of drives. It could also certify individuals in chain-of-evidence procedures for data forensic investigations. In addition, it could direct military, law enforcement, and intelligence agencies to the companies that have the highest success rates for particular types of intentional drive damage.

An easier step toward transparency is for the best data recovery companies to set the standard for the industry by example. This might include providing more information about actual recovery techniques and the true nature of most recoveries. It might also include listing success rates by model or by failure type on their websites. However, unless this is adopted by the top companies at the same time, such disclosures can give a negative impression to the end-user who is searching for a data recovery company for the first (and only) time. They are likely to feel more comfortable with the company claiming to have “magic machines” and “proprietary processes” that yield a success rate of over 90%.

Submit claims of extraordinary data recovery capabilities to the peer-reviewed scrutiny of the *IEEE Transactions on Magnetics*.

As a first step for the industry, all data recovery companies that make a claim of special machines or “proprietary processes” that go beyond the standard replacement of failed parts should present their unique capabilities for independent review. A straightforward way to accomplish this is to submit a paper to the most important journal of peer-reviewed technical papers for the disk drive industry – the *IEEE Transactions on Magnetics* [7]. I call on all data recovery companies to submit their claims of extraordinary capabilities to the scrutiny of this refereed journal.

4. Data Recovery Technology

This white paper focuses on hardware methods of data recovery. Regardless of how the sectors of data on a failed drive are read, the data must be reassembled into useful files. This is done on a computer by special software. Then the files are written to another medium for return to the user.

Magnetic material has been worn away from both sides of this glass disk. Notice that the yellow paper is visible *through* a portion of the top disk.

In a failed hard disk drive, the disk surface may or may not be damaged. If the disk is not physically damaged, the user’s data is still there, unless it has been overwritten. If the disk is physically damaged, there is no data left wherever the magnetic material of the disk is removed. The magnetic layer that contains the data is only about a microinch thick. So *any* scratch is likely to have completely removed the magnetic material in that area. The heads do not scratch the disk in normal operation because they are actually flying over the surface – although the "flight" is at a spacing of less than 1 microinch! If the disk is bent so that the heads can no longer fly, there is no documented method for commercially viable recovery.

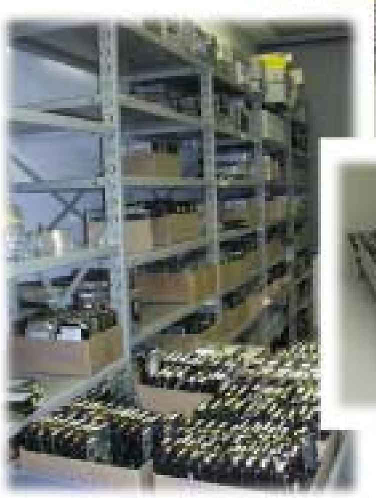
4.1 Traditional Hardware Replacement Methods

The most advanced, commercially viable technique for recovering data from a hardware-failed disk drive is careful replacement of the failed parts. If the part to be replaced in inside the head/disk assembly (HDA), the replacement should be performed in a clean environment, as shown in the photo to the left. Remember that the head must fly about a microinch above the surface of the disk, so a greasy fingerprint or a stuck particle can cause the repaired drive to crash. This is likely to result in even more damage to the data on the disk.

For part-replacement to be successful, spare parts must be available for the specific drive. Drive companies and their component suppliers do not supply spare parts. The parts must come from new "donor" drives of the same type. However, the tight matching of the head with the disk and the hyper-tuning of the system parameters means that it is less likely that a *similar* drive's parts will work. The parts must come from the *same* drive model.

Since most drives that are sent for data recovery are a few years old, they are typically no longer available for purchase. For a data recovery company to have the parts on hand for replacements, they must maintain an inventory of popular drives from the recent past. The pictures to the left show some of ActionFront's inventory of donor drives for part-replacement.

An inventory of donor drives is needed for successful part-replacement.

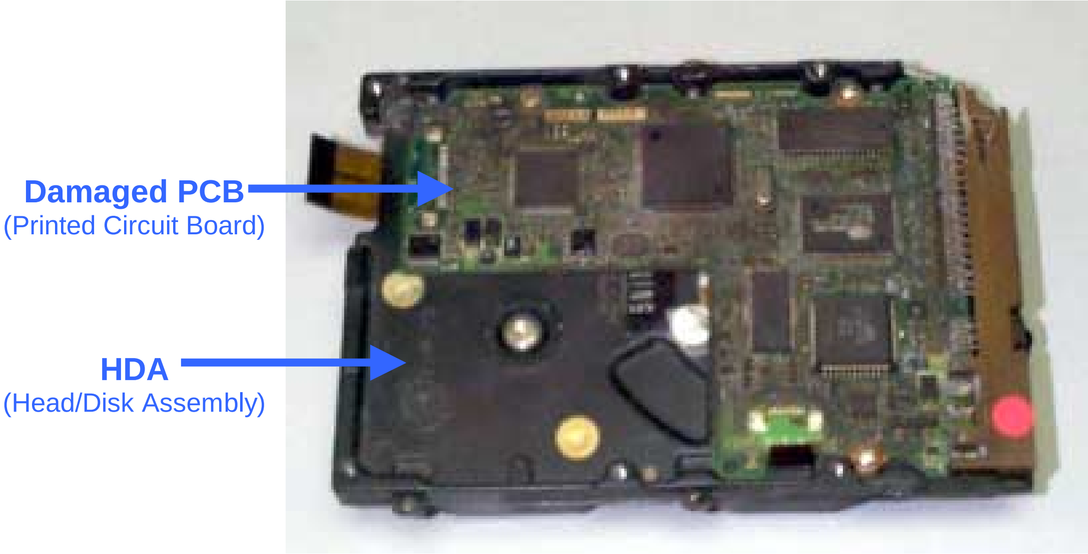


4.1.1 Replace the PCB

The simplest hardware repair is to remove the failed PCB (printed circuit board) and replace it with one from a functioning drive. This can be done outside the cleanroom because there is no need to open the HDA.

