Design and Implementation of ECC Module Based on BCH Code in SSD

Yifei Niu, Songyan Liu, Yanlin Chen, Xiaowen Wang and Huan Liu

Abstract—Error Correction Code (ECC) is an effective method to ensure the correctness of data stored in Solid State Disk (SSD). At present, Bose, Chaudhuri, Hocquenghem (BCH) code is the most widely used in ECC. However, how to improve performance of encoding and decoding has always been a problem. A new pipeline operation is proposed by this paper to improve the performance of ECC module based on BCH code in SSD. Pipeline operation is used in I/O transmission, and the ECC process submerges in pipeline. The 64-bit parallel architecture is adopted to complete the encoding. The 3-stage pipeline structure is adopted in decoding. Therefore, the efficiency of encoding and decoding is improved and the latency is reduced. ECC module is able to support multiple error correction capabilities. The capability is configurable to 24-bit, 40-bit, and 56-bit for 512Bytes. The data throughput of the ECC module can reach 7.68Gbps.

Keywords—BCH code, Decoding, ECC, Encoding, SSD.

I. INTRODUCTION

As the data processing speed of the Central Processing Unit (CPU) is getting more and more fast, the performance gap between the traditional Hard Disk Drive (HDD) and CPU is also increasing [1]. SSD based on NAND Flash has characteristics of fast reading and writing speed as well as low power consumption compared with HDD, which replaces HDD gradually [2]. Flash bit in SSD may flip due to program/erase (P/E) cycles, read disturb, program disturb, data retention, etc., so that the data stored in flash will be incorrect [3]. Similar to magnetic storage and optical storage, flash also requires error control technology to ensure the integrity and reliability of data [4]. To further increase storage density, MLC (Multi-Level Cell) and TLC (Triple-Level Cell) are becoming mainstream [5].

The earliest ECC code used in NAND Flash is Hamming code. Its principle and implementation are simple, but only single bit error can be corrected. As the Hamming code is not enough to process a higher bit error rate, the industry began to introduce RS (Reed-Solomon) code as ECC code. The RS code has strong ability to correct burst errors, whereas NAND Flash is more likely to occur random independent errors. However, the errors within NAND flash are properly solved with the BCH code [6]. BCH code is the most common and wide ECC code for NAND Flash [7].

BCH code is an important type of cyclic code. Previous papers accomplished plentiful research on NAND Flash. Arul K. Subbiah [8] presents a novel method to reduce the area of the BCH multimode encoder based on a re-encoding scheme. Ping Chen [9] implements a high performance low complexity ECC module circuit for SSD controllers by BCH code. Sheyang Ning [10] proposes an advanced bit flip scheme to correct major program errors. Byeonggil Park [11] presents a novel folding technique for BCH decoders, the regularly structured GF multiplier which is efficiently folded to reduce the complexity and the critical delay. In previous papers, almost all of the authors implement by verilog language or other applications, and the parallel bits are 8, 16 or 32 bit. This article implements ECC module based on FPGA (Field Programmable Gate Array) and tests in practical SSD. The parallel bits are 64. The reading and writing pipeline are adopted in I/O transmission, and ECC is completed in I/O transmission, so as to further improve the performance of the system.

The paper is organized as follows. Section II presents BCH encoding and decoding design within ECC module. Section III explains the hardware implementation. Section IV depict the test results. Finally, section V is the conclusion.

II. BCH ENCODING AND DECODING

A. BCH Encoding

For a BCH (n, k, t) code, where the code length, message length and maximum error correction number are denoted by n, k, and t, the BCH code is calculated by equation (1) and (2) in Galois Filed(GF):

\[ c(x) = x^{n-k}m(x) + r(x) \]  

\[ r(x) = x^{n-k}m(x) \mod g(x) \]

Where \( m(x) = m_{k-1}x^{k-1} + m_{k-2}x^{k-2} + \ldots + m_1x + m_0 \) stand message polynomial, \( c(x) \) is code polynomial, \( g(x) \) means generator polynomial and \( r(x) \) is the remainder polynomial of BCH code.
Since the BCH code is cycle code, the Linear Feedback Shift Register (LFSR) can be used to implement the BCH encoder. The circuit of serial LFSR is as Fig. 1, where \(g_0 = 1, g_{n-k} = 1\); \(g_i = 1(1 \leq i \leq n-k-1)\) presents the line is connect, \(g_i = 0\) means the line is disconnected. Initially, the switch \(K_1\) is closed, \(K_2\) is off, and when the LFSR outputs the check bit, \(K_1\) is switched off and \(K_2\) is closed. The information bits are input from the right edge, which is equivalent to the dividend multiplied by \(x^{n-k}\) and then divided by \(g(x)\). The hardware implementation of traditional serial encoder is relatively simple, but the clock cycle of encoding is long, which cannot meet the application requirements of NAND Flash, so that we designed the encoder with parallel structure.

