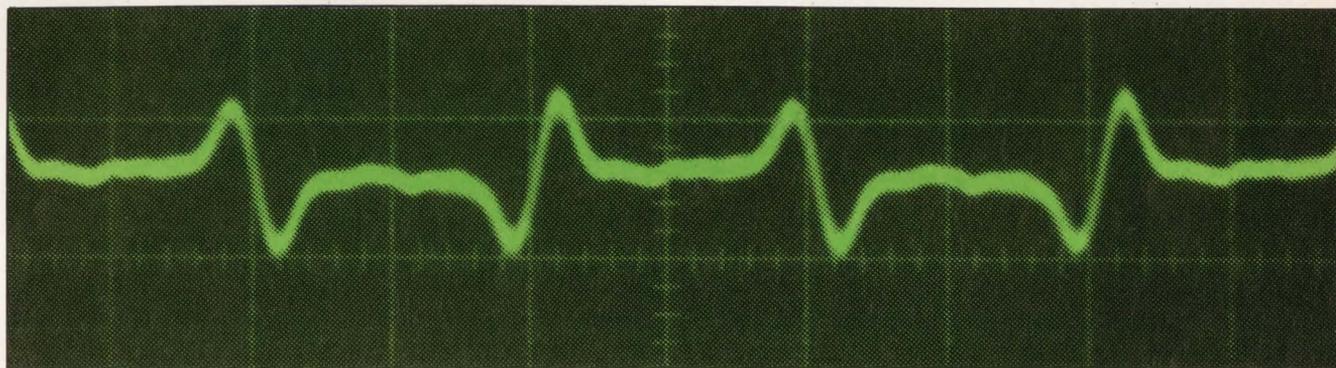


IMPROVING MASS STORAGE DATA INTEGRITY

Designers can do a lot to improve the reliability of the data being read from Winchester disks and other mass storage media. In fact, they should be doing more.



by Max Roth

Mass storage performance and reliability are tied to data integrity. However, even with error correction schemes, many recoverable errors impede throughput. Lost data affect performance even more dramatically. As 5¼" Winchester drive capacities and bit and track density increase, the read channel must be enhanced to maintain acceptable disk drive error rates.

Primarily, acceptable disk drive error rates are determined (somewhat arbitrarily) by disk technology rather than by predefined system requirements. Most 5¼" Winchester products support error rates of 1 in 10^{10} bits for recoverable (soft) errors. For nonrecoverable (hard) errors, 1 in 10^{12} bits is the norm. By definition, soft errors are recoverable by multiple read retries and therefore do not necessarily affect data reliability. Excessive

soft errors, however, may degrade throughput because of multiple read retries. Nonrecoverable hard errors occur primarily as a result of defects in a disk's recording surface.

Data integrity can be maintained two ways. First, the system designer can maintain low error rates by using error correction code (ECC). ECC solutions vary dramatically in terms of redundancy required, correction capability, and miscorrection errors. Both hard and soft errors are correctable with ECC. However, it should not be used to correct soft errors where a particular code's miscorrection is greater than the probability of recovery from multiple read retries. On the other hand, ECC does offer significant data reliability improvement for detecting and correcting hard errors. Historically, though, few system designers have accepted the responsibility of maintaining data integrity.

Relying on disk drive manufacturers to assure low error rates is the second and traditional way to maintain data integrity. But many in the computer industry believe that disk drive manufacturers are unnecessarily carrying the primary burden for maintaining disk system data reliability. By providing nearly perfect disk media and sophisticated read/write (R/W) channels, disk manufacturers have maintained very low disk drive error rates.

Max Roth is a member of the technical staff at Vertex Peripherals Corp, 2150 Bering Dr, San Jose, CA 95131, where he is responsible for read/write channel and spindle motor control. Mr Roth holds a BS from the Technion Institute of Technology, Haifa, Israel.

System reliability could easily be maintained if the controller manufacturers/designers improve their own error correction schemes by using greater redundancy. They could also improve reliability by selecting ECC codes that increase correction capability while minimizing miscorrection. This would even be true with a significant increase in raw disk drive error rates. In any case, system manufacturers should implement ECC to guard against error rate variations among disk drives.

