DEVELOPMENT STANDARD

QIC-112
Revision A
27 Oct 87

ECC FORMAT

**Quarter-Inch
Cartridge
Drive Standards, Inc.**

311 East Carrillo Street
Santa Barbara, California 93101
Telephone (805) 963-3853
Fax (805) 962-1541
www.qic.org

(See important notices on the following page)

QIC 112
Revision A
Octoer 27, 1987

ECC Format

ERROR CONTROL

The purpose of this document is to specify how error correction may be added to tapes written on drives compatible with the QIC 24/120/150 formats in order to allow reliable data interchange between drives from different vendors and/or different hosts. The ECC format described herein shall be applied on top of any existing physical format(s), but below the host interchange forrmat level (e.g., QIC 87-22) This ECC format is *not* intended to be part of the physical format; i.e., the drive is not required to perform any ECC operations, which operations may be entirely implemented by the host. This document does not include a mechanism for automatic ECC format identification.

Hard errors are caused by media defects in the recording area. Two strategies shall be used to defeat hard errors:

(1) elimination of blocks with defects by performing read-after-write checking and rewriting all incorrect blocks, and

(2) correction of blocks with hard errors after use.

1. Elimination of Blocks With Defects

Blocks containing initial media defects shall be excluded from use via rewriting as required in the physical format specification (e.g., QIC-150).

2. Error Correction

The cartridge data recording operation shall include two 512-byte blocks of ECC information in each frame which shall be used during the data reading operation to reconstruct blocks in error. Within each frame, the error-correction algorithm shall correct up to two CRC errors (pointers) or a single CRC failure (i.e., an error undetected by the CRC). The error correction algorithm can also detect, but not correct, almost all other errors in each data frame with an extremely high reliability.

2.1 Frame Format

A frame is the smallest unit which can be read from or written to the tape. A frame consists of 16K bytes arranged in 32 blocks. The bytes in a frame are considered to be arranged in a 32 x 512 matrix. The parity rows (i.e., blocks 30 and 31) shall be chosen so that each column of the matrix forms an independent Reed-Solomon codeword of redundancy two, with 8-bit characters, as shown in figure 1. Data shall be written on the tape row by row, starting with row 0, and within each row (i.e., block) the bytes shall be written starting with column 0.

Frame boundaries on the tape shall be derived implicitly from the physical block numbers. The first block on the tape shall be a control block [Note: possibly repeated several times for error tolerance]. All other blocks on the tape shall be data blocks, and block N shall be handled by the ECC as logical block number ((N-1) MOD 32) in frame number ((N-1) DIV 32). It is not possible to have shortened frames in this format.

```
                            COLUMNS
         0     1     2        ...         510   511
      -------------------------------------------------
   0 |     |     |     |              |     |     | CRC |  D
      -----+-----+-----+--------------+-----+-----+-----   A
   1 |     |     |     |              |     |     | CRC |  T
      -----+-----+-----+--------------+-----+-----+-----   A
     |     |     |     |              |     |     |     |
     |     |     |     |              |     |     |     |  R
     |     |     |     |              |     |     |     |  O
      -----+-----+-----+--------------+-----+-----+-----   W
  28 |     |     |     |              |     |     | CRC |  S
      -----+-----+-----+--------------+-----+-----+-----
  29 |     |     |     |              |     |     | CRC |
      =================================================
  30 |     |     |     |              |     |     | CRC |  PARITY
      -----+-----+-----+--------------+-----+-----+-----
  31 |     |     |     |              |     |     | CRC |  ROWS
      -----+-----+-----+--------------+-----+-----+-----
```

Figure 1. ECC Frame Format

2.2 Field Representation

GF(256) is the field with 256 elements. Each field element a has the form

$$a = a_7x^7 + a_6x^6 + a_5x^5 + a_4x^4 + a_3x^3 + a_2x^2 + a_1x + a_0 .$$

where each a_i is either 0 or 1. A field element a shall be represented by a byte as shown in figure 2 below.

```
  MSB                                                   LSB
 ---------------------------------------------------------
|      |      |      |      |      |      |      |      |
|  a7  |  a6  |  a5  |  a4  |  a3  |  a2  |  a1  |  a0  |
|      |      |      |      |      |      |      |      |
 ---------------------------------------------------------
```