![Fig. 1. Circuit of serial LFSR](image)

In this paper, the unit of checking data is sector, the system provides several different ECC abilities, including 24bit, 40bit and 56bit. The length of checksum for different ECC ability is shown in Table 1.

<table>
<thead>
<tr>
<th>Error Correction (bits)</th>
<th>Galois field GF(2^m)</th>
<th>Length of checksum (Byte)</th>
<th>Coding efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>24</td>
<td>GF(2^5)</td>
<td>39</td>
<td>0.93</td>
</tr>
<tr>
<td>40</td>
<td>GF(2^13)</td>
<td>65</td>
<td>0.89</td>
</tr>
<tr>
<td>56</td>
<td>GF(2^13)</td>
<td>91</td>
<td>0.85</td>
</tr>
</tbody>
</table>

### B. BCH Decoding

For the transmission code \(c(x)\), it is defined that the received code is \(R(x)\). If the original code is interfered during transmission, an error pattern of the following equation (3) will be generated:

\[
e(x) = e_0 + e_1 x + ... + e_{n-1} x^{n-1}
\]

(3)

Therefore, the receiving polynomial \(R(x)\) as equation(4) can be represented by the transmission code and error pattern.

\[
R(x) = e(x) + c(x)
\]

(4)

The decoding of BCH code is as follows:

1. Calculating syndrome polynomial \(S_1, S_2, ..., S_{2t}\) by \(R(x)\), then judging whether the received code has an error according to the value of syndrome. The circuit of calculating syndrome is as Fig. 2:

![Fig. 2. The circuit of calculating syndromes](image)

(2) Calculating the coefficient of error location polynomial \(\sigma(x)\) according to syndrome;

(3) Calculating the roots of error location polynomial \(\sigma(x)\) and then getting the error locations and correct errors.

Syndrome polynomial is calculated by equation (5):

\[
S = (S_1, S_2, ..., S_{2t}) = R \times H^T
\]

(5)

It is provided that \(e(x)\) has \(v\) errors at positions \(j_1, j_2, ..., j_v\), where \(0 \leq j_1 < j_2 < ... < j_v < n\), and then we can get \(e(x)\) as equation (6):

\[
e(x) = x^{j_1} + x^{j_2} + ... + x^{j_v}
\]

(6)

For calculating the syndrome, it is to substitute \(2t\) roots \(\alpha' (i = 1, 2, ..., 2t)\) of \(g(x)\) into \(R(x)\), and then getting equation (7):

\[
S_i = R(\alpha') = R_0 + R_1\alpha' + ... + R_{n-1}(\alpha')^{n-1}
\]

(7)

Since \(R(\alpha') = e(\alpha') + c(\alpha')\), \(c(\alpha') = 0\), we can get equation (8):

\[
S_i = e(\alpha') = (\alpha'^{j_1}) + (\alpha'^{j_2}) + ... + (\alpha'^{j_v})
\]

(8)

For \(1 \leq l \leq v\), we define \(\beta_l = \alpha'^{j_l}\), and equation (8) can be abbreviated as equation(9):

\[
S_i = \beta_1 + \beta_2 + ... + \beta_v
\]

(9)

Define a polynomial \(\sigma(x)\) of degree \(v\) on \(GF(2^n)\):

\[
\sigma(x) = (1 + \beta_1 x)(1 + \beta_2 x) ... (1 + \beta_v x)
\]

\[
= \sigma_0 + \sigma_1 x + ... + \sigma_v x
\]

(10)

Where \(\sigma_i\) is the coefficient of error location polynomial, if we know the value of \(\beta_i\), we can get the error location in polynomial \(R(x)\), and \(\beta^{-1}, \beta^{-2}, ..., \beta^{-v}\) are the roots of \(\sigma(x)\).