If this method is successful, and there are is no other hardware problem, the drive can be powered on. It will spin-up, the heads will seek, the firmware and parameters tables will be read from the system area of the disk and all of the data should be accessible.

This method can fail if the drive stores the parameters tables on non-volative memory on the PCB. Howver, it is possible in some cases to transplant this memeory chip from the failed drive's PCB to the donor PCB. This method can fail if the donor drive's PCB does not contain a very similar version of the PRML read channel to the one on the failed drive. This is because the channel settings in the parameters table might not work for the new chip. This method can also fail if there is additional damage to the drive that prevents it from seeking to the system area and reading the drive's firmware and parameter tables.



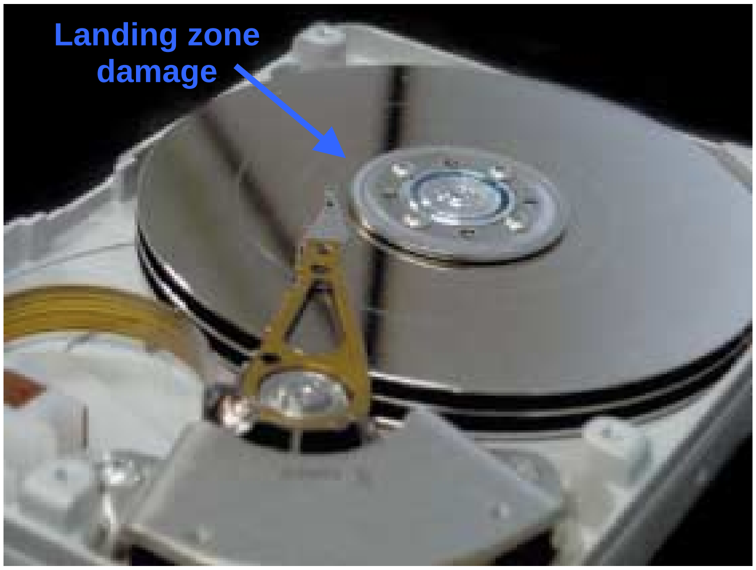
4.1.2 Replace the Firmware

In a working drive, the power-on sequence is usually similar to the following steps.

1. The chips return basic power-on status information.
2. If everything is all right, the spindle motor spins up to its target RPM (revolutions per minute), such as 10,000. Up to this point, the heads have been *parked* and held in place by the *actuator latch* against a *crash stop*. Parking may be in the *landing zone*, which is a portion of the disk surface at the extreme ID (inner diameter). Or the heads may be parked off the extreme OD of the disk on a ramp in *ramp-loading* drives.
3. When the target RPM is reached, the actuator latch disengages and the electronic subsystems start reading the signal coming back from one of the heads. The signal is searched for the servo wedges. Recall that these repeat 50 to 200 times on each revolution of the disk. At this point, the drive has been under the control of relatively simple programs (firmware) stored in ROM on the PCB.
4. When the servo has been found and synchronized to, the drive is now able to seek to the system area of the drive. This is typically at the extreme OD of at least one of the disk surfaces.
5. The firmware ("disk ware") that resides here is then read. It contains larger programs, which control the drive more precisely, that are to be written to and executed from the drive's RAM. This area also contains the parameters tables that provide information about the physical characteristics of the heads, the optimized channel settings, the layout of the data, and where any defective sectors have been re-mapped.

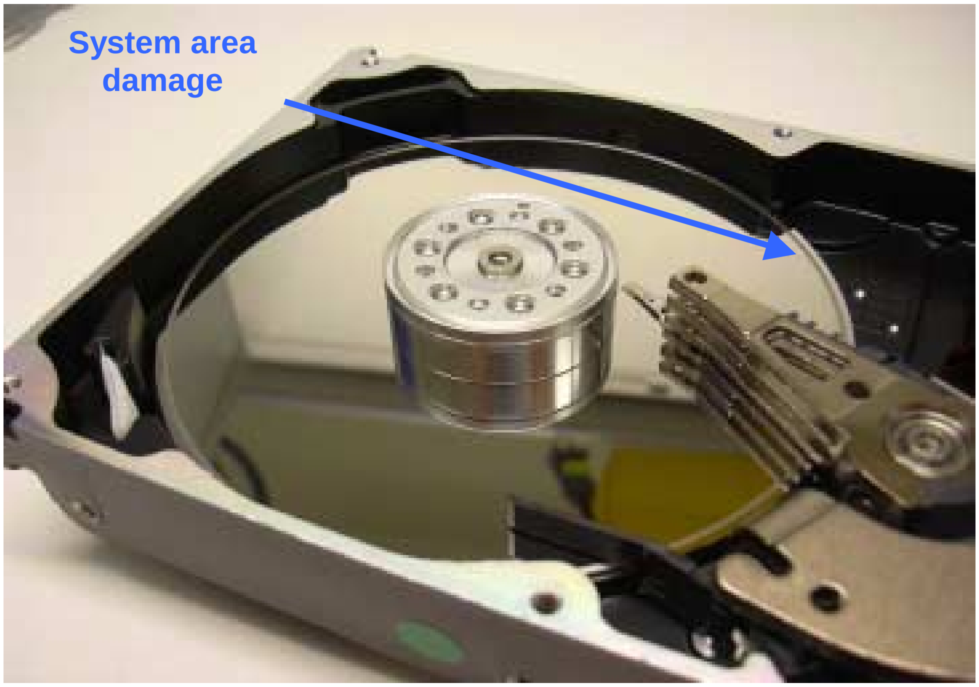
6. At this point, the drive may do some additional self-tests and re-calibrations. It will then signal that it is ready to accept commands.

