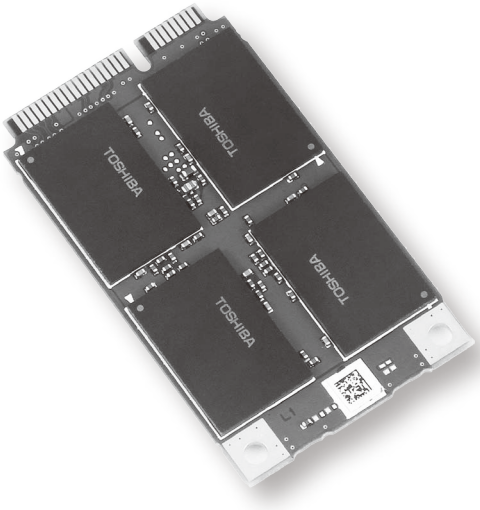
High-Performance Control Technologies for SSDs Using 19 nm Generation NAND Flash Memory for PC Market

● ASANO Shigehiro ● OSHIMA Takashi ● OSANAI Takeki



Solid-state drives (SSDs) using NAND flash memory have several advantages, including higher data transfer rates, higher shock resistance, and lower power consumption, when compared to hard disk drives (HDDs). As the capacity per unit cost of SSDs is still currently lower than that of HDDs, increasing capacity per unit cost with the shrinking of the manufacturing process for NAND flash memory is essential for the adoption of SSDs in the mainstream PC market.

In keeping with this trend, Toshiba has developed an SSD for the PC market using its 19 nm geometryNAND flash memory. To overcome the degradation of reliability and data transmission speed resulting from miniaturization, Toshiba has developed advanced SSD controller technologies such as advanced error correction coding, parallel data processing methods, and energy-saving features.

1. Introduction

In recent years, NAND memories have been the driving force behind the miniaturization of semiconductor processes. Toshiba has already begun volume production of 19 nm generation NAND memory devices, and is developing a range of SSDs using NAND memory as the storage medium. When compared with an HDD, an SSD is more compact and lightweight, offers higher read speeds , and has greater resistance to vibration- and shock and lower power consumption. For mobile applications in particular, even more widespread adoption is anticipated, but this requires a reduction in the unit cost of capacity, achieved through a further increase in density. However, with an increase in density, a decline in the reliability and data transfer rate of NAND memory is anticipated, and therefore high-performance controller technology is required to solve these problems at the system level for SSDs.

Table 1. Specifications of SSD controller

Item		Specification
Host interface		SATA 6 Gbit/s (Gen-3)
NAND flash memory		19 nm MLC, toggle2
NAND interface		8 channels, 4-way interleaved
Performance	Sequential read	520 Mbyte/s
	Sequential write	320 Mbyte/s
	Random read	75 k IOPS (for 4 Kbyte access)
	Random write	30 k IOPS (for 4 Kbyte access)
Process		Low power 40 nm, 7 layers

MLC: Multi-Level Cell IOPS: Input/Output per Second

Toshiba has developed a high-performance controller equipped with parallelization technology to enable increased speed, along with powerful error correcting codes (ECC), and technology to reduce power consumption compared to prior Toshiba models (THNSNC series). This paper describes these controller technologies developed to implement a client SSD using 19 nm generation NAND memory.

2. Controller architecture

Toshiba's SSD controller described in this paper uses Serial Advanced Technology Attachment (SATA) Generation 3 as the host interface. Table 1 lists the specifications, and Fig. 1 shows a block diagram.

This controller can be divided, principally, into a front end consisting of a SATA host interface and a SATA controller, and a back end for carrying out command processing, including access to the NAND memory. A multiprocessor configuration is adopted, with each of the front and back ends equipped with a single CPU, and parallelization is used to achieve increased speed. The back end has a command queue holding multiple commands received from the front end, and a data buffer carrying out data buffering.

The command queue uses Native Command Queuing (NCQ) defined for the SATA interface to achieve increased speed, by receiving and executing multiple commands non-sequentially. The data buffer fulfills the function of speed matching between accesses from the host and the NAND memory, and also of writing multiple data sectors together into NAND memory.