Since the coefficient of syndrome polynomial \(S_1, S_2, ..., S_{2t}\) and the error position polynomial \(\sigma(x)\) are linked by the number of error positions \(\beta_1, \beta_2, ..., \beta^{-v}\), the following equation (11) can be obtained by associating the polynomial \(\sigma(x)\) and \(R(x)\).

\[
S_1 + \sigma_1 = 0
\]

\[
S_2 + \sigma_1 S_1 + 2\sigma_2 = 0
\]

...  \(\sigma_v S_{v-1} + \sigma_{v-1} S_v + v\sigma_v = 0\)

(11)

Equation (11) is also named Newton identity. We can find the identity that matches Newton identity and the time of numbers are less, that is \(\sigma(x)\). The \(\sigma(x)\) can be calculated by \(\sigma_i\), and we can determine the error location by calculating the root.
of $\sigma(x)$.

It can be implemented by iteratively through an efficient algebraic hard decision decoding algorithm, which is usually called Berlekamp-Massey (BM) algorithm [12]. The process of BM algorithm is as follows:

Firstly, determining the appropriate $\sigma^{(1)}(x)$ to satisfy the first equation of Newton identity. Then, provided that $\sigma_2$ is equal to $\sigma_1$, we need to judge if this matches the second equation of Newton identity. If it is satisfied, we can get $\sigma^{(2)}(x) = \sigma^{(1)}(x)$. Otherwise, a correction value is added in $\sigma^{(1)}(x)$ to make the second equation be established. $\sigma^{(2)}(x)$ is obtained after $2\ell$ iterations, and then $\sigma(x) = \sigma^{(2)}(x)$ can be determined. The reciprocal of roots for error location polynomial are error locations. From the mathematical perspective, the problem of finding the error location has been solved, but seeking root is not so easy to achieve in engineering.

Qian Wentian proposed a method of calculating roots in 1964, which was called Chien search, so that lots of problems in engineering were solved. The idea of Chien algorithm is: There is a primitive element $a$ in $GF(2^m)$, and the reciprocal of $1, a, a^2, \ldots, a^{m-1}$ are substituted into $\sigma(x)$, we can get equation (12):

$$\sigma(a^{-i}) = \sigma_0 + \sigma_1 a^{-i} + \sigma_2 (a^{-2i}) + \ldots + \sigma_{\ell}(a^{-\ell i})$$

(12)

Where $a^{-i}$ is the root of $\sigma(x)$ if $\sigma(a^{-i}) = 0$, and then whether the $i$-th bit in the received code has an error will be known. If there is a bit error in the $i$-th bit, and the code of this position is corrected, the result of the decoding is $R(x) + e(x)$. If it is not zero, it indicates that there is no error in the corresponding position.

### III. HARDWARE IMPLEMENTATION

The ECC module of this paper is based on SSD. The SSD is mainly composed of a flash array and a controller. The flash array is divided into 3 logical channels, each of which contains a flash array of 8 NAND Flash chips, a cortex-M1 processor and an NFI (NAND Flash Interface) unit. The structure of hardware is shown in Fig. 3, where NAC is NAND Array Controller, DMAC is Directly Memory Address Controller.

The flash memory chip uses MTL NAND Flash of Micron’s MT29F256G08C8EAB, which has a capacity of 32G per chip and contains 2 wafers. Each wafer contains 2 planes. Each plane

contains 1024 blocks, and each block contains 512 pages. The page has a data space of 16KB and a reserved space of 1216 bytes. Each NAND Flash is 8-bit width, and ECC operates 8 NAND Flash in parallel, each is called a physical channel.

In this paper, the ECC module includes encoding and decoding. Encoding and decoding are completed in a byte stream. When writing data, original data will be written to NFI and LFSR respectively. LFSR will calculate check code, then add binary to the behind of original data, and finally write it to NAND Flash. When reading data from NAND Flash, it will be read to page buffer and LFSR respectively, then it will be detected whether an error has occurred by ECC error detection register. If an error occurs, the syndromes are popped from LFSR, and the error locations and numbers are calculated by ECC core in the end. Next, ECC correction reads and corrects the errors from page buffer and returns the status to the ECC correction FIFO. There are 8 page buffers, and one of the page buffers is for a physical channel. The structure of ECC for one logical channel is shown as Fig. 4.

#### A. Write Pipeline

The unit of writing data is page, and 8 physical channels can be written in parallel at the same time. Therefore, the MC (Master Controller) waits for message in data buffer to reach the size of 8 pages, and then performs the write operation.

![Fig. 5. Write Pipeline](image-url)
chip data transmission are shared. Therefore, after executing the I/O transmission of a command and starting the programming operation, NAC does not wait for the end of programming. However, it starts preparing to receive the next command, and ECC is completed in I/O transmission. When the same logical unit receives the second write command, it waits for the completion of previous operation and checks whether the previous operation was successful. To further optimize performance, we used a write operation with a cache. Fig. 5 shows the results of the pipeline of the write operation, it takes two channels for example. Where D represents the process of DMA transmission from the message processing unit to page buffer, and I/O represents the process of data transmission in the ONFI bus, ECC presents the ECC process, P represents the programming process of internal flash unit.