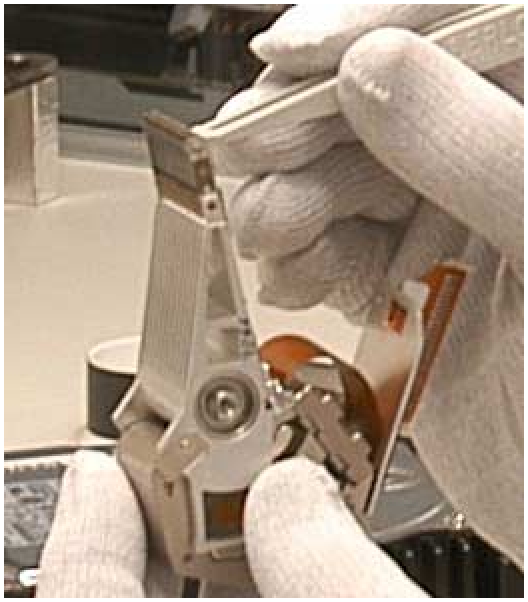
Clearly, anything that stops the drive from reading the firmware complicates part-replacement considerably. If the media near the landing zone is damaged, the servo might not be acquired. If there is a scratch at the OD that destroys the system area, the parameters for the hyper-tuned drive are lost, along with the firmware and defect management information.



One technique that sometimes works on failed drives is to power-on a good drive and let it load the firmware and tables into RAM. Then this “hot” PCB is connected to the HDA of the failed drive. It is also possible to get a “snapshot” of the donor drive’s RAM contents and write this to the RAM of the failed drive.

For either method of getting the firmware loaded into the failed drive, the best outcome is that the drive will spin up, its own servo will be re-synchronized, and seeking will be possible. Of course, the wrong “defects” will probably be mapped out because the defect table from the *donor* drive will be in RAM.





4.1.3 Replace the Head Stack

The highest level of skill is required to replace damaged heads – without damaging something else in the process. The heads are connected via suspensions and arms to an E-block that has bearings in its hub. This assembly, with a portion of the voice-coil motor actuator, is called the *head stack*. When heads are replaced, it is easiest to replace the entire head stack. This is also referred to as a *head transplant*.

Great care is necessary with the replacement head stack to ensure that the *air bearing surfaces* of the heads (actually the “sliders”) do not touch each other. This is because they can easily damage each other or even stick together. When removing the damaged heads, it is important not to drag them across the disk and cause more damage. Ideally custom tooling, such as a special *comb*, will be used to lift the sliders off the disk surface before removing them.

Equal or greater care is necessary when loading the good heads back onto the damaged drive’s media.

4.1.4 Move the Disks to Another Drive

If the base casting is badly damaged, or the spindle motor is burned out, or the spindle bearings have seized, it is necessary to remove the disks from the failed drive. These disks must be re-mounted on the spindle motor of a good drive. This procedure requires all the skill of head replacement with the additional skill of remounting disks without further damaging them.

It is very important to preserve the spacing between the disks and their rotational alignment to each other. This makes the possibility of servoing on the remounted disks much more likely. Furthermore, if two highly polished surfaces, such as those on disks or heads, touch they can become bonded together. This is usually called *stiction*. If two disk surfaces become bonded in this manner, it is usually impossible to separate them without causing excessive, but microscopic, damage.

Once the good heads are loaded onto the remounted disks, the power-on procedure can begin.

4.2 “Magic Machines” and “Proprietary Processes”

Reading some data recovery websites can lead one to believe that they have “Magic Machines” that routinely recover data from failed drives. I saw no evidence or independent verification that such devices exist for *commercially viable* data recovery. If they do have a magic machine it may have been created for a high-value job in the past, and probably only worked marginally.

It is very telling that the US Department of Defense's Combating Terrorism Technology Support Office recently placed a "Broad Agency Announcement" *seeking* just such a magic machine for damaged, erased, or overwritten media [8].

Any "Proprietary Processes" cited by data recovery companies are likely to be custom fixtures, such as combs, and handling procedures for replacing failed parts without causing additional damage. Companies may also have written their own proprietary software tools for re-assembling recovered sectors into useful files.

However, there *are* very special machines used by drive manufacturers for the design and analysis of drive components. It is often suggested that these precision instruments, spin-stand testers and magnetic force microscopes (MFMs), can be used for data recovery.

4.2.1 Spin-Stand Testers

Hard disk drive manufacturers and their head, media, preamplifier, and read channel suppliers do have very accurate, very expensive "magic machines," called spin-stands [9]. These are used for testing and experimenting with heads and disks. They are used mostly by research and development departments and by incoming inspection, production testing, and quality control personnel.

Spin-stand testers are accurate, flexible instruments that illustrate the benefits of drive-independent test equipment.

Spin-stands are very accurate and flexible – for analyzing raw disks. Virtually any data pattern can be written and the positioning accuracy and repeatability are in the nanometer range. However, this typically requires that the tester write its own servo pattern. Reading a disk that has been written by a drive is more problematic.

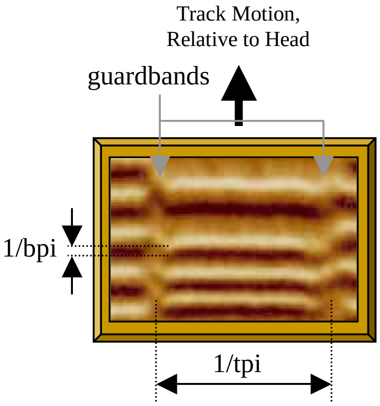
First the disk and head must be aligned as close to their relationship in the disk drive as possible. Then the electronics and software must be programmed to utilize the servo pattern written on the disk. If the servo can be followed, the parameters for the head and channel still need to be optimized. Assuming that is possible, the data *written* to the disk should be readable.