While system designers continue to rely on disk drive manufacturers to provide low error rates, the manufacturers must maintain these low error rates by using nearly perfect recording surfaces and a capable read channel. It is worth noting that data separation is an important read channel function. This feature is implemented in the disk controller for present 5¼" Winchester disk drives. Because the data separator's performance can contribute to data errors, future interface standards are incorporating data separation in the disk drive electronics rather than in the controller. This places full responsibility for data integrity on the disk drive manufacturer.

Causes of data errors

A data bit from the disk drive must be detected within a fixed length of time, known as the decision window. The data transfer rate and the encoding scheme determine the duration of the decision window. In the case of standard 5¼" Winchester

drives, using a 5M-bps data transfer rate and modified frequency modulation (MFM) encoding, the decision window is 100 ns wide. Any bit undetected within this time period is considered an error.

In the ideal world where there is no interference and no noise, all data bits are centered within the decision window and no data errors occur. If intersymbol interference (pulse interaction) is considered, a bit shift or jitter occurs (no pulses are exactly centered but all are within the decision window). The amount of bit shift can be determined by superimposing adjacent bits. This amount is primarily a function of the encoding scheme used.

Noise and jitter intrinsic to the data separator also induce bit shift. Noise includes media head and electronic noise, overwrite modulation, and adjacent track interference caused by mispositioning of the read head in the data track. The result is a bit distribution superimposed on the bit shift due to pulse interaction. Mathematically, the bit shift (J = bit shift jitter) can be expressed as

$$J = X_1 + X_2 + X_3 + X_4$$

where

X_1 = zero-crossing jitter caused by noise

X_2 = intersymbol interference

X_3 = data separator window jitter

X_4 = mispositioning jitter

The nominal value of J is zero, as in the ideal case. Deviation from nominal is due to the variables X_1 , X_2 , X_3 , and X_4 . These variables have

Just what is ECC?

by Neal Glover
Data Systems Technology

Error correction code (ECC) allows data to be accurately reconstructed from encoded data that contain errors. These codes are used in computer and communication systems to increase data storage and transmission integrity.

An encoded data record typically contains a data segment that is identical to the raw data, and a redundant segment that is generated from the raw data by a generator polynomial. Dividing the encoded data by the generator polynomial is the first step in decoding. Data are assumed to be error free only if the remainder is zero. Usually, a nonzero remainder has enough information to allow the accurate reconstruction of the original data, provided that errors within the encoded record do not exceed the capability of the code being used.

ECC has guaranteed correction and detection abilities. Errors that exceed a code's guaranteed correction and detection capacities are subject to miscorrection.

Miscorrection probability is a function of record length, correction ability, and redundancy. If some error types are more probable than others, polynomial selection can influence miscorrection probability.

Historically, most magnetic disk controllers have used single-burst correcting codes. Most of these controllers use reread to recover from temporary errors, and error correction to recover from hard errors. This technique maintains data accuracy with codes that use only a modest amount of redundancy.

Some new controller designs use more redundancy to implement more powerful codes, including multiple-burst correcting codes. In addition, error-tolerant techniques are being used for address marks, sync marks, and header information. Multiple-burst correcting codes and other error-tolerant techniques will likely be widely used in future disk controllers due to new pushes in disk technology, new defect philosophies, and the lower cost of large scale integration.