From Fig. 5, we can see if we use internal cache buffer of flash chip, the I/O transmission and ECC would be hidden in the programming. In the main stage of writing, the internal programming of chip takes the longest time and becomes a performance bottleneck. The time of DMA, I/O and ECC is short, which can be hidden in the pipeline. Finally, the overall performance of writing will depend on the bandwidth of internal programming, which can be expressed as equation (13):

$$ P_{\text{write}} = N_{\text{logic chns}} \times N_{\text{LUNs}} \times \frac{C_{\text{page}} \times N_{\text{phys chns}}}{T_{\text{program}}} (MB/s) $$ (13)

Where $N_{\text{logic chns}}$ represents the number of logical channels, $N_{\text{LUNs}}$ represents the number of logical units in each chip, $C_{\text{page}}$ presents the size of valid value in a page, $N_{\text{phys chns}}$ represents the number of physical channels in each logical channel, $T_{\text{program}}$ represents the time spent in internal page programming in milliseconds.

**B. Read Pipeline**

Unlike the internal programming process of writing operation, which is the most time-consuming and becomes a performance bottleneck. The internal read and DMA process in the read operation are fast, and the I/O transfer process of chip becomes a performance bottleneck. In other words, the overall read performance will depend on the chip's I/O bandwidth. Fig. 6 illustrates the pipeline process of reading operation in two logical channels.

Since two logic units can perform internal read operation simultaneously, this process is hidden during the I/O transmission. For DMA and ECC, the time they cost is less than I/O transfer, so they are hidden in I/O transfer.

However, there is still a factor to be considered for the reading operation. Each sector in the flash chip also stores redundancy check information such as ECC and CRC, and the actual amount of data obtained is less than the amount of data passing through the I/O of the flash chip. Ultimately, the overall read performance will be expressed as equation (14):

$$ P_{\text{read}} = N_{\text{logic chns}} \times N_{\text{phys chns}} \times B_{1/O} \times \frac{L_{\text{data}} + L_{\text{OOB}} + L_{\text{CRC}} + L_{\text{ECC}}}{(MB/s)} $$ (14)

Where $N_{\text{logic chns}}$ represents the number of logical channels and $N_{\text{phys chns}}$ denotes the number of physical channels in each logical channel. $B_{1/O}$ represents the bandwidth of chip bus, the length of data, out-of-band information, CRC and ECC check code in each sector are $L_{\text{OOB}}$, $L_{\text{CRC}}$, and $L_{\text{ECC}}$, respectively.

**C. LFSR and Syndrome FIFOs**

There are 8 LFSR modules, ECC code and pre-syndromes are generated in them, each module is responsible for one physical channel. The width of a LFSR is 8 bits, and 8 physical channel is parallel. The principle of LFSR has been introduced in previous section.

When programming data, NFI will wait for LFSRs to be idle after all payload data are shifted into LFSRs, and then it will shift out the ECC code from LFSRs and store them to NAND Flash. Similarly, while reading data, NFI will shift all data into LFSRs, including ECC code. After that, when the ECC Core module needs the pre-syndromes to calculate the error numbers and the error locations, NFI will wait for LFSRs to be idle and then shift out the pre-syndromes from LFSRs.

There are 8 syndrome FIFOs, the generated pre-syndromes by LFSR are pushed into them when the ECC correction ability is open and some LFSR finds errors in the receiving data and the read-back page is valid. Each of syndrome FIFOs is 28-bit width and 32-level depth. One FIFO only for a particular LFSR, and a
FIFO is not available for any other LFSRs even it’s empty. It will take at least 32 clock cycles to push all pre-syndromes into FIFO, or more clock cycles if any FIFO is not empty at the moment. Then LFSR can process next sector data.

D. Arbiter and ECC Core

In this article, there are 2 ECC cores, however, we have 8 Syndrome FIFOs, so that an arbiter is needed to determine which data will be pushed into ECC core. When arbiter allows ECC core to read data from any of syndrome FIFOs, it also records the request sequence of 2 ECC cores. There are two levels as arbitration mechanism in order to hold the order of incoming operations: If requests arrive at different time, the policy is to determine according to which request comes first; If requests arrive at the same time, it adopts the policy of fixed priority.