However, unless the exact read channel and its coding options are available for the tester, all that will be delivered is scrambled, RLL encoded, ECC codewords at best. These must still be decoded and then assembled into useful files. Note also that the head will be flying over the disk surface, so the disk must not be significantly damaged.

In reality, the scenario above is very difficult to successfully implement even for a drive manufacturer. It takes a great deal of trial-and-error investigation by a very knowledgeable operator. It would be much more difficult for a data recovery company to implement this technique successfully across virtually all manufacturer's drives *cost-effectively*.

However, the drive-independent nature of the spin-stand is a very appealing and necessary feature for a general data recovery tool. What is needed is a device that offers similar flexibility, can detect and *decode* the user's data, is much less costly, ideally works for every drive made, and will continue to work (with modifications) for future drives.

4.2.2 Magnetic Force Microscopes (MFM)

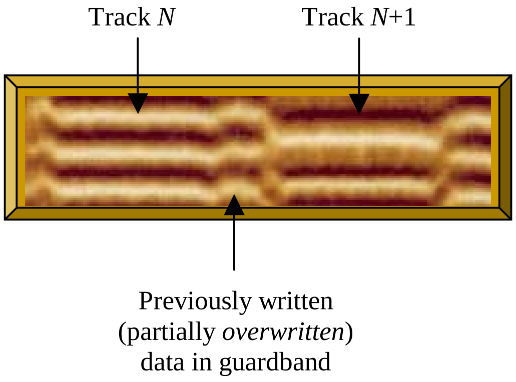


The ultimate tool for analyzing the magnetic data on disks is the MFM. It is related to the atomic force microscope (AFM), except it responds to the magnetic force of the disk's data and servo patterns [10]. Typically the instrument offers both AFM and MFM capabilities. It provides phenomenal images of the topology and magnetization of the disk.

The figure to the left is an MFM image of a portion of a track of data. The dark and light horizontal lines are the individual transitions. Assuming the transitions are 1s, the spaces in between the transitions are the 0s. The detail clearly reveals the guardbands between tracks and even the curl at the edges of the written track due to the shape of the write field.

The MFM probe must be very close to the disk surface in order to get these images. Therefore it cannot easily follow a badly damaged (*e.g.*, bent) disk. The biggest drawback, however, is its speed. The MFM scans about a 100 micron by 100 micron area at a time, then the sample must be moved and the next area scanned.

As a very rough approximation, if a 3 1/2" disk is to be imaged and the MFM can scan and move to the next area in one minute (quite fast!). It would take about 60 weeks of 24 hour/day operation to scan one surface. If the disk surface holds 50GB of data, for example, the image files that would be generated from the MFM would be many times this amount – perhaps generating tens of terabytes of image information to *analyze*. For example, all of these individual images would need to be stitched together into a complete disk image and a software image processing algorithm would need to be used to 1) servo on each track and 2) generate the read gate to indicate the beginning and ending of each sector. Finally a signal from the center of the track image would need to be generated as a readback signal, detected, decoded and assembled into useful files.



The most intriguing possibility for magnetic force microscopy as a data recovery tool is reading overwritten data [11]. As shown in the image to the left, when a track is overwritten there is often a portion of the previously written data remaining. This is due to small variations in the servo's placement of the write element as well as the effects of spindle runout. It is theoretically possible to take all the steps listed above but generate the readback signal from in between tracks rather than from track center. This procedure will have about the same level of difficulty, but the error rate of the readback signal will be much worse. Also the overwritten signal will be slowly fading in and out due to non-repeatable spindle runout that occurs during writing. Such an effort could only be afforded for a small amount of the most important data for national security.

4.2.3 The Spin-Stand MFM?

To over come the time of image acquisition with an MFM, it has been demonstrated [12] that magnetic recording heads can be used on a spin-stand tester to create an *image* of the magnetic pattern on the disk. That is, a flying

GMR head is used in place of an MFM probe. This has the advantage of being able to image a disk in a few hours, depending on the resolution desired.

However, it still leaves all the problems of analyzing (quickly) the many terabytes of image data generated. The images must be arranged in the correct spatial pattern and the tracks followed by some image processing servo routine. The readback signal from the track center (or guardband) must be generated. And finally the data must be detected, decoded, and assembled into useful files. An improvement on this system would be to servo the imaging head during the scan by using the magnetic patterns written on the disk.

4.2.4 Exotic Recovery

It is *theoretically* possible to read some overwritten data.

Although such exotic methods of data recovery are theoretically possible, and have even been discussed in the peer-reviewed literature [11, 12], I have found no evidence of *commercially viable* recoveries being performed with them. Furthermore, I have seen no public demonstrations of any of these methods that show the recovery of files or even user data – only images or raw encoded data.

5. The Frontiers of What's Possible: What Makes Data Unrecoverable?

From the preceding descriptions of hard disk drive technology it should be clear that part-replacement for data recovery is difficult now and likely to get more difficult in the future. Part-replacement fails for a variety of reasons, but most of them reflect the hyper-tuning drives undergo to achieve high manufacturing yields combined with high data density.

The drives optimize the particular head/media/electronics combination they have as well as adapt to the precise physical relationships between the positions of the read element, write element, spindle center, and head stack pivot point. Because of hyper-tuning, the range of parameters over which a drive can operate is very small and likely to get even smaller. Part-replacement, by its nature, succeeds most often in drives that work over a wider range of parameter values.

5.1 When Firmware Replacement Fails

The drive and its firmware are optimized for high production yields, reliability, and data integrity – not for data recovery after years in the field. The optimization of drive parameters for the specific head/media/electronics combination in the drive results in a hyper-tuned product that works very well, but only over a narrow range of system parameters.

Because of this, firmware replacement is likely to fail for one or more of the following reasons.