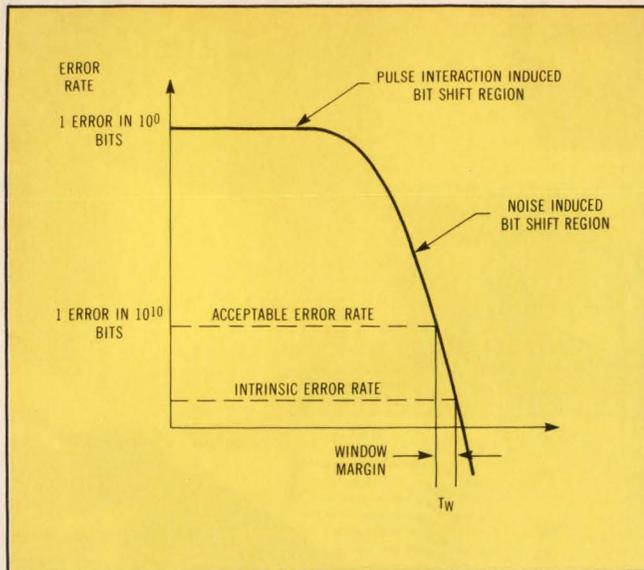


Fig 1 A marginal variable frequency oscillator (MVFO) plot is useful in determining error rate and margin information. The bit shift's relation to error rates and time is plotted with the goal of optimizing window margin (T_w) times.

mean value n (equal zero) and a symmetrical (with respect to n) probability density function (PDF). In most cases, the PDF can be approximated by a Gaussian distribution.

If the total bit shift jitter, J , is to satisfy a stringent requirement, the same holds for variables X_1 , X_2 , etc. A practical measure for the total jitter would be the variance of this function as derived from the variance of the variables.

Assuming that the input variables are not correlated and are normally distributed, the PDF for J will be a normal distribution. Since a read error occurs any time the zero-crossing or bit detection is outside the decision window, the probability of making an error is the area from a time (T_w) to infinity if the PDF of the zero-crossing jitter is Gaussian. (See the Bibliography for publications that support this assumption.)

If the decision window is varied, the number of errors will also vary. Errors can easily be plotted against variable decision window time. If the log of errors is plotted against a variable decision window, however, the resultant curve is referred to as a marginal variable frequency oscillator (MVFO) plot (Fig 1).

MVFO gets its name from the technique of varying the data separator window or decision window by marginal control of the data separator's clock frequency. Test equipment can plot MVFO curves for disk drives. The MVFO plot in Fig 1 shows the relative bit shift induced by both pulse interaction and noise.

Error rates are essentially 100% when the decision window is reduced to less than the bit shift time induced by pulse interaction. The intrinsic error rate is the best error rate achieved when using

the maximum available decision window. By reducing the decision window until the acceptable error rate is measured, a window margin factor can be obtained. Note that any chosen acceptable error rate must be higher than the read channel's intrinsic error rate. An MVFO plot can be generated for any disk drive. It is also a useful tool in characterizing error-rate performance.

If data rates are fixed, as in present 5¼" Winchesters, then higher track density achieves increased capacity. Thermal expansion between servo surface and data surfaces, bearing noise (referred to as "non-servoable" mechanical errors), and servo tracking errors limit track density. The density of recently introduced 5¼" Winchesters is approximately 1000 tpi, as opposed to 300 to 400 tpi in low capacity drives.

Ontrack and offtrack adjacent interference are primary contributors to the bit shift noise component. Fig 2 illustrates the ontrack and offtrack interference mechanism. Offtrack interference [Fig 2(a)] is due to signal noise picked up by the R/W head from the adjacent track as a result of head-to-track misposition. Ontrack interference [Fig 2(b)]

