RTMP Is eopy

ERROR RECOVERY DATA STORAGE SYSTEM AND METHOD

DESCRIPTION

The present invention relates to a data storage system having a wide write element and a narrow read element.

Direct access storage devices (DASD) or disk drives, store information on concentric tracks of a rotatable magnetic recording disk. A read/write head is moved from track to track in order to read and record the desired information. Typically, the head is positioned on an air bearing slider which flies above the surface of the disk as the disk rotates. In some disk drives, the slider rides on a liquid film or bearing on the disk. A head suspension assembly connects the slider to a rotary or linear actuator which moves the head to different radial positions along the disk.

During operation of the DASD units, it is important that the head stays positioned over the center of the track on which it is writing or reading data. This has become increasingly difficult as the tracks have become smaller in width and are spaced closer together in order to increase the overall data density of the DASD. The head may be off centre due to differences in thermal expansion of the different parts of the DASD, inaccuracies in the head position actuator, imperfect axial rotation of the spindle motor which rotates the disk, or any of a number of other factors.

One solution to the head mispositioning problem is to move the head in incremental steps across the track and try to determine the actual centre of the track. The references which discuss this include US4,485,418 issued November 27, 1984; European Patent Application 0220382 published June 5, 1987; IBM Technical Disclosure Bulletin, Vol. 35, No.4B, p.303, Sept. 1992; IBM Technical Disclosure Bulletin, Vol.29, No.2, p.586, July 1986.

Incomplete erasure of a previously recorded data track is another problem which occurs due to mispositioning of the head. When a new data field is written on a track, it overwrites the previous old data field. However, if the head is not properly centreed on the track, then a slice

of the old data may remain at the edge of the new data field. This may cause errors in reading.

One approach to solving this problem has been to provide erase bands between the data tracks. The head may have additional separate erase elements on either side of the write element in order to insure that the edge of each newly written field does not contain any data. Alternatively, the write element may be used to create the erase bands on either side of a newly written data field. Examples of this include US 4,858,048 issued August 15, 1989; US 4,771,346 issued September 13, 1988; US 4,644,421 issued February 17, 1987; Japanese application 03-157805 published July 5, 1991; German patent application DE2654577 published June 8, 1978; IBM Technical Disclosure Bulletin, Vol. 32, No. 10B, p. 90, March 1990; IBM Technical Disclosure Bulletin, Vol. 28, No. 1, p. 203, June 1985.

Magneto-resistive heads have recently been used in DASD devices. These heads contain separate write and read elements. The wide write element is a transducer element which is optimized for writing. The narrower read element is a magneto-resistive element which is optimized for reading. The magneto-resistive head allows smaller and more closely spaced data tracks to be written and read, thereby increasing the overall capacity of the DASD.

It has been discovered that because the magneto-resistive read element is much narrower than the write element, there is a good chance that a previously written field which has not been completely erased may be mistakenly read. The probability of reading an old data field is further increased by the need to hunt across the track width in some circumstances in order to successfully read a data field. This hunting only increases the chance that an old unerased data field at the edge of the track may be mistakenly read. The addition of erase bands on either side of the data track is not practical due to the close spacing of the tracks. The edge of a data track cannot be erased without also erasing a portion of a neighboring data track.

What is needed is a DASD system and method which does not need erase bands and which will not mistakenly read incompletely erased previous data fields.

Accordingly, the present invention provides

a method for reading data from a data track on a storage medium, the method comprising the steps of positioning a read head at a first location relative to the data track; reading the data stored at the first location; moving the read head to a second location relative to the data track, the second location being a predetermined distance D from the first location; reading the data stored at the second location; comparing the data stored at the first and second locations; concluding that the data read from the data track is correct if the comparision indicates that the data stored at the first and second locations are identical.

The present invention further provides a system for reading data from a data track on a storage medium, the system comprising means for positioning a read head at a first location relative to the data track; means for reading the data stored at the first location; means for moving the read head to a second location relative to the data track, the second location being a predetermined distance D from the first location; means for reading the data stored at the second location; means for comparing the data stored at the first and second locations; means for concluding that the data read from the data track is correct if the comparision indicates that the data stored at the first and second locations are identical.