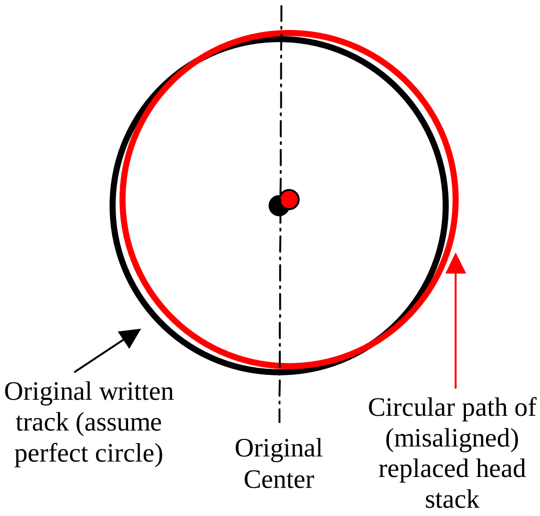
- Channel settings for servo signal detection are too far off for good servo reading
- Servo synchronization is achieved but the head offset measurements are too far off to yield proper seeking
- The zone table information that identifies the layout of the drive and the adaptive format information, such as bpi and tpi, is completely different form the drive and no data can be read

- Sectors needed for crucial files are listed as defective (from the donor drive’s G-list and P-list) and defective sectors are listed as good
- The channel settings for data are too far off to get good read error rate performance

5.2 When Head Stack Replacement Fails

In a 100ktpi drive, each track of data is less than 10 microinches wide. A typical specification for servo track following is to be within 10% (1 microinch) of track center. The ID to OD *full-stroke* seek is about 1 inch for a 3 1/2” drive. This means that the servo control must operate over 6 orders of magnitude.

To make this possible, certain key physical parameters of the drive are measured, or calibrated, in the factory. For example, one such parameter is the offset between the read and write elements on each head, and how the relationship changes from track-to-track due to the effects of skew angle. Another parameter is related to the fact that the tracks are not perfect circles. This is called *eccentricity* and its effect is referred to as repeatable runout (RRO). RRO can be measured on each surface and a periodic term added to the servo algorithm to compensate for this predicable movement of the track relative to the head during each revolution. For clarity, the figure to the left illustrates RRO caused by a shift in the disk center relative to the location of the pivot point for the new replacement head stack.



The readback signal from the preamplifier depends on both the medium and the head. The parameters of the preamp and the read channel are optimized for the signal produced by this particular combination.

Head stack replacement is likely to fail for one or more of the following reasons.

- The head’s flying height is significantly different, resulting in a changed pulse shape, a weaker signal, or a saturated signal
- The head’s sensitivity relative to the medium’s magnetic strength (“$M_r t$”) is significantly different, resulting in a changed pulse shape, a weaker signal, or a saturated signal
- Head stack is mis-aligned on one or more surface from the positions of the original heads leading to excessive eccentricity that cannot be tracked out by the servo (or at least not with its factory-determined parameters)
- The spacing of the heads might be different relative to the spacing of the disks in the disk pack, which can make head reloading difficult, possibly resulting in disk damage

5.3 When Disk Remounting Fails

Removing the disks from a failed drive and remounting them onto the spindle motor of another base casting has all of the same problems with magnetic matching and physical alignment as head transplant. However it also has two additional problems.

A user's file may be spread over multiple surfaces. It is required then that servoing with one head be seamlessly transferred to servoing with another head. To accomplish this, it is necessary that the servo wedge timing relationships between the surfaces be known. If a disk rotates *relative to* the others in the disk pack (a condition known as *disk slip*) the servo timing relationship is altered.

If this happens in the field (perhaps due to a rotational shock), the worst consequence is that a write will be executed, based on the wrong servo timing when a *servo wedge* is under the write element. This write operation will destroy the servo information. If this continues for even a few wedges, the surface (and hence the drive) is likely to be unreadable by normal means.

HDD manufacturers typically use carefully controlled robots to place the disks and spacers in the pack, balance the disks (if needed), and torque the retaining screws precisely. The disk can warp (or "potato chip") if retaining screws do not provide even pressure or if a spacer is not very flat. The head can follow a certain amount of warpage as it flies, but it is possible to have excessive potato chipping that results in erratic flying. It is also possible that the disk's motion and the windage that it generates can excite certain mechanical resonances in the suspension and arm, which can make precise servoing very difficult.

Disk remounting is likely to fail for one or more of the following reasons.

- Disk slip
- The centers of the disks line up differently, resulting in disk-to-disk eccentricities that are different than the servo is programmed to correct
- Disk warp
- The spacing of the disks might be different relative to the spacing of the heads in the head stack; this can result in load force differences that cause excessive flying height differences
- The spacing of the disks might be different relative to the spacing of the heads in the head stack, which can make head reloading difficult, possibly resulting in disk damage

5.4 When the DATA Fails

Paintings on cave walls, images carved in stone, and pigments on leather have lasted through the millennia. Magnetically recorded data will not fare as well. Currently, most of our data is "born digital" and lives on magnetic disks and tapes. In the past, one could expect magnetically recorded data to last about 50 to 100 years, under normal conditions. However, there are two unique problems with digital data.

***Super-paramagnetism* refers to the gradual thermal decay of magnetically stored data**

First, digital information tends to be all or nothing in terms of recovery. For example, written, painted, or carved works are likely to degrade slowly over time. Once digital data degrades to the point that ECC can no longer correct it, the data is lost. Second, magnetic data is not inherently "human readable." That is, machines are needed to read the magnetic data. This is in clear contrast to paintings, carvings, and writings that are read with the naked eye. If the machine designed to read a particular medium is broken, the data is effectively lost even if the magnetic pattern is still intact. Furthermore, even if the drive can provide the data perfectly, the programs that use that data and the machines that they run on must also be available.