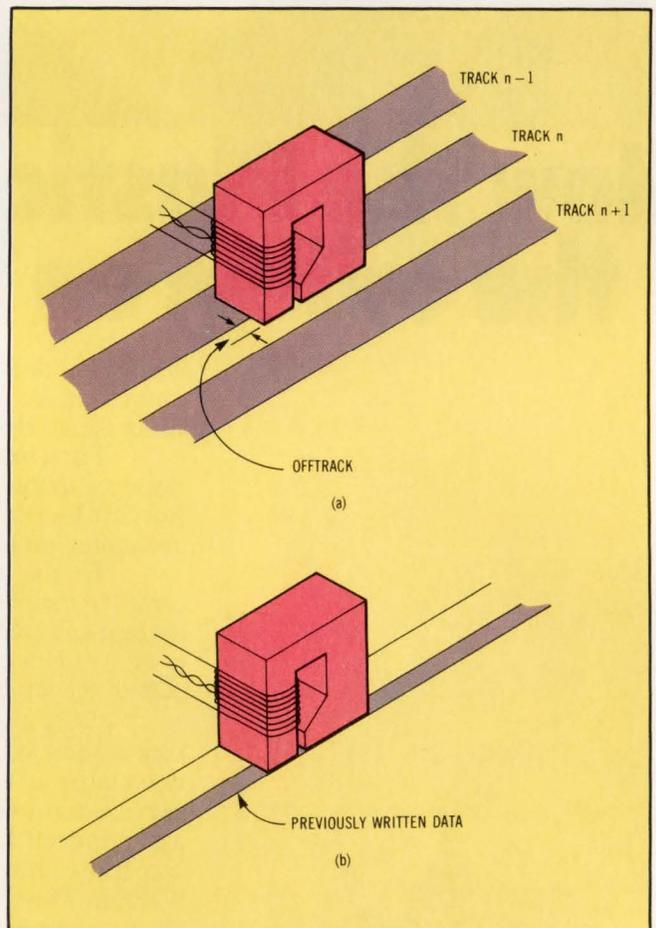


Fig 2 Two types of interference plague read heads. Offtrack interference (a) is caused by the pickup of adjacent track data. Ontrack interference (b) is caused by previously recorded information that was not precisely centered on track.

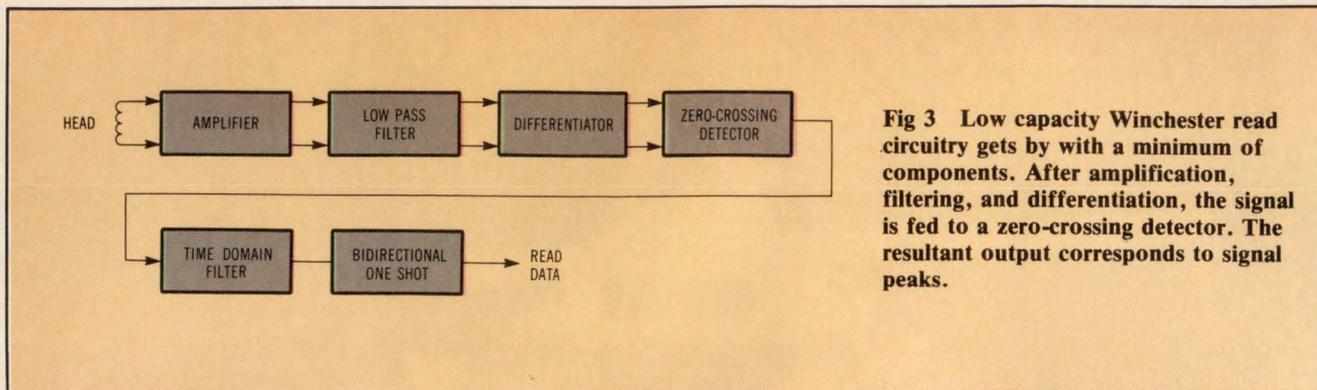


Fig 3 Low capacity Winchester read circuitry gets by with a minimum of components. After amplification, filtering, and differentiation, the signal is fed to a zero-crossing detector. The resultant output corresponds to signal peaks.

is due to noise picked up from previously written data on the same track, where the overwrite was not precisely centered. The read channel must accommodate approximately 12% misposition of head to track to maintain an acceptable error rate. It is up to the disk drive mechanical and servo systems to ensure that head to track misposition is maintained at less than 12% of track pitch.

A read channel rationale

In simple form, a drive's read channel amplifies the signal emanating from the magnetic read heads. However, read channels are becoming increasingly complex and vital to data reliability as 5¼" Winchesters grow in capacity and performance. Today's high capacity drives operate at up to 960 tpi. They use much less signal energy and have higher noise interference susceptibility. For example, the Vertex V100 series generates a signal of 0.4 to 0.8 mV peak to peak, versus 2 to 5 mV for a typical low capacity drive.

Obviously, read channel design becomes more critical in the higher capacity unit. In addition, many high capacity drives operate with varied encoding schemes, such as run length limited (RLL). These schemes are more efficient, but they also require greater bandwidth than MFM encoding. Unfortunately, greater bandwidth channels are more susceptible to noise.

