

So lassen sich die Vorteile von Flash richtig nutzen



von Hermann Strass

eine Publikation von

speicherguide.de
Das Storage-Magazin

powered by



So lassen sich die Vorteile von Flash richtig nutzen

Flash für Profis: Der große Vorteil von Flash-Speichern besteht darin, dass sie wesentlich schneller als Festplatten sind – aber auch teurer. Was Administratoren beim Einsatz von Flash-Speichern wissen sollten, und wie sie sich optimal in Speichersystemen einsetzen lassen.

Flash ist eine nichtflüchtige Speichertechnik auf Halbleiterbasis mit wahlfreiem Zugriff, ähnlich wie RAM (Random Access Memory), einer flüchtigen Speichertechnik ebenfalls auf Halbleiterbasis. Derzeit sind noch die meisten Speichersysteme aus Festplatten mit rotierenden Magnetscheiben zur blockweisen Speicherung und aus Magnetbandsystemen mit kontinuierlicher serieller Speicherung zur Archivierung aufgebaut. Flash aus Halbleitermaterial ohne mechanisch bewegliche Teile arbeitet erheblich schneller und dank elektronisch wahlfreiem Zugriff fast ganz ohne Verzögerungszeiten (Latenz). Weitere technische

Aspekte sind zu beachten wenn ein optimales Speichersystem geplant wird.

1 Flash-Technik-Grundlagen: Derzeit gibt es drei Varianten der Flash-Technik. Der Unterschied bezieht sich auf die Anzahl der gespeicherten Bits innerhalb einer Flash-Speicherzelle. Eine »Single Level Cell« (SLC) enthält ein Bit, das zwei logische Zustände (0 und 1) in zwei Spannungsbereichen darstellt, die von drei Spannungsebenen eingegrenzt werden. Entspricht die Spannung in der Zelle dem niedrigeren Spannungsebene, dann repräsentiert dies einen logischen Zustand. Entspricht die Spannung in der

Zelle dem höheren Spannungsebene, dann repräsentiert dies den anderen logischen Zustand.

In einer »Multiple Level Cell« (MLC), die mathematisch korrekt »Double Level Cell« heißen sollte, gibt es zwei Bits, die vier logische Zustände (00, 01, 10, 11) auf vier Spannungsebenen darstellen. In einer »Triple Level Cell« (TLC) gibt es logischerweise drei Bits, die acht logische Zustände durch acht Spannungsebenen darstellen. Der gesamte Spannungsbereich für alle drei Varianten ist fast gleich. Das bedeutet, dass der jeweilige Spannungsbereich für einen logischen Zustand abnehmend kleiner ist. Mehr Bits innerhalb einer Zelle bedeutet langsamerer Zugriff und weniger Fehlertoleranz.

Für TLC werden sehr viel aufwendigere Fehlerkorrekturverfahren benötigt als für MLC oder SLC. Das erhöht die Komplexität und benötigt Verarbeitungszeit, womit die Latenz (Zugriffszeit) erhöht wird. Beim

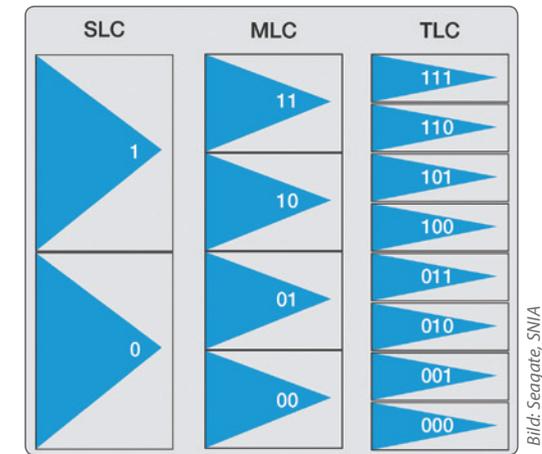


Bild: Seagate, SNIA

Die Spannungsbereiche werden kleiner, wenn mehr Bits je Zelle gespeichert werden.

