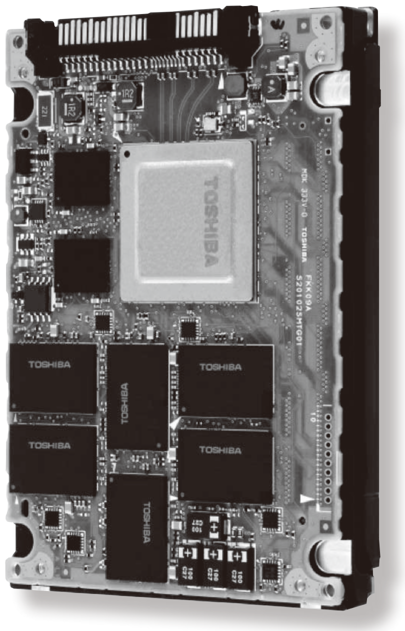
1.6 Tbyte SSD for Enterprise Use Applying MLC NAND Flash Memory

●KIMURA Akihiro ●MORO Hiroyuki ●MATSUSHITA Hiroki



The use of solid state drives (SSDs) in enterprise servers and storage systems for network computing has been expanding due to superior performance, such as ultrahigh levels of input/output operations per second (IOPS) and IOPS per watt, compared with those of conventional hard disk drives (HDDs). However, the higher bit cost of the NAND flash memory used in SSDs compared with that of the magnetic disks used in HDDs, as well as the use of single-level cell (SLC) NAND flash memories in conventional SSDs (to ensure the high data integrity required for enterprise use) are significant issues that need to be resolved in order to achieve large-capacity SSDs.

As a solution to this issue, Toshiba has developed a 2.5-inch SSD for enterprise applications achieving the highest capacity(*1) in the industry of 1.6 Tbytes[1], applying multi-level cell (MLC) NAND flash memory. This SSD offers not only a lower bit cost but also high performance and reliability by means of a newly developed SSD controller.

(*1) As of June 2013, for 2.5-inch enterprise SSDs. As researched by Toshiba

1. Introduction

Computer servers and storage systems for enterprise applications require SSDs with higher performance, reliability and capacity than ever before.

In response, Toshiba has developed the PX02SMB160 2.5-inch SSD for enterprise applications with a capacity of 1.6 Tbytes. The new SSD utilizes MLC NAND flash memory that can store two bits of information per cell, and a unique controller technology. This article provides an overview of the PX02SMB160 as well as the underlying technologies used to achieve high performance and reliability.

2. Product Overview

The PX02SMB160 (Fig. 1) is Toshiba's second-generation enterprise SSD. The new SSD storage medium is an MLC NAND flash memory fabricated with a 24 nanometer (nm) process to achieve lower bit cost. The downside of the MLC NAND is lower read and write performance compared to SLC NAND of a similar process technology. MLC NAND also requires a robust error-correction algorithm to compensate for a larger bit error rate compared to SLC NAND. To address these issues, Toshiba has developed unique controller technology[1]. The new controller embedded in the PX02SMB160 combines access performance, excellent reliability, and data integrity required for enterprise SSDs.

The PX02SMB160 incorporates twice as many NAND flash memory devices as Toshiba's first-generation SSD to achieve higher capacity. To accomplish this goal, high-density board assembly and optimal thermal dissipation designs were examined. As a result, the PX02SMB160 delivers higher storage capacity, higher access performance and higher power efficiency than the first-generation MK4001GRZB(2) that relies on SLC NAND (Table 1).

Furthermore, the PX02SMB160 also provides an SAS*-3 interface (with a data rate of 12 Gbits per second[2]) and self-encrypting capability.

* SAS: Serial Attached SCSI (Small Computer System Interface)
Note 1: As of June 2013, for 2.5-inch enterprise SSDs. As researched by Toshiba

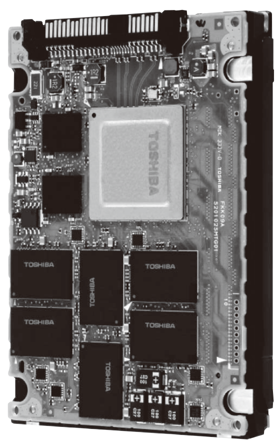


Fig. 1 PX02SMB160 SSD with 1.6 Tbyte capacity for enterprise use—
Built on the first-generation MK4001GRZB, the PX02SMB160 uses MLC NAND flash memory for lower bit cost and higher performance and reliability