The ECC Core will pop pre-syndromes from syndrome FIFOs, and then pre-syndromes will be expanded to syndromes. In an ECC core, there’re 3 stages of pipeline, precisely, syndrome expand, Berlekamp and Chien search. The structure of ECC core is depicted in Fig. 7. Finally, calculating error numbers and error locations can be done by any of ECC cores, and error status will be returned to the next module.

E. ECC Correction and ECC Correction FIFO

The error information from 2 ECC cores will be pushed into ECC correction. Then ECC correction reads error data from page buffer and corrects them. The error correction sequence is consistent with the order allowed by arbiter. After processing all error data of one sector, ECC correction module will push an entry to the ECC correction FIFO.

The ECC Correction FIFO is 16-level depth, and each FIFO entry reports correction status of one sector for one physical channel after it finishes ECC process of a sector. And the messages of ECC Correction FIFO can be read by software. The FIFO contains command pool index, lane number and sector number, ECC warning flag, error bit quantity, as well as finish flag.

F. ECC data pipeline

There’re several stages of pipeline for the total process of ECC module. Firstly, when the 8 syndrome FIFOs are all empty, syndromes from 8 LFSR modules will be pushed into 8 syndrome FIFOs. If the syndromes of previous sector in 8 channels have been entered into ECC cores, next syndrome data will be shift to syndrome FIFOs. This rule reduces the complexity of design at the cost of efficiency.

Secondly, the arbiter will grant an ECC core to read the data from one of 8 syndrome FIFOs, when any of ECC cores is available and some of 8 syndrome FIFOs are not empty. When other data of channels have flowed into ECC cores, some syndromes in some channels will remain in the FIFOs. Because we have 8 syndrome FIFOs, among which there are only 2 ECC cores.

Thirdly, the ECC core will calculate error numbers and error locations. An ECC core may receive syndrome data of a few channels in a short time. However, they may not be able to receive syndromes from syndrome FIFOs when they are all busy in correcting.

IV. Test

For the design proposed in this paper, it has calculated in theory and tested on the SSD in practice. The theoretical and test value of reading and writing performance in different error correction capabilities are shown in Fig. 8 and Fig. 9.

![Fig. 8. Theoretical value for reading and writing performance in different error correction capabilities](image)

![Fig. 9. Test value for reading and writing performance in different error correction capabilities](image)

As shown in Fig. 8 and Fig. 9, reading and writing speed will decrease along with the increase of error correction bits, and the test value is lower than the theoretical one due to system power consumption, etc. From the result analysis, the performance bottleneck of writing operation is the bandwidth of programming operation. The performance bottleneck of reading data is the bandwidth of data transmission of the chip bus, and the ECC operation is submerged in pipeline, which is not the performance bottleneck of the system. Here is a 40-bit error correction capability used for test in SSD, the test results are compared with other papers as shown in Table 2.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>24-bit ECC</td>
<td>607.4</td>
<td>422.8</td>
<td>407.4</td>
</tr>
<tr>
<td>40-bit ECC</td>
<td>565.7</td>
<td>401.5</td>
<td>381.9</td>
</tr>
<tr>
<td>56-bit ECC</td>
<td>599.9</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
paper [9] applies a 32-bit parallel architecture in encoding and 2-stage pipeline to accomplish the decoder with strong error correction capability and long decoding latency. For paper [12], it proposes a parity-check block as an additional decoding step, and a novel memory based syndrome updating method effectively improves the energy efficiency as well as the decoding latency. Despite the decoding latency is short, the error correction capability is not strong in such a method. As shown in Table 2, by increasing the degree of parallelism and using pipeline scheduling adequately, we can effectively increase data throughput and reduce decoding latency.

V. CONCLUSION

This paper improves the reading and writing performance of the system by designing the pipeline of the ECC module and increasing the degree of parallelism. The ECC module can support different error correction bits, having been tested on the physical SSD, whose test results have reached the anticipated effect. It supports 7.68Gbps with 64-bit parallel encoding and 3-stage pipeline decoding structure. Planar-level parallelism is not used in this design because of the more demanding constraints, which will result in the increase in the complexity of system, overhead of pipeline scheduling and the power consumption of system. For writing operation, if the chip is shared by I/O bus and DMA between the planes, the time of I/O, ECC and DMA will not be submerged adequately by the programming operation. Therefore, we will improve the pipeline scheduling mode to achieve the plane level in order to further improve system performance in the further work. Considering there are only two ECC cores, the decoding efficiency may be impacted, we will also perfect this part to further improve the performance of the system.

REFERENCES