The lookup table (LUT) determines the relationship between

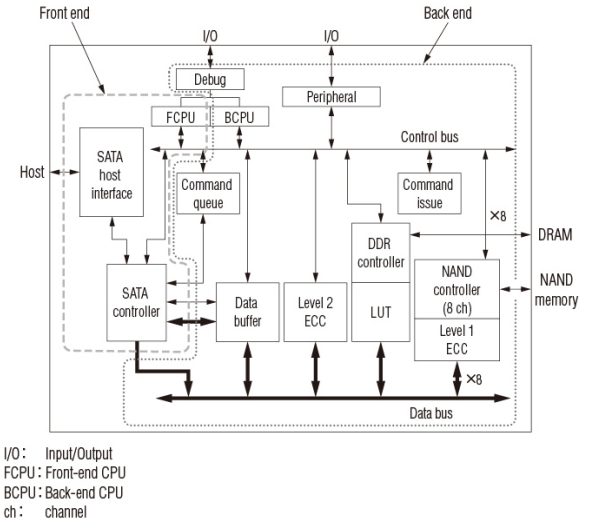


Fig. 1 Block diagram of SSD controller —
The controller is comprised of multiple processors, a hardware engine, the interface to NAND memory, and a host interface, connected by a bus.

the logical block address (LBA) and the physical address in NAND memory. The address management table is stored in DRAM connected externally to the controller, accessed through a built-in Double Data Rate (DDR) controller. The NAND controller for access to NAND memory supports 19 nm generation Multi-Level Cell (MLC) NAND memory and Toggle 2.0 interface. To improve the data transfer bandwidth, it supports eight channels and four-way interleave. The NAND controller has a built-in level 1 ECC block to protect data stored in the NAND memory. A level 2 ECC block is provided externally to the NAND controller.
The operation of these ECC blocks is described in Section 4.
LBA and NAND addresses are not required to be the same, and a mechanism is used to map LBA and NAND addresses. In general, the LBA access pattern is not necessarily uniform. If the LBA and NAND memory addresses are made to coincide, writing only to particular LBAs leads to increased wear to a particular part of NAND memory. To enhance reliability, an upper limit is placed on the number of writes to NAND memory, and it is undesirable to have a concentration of writes to a particular part of NAND memory. The technique of "wear leveling" is applied to ensure uniform overall write counts to NAND memory. Wear leveling requires the use of a lookup table that provides a mapping function between LBA and NAND memory.

Next, we describe the basic processing flow for read and write operations. A SATA read command from the host passes through the front end of the controller to the command queue. A read command that has entered the command queue is sent to the LUT to get the mapping between the LBA specified read address and the NAND memory physical address. Then the read command is sent to the appropriate NAND controller channel corresponding to the physical address, and the data is read out. The data read is transferred to the data buffer, and then sent to the host through the front end.
A write command from the host passes through the front end, and the write data, LBA, and data attributes are sent to the data buffer. Next, the LUT provides the mapping between LBA and NAND memory physical address, and data is written from the data buffer to the NAND memory location pointed by the physical address. Level 1 and level 2 error-correcting codes are generated and written with the write data to NAND memory.

3. Technology for increased speed

To achieve high data transfer rates, Toshiba's controller implements pipelined command processing. In other words, the controller processing is internally divided into a progression of blocks, through which each command is passed.. At this point, the amount of processing allocated to each block is approximately equal. The input/output units of each block are provided with internal buffers so that when the processing of a command completes in a block, it is passed to the following block, and another command passed from the preceding block can start processing.
This means that the blocks can simultaneously process different commands and that the overall system can process multiple commands in parallel. In Fig. 2, the pipeline concept is shown with an example of how multiple read commands are processed continuously.
The command pipeline processing is carried out using SATA NCQ, in such a way that before the execution of one SATA command completes, the next SATA command can be issued by the host.
With this controller, a sequential read rate of 520 Mbyte/s and sequential write rate of 320 Mbyte/s [1]can be achieved, while random read rate of 75 k IOPS (Input/Output per Second) and random write rate of 30 k IOPS can be achieved for 4 Kbyte access.

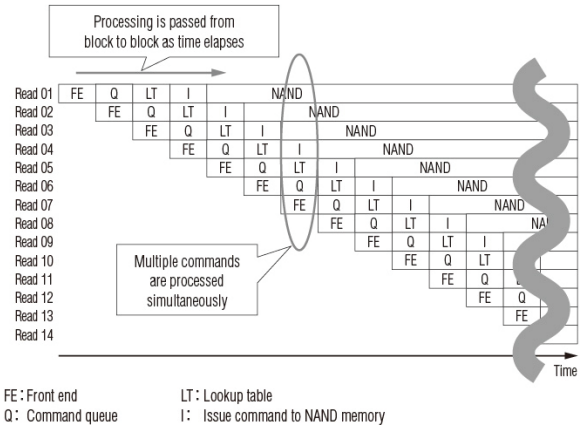
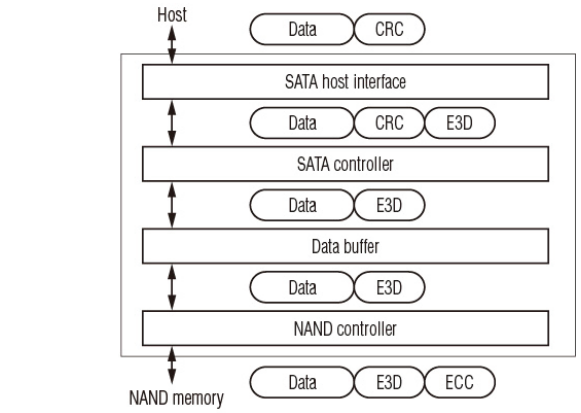