Fig 3 illustrates a typical low capacity 5¼" Winchester read channel control circuit. The readback signal picked up by the head is amplified and then differentiated after it is filtered by the low pass filter. The zero-crossing detector output corresponds to the readback signal peaks. In the high resolution case, false detection can occur due to a droop in the differentiated signal. A time domain filter solves the problem by ignoring those pulses. However, this solution is sensitive to the encoding scheme used.

To achieve low error rates and allow flexible selection of recording codes, several circuit elements can be incorporated into modern Winchester hardware. Signal preamplifiers located on the actuator arm, near the R/W heads, improve the signal-to-noise ratio by amplifying the signal before additive

noise is introduced. Automatic gain control (AGC) compensates for variations in head signal amplitude. These variations are caused by normal variations in head flying height and differences in the efficiency of media and heads. Signal qualifier circuits ensure accurate data detection.

The signal qualifier circuit couples a peak detector with a zero-crossing detector. The peak detector accurately determines the position of the peaks in the time domain. By correlating those peaks with the zero-crossing detector output, the pulses that are caused by noise can be rejected. As the frequency span of the recording code is enlarged, the probability of false zero-crossing detection becomes higher.

While the MFM frequency ratio is 1:2, it can be as high as 1:4 in some RLL codes. Therefore, a peak detector scheme is less code dependent than a time domain filter. Fig 4 illustrates the qualifier circuit's effectiveness and shows an actual signal from the R/W head after preamplification. The differentiated signal shows significant deflections near the zero crossing.

If a zero-crossing detector circuit is used without signal qualification, the resultant data out are not usable. With signal qualification, nondata induced zero-crossing pulses are eliminated, leaving usable data out. Although these circuit elements are not generally found in low capacity 5¼" Winchesters,

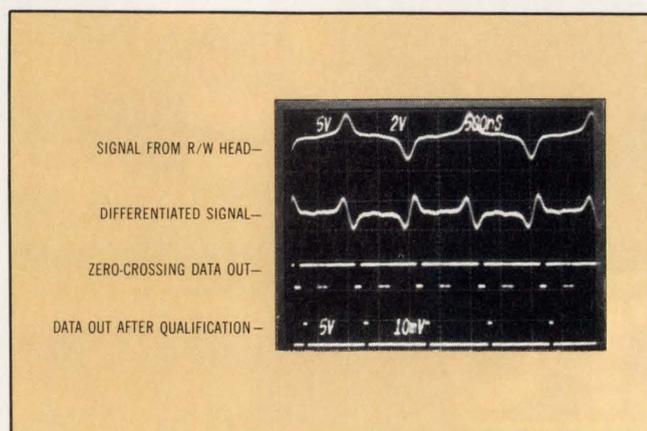


Fig 4 By eliminating false zero-crossing frequency components, reliable data are ensured.