Table 1. Main PX02SMB160 specifications

Characteristics			Specifications	
			PX02SMB160	MK4001GRZB
Storage capacity			1.6 Tbytes	400 Gbytes
NAND process			24 nm MLC	32 nm SLC
Host interface			SAS-3 (12 Gbits/s)	SAS-2 (6 Gbits/s)
Access performance	Sequential reads	(Mibytes/s)	935	500
	Sequential writes	(Mibytes/s)	410	250
	Random reads	(kIOPS)	130	90
	Random writes	(kIOPS)	27	16
Power efficiency		(kIOPS/W)	18.5	13.8
Data integrity		(bit error rate)	1×10^{-17}	1×10^{-16}

IOPS: Input/Output Per Second
$Mi = 2^{20}$

3. Controller Design

Fig. 2 shows a block diagram of the new controller specifically designed for enterprise SSDs, and the interconnections between the controller and the NAND flash memory.

The PX02SMB160 enterprise SSD provides 16 channels of NAND flash memory interfaces, twice as many as the Toshiba client SSD which has eight. The PX02SMB160 has an interleaved memory access architecture in which each channel is associated with two or more NAND flash memory chips (each with 8-Gbyte capacity). This means that while one memory chip is reading data from or writing data to memory cells, the controller can perform a data transfer with another memory chip simultaneously. The increased concurrency among NAND flash memory chips not only compensates for the relatively slow read and write speeds of the MLC NAND but also improves the overall SSD access performance. Furthermore, each NAND flash memory interface has a 1-to-2 multiplexer to double the number of accessible NAND memory devices. These design techniques combine to realize a high storage capacity of 1.6 Tbytes with high performance.

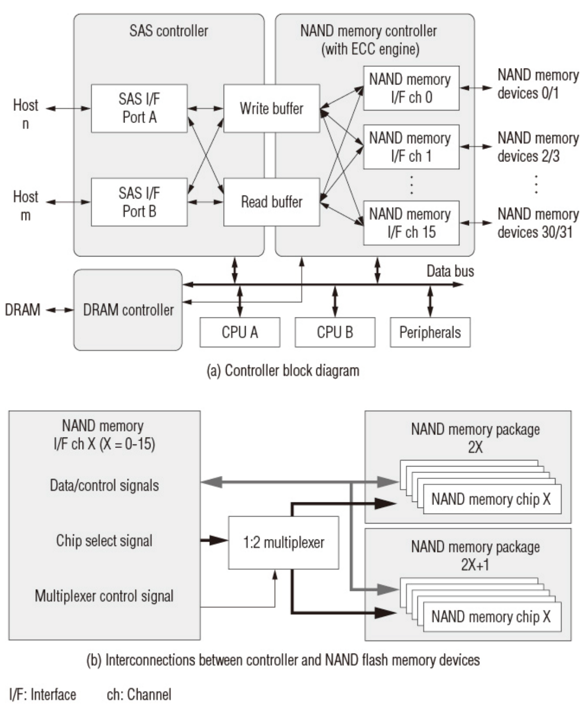


Fig. 2 Block diagram of the newly developed SSD controller and topology between controller and NAND flash memory devices—
The PX02SMB160 provides outstanding access performance with increased concurrency due to its 16-channel NAND flash memory interface and an interleaved memory structure. The 1-to-2 multiplexer in each NAND flash memory interface doubles the number of accessible NAND flash memory devices, achieving storage capacity of 1.6 Tbytes.

The PX02SMB160 incorporates Quadruple Swing-By Code (QSBC™), Toshiba's unique error correcting code (ECC), which incorporates enhanced ECC Level 1 and 2 to improve data integrity protection. While using MLC NAND, the PX02SMB160 meets the reliability requirements for enterprise SSDs: 24-hour operation for five years and daily data writing of 16 Tbytes, 10 times the drive capacity[3].

4. High-Density Board Assembly

While Toshiba's first-generation SSD contains 16 NAND flash memory devices in a 15-mm-thick casing with the 2.5-inch form factor, the second-generation SSD was designed to house 32 NAND flash memory devices for a storage capacity of 1.6 Tbytes. Although the first-generation SSD has NAND flash memory devices only on one side of printed circuit boards, we considered double-sided mounting for the second generation (Fig. 3).