Fig. 2 Concept of read pipeline —
By means of pipeline processing, multiple commands can be executed simultaneously, increasing the throughput.

4. Technology to improve reliability

To enhance reliability of the data stored in NAND memory, level 1 and level 2 error correction functions are provided. When an error is detected, the data is first passed through the low latency level 1 error correction . If the error cannot be corrected by level 1 error correction, the data is next passed to the higher latency level 2 error correction.
As density continues to increase, the susceptibility of DRAMs to soft errors has been cited as a problem[(1)]. To enhance reliability, in addition to block code protection of data in word units, parity bits are used to identify erroneous bits to correct the error. By using this method, higher reliability can be achieved, while limiting the reduction in DRAM usage rates due to the error correcting codes.



CRC: Cyclic Redundancy Check

Fig. 3 Structure of end-to-end error detection (E3D) coding —
E3D codes are added on the data path, and finally a recheck is made to ensure that there are no errors on the data path.

Error correction codes are also added to the SRAM within the controller in word units so that the occurrence of an error can be detected and corrected.
For user data, End to End Error Detection (E3D) codes are added so that an error occurring at the point of data transfer can be detected so that user data with errors is never output externally.
The E3D code addition process is shown in Fig. 3.

5. Technology to reduce power consumption

To limit the increase in power consumption due to the increased performance, Toshiba developed the following technology.
To reduce the leakage current in the controller when on standby, High Vth cells are the basis of the design. For each block or function, a separate clock gate is provided so that non-operating circuits can be shut down at the lowest level possible. To switch the whole system to a low power consumption state, the phase-locked loop (PLL) itself can be shut down.

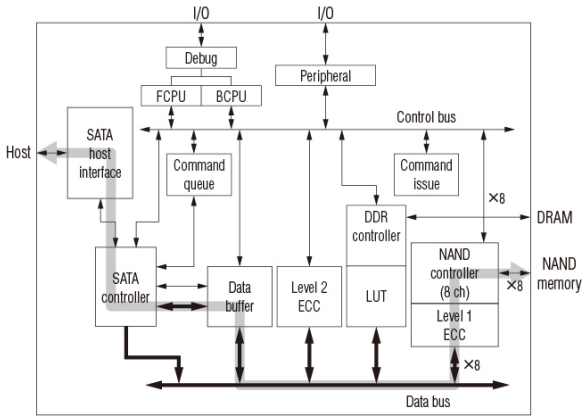


Fig. 4 User data path in SSD controller —
The transfer path for user data does not pass through the DRAM, enabling a reduction in power consumption.

Further, during data transfer, as shown in Fig. 4, Toshiba adopted an architecture such that user data can be transferred through internal buffers only to NAND memory without needing to pass through DRAM. This allows the power consumption of the DRAM and the DRAM interface to be reduced.
The NAND memory interface is equipped with a dedicated PLL which can be controlled from the controller firmware. By controlling the interface speed according to the characteristics of the SSD system and the NAND memory, it benefits the trade-off between performance and power consumption.

To achieve a lower power consumption for the SSD system during Standby, a control function is provided to shut down the supply of power to the DRAM and NAND memory chips. When the SSD is on Standby the power consumption is also reduced for the controller peripheral chips which are not in use.

6. Transmission line design

For DDR support with the NAND interface, transmission line simulations were carried out for 64 Gbyte through 1 Tbyte[2] (terabyte: 10^{12} bytes) models to verify signal integrity of DDR operation.
The transmission line simulation waveforms and measured waveforms were confirmed to be consistent (Fig. 5).

7. LSI implementation

The Toshiba SATA SSD controller is a standard cell design using a 40 nm generation low power process with seven interconnect layers. As described in Section 5, High Vth cells with low leakage current are the basis of the design. For paths with extreme timing requirements, a secondary high-speed cell with a reduced threshold voltage Vth is used.

