

Swissbit: Wie man Flash für Industrieanwendungen auswählt

# Langlebiger Flash-Speicher

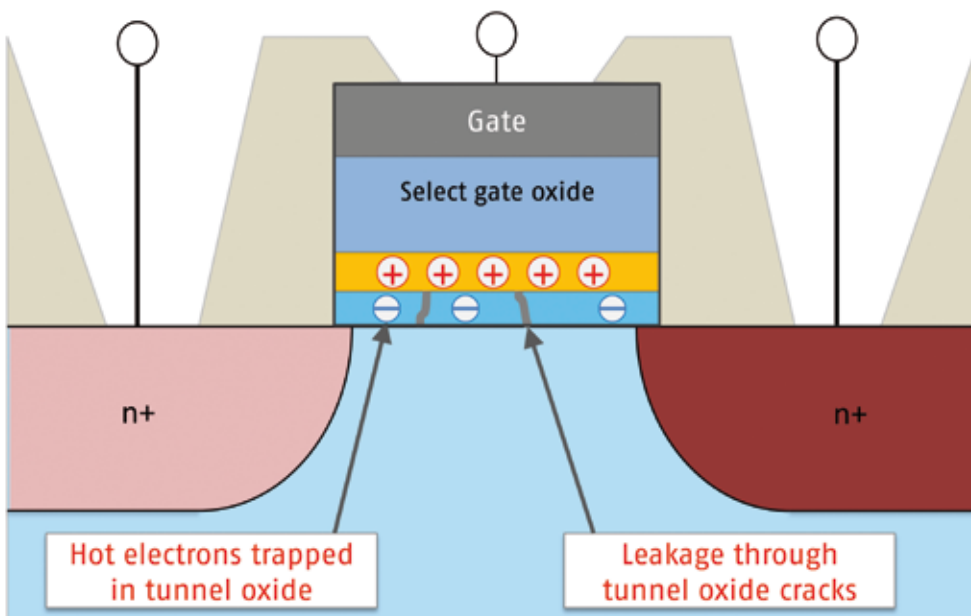
*Den passenden, langlebigen Flash-Speicher für Module oder SSDs in industriellen Anwendungen auszuwählen, ist eine komplexe Angelegenheit. Preis-Leistungs-Vergleiche scheitern, weil schon die anforderungsgerechte Definition von Leistung von vielen verschiedenen Faktoren abhängt.*

VON ULRICH BRANDT,  
DIRECTOR MARKETING DER SWISSBIT AG

In normalen IT-Anwendungen kann der Preis pro GByte sowie Schreib- und Lesegeschwindigkeit als Einkaufskriterium ausreichen. Hingegen hat andere Prioritäten, wer SSDs für den Einsatz in industriellen Anlagen sucht oder Speicherkarten für eine Telemekoanwendung im Freien beschaffen soll: Kapazität ist oft überhaupt kein Thema, dafür aber Robustheit und Langlebigkeit. Endurance (die Lebensdauer von Flash-Medien) und Retention (die Lebensdauer der gespeicherten Daten) sind komplexe Themen. Das wird auch durch die Anwendung beeinflusst, also: Wie werden Daten geschrieben? Oder geht es mehr ums Lesen?

## Alterungserscheinungen

Die Zellen eines NAND-Flashs überstehen nur eine begrenzte Zahl von Löschungen, weil sich in der Oxidschicht, die verhindert, dass die Elektronen aus dem Floating Gate abfließen, bei jeder Erzeugung des Tunneleffekts durch die Löschspannung Elektronen mit erhöhtem Energieniveau ansammeln. Damit verschiebt sich auf Dauer die Schwellwertspannung, bis die Zelle nicht mehr lesbar ist. Einen weiteren Alterungseffekt verursacht das Entstehen leitender Pfade durch die Oxidschicht. Hierdurch verliert die Zelle allmählich ihren Ladungszu-



Eine alternde Zelle: In der Tunneloxidschicht sammeln sich Elektronen, wodurch sich allmählich die Schwellwertspannung verschiebt. Risse im Tunneloxid schaffen Leckstrompfade, durch die Ladung abfließen kann. Die Lesefehler nehmen zu, bis der gesamte Block als „Bad Block“ ausgemustert werden muss.