In accordance with a preferred embodiment of the present invention, a data storage system comprises a rotatable magnetic disk having data tracks of width W, a magneto-resistive head attached to an actuator located over the data tracks, and a control unit connected to the head and actuator. If a data field is read with the head positioned within a certain small spacing of the centre of the data track, the data is deemed to be correct and is sent to the host system 70. If the data cannot be read within this narrow band at the track centre, the head is moved in incremental steps (both positive and negative) across the track. Once the data is successfully read, the head is moved a distance D across the track and the field is again read. If the data read at the two locations is equal, then the data is deemed to be correct and is sent to the host system 70. The distance D is selected to be larger than the width of any previously unerased old data. This ensures that the old data is never read by mistake. In the preferred embodiment $W/4 \leq D \leq W/2$ and D is optimally equal to approximately W/3.

This Sheet Has Been Replaced



An embodiment of the invention will now be described, by way of example only, with reference to the accompanying drawings in which:

Fig. 1 is a schematic diagram of a data storage system of the present invention;

Fig. 2 is a top view of a portion of the system of Fig. 1;

Fig. 3 is a schematic diagram of a magneto-resistive head of the present invention;

Fig. 4 is a schematic diagram of a data track of the system of Fig. 1;

Fig. 5 is a schematic diagram of three adjoining data tracks of the present invention and illustrates the reading and writing of a track;

Figs. 6, 7, 8, 9, 10, and 11 are flow chart diagrams which illustrate the operation of the system of the present invention.

Figs. 1 and 2 show schematic diagrams of a data storage system of the present invention which is designated by the general reference number 10. System 10 comprises a magnetic recording disk 12 which has a magnetic recording layer and which has a plurality of concentric data tracks 14. Disk 12 is attached to a spindle 16. Spindle 16 is connected to a spindle motor 18 which is mounted to a chassis 20.

A read/write head 30 is positioned over disk 12. Head 30 is attached to a suspension 32 which in turn is attached to an arm 34. Arm 34 is connected to a linear actuator 36 which moves head 30 to different radial positions over disk 12. Actuator 36 may be a voice coil motor or alternatively, the actuator 36 may be a rotary actuator. Actuator 36 is also mounted to chassis 20.

A controller unit 50 provides overall control to system 10. Controller unit 50 typically contains a central processing unit (CPU) 52 and a memory (MEM) 54 along with other digital circuitry as is known in the art. Memory 54 may comprise dynamic random access memory (DRAM) or read only memory (ROM) integrated circuits. Controller unit 50 is connected to an actuator control/drive unit 56 which in turn is connected to actuator 36. The controller unit 50 controls the movement of head 30 over disk 12. The controller unit 50 is connected to a read/write channel 58 which in turn is connected to the head 30. The controller unit 50 can send data to and receive data from the disk 12. Controller unit 50 is connected to a spindle control/drive unit 60 which in turn is

connected to spindle motor 18. The controller unit 50 also controls the rotation of disk 12. A host system 70 70, which is typically a computer system, is connected to the controller unit 50. System 70 may send digital data to controller unit 50 to be stored on disk 12, or may request that digital data be retrieved from the disk. The basic operation of DASD units is well known in the art and is described in more detail in Magnetic Recording Handbook, C. Dennis Mee and Eric D. Daniel, McGraw & Hill Book Company, 1990.

Fig. 3 shows a schematic diagram of head 30 of Fig. 1. Head 30 comprises a slider element 100. Slider element 100 has an air bearing surface 102 which faces the surface of disk 12. The surface 102 is designed to create an air bearing between itself and the rotating disk 12 to allow the slider 100 to fly above the surface of the disk 12. A read/write element 106 is attached to a side surface of slider 100. The element 106 comprises an inductive write element 110 and a magneto-resistive read element 112. The write element 110 is much wider than the read element 112. The write element 110 and the read element 112 are connected to the read/write channel 58. The read/write element 106 may be similar to that described in US patent 3,908,194 assigned to International Business Machines Corporation. The width (W) of the data tracks 14 is approximately equal to the width of the write element 110 and is approximately equal to 300 microinches in a preferred embodiment.

Fig. 4 shows a schematic diagram of a data track 14 of disk 12. Track 14 comprises blocks of magnetically recorded data. The blocks include a home address 150, a record zero block 152, and a plurality of record blocks 154. Each block is separated by a gap 156. The home address 150 contains information on each track and the record zero block 152 contains information on the number and length of the records that follow.

Each record block 154 comprises a count field 170, an optional key field 172 and a data field 174. The count field 170 contains data which identifies the particular record and the lengths of the fields, the key field 172 contains data on the type of information contained in the following data field, and the data field 174 contains the actual data which is to be stored. All of these fields are separated by gaps 176. Each field also contains an error correction code (ECC) 180 and a cyclical redundancy check (CRC) code 182. The codes are used by the

controller unit 50 to detect errors in the preceding data in the field and provide error correction. A more detailed description of such coding is given in Magnetic Recording Handbook, supra, Part 2, Chapter 8, page 1182.