Übergang auf 3D-Flash wird die einzelne Zelle wieder vergrößert und damit die Zuverlässigkeit verbessert. Beim Einsatz von angepassten Fehlerkorrekturalgorithmen könnte TLC auch für besonders leseintensive professionelle Anwendungen genutzt werden.

Wichtige Flash-Parameter

SLC ist die schnellste, zuverlässigste aber auch teuerste Technik. Sie wird nur für die anspruchsvollsten professionellen Anwendungen eingesetzt. MLC ist die am meisten

Variante	Bits/Zelle	Spannungsbereiche	W/E-Zyklen	Geschwindigkeit	Kosten
SLC	1	2	100 000	sehr schnell	sehr hoch
eMLC	2	4	10 000	etwas über Durchschnitt	mittel
MLC	2	4	5 000	etwas unter Durchschnitt	unter Durchschnitt
TLC	3	8	1 000	langsam	niedrig

Technischen Fakten und Merkmale von Flash

Quelle: Seagate/Hermann Strass

eingesetzte Technik für den Einsatz im Büro und zu Hause. Bei MLC gibt es die eMLC-Untervariante (enterprise oder enhanced MLC), die gegenüber »normalem« MLC verbessert und etwas teurer ist. Diese eMLC-Variante wird in vielen professionellen und anspruchsvollen Büro-Anwendungen genutzt.

Der Begriff W/E-Zyklen (Write/Erase) bedeutet Schreib- und Löschyklen bei denen Elektronen in die »Floating Gate«-Zellen hinein gespeichert oder heraus gelöscht werden. Dadurch werden die Zellen im Laufe der Zeit geschädigt, so dass üblicherweise nur etwa die in obiger Tabelle angegebene Zahl an Zyklen erreicht wird. Es gibt viele weitere Parameter, welche die Nutzungsdauer der Flash-Zellen beeinflussen. Deswegen sind die Angaben in den drei rechten Spalten der Tabelle nur grobe Schätzungen.

Es gibt weitere Parameter um die Nutzungsdauer, mögliche Fehlerraten oder Zuverlässigkeit zu messen. Ein Parameter heißt »Failure in Time« (FIT), wobei die Zeit (time) typischerweise mit einer Milliarde Stunden (10E9) angenommen wird. In amerikanischer Zählweise heißt diese Zeitspanne Trillion. FIT ist ein reiner Statistikwert der auch Lambda (λ) genannt wird, wobei niedrigere Zahlen besser sind. Mit einiger Vor-

sicht kann der Wert $1/\lambda$ mit MTBF (Mean Time Between Failures) gleichgesetzt werden. Der FIT-Wert eines Systems ist die Summe der FIT-Werte aller Komponenten.

In einem professionellen Flash-System ist der FIT-Wert der Netzausfallkomponenten besonders wichtig, weil er großen Einfluss auf die Fehlerrate und deshalb auf die Zuverlässigkeit hat. Elektrolyt-Kondensatoren und Spulen sind sehr fehleranfällige Komponenten. Die Hersteller von professionellen Systemen, wie beispielsweise **Seagate**, kennen diese Probleme und sorgen entsprechend für zuverlässige Systeme. In Systemen für zu Hause oder für das Büro werden diese möglichen Probleme ignoriert, weil das die Kosten in die Höhe treibt.

NAND oder NOR – eine Einsatzfrage

In kommerziellen Speicheranwendungen für Büro und Rechenzentrum wird fast ausschließlich NAND-Technik genutzt. Die NOR-Technik wird hauptsächlich in industriellen Anwendungen für Steuerungen (Controller) und in Echtzeitanwendungen mit »execute-in-place« (XIP) eingesetzt. Die wichtigsten Unterscheidungsmerkmale sind Kapazität, XIP und Kosten.