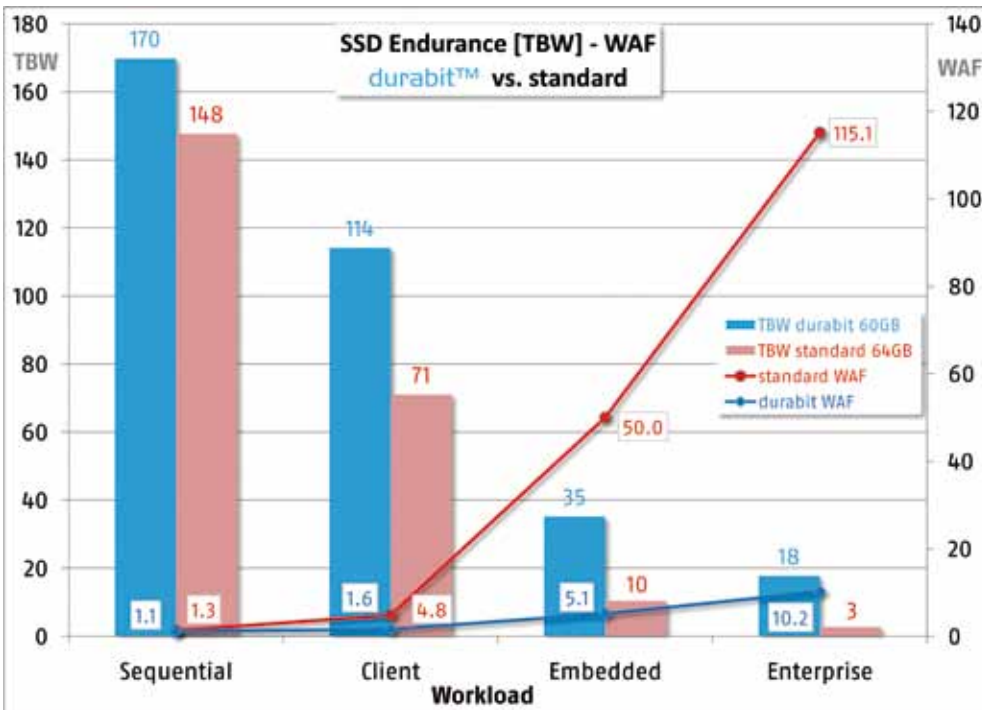
Ulrich Brandt, Swissbit

„Messungen mit einer 480-GB-SSD von Swissbit ergaben je nach Messverfahren eine Lebensdauer von 1360, 912 oder 140 TBW (Terabyte Written).“

stand und damit das gespeicherte Bit. Dieser Effekt, der durch hohe Temperaturen verstärkt wird, tritt besonders dann auf, wenn die zulässigen P/E-Zyklen (Programmier-/Lösch-Zyklen) zur Neige gehen. Die Retention nimmt in dieser Phase massiv ab. Kann man sowohl bei einem Single-Level-Cell-NAND wie bei einem Multi-Level-Cell-NAND ursprünglich von einer Retention von zehn Jahren ausgehen, so liegt sie am Ende des Lebenszyklus nur noch bei einem Jahr. Dieser Punkt ist beim MLC nach 3000 P/E-Zyklen erreicht, beim SLC erst nach 100.000 P/E-Zyklen. Deshalb ist SLC in besonders anspruchsvollen Anwendungen so beliebt – hingegen kommen kostengünstige Tripel-Level-Cell-NAND-Chips (TLC) wegen der niedrigen P/E-Zyklen für langlebige Speicher nicht in Frage: Hier müssen acht verschiedene Ladungsniveaus unterscheidbar sein, um 3 Bit pro Zelle zu schreiben. Folglich machen sich die degenerativen Effekte viel schneller bemerkbar: Bei TLC sinkt die ursprüngliche Retention von einem Jahr nach 500 P/E-Zyklen auf nurmehr drei Monate.

## Maßnahmen

Die Folgen der physischen Degeneration der Chips werden über verschiedene Mechanismen abgemildert. Weil beim Ausfall von Zellen der ganze Block als „Bad Block“ markiert werden muss, wird mit dem Wear Leveling eine gleichmäßige Nutzung der physischen Speicheradressen sichergestellt. Doch Lesefehler deuten



F-60 durabit erweist sich im Vergleichstest mit Standardprodukten deutlich langlebiger und hat geringeren WAF. Der Effekt wird durch FTL im DRAM und erhöhtes Over Provisioning erreicht.

nicht nur auf Abnutzung hin. Bei jedem Schreiben werden die Zellen in der Umgebung der zu programmierenden Zelle gestresst, das heißt, sie weisen eine leicht erhöhte Spannung auf (Program Disturb). Mit der Zeit erhöht sich in diesen Zellen die Schwellwertspannung, und es kommt zu Lesefehlern, die nach dem Löschen des Blocks wieder verschwinden. Desgleichen führt auch das Lesen zu Stress (Read Disturb). Hier sind es die benachbarten Pages, die Ladung ansammeln. Ein interessanter Aspekt: Besonders stark ist der Effekt bei Anwendungen, die immer wieder dieselben Daten lesen. Das heißt auch im Inneren eines Speichers, der nur gelesen wird, müssen zur Fehlerkorrektur immer wieder Blöcke gelöscht und Pages geschrieben werden.