Fig. 5 shows a schematic diagram of three adjoining data tracks of the system 10 and illustrates the problem of reading old data by mistake. A centre track N is located between tracks N-1 and N+1. A dotted line 200 represents the ideal centre of track N. First, an old data field 202 is written on track N. The controller unit 50 sends digital data to the read/write channel 58 which converts it to an analog signal which drives the write element 110. The write element 110 generates magnetic fields which record the data as field 202 on the magnetic disk 12. Due to mispositioning of the head, the field 202 is written slightly to the right of the track centre line 200. Next, the old data field is to be replaced with a new data field 204. Field 204 is written on track N. However, the head is offset to the left of the centre line 200. A significant portionof the old data field 202 remains unerased at the edge of the new data field 204. Due to the narrow width of the read element 112, there is a possibility that this old field 202 may be read and mistaken for the new field 204. This would result in faulty data being sent to the host system 70 70. The present invention solves this problem by verifying that the data read is not from the old field.

Figs. 6-11 are flowchart diagrams which illustrate the operation and method of system 10 of the present invention and is designated by the general reference number 300. This method is implemented as stored digital data in the controller unit 50 in the preferred embodiment. Alternatively, the method could be implemented as discrete electronic circuits. At step 302, system 10 reads a field of a data track of disk 12. The current position of the head for the first read attempt is known as the zero offset position. In other words, the controller unit 50 moves the head 30 to a position where the controller unit 50 estimates that the centre of the desired track is located. This zero offset position may or may not be at the actual track centre 200. The read element 112 senses the magnetically recorded data on the data track 14 and sends an analog signal to the read/write channel 58. The read/write channel 58 converts the analog signal to a digital signal which is then received by the controller unit 50.

At step 304, it is determined whether or not the read was successful. This is determined by the controller unit 50 by checking the ECC and CRC codes of the field by methods which are well known in the art. If the read was successful, then at step 306 the data which was read from the disk is sent to the host system 70 70.

If a successful read is not achieved, then the field is reread with the head 30 positioned at offset zero at step 308 and is checked for a successful read at step 310. If the read is successful, then the data is sent at step 306. If the read is not successful, then the length of the record is checked at step 312. The controller unit 50 has stored data on the length of the count fields which are all the same length. Once a count field is read, the controller unit 50 will have information on the lengths of the following key and data fields. If the particular field length of the field which is to be read is less than or equal to L1 bytes, the field is a short field, then at step 314 L1 values are set. If the field length is greater than L1, then the operation goes to step 316. At step 316 it is determined if the field length is greater than or equal to L3 bytes. If so, the field is a long field and the operation goes to step 318, where L3 values are set and, if not, the field is a medium field and the operation goes to step 320 where L2 values are set.

The reason for checking the length is that if the record is very short, there is a possibility that the error correction code methods may be able to correct the data even though the data was read from an undesired unerased old record. For a very short field, a lower level of error correction will be used. Also, the distance D that the head moves across the track will vary depending on the field length. In a preferred embodiment L1 equals 2048 bytes and L3 equals 6468 bytes. The following table shows the values which will be set.

Field Length	ECC Check	imax	S(1)	D(1)	S(2)	D(2)	S(3)	D(3)	DF
Short	Type 1	3	15	-	30	110	40	100	100
Medium	Type 2	3	15	-	30	90	40	75	100
Long	Type 2	3	15	-	30	-	40	-	-

The ECC type varies depending upon the record length. For short records a lower level type 1 ECC is used. For long records a higher level type 2 ECC is used. These high and low level ECC are described in US Patent 4,706,250 issued November 10, 1987 to A.M. Patel. The imax value is an integer value which governs the maximum number of offsets positions in a positive or negative direction in which the head can be moved across a track. The D(1), D(2), D(3) and DF values are given in microinches and represent the distance apart at which two successful reads must be made in order to confirm the data. The values S(1), S(2), S(3), give the distances, in terms of microinches, that the head will be moved away from the zero offset value. The use of these values will become more apparent from the description below.

At step 322 the value i is set to one. The value i indexes the offset position to which the head is moved relative to the zero offset. At a step 324, the head 30 is moved to offset S(i); S(1) in this first case. This is the distance which the head 30 is moved in a positive direction relative to the zero offset and this distance is obtained from the above table. The field is read at step 326 and at step 328 it is determined if the read was successful. If the read was successful, then the data is sent at a step 330. If not, at step 332 the head is moved to offset value negative S(i); negative S(1) in this first case. The field is read at step 334 and is checked at step 336. If the read is successful, then the data is sent at step 330 and if not, then the operation moves to step A of Fig. 7.