NAND-Speicherzellen sind entlang ihrer Bit-Leitungen aufgereiht. Damit ist dieser

Speicher kompakter und langsamer, bei höherer Kapazität und geringeren Kosten/Preisen für kommerzielle Anwendungen sowie in Wegwerf-Produkten (USB-Sticks usw.). NOR-Speicherzellen sind parallel zu ihren Bitleitungen aufgereiht. Dafür wird mehr Platz auf den Speicher-Chips benötigt. Deshalb sind aber Lesezugriffe besonders schnell. Das ist nötig für XIP und für Echtzeitanwendungen. Die Speicherdichte ist geringer als mit der NAND-Technik, deshalb sind die Kosten/Preise höher.

2 Besondere Merkmale der Flash-Technik: Die Übertragungsraten bei Flash-Speichern sind deutlich höher und die Latenzzeiten sind deutlich kürzer als bei herkömmlichen magneto-mechanischen Speichern (HDDs). Flash benötigt weniger Energie für den Betrieb und für die Kühlung als Festplatten (HDDs). Demgegenüber stehen aber die erheblich höheren Kosten für Flash-Chips, Geräte (Laufwerke) und Systeme. In optimal konfigurierten größeren Systemen wird es je nach den Anwendungsanforderungen eine Mischung der Techniken geben, unter Berücksichtigung der jeweiligen Vorteile.

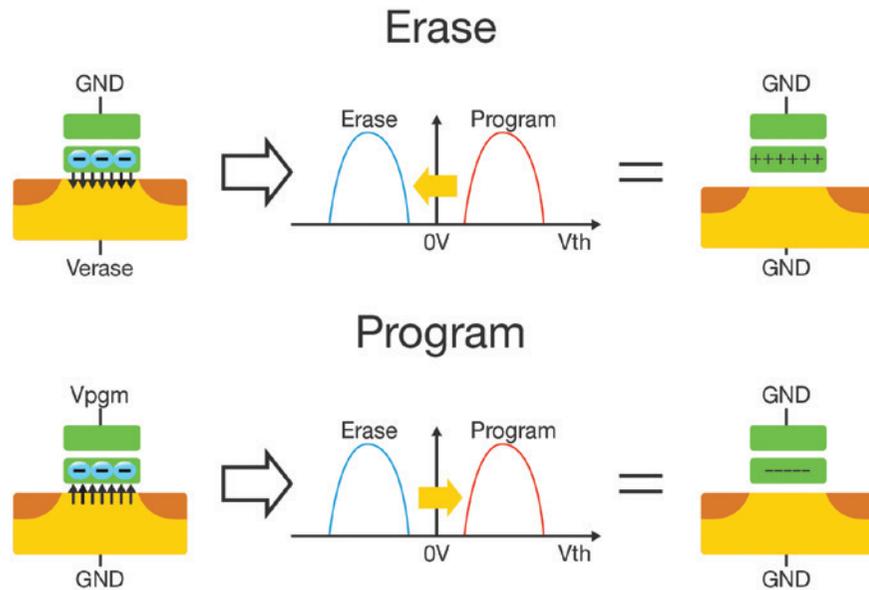
Wenn unternehmenskritische Daten auf einem Flash-Speichersystem gespeichert sind, dann müssen diese besonders gründ-

lich gelöscht werden, wenn sie nicht mehr gebraucht werden. Der Löschvorgang dauert sehr lange. Deshalb bietet Seagate ohne Extrakosten eine Selbstverschlüsselung (SED = Self-Encryption Device) für Flash-Laufwerke an. Dabei werden alle Daten automatisch beim Speichern ohne Zeitverzögerung verschlüsselt. Wird der dazu nötige Schlüssel gelöscht, dann sind die Daten ohne Zeitaufwand komplett gelöscht.