The gradual degradation of magnetic information is often referred to as *thermal decay*. Briefly, there may be millions of atoms that are magnetically oriented in a certain direction in one bit. Over the years, thermal energy, *i.e.* heat, will make some of the atoms “forget” their magnetic orientation in the bit. Now there are fewer atoms maintaining the data, which means that it takes even less thermal energy to get a few more atoms to forget the data. After some period of time an “avalanche point” is reached resulting in a random orientation for the bit and the data is lost – it has decayed. This is referred to as the *superparamagnetic effect*.

The thermal stability of bits drops rapidly as the areal density increases – there are fewer atoms in each bit to retain the magnetic orientation. High temperature environments can make thermal decay worse. It has also been demonstrated that writing to a particular track can cause degradation in the bits in the adjacent tracks. It is not widely known, but many modern drives routinely check for thermal decay of bits in the field and rewrite the sectors in which degradation is identified.

Unfortunately, there are likely to be many memories lost in the future as home videos, long forgotten in a hot attic, are replayed only to find the image degraded or lost due to thermal decay.

6. Future Success Depends upon Developing Drive-Independent Data Recovery Capabilities

Recovering currently unrecoverable data requires the development of drive-independent data recovery techniques. These techniques must return user data, cost-effectively, from most drive models. The requirements for such a capability are listed below.

Minimum Requirements for Commercially Viable Drive-Independent Data Recovery

- Must be economical in time, personnel, equipment, and cost to the end-user
- Must be able to read individually identifiable sectors of data from anywhere on the disk surface
- Must be able to continually, incrementally improve to provide good error rate performance at the lower signal-to-noise ratios of newer drives
- Must be flexible to accommodate the differences between the signal processing and coding used by each drive manufacturer, each drive model, and even each drive
- Must be combined with off-line methods for determining the drive layout, adaptive format, servo pattern, and detection/decoding parameters used by each drive

Additional Desirable Features for Drive-Independent Data Recovery

- Compatibility with other methods of acquiring signals from deliberately damaged disks, as an aid to worldwide counter-terrorism activities and law enforcement
- Capability to recover thermally decayed data for historically significant recoveries that might be needed in the future
- Capability to recover long-term archived data from very old drives that are no longer serviceable

7. The FIRST Public Demonstration of Drive-Independent Data Recovery: ActionFront's SignalTrace^{TM} Technology

In a public demonstration on the exhibit floor of the 2004 IEEE NASA Mass Storage Systems and Technologies Conference (Adelphi, Maryland), ActionFront Data Recovery Labs, Inc. demonstrated the successful drive-independent recovery of user data with their prototype system employing SignalTrace^{TM} technology. ChannelScience assisted with some portions of the development of SignalTrace^{TM} technology.

The prototype system shown in the picture to the left replaces the PCB electronics of the drive. The complete system consists of a drive (to be recovered) without its PCB; a power supply; a differential probe; a digitizer to capture the head's readback signal; the algorithms for servo acquisition, synchronization, seeking, and following; the electronics boards shown and algorithms for the detection, decoding, descrambling and ECC checking of the data; and a PC for control of the system. For the prototype, the PC also performs the detection and decoding operations.

The recovery demonstration used a working Western Digital Enterprise (WDE4360) drive. To prepare the drive for the demonstration, a special utility was used to write a user-specified pattern of text to every sector of the drive. The utility also appended to this text message a unique LBA. On readback, this LBA identifier verifies that the drive was controlled to read the correct sector. After writing, the drive's PCB was completely removed.

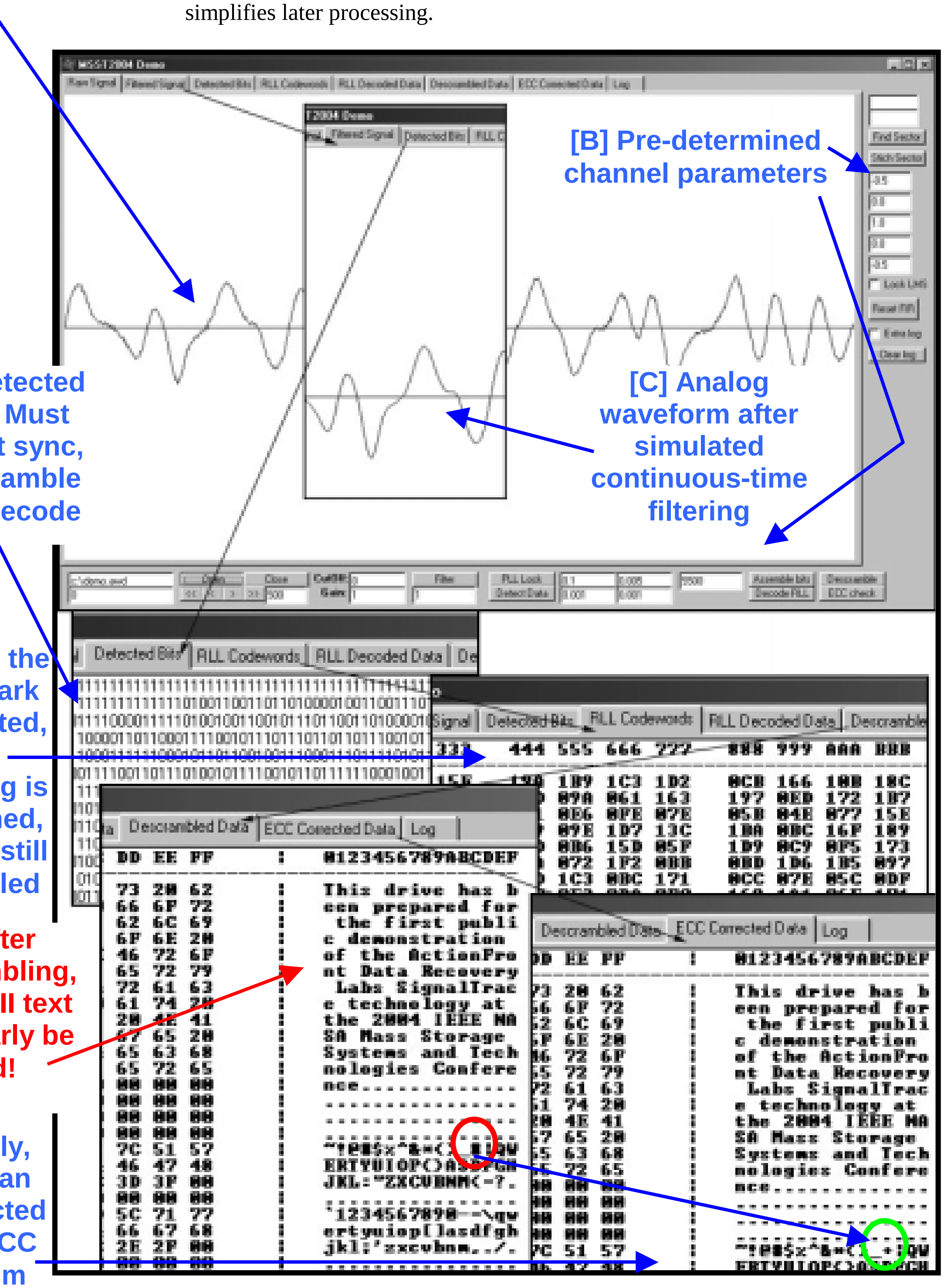
The demonstration of drive-independent data recovery using ActionFront's prototype system proceeds as follows. The SignalTrace^{TM} motor controller spins up the drive to the approximate RPM. The voice coil motor (VCM) that moves the head stack is sent a strong current to pull away from the drive's magnetic actuator latch. Then the current is reduced to slowly move the head out toward the OD until the force of the VCM is counterbalanced by the force of the windage and flex cable.