The overall operation of the steps in Fig. 6 may now be summarized assuming the field was not able to be read at the zero offset position.

If the record can be read at a single small step S(1) on either side of offset zero, then the data is deemed to be correct and is sent out even though there was only one successful read. However, if the field must be read at a distance greater than S(1) distance away from offset zero, then an additional confirming read at a different head position is needed.

In Fig. 7 step A leads to step 400. At step 400, i is incremented by one. At step 402, i is compared to imax. If i is greater than imax, then the maximum number of offsets positions have already been taken and the operation moves to step X of Fig. 10. If i is less than or equal to imax, then at step 404 the head 30 is moved to offset to S(i). At step 406 and 408, the field is read and checked respectfully. If the read was successful, then at step 410 the field length is checked. If the field length is greater than or equal to L3, then the operation goes to step 412 and the data is sent to the host. If the field length is less than L3, then the next step is B of Fig. 8. If the read at step 408 was not successful, then the head 30 is moved to negative S(i) at step 414. The field is read and checked at step 416 and 418, respectfully. If the read is not successful, then the operation returns to step 400. If the read is successful, then at step 420 the length of the field is checked. If the field length is greater than or equal to L3, then the data is sent to the host at step 422. If the field length is less than L3, then the next step is C of Fig. 9.

The overall operation of Fig. 7 may now be summarized. The head 30 is alternatively moved the same positive and negative distance away from the zero offset position across the track in order to try and read the record. The step distance is progressively increased. If a successful read is achieved at a positive offset position and the field length is less than L3, the operation goes to step B for a confirming read. If a successful read is achieved at a negative offset position and the field length is less than L3, the operation goes to step C for a confirming read. If the field length is greater than or equal to L3, and a successful read has been made, then the chance that the data read is from an incompletely erased old record is very small and the data is deemed to be correct.

Alternatively, fields greater or equal to L3 may also have a confirming read. In this case steps 410, 412, 420 and 422 may be eliminated and steps 408 and 418, would lead directly to steps B and C.

respectively, if a successful read was made. Appropriate values for D (such as 100 microinches) for L3 could be added to the table above.

Fig. 8 shows that step B leads to step 500. At step 500, the CRC value of the record which has just been read at offset S(i) is stored as a value CRC[S(i)]. At step 502 the head 30 is moved to offset S(i)+D(i). D(i) is a distance selected to be greater than that of the width any unerased previous old record. In a preferred embodiment, $W/4 \leq D(i) \leq W/2$ and D is equal to approximately W/3, where W is equal to the width of the track. At steps 504 and 506 the field is respectfully read and checked. If the field was not read successfully, then the operation goes to a step D of Fig. 7. If the record is successfully read, then at step 510 the CRC value of the record read at S(i)+D(i) is stored as a value CRC[.S(i)+D(i)]. At step 512 the CRC[S(i)+D(i)] value is compared to the CRC[S(i)] value. If the values are equal, then it is determined that this is a new record, not an old unerased record, and the data is sent at step 514.

Fig. 9 shows step C leads to step 700. At step 700, the CRC value of the record which has just been read at offset -S(i) is stored as a value CRC[-S(i)]. At step 702 the head 30 is moved to offset -S(i)-D(i). At steps 704 and 706 the field is respectfully read and checked. If the field was not read successfully read, the operation goes to step E of Fig. 7. If the field is successfully read, then at step 710 the CRC value of the record read at -S(i)-D(i) is stored as value CRC[-S(i)-D(i)]. At step 712 the value CRC[-S(i)-D(i)] is compared to the value CRC[S(i)]. If the values are equal, then it is determined that this is a new record and is not an old unerased record and the data is sent at step 714.

Therefore, if a first successful read is achieved a certain distance S(i) beyond the centre of a track, a second verifying read of the same data is attempted at least a distance D(i) away from the first read.

In Fig. 10, step X leads to step 800. The head 30 is moved to offset DF/2. The field is read and checked at steps 802 and 804, respectively. If the record is not read successfully, then the operation goes to step Z of Fig. 11. If the read is successful, then the CRC value is stored as value CRC[DF/2] at step 806. At step 808 the head 30 is moved to offset -DF/2. At steps 810 and 812, the field is read and

checked, respectively. If the field is not read successfully, then the operation goes to step Z of Fig. 11. If the field is successfully read, then the CRC value is stored as value CRC[-DF/2] at step 814. At step 816, the value CRC[DF/2] is compared to the value CRC[-DF/2]. If the values are equal then the data is sent at step 818, and if not, the operation goes to step Z.