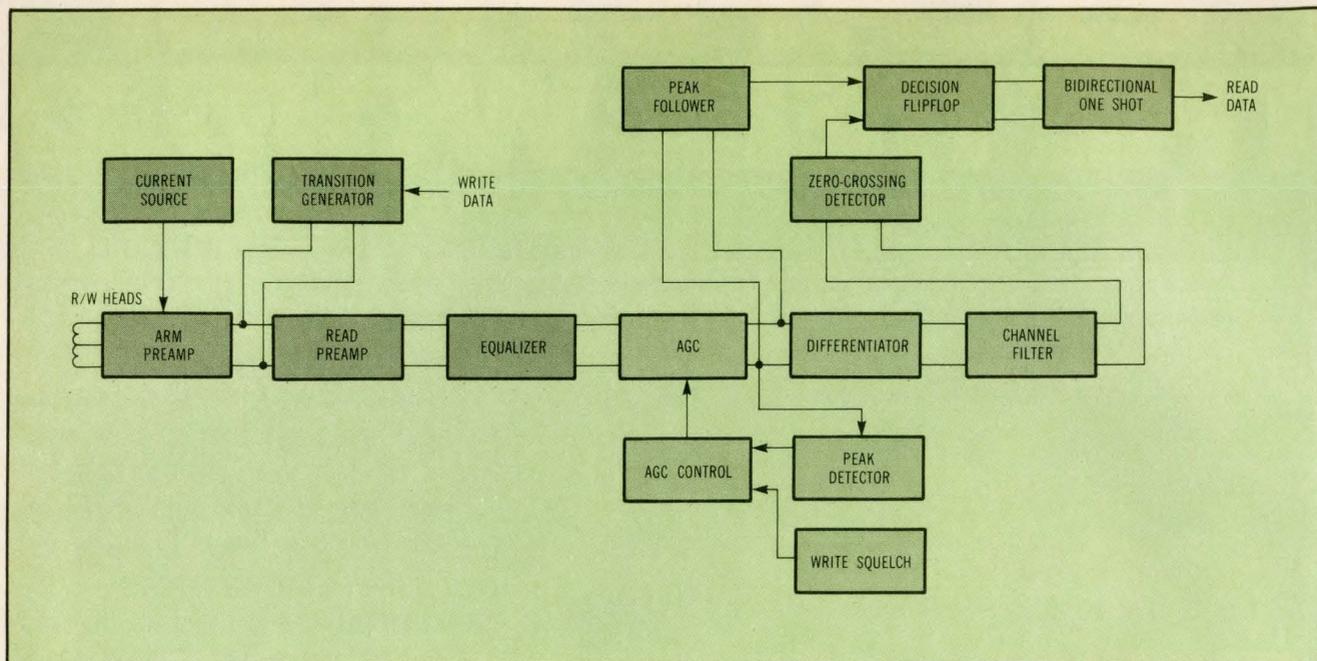


Fig 5 Vertex's high capacity Winchester R/W control takes an equalized signal output of the AGC amplifier and differentiates, filters, and digitizes it. A peak follower circuit processes this signal and, in conjunction with a zero-crossing detector, generates read data.

they are common in high capacity, high performance disk drives such as the IBM 3380.

The control block diagram of the V100 series read channel incorporates the elements described (Fig 5). In the V100 read channel, a preamp located close to the head itself amplifies the head signal. The amplified signal is applied to the read preamp via a balanced transmission line. After amplification, an equalizer modifies the signal spectrum. An AGC amplifier holds the signal at a constant amplitude to allow for head output variation and amplifier tolerances. AGC amp output is differentiated, filtered, and digitized by the zero-crossing detector. Then, after qualification by the peak follower, these data are available for reading.

Sophisticated read channel implementations will become increasingly important as 5¼" Winchester drive capacity and performance increase. Track and bit densities will continue to rise with improvements in magnetic head and media technologies. New encoding schemes and higher transfer rates will further enhance 5¼" drive performance. However, read channel implementation will make or break small drive data reliability in the near term, by amplifying and accurately differentiating the signal from the drive's R/W heads. The exceptions will occur when controller manufacturers take on an enlarged responsibility for implementing ECC.

In the foreseeable future, 5¼" drives will use enhanced equalization along with the ability to modify particular frequencies of the readback signal spectrum. This capability is essential to vertical recording schemes. Adaptive filtering will also be implemented using a microprocessor to change the

bandwidth of a channel as a function of the head's position on the disk. Finally, enhanced signal processing, using correlation techniques adapted from radar technology, will allow accurate processing of signals in even worst-case signal-to-noise conditions.

Bibliography

- A. S. Hoagland, *Digital Magnetic Recording*, Wiley & Sons, New York, NY, 1963.
- G. G. Hughes and R. K. Schmidt, "On Noise in Digital Recording," *IEEE Transactions on Magnetics*, vol 12, no 6, Nov 1976.
- E. R. Katz and T. G. Campbell, "Effect of Bitshift Distribution on Error Rate in Magnetic Recording," *IEEE Transactions on Magnetics*, vol 14, no 6, Sept 1978.
- H. E. Melbye and C. S. Chi, "Non-linearities in High Density Digital Recording," *IEEE Transactions on Magnetics*, vol 14, no 6, Sept 1978.

Please rate the value of this article to you by circling the appropriate number in the "Editorial Score Box" on the Inquiry Card.

High 701

Average 702

Low 703