3 Flash in Speichersystemen: Oft wird versucht, Flash als direkten Ersatz für alle magnetischen Festplatten (HDDs) einzusetzen. Dies ist derzeit noch keine optimale Lösung. Flash ist erheblich teurer als Festplatten. Also macht der Einsatz von Flash nur dort Sinn wo die höhere Geschwindigkeit benötigt wird oder wo eine Anwendung mit Festplattentechnik nicht machbar ist. Die höhere Geschwindigkeit von Flash zeigt sich besonders im Lesebetrieb. Schreiben ist deutlich langsamer als Lesen, weil ein Speicherbereich vor dem Schreiben in einem eigenen Arbeitsgang zuerst gelöscht werden muss.

Eine Flashzelle wird durch den Betrieb abgenutzt, insbesondere bei Schreib- und Löschvorgängen. Deshalb muss die Anzahl der Schreibvorgänge auf ein Minimum reduziert werden, damit eine längere Nut-

All Flash requires Erase before Programming



Eine Flashzelle wird gelöscht, indem Elektronen aus dem Floating-Gate herausgezogen werden.

zungsdauer möglich wird. Schreibvorgänge mit nur wenigen Daten in unterschiedlichen Bereichen müssen vermieden werden. Zu schreibende Daten werden in einem Zwischenspeicher gesammelt und erst dann geschrieben, wenn mindestens eine Menge vorhanden ist, die einen minimalen Schreibbereich (Block) füllen kann.

Löschen und Schreiben im Flash-Speicher kann aus technischen Gründen nur blockweise (Löschen) oder seitenweise (Lesen, Schreiben) erfolgen. Die Block- oder Seitengrößen sind abhängig von der Architektur der im Gerät verwendeten Chips. Die Blöcke (für Löschen) und Seiten (für Lesen und Schreiben) sind üblicherweise unter-

schiedlich groß. In professionellen Systemen werden diese Parameter optimal angepasst. So werden Größen als Multiple von 4 KByte gewählt, wie sie in Betriebssystemen und Anwendungssystemen üblich sind.

USB-Stick: höchste Kapazität bei kürzester Nutzungsdauer

Aus Kostengründen werden nicht immer alle Parameter in allen Flash-Systemen berücksichtigt. Ein USB-Stick als Wegwerfprodukt bietet höchste Kapazität bei kürzester Nutzungsdauer (mit TLC-Chips) zu niedrigsten Preisen, ohne besonderen Aufwand für hohe Qualität oder lange Nutzungsdauer. Speichersysteme für Heim- und Büro-Anwendungen bieten mittlere Kapazität bei mittlerer Nutzungsdauer (mit MLC-Chips). In professionellen Geräten oder Systemen

werden SLC-Chips verwendet um maximale Leistung und höchste Nutzungsdauer zu erhalten. Für etwas reduzierte Anforderungen werden eMLC-Chips eingesetzt.

Diese unterschiedlichen Anforderungen werden von den Experten bei der Produktentwicklung berücksichtigt. Üblicherweise wird für mittlere Anforderungen (Heim und Büro) eine Schreibleistung von 40 GByte an Daten pro Tag über eine Zeit von drei Jahren eingeplant. Bei professionellen Speichersystemen rechnet man mit 8 TByte (8.000 GByte) an Daten je Tag für einen Zeitraum von drei Jahren. Die Anforderungen eines professionellen Systems liegen also um den Faktor 200 über den Heim- und Bürosystemen. So erklären sich die großen Unterschiede in den Entwicklungs- und Produktionskosten sowie die nennenswerten Preisunterschiede bei die-

Application Class	Workload (see JESD219)	Active Use (power on)	Retention Use (power off)	Functional Failure Requirement (FFR)	UBER Requirement
Client	Client	40° C 8 hrs/day	30° C 1 year	≤3%	≤10 ⁻¹⁵
Enterprise	Enterprise	55° C 24 hrs/day	40° C 3 months	≤3%	≤10 ⁻¹⁵