#### Wie wird Endurance gemessen?

Es gibt zwei Maße, mit denen Hersteller die Lebensdauer von SSDs angeben: Terabyte Written (TBW) und Drive Writes Per Days (DWPD). TBW sagt aus, wie viele Daten über die Lebenszeit insgesamt geschrieben werden können. DWPD gibt an, wie viele Daten innerhalb der Garantiezeit pro Tag maximal geschrieben werden können. Das Problem: Entsprechende Benchmarks sind sehr aufwendig und man ist auf die Angaben des Herstellers angewiesen. Ob sie allerdings eine Aussagekraft für die Anwendung haben, für die man einen Datenträger sucht, ist ungewiss, denn tatsächlich hängt der Wert extrem vom Wor-

kload des Tests ab. Messungen mit einer 480-GB-SSD von Swissbit ergaben je nach Messverfahren eine Lebensdauer von 1360, 912 oder 140 TBW. Den beeindruckendsten Wert bekommt man beim sequenziellen Schreiben. Der zweite Wert ergibt sich bei »Client Workload«, der dritte bei »Enterprise Workload«. Bei diesen beiden Belastungstests handelt es sich um Vorgaben der Standardisierungsorganisation JEDEC. Der Client Workload orientiert sich am Nutzerverhalten eines PC-Anwenders und erzeugt zum größeren Teil sequenzielle Zugriffe. Der Enterprise Workload simuliert das Verhalten eines Servers mit Multi-User-Umgebung, die zu 80 Prozent Random-Zugriffe erzeugt. Mit diesen Standards wäre eigentlich eine Vergleichbarkeit gegeben, allerdings machen viele Hersteller keine Angaben zum zugrunde liegenden Workload und werben mit sequenziellen Schreibwerten. Weil, wie das Beispiel zeigt, zwischen diesem und dem Enterprise Workload der Faktor 10 liegen kann, ist äußerste Vorsicht bei nicht näher spezifizierten, auffällig hohen Endurance-Werten angezeit.

#### Schreibverstärkung reduzieren

Das Mapping zwischen logischer und physischer Adresse, ECC und weitere Mechanismen wie die Garbage Collection zur Freigabe von Blöcken sind alles relevante Vorgänge, um die Funktion und Leistung von Flash-Speicher zu verstehen und einzuordnen. Ein wichtiger Be-

griff ist in diesem Zusammenhang der Write Amplification Factor (WAF), zu Deutsch die Schreibverstärkung. Er gibt das Verhältnis zwischen den vom Host kommenden Anwenderdaten und der tatsächlich ins Flash geschriebenen Menge an Daten wieder. Er ist ein Maß für die Effizienz der Arbeit eines Flash-Controllers. Die Schreibverstärkung zu reduzieren, ist einer der Schlüssel für längere Endurance. Workload-Faktoren, die den WAF beeinflussen, sind der Unterschied zwischen sequenziellen und Random-Zugriffen oder die Größe der Dateiblöcke im Verhältnis zu Seiten und Blockgrößen. Die beiden Grundbedingungen: Seiten müssen nacheinander beschrieben werden und Blöcke als Ganzes gelöscht werden. Das Mapping zwischen logischer und physischer Adresse bezieht sich im Standardverfahren auf Blöcke. Sehr effizient ist dies bei sequenziellen Daten, weil dann die Seiten eines Blocks hintereinander weg geschrieben werden. Ein Anwendungsbeispiel sind kontinuierlich gesammelte Videodaten. Bei Random-Daten werden hingegen Seiten in vielen verschiedenen Blöcken geschrieben, sodass bei jedem internen Umschreiben pro Page ein ganzer Block gelöscht werden muss. Die Folge sind hoher WAF und sinkende Lebensdauer. Für nicht sequenzielle Daten ist daher Page-based Mapping besser geeignet, das heißt, die Firmware sorgt dafür, dass Daten verschiedenen Ursprungs sequenziell in Seiten eines Blocks abgelegt werden können. Die Zahl der Löschungen sinkt und die Schreibperformanz steigt. Nachteil dieser Methode ist die größere Zuordnungstabelle des FTL (Flash Translation Layer), doch dies kann mit einem integrierten DRAM ausgeglichen werden.

Ein weiterer Faktor, der die Schreibverstärkung in die Höhe treibt, ist der Nutzungsgrad des Datenträgers. Je mehr Daten auf einem Flash-Medium gespeichert sind, desto mehr muss die Firmware Bits umschaukeln. Auch hier ist Page-based Mapping von Vorteil. Doch die Hersteller können eine weitere Stellschraube nutzen: das Over Provisioning. Damit ist ein Bereich des Flashs gemeint, der nur für die internen Aktivitäten reserviert ist. Üblicherweise sind das 7 Prozent – der Unterschied bei Gigabyte-Angaben zwischen dem binären und dem dezimalen Wert. Werden statt 7 Prozent 12 Prozent als Over Provisioning reserviert, hat das einen erstaunlichen Effekt: Bei einem Endurance-Vergleich (TBW bei Enterprise Workload) zwischen zwei vom MLC-NAND-Chip her gleichen SSDs erreichte das Swissbit-Modell F-60 durabit mit 60 GByte einen 6,6-fach höheren Wert als das Modell F-50 mit 64 GByte. Bei den Versionen 240 GB/265 GB war der Wert sogar 10-mal höher. (es)