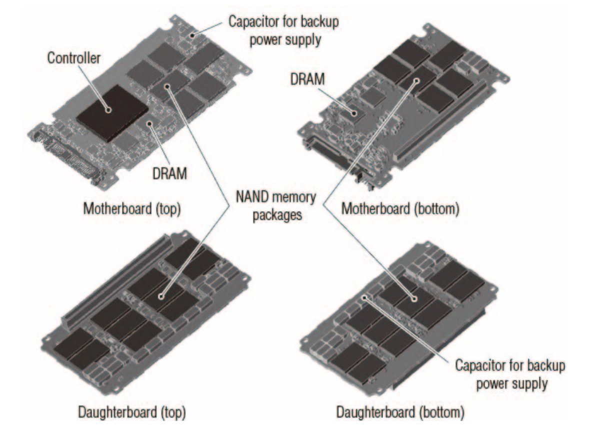


Fig. 3 Printed circuit board (PCB) assemblies with NAND flash memory packages on double surfaces—
The PX02SMB160 has NAND flash memory devices on both sides of two PCBs.

However, because there is a significant mismatch in the linear coefficient of thermal expansion between the silicon substrate of NAND flash memory chips and the printed circuit board, double-sided mounting is generally known to have lower heat shock resistance (soldering reliability) than single-sided mounting. We utilized our originally developed thermal distortion simulation technique using three-dimensional (3D) models to achieve optimal placement of NAND flash memory devices on the board. The simulation indicated that placing NAND flash memory devices on both sides of the board facing each other would help reduce the mismatch in the linear coefficient of thermal expansion and thus provide heat shock resistance equivalent to that of the first-generation SSD.

5. Thermal Design

Another issue we had to solve to achieve high-density board assembly was the increased temperature caused by self-heating. The challenge here was to place the controller and NAND flash memory close together, although the controller generates far more heat than other parts and the maximum operating temperature of the NAND flash memory is relatively low.

Conversely, a preferable aspect is that the major applications of enterprise SSDs are computer servers and storage systems equipped with fans to cool the entire system. We therefore performed thermal flow analysis to simulate possible temperature rises (Fig. 4) to find how best to utilize the fan air to dissipate the heat generated by the controller from the SSD casing, thereby reducing heat propagation to the NAND flash memory.

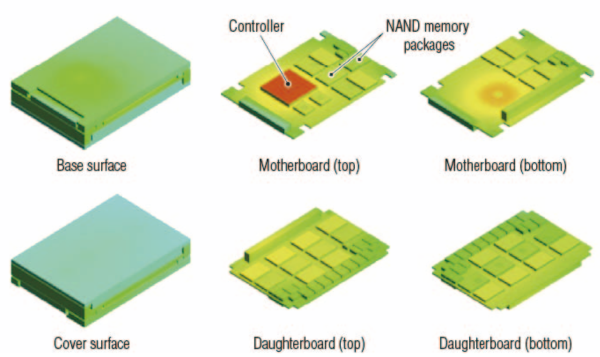


Fig. 4 Temperature profile of PCB assemblies and outer cases obtained by thermal flow analysis—
The new SSD optimally exploits the system fan air to eliminate controller-generated heat from the SSD casing, thereby reducing the temperature rise of the adjacent NAND memory.

As the result of this thermal study, the temperature rise of the controller was reduced by 25% thanks to the improved thermal conductivity of the thermal conductive sheet placed between the controller and the SSD casing.

6. Data Integrity Protection

Because the new enterprise SSD uses MLC NAND, its controller incorporates an enhanced error correcting code, as previously described. Furthermore, the new SSD provides background patrol, or a function to check data integrity and correct errors as they happen; operating autonomously while the drive is on. Stable access performance is realized during 5-year, 24-hour operation[3] by combining background patrol with conventional wear leveling that equalizes program/erase cycles across the NAND flash memory.

Furthermore, the PX02SMB160 is equipped with a backup power supply comprising large-value capacitors designed to flush the cached user data to the NAND memory in the event of an unexpected power loss.

7. Conclusion

The market demands ever-greater capacity and access performance for SSDs using NAND flash memory as a storage medium, especially in enterprise applications such as computer servers and storage systems. Conversely, the NAND flash memory will suffer from data integrity degradations with shrinking process geometries, and the controller will require even more power to achieve further-enhanced error correction.

We will leverage our accumulated expertise to continually improve our controller design and board assembly technologies to further enhance the performance and reliability of SSDs.

[1] 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."

[2] Read and write speed may vary depending on the host device, read and write conditions, and file size.

[3] Toshiba research

References

(1) ASANO Shigehiro. et.al. High-Performance Control Technologies for SSD Using 19 nm-Generation NAND Flash Memory
TOSHIBA Review 67, 12, 2012, P.43-46.

(2) KIUCHI Hidemichi. MK4001GRZB Solid-State Drive forEnterprise Use Achieving High Performance and High Reliability
TOSHIBA Review 66, 8, 2011, P.40-43.

KIMURA Akihiro
Storage Products Div.

MORO Hiroyuki
Storage Products Div.

MATSUSHITA Hiroki
Storage Products Div.