For large-scale design using the 40 nm generation process, interconnect delays can be a problem. Toshiba's design reduces interconnect delays and increases the degrees of freedom for automatic connections during layout by providing a large number of flip-flops, while also segmenting the connections for the bus matrix in which longer connections are concentrated. (Compared to Toshiba THNSNC series)

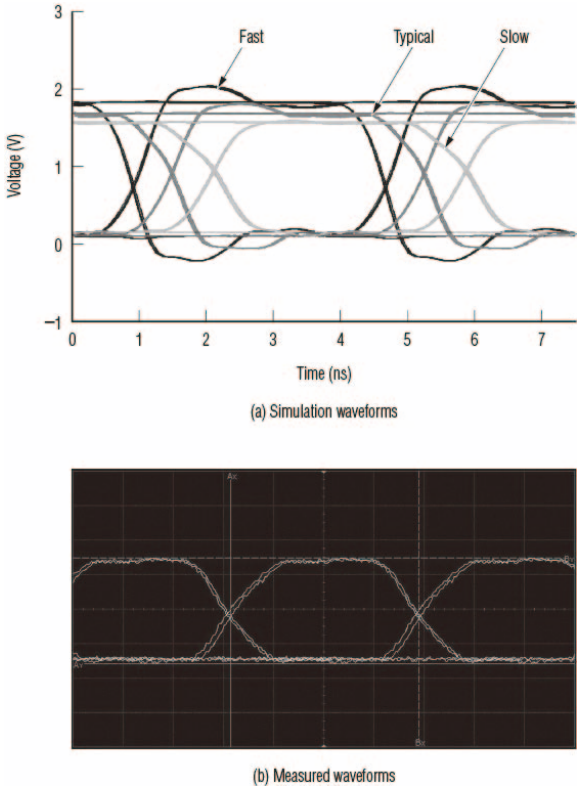


Fig. 5 Results of simulation and measurement of NAND transmissionline signal waveforms —
The simulation waveforms and measured waveforms for the NAND transmission line were confirmed to be in agreement.

Fig. 6 shows a chip photograph of the Toshiba SATA SSD controller. Three sides of the chip are occupied by eight channels of NAND memory interface and the DDR interface.

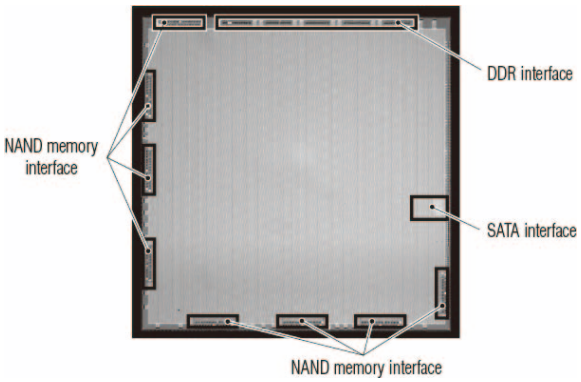


Fig. 6 SSD controller chips —
The NAND memory, DDR, and SATA interfaces are dispersed around the periphery.

8. Conclusion

This paper has provided descriptions of SSD controller parallelization technology, advanced reliability technology, and low power consumption technology to support increased speeds for 19 nm generation NAND memory.
With this technology as a base, further future increases in density will be possible with strengthening of the error correction capability. Although this paper describes Toshiba's SATA SSD controller, this technology can also be applied to future SSD interfaces beyond SATA.

[1] Read and write speed may vary depending on the host device, read and write conditions, and file size.
[2] Definition of capacity: Toshiba defines a megabyte (MB) as 1,000,000 bytes, a gigabyte (GB) as 1,000,000,000 bytes and a terabyte (TB) as 1,000,000,000,000 bytes. A computer operating system, however, reports storage capacity using powers of 2 for the definition of 1GB = 2^{30} = 1,073,741,824 bytes and therefore shows less storage capacity. Available storage capacity (including examples of various media files) will vary based on file size, formatting, settings, software and operating system, such as Microsoft Operating System and/or pre-installed software applications, or media content. Actual formatted capacity may vary."

References

(1) Shroeder, B. et al. "DRAM erros in the wild: a large-scale fi eld study". Proceedings of the eleventh international joint conference on Measurement and modeling of computer systems (SIGMETRICS '09). Seattle, WA, USA, 2009-06, ACM. New York NY, USA, ACM, 2009, p.193–204.


ASANO Shigehiro
Computer Architecrure & Security Systems Lab.


OSHIMA Takashi
Storage Products Div.


OSANAI Takeki
Storage Products Div.