At this point, the prototype SignalTrace^{TM} system is acquiring the servo wedges and synchronizing to them. It commands the motor controller to make fine adjustments to the RPM as needed, based on the servo wedge timing. It also finds the once-around spindle index provided by the servo. It does this with whichever head is selected by the SignalTrace^{TM} software in the control PC.

***SignalTrace^{TM}* controls the drive to seek to any track on any surface.**

From the PC, the drive can be commanded to seek to any track (with any head). The seek-and-settle time is 5 to 10 seconds. A single-track seek takes less than 1 second. The servo control algorithm is implemented in a Motorola MCS5407 ColdFire microprocessor. Note that the system must be pre-programmed with the servo layout, the zone frequencies, the channel parameters, and the codes used. These are determined off-line.

The head/preamp readback signal for all of a track, or only a portion of it, is captured by a high-speed digitizer, which is connected through a differential oscilloscope probe. The signal is oversampled about 10 times over the data rate of the zone to capture the original noisy analog waveform [see A in figure below]. Only the non-servo wedge portions of a track are captured. This simplifies later processing.



[A] Digitized noisy analog waveform

[B] Pre-determined channel parameters

[C] Analog waveform after simulated continuous-time filtering

[D] Detected bits: Must detect sync, descramble and decode

[E] After the sync mark is detected, RLL decoding is performed, but it is still scrambled

[F] After descrambling, the ASCII text can clearly be read!

[G] Lastly, errors can be corrected by the ECC algorithm

In the production system, the processing, detection, and decoding of this waveform will be performed in hardware, using an FPGA. In the prototype, the host PC performs these steps. The detection algorithm is preset for the filtering, gain and timing loop parameters, and the detection targets. These parameters were pre-determined offline and are set as shown [see B].

***SignalTrace^{TM}* has demonstrated the complete control of a disk drive and the successful retrieval of user data.**

The results for simulated continuous-time filtering are shown in [C]. The rest of the channel signal processing is not shown explicitly. However, automatic gain control (AGC), phase-locked loop (PLL) controlled sampling, and adaptive finite impulse response (FIR) filtering are performed in the PC. The detection of 1s and 0s is performed and the results are displayed [D].

The sync mark must be found in this sequence before RLL decoding can proceed [E]. After the RLL decoding, the data must still be descrambled before the ASCII text and unique LBA written by the utility program can be seen [F]. Note, the requested LBA was found, but that portion of the decoded sector is not shown in the figure above. A few errors were made in the detection process (not on purpose). These were corrected by the ECC [G].

This is a milestone for the data recovery industry.

This demonstrates the complete control of a disk drive and the returning of corrected user data without relying on any electronics (except the preamp inside the HDA) or signal processing from the drive itself.

ActionFront and ChannelScience worked together to overcome many long-standing challenges in order to achieve this milestone in data recovery history. An especially important advancement is the cryptographic procedures employed by the research staff at ActionFront to descramble, RLL decode, and ECC correct the raw detected data. This was reverse engineered, based on first-principles analysis of a good drive of the same model. These highly specialized techniques – as well as the determination of many channel parameters, servo layout and data layout – must be applied to each new drive model before recovery can be attempted. This is because the needed information for drive-independent data recovery is not readily available from the drive and channel companies.

To further document and verify this milestone for the data recovery industry, ActionFront has agreed to submit a paper on SignalTrace^{TM} technology to a refereed technical journal.

***SignalTrace^{TM} Technology* will provide an important tool for law-enforcement and counter-terrorism professionals.**

Drives continue to evolve, getting more sophisticated, adaptive, and hyper-tuned. For data recovery of hardware-failed drives to continue to be successful, drive-independent data recovery techniques, such as SignalTrace^{TM} technology, must be made commercially viable. Furthermore, they must work for most popular drive models and they must continue to accommodate the relevant new innovations in HDDs. An important additional benefit of drive-independent data recovery is that it can be compatible with exotic data acquisition techniques for retrieving readback signals from intentionally damaged disks. This can be a significant tool for law enforcement and counter-terrorism professionals.