The operation illustrated in Fig. 10 may now be understood. If i>imax, then the operation starting at step X is used. An attempt is made to successfully read the field at two positions which are located a distance DF (D Final) apart. This is done by moving the head first to a position positive DF/2 and then to a position negative DF/2 from the zero offset position. If the confirming read cannot be made, then the operation goes to step Z of Fig. 11.

Fig. 11 shows step Z leading to a step 900 at which the head 30 is moved back to the zero offset position. At steps 902 and 904 the field is read and checked, respectively. If there is a successful read, then the data is sent at step 906. If not, then an error message is sent to the host system 70 at step 908. These last steps of Fig. 11 are implemented when no successful read and verification of the record has been achieved and the host system 70 must be notified that the desired record cannot be accessed.

In summary, the present invention discloses a data storage system which verifies that the data read is correct by comparing the data read at two positions a distance D apart. In a preferred embodiment, this is done by comparing the CRC values of both data reads. However, the invention could also be practiced by comparing other portions of the fields, such as the error correction code. Alternatively, the fields may be compared in their entirety.

Other embodiments of the present invention are possible. For example, instead of starting at offset zero, the head could be positioned at an outer position initially and the moved incrementally closer to the zero offset, alternating from side to side. Also, the exact value of D may be adjusted depending upon the disk size or radial position of the desired track. Additional re-read steps may also be added so that a read is attempted multiple times at each head position.

Although the present invention uses a variable data field length, the invention can also be applied to fixed block systems. These systems are known in the art and have data fields which are sectored or fixed in length. In such a case, the determination of the data field length may not be necessary.

* C L M S *

THIS SHEET HAS BEEN REPLACED

CLAIMS

1. A method for reading data from a data track on a storage medium, the method comprising the steps of

positioning a read head at a first location relative to the data track;

reading the data stored at the first location;

moving the read head to a second location relative to the data track, the second location being a predetermined distance D from the first location;

reading the data stored at the second location;

comparing the data stored at the first and second locations;

concluding that the data read from the data track is correct if the comparision indicates that the data stored at the first and second locations are identical.

2. A method as claimed in claim 1 wherein $W/4 \leq D \leq W/2$, where W is the width of the data track.

3. A method as claimed in claim 2 wherein D is substantially equal to W/3, where W is the width of the data track

4. A method as claimed in any preceding claim, wherein the step of comparing comprises comparing portions of the data stored at the first and second locations.

5. A method as claimed in any preceding claim, wherein the step of comparing comprises comparing the respective cyclic redundancy check codes of the data stored at the first and second locations.

6. A method as claimed in any preceding claim, wherein the second location is disposed a distance from the first location according to the length of the data to be read.

7. A system for reading data from a data track on a storage medium, the system comprising

This Sheet Has Been Replaced

* C L M S *

THIS SHEET HAS BEEN REPLACED

means for positioning a read head at a first location relative to the data track;

means for reading the data stored at the first location;

means for moving the read head to a second location relative to the data track, the second location being a predetermined distance D from the first location;

means for reading the data stored at the second location;

means for comparing the data stored at the first and second locations;

means for concluding that the data read from the data track is correct if the comparision indicates that the data stored at the first and second locations are identical.

8. A system as claimed in claim 7, wherein $W/4 \leq D \leq W/2$, where W is the width of the data track.

9. A system as claimed in claim 7, wherein D is substantially equal to W/3, where W is the width of the data track.

10. A system as claimed in any of claims 7 to 9, wherein the means for comparing comprises means for comparing portions of the data stored at the first and second locations.

11. A system as claimed in any of claims 7 to 10, wherein the means for comparing comprises means for comparing the respective cyclic redundancy check codes of the data stored at the first and second locations.

12. A system as claimed in any of claims 7 to 11, wherein the second location is disposed a distance from the first location according to the length of the data to be read.

13. A system as claimed in any of claims 7 to 12, wherein the means for moving causes the read element to be moved in incremental distances away from the first location in order to obtain a first succesful read.

This Sheet Has Been Replaced

* C L M S *

THIS SHEET HAS BEEN REPLACED

14. A system as claimed in claim 13, wherein the distance D varies according to the number of incremental distances which the head element has been moved.

15. A system as claimed in any of claims 7 to 14, further comprising a write element having a width greater than the means for reading the data.

16 A system as claimed in any preceding claim, wherein the means for reading the data is a magneto-resistive element.

This Sheet Has Been Replaced