Figure 2. Bit Numbering Convention

Field math operations (addition, multiplication, division) are defined to be polynomial math modulo an irreducible binary polynomial of degree eight, f(x), where binary addition is the logical exclusive-or operation and binary multiplication is the AND operation. The irreducible polynomial used to generate the field GF(256) shall be:

$$f(x) = x^8 + x^7 + x^2 + x + 1 .$$

2.3 Code Generator Polynomial

Let r be a root of f(x). The generator polynomial for the Reed-Solomon code shall be

$$g(x) = \prod_{i=0}^{1} (x - r^i) = x^2 + (r+1)\,x + r = x^2 + 3_{hex}\,x + 2_{hex}$$

Encoding a frame shall be accomplished as follows. In a full frame, each column contains data bytes do to d3l, numbered as in figure 1. The parity bytes d3O to d3l in each column shall be chosen so that the polynomial

$$d(x) = \sum_{i=O}^{31} d_{3l-i} \; x^{i}$$

is divisible by g(x), using polynomial division over GF(256).

2.4 Example Codewords

The folIowing columns of bytes are codewords for the polynomials defined in the preceding sections, using hex notation for the field elements.

Row #						
0:	00	00	00	00	00	00
…						
27:	00	00	00	00	00	00
28:	00	00	01	02	04	07
29:	01	10	00	04	08	0c
30:	03	30	07	02	04	01
31:	02	20	06	04	08	0A

APPENDIX A. ECC Performance Analysis

In this section, the improvement in bit error rate due to the ECC will be informally examined for a QIC- 150 product, ignoring the possibility that a single dropout can cause two block errors [Note: this may not be negligible, but it is difficult to analyze without better error modeling.]. Let q be the input bit error event rate coming off of the tape. If L is the block size in bytes (here, L = 512), then the probability that a block is in error (i.e., the input block error rate) is

$$P = 1 - (1-q)^{8L} \approx 8Lq \; .$$

The probability that exactly k block errors occur in 32 blocks is:

$$p(k) = C(32,k)\,(1 - P)^{32-k}\,P^k \approx C(32,k)\,P^k \, ,$$

where C(n,k) = n!/(k!(n-k)!) is the combinatorial "choose" function.

Since the Reed-Solomon ECC can correct up to 2 CRC errors, the probability that a given frame has an unrecoverable error due to CRC errors is simply the probability that three or more block CRC errors occur. To first order, this probability is simply the probability that exactly three such errors occur. That is, the probability of an uncorrectable error pattern (frame error) due to CRC error is

$$Ql \approx C(32,3)\,P^3$$

The probability of block CRC failure (using a 16-bit CRC) is no worse than P/65536, but in practice the GCR code aids considerably in detecting CRC failures. A frame error could occur with two CRC failures or one CRC error and one failure. The probabilities of these two error mechanisms are

$$Q2 \approx C(32,2)\,(P/65536)^2 \text{ and } Q3 \approx 2\,C(32,2)\,(P^2/65536).$$

The overall error rate is (again to first order) Q ≈ Ql + Q2 + Q3.

As an example, suppose q = 1E-8 (obviously a worst-case end of life condition).
Then, P ≈ 4E-5, and

Ql ≈ 4E-10, Q2 ≈ 2E-16, Q3 ≈ 3E-11

so the dominating term is Ql, as we would expect. However, notice that if the ECC did not handle the case of a single CRC failure, the frame error rate would be given (pessimistically) by C(32,I) P/65536 ≈ 2E-8. In other words, it is important that the CRC failure mode not be ignored, unless the GCR code will detect at least 99.5% of what would otherwise be CRC failures. Thus, depending on the amount of GCR detection available, a host may elect to ignore the single CRC failure and perform no ECC operations unless a CRC error is reported.

Given the above, Q ≈ Q 1 ≈ 4E- 10, and the output bit error rate is

$$Po \approx Q/(32x8xL) \approx 3E\text{-}15,$$

From a user's standpoint, a much more important figure is the tape error rate. Since a tape with 135 MB of data has roughly 8600 frames, the tape error rate is 2xC(8600,1)xQ = 8E-6. In other words, in only one of 125,000 full reads of a tape would an uncorrectable error occur. The corresponding number without ECC is roughly 10, so on the average ten uncorrectable errors would occur on each full tape read without ECC, given this raw error rate.