8. Conclusions

Backup often and *test* the backups, or …

The majority of drives that are sent to data recovery companies for hardware failure are a few years old. While some of them still respond well to traditional part-replacement, there are some that are almost never recoverable. These may have been "hyper-tuned" in the factory so that high data density can be achieved together with high manufacturing yields, and/or they may have corrupted *system-areas* on the disk where drive parameters tables are stored. Such drives require a very precise matching of the characteristics of the head, disk surface, and the system parameters that is not possible with traditional part-replacement. As data density continues its rapid increase, it is expected that fewer hardware-failed drives will be recoverable with traditional part-replacement.

Drive-independent data recovery methods need to be developed that re-optimize the replaced head, disk location, electronics, and/or firmware and parameters table for the media from the failed drive. This requires recreating much of the drive's optimization routines as well as mimicking the drive's own methods of seeking to a disk location and track-following. This can vary from drive-to-drive – even within the same model of drive. Creating a cost-effective, reliable data recovery method that works across many drives now and in the future requires constant R&D at the edge of the state-of-the-art in disk drives.

Therefore, it is likely that the capability to recover data from almost all of the latest drives will only be available from the best of the best -- the data recovery companies that other data recovery companies turn to for their most challenging tasks. Drive manufacturers could help data recovery efforts by providing features such as special commands to load and run optimization routines that allow part replacement to work better. However this is unlikely given the effort that drive companies must devote to increasing areal density, manufacturing yields, and reliability.

ActionFront Data Recovery Labs Inc. is the first and only company to publicly demonstrate the capability of drive-independent data recovery, with its SignalTrace™ Technology. ChannelScience assisted with portions of the development of this capability. SignalTrace™ Technology has demonstrated the capabilities needed to recover data that is currently unrecoverable by traditional part-replacement. Furthermore, the business goal is to make this technology-intense method commercially viable, so that it is within reach of the individual end-user of hard disk drives – not exclusively large corporations and government agencies.

Furthermore, ActionFront has answered my call for independent verification of data recovery claims. Nick Majors, president of ActionFront, has committed his research staff to submitting a technical paper on SignalTrace™ technology to a peer-reviewed journal. When complete, the paper will be available on the ActionFront website.

ACTION FRONT
DATA RECOVERY LABS

9. References

[1] Hitachi Global Storage Technologies website, www.hgst.com, follow links to "About Us," "Company Background," "Technology Timeline."

[2] R.A. Cideciyan, *et al.*, "A PRML System for Digital Magnetic Recording", *IEEE Journal on Selected Areas in Communications*, January 1992, pp. 38-56.

[3] R.A. Cideciyan, *et al.*, "Noise Predictive Maximum Likelihood Detection Combined with Parity-Based Post-Processing", *IEEE Transactions on Magnetics*, March 2001, pp. 714-720.

[4] A 30-day Free Trial version of PRML*pro*™ can be downloaded from www.ChannelScience.com.

[5] Professional data recovery definition is from ActionFront Data Recovery Labs' *Data Emergency Guide*.

[6] J. Donovan and J. Kim of TrendFocus, "Growth Returns to a Leaner, Healthier HDD Industry, *Insight*, January-February, 2004. Available at www.idema.org.

[7] More information on the *IEEE Transactions on Magnetics* is available at www.ieee.org.

[8] US Department of Defense Broad Agency Announcement (BAA), May 14, 2003: https://www.bids.tswg.gov/tswg/bids.nsf/(BAAView)/24C936C9BBF67CCD85256CDF007C8FC3/$FILE/DAAD_03_T_0024.pdf

[9] Guzik Technical Enterprises Inc. makes spin-stands that are very widely used throughout the HDD industry, www.guzik.com

[10] Veeco Inc. makes popular MFMs and other tools for nanoscale metrology, www.veeco.com

[11] R. Gomez, A. Adly, I. Mayergoyz, and E. Burke, "Magnetic Force Scanning Tunnelling Microscope Imaging of Overwritten Data", *IEEE Transactions on Magnetics*, September 1992, pp. 3141-3143.

[12] I.D.Mayergoyz, C. Serpico , C. Krafft, and C. Tse, "Magnetic Imaging on a Spin-Stand", *Journal of Applied Physics*, Vol. 87, No.9, May 1, 2000, pp.6824-6826.

About the Author

Charles Sobey has worked in the data detection and data storage industries for two decades. His projects have included military communication systems, medical image processing systems, design and manufacture of magnetic recording heads, and modeling and development of detection methods, such as those found in PRML (partial-response maximum likelihood) read channels for data storage devices. He is a senior member of the IEEE and a member of the board of directors of the 3rd largest IEEE chapter in the world, the Dallas Texas Communications and Vehicular Technology society.

Chuck is Chief Scientist of ChannelScience and the creator of PRML*pro*™, a commercial software tool that replicates most of the detection-related signal processing that occurs in magnetic disk, tape, and optical data storage devices. His technical interests include nanotechnolgy-based data storage methods, integrating hard disk drives into non-traditional applications, physiological monitoring, and applying advanced adaptive data detection techniques to chemical and biological warfare sensors for homeland security.

A favorite challenge is helping companies prepare their new technologies for acceptance in the hard disk drive industry. Chuck can also be found on the other side of the table, evaluating the technologies of companies that are investment candidates or merger and acquisition targets. He also assists large, small, and emerging companies with technology roadmap assessment and technical marketing.

Chuck has five issued US patents and several published papers and articles. He has authored and taught very popular seminars around the world for data storage professionals through KnowledgeTek Inc. These include such topics as hard disk drive technology, servo positioning, PRML detection, error correction coding (ECC) and iterative detection. A free download of the 30-day trial version of PRML*pro*™ is available at www.ChannelScience.com.

Please Note

This white paper is for information only and is provided as is. The reader is encouraged to refer to original sources and to check the patent ownership of ideas or technology presented herein before using any of the information provided. Not all data recoveries are successful, regardless of the company or method used. The author is not responsible for any direct or consequential losses due to the use of the information contained herein.