From JESD218A, Copyright JEDEC. Reproduced with permission by JEDEC

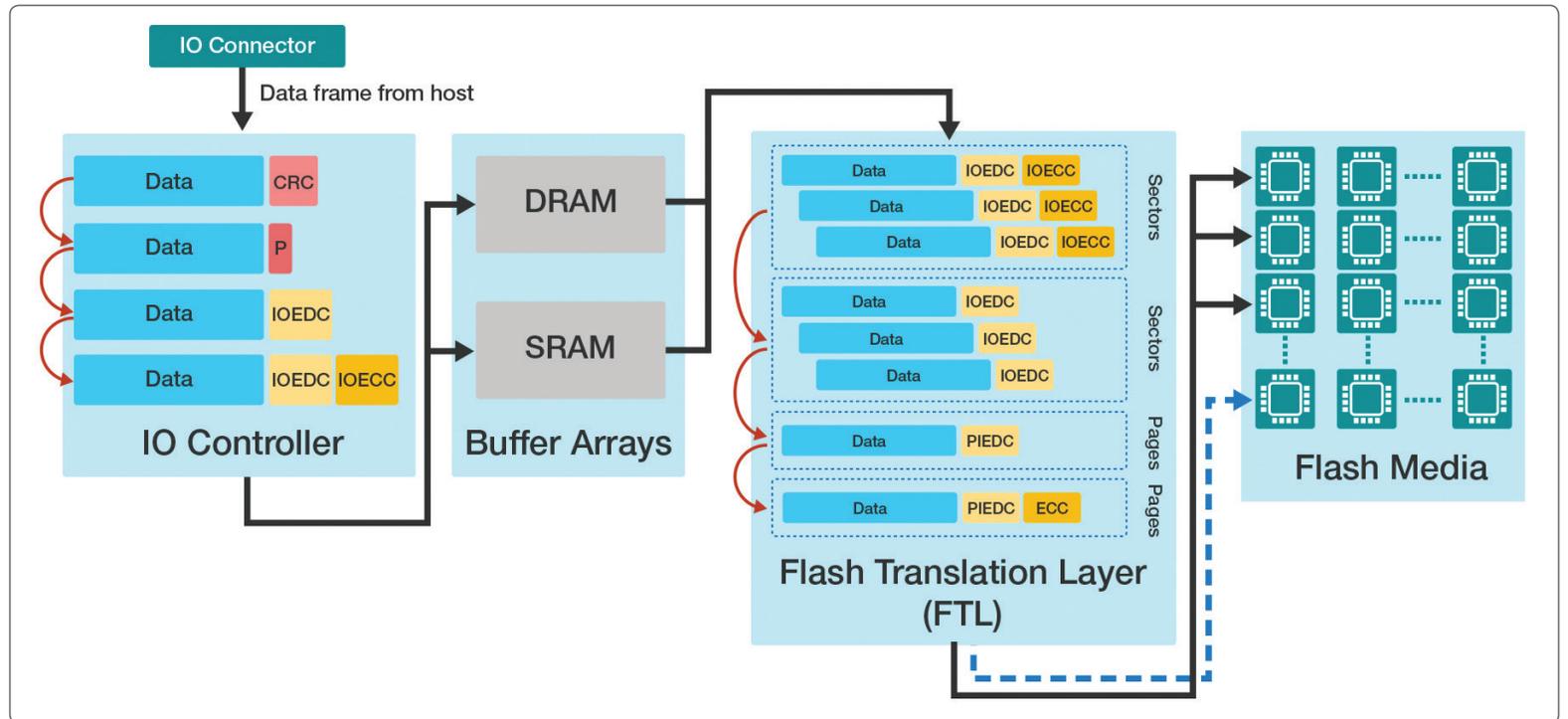
Unterschiedliche Anforderungen an Consumer-class- (Client, Büro und Heim) und Professional-class- (Enterprise, professionelle Rechenzentren) Produkte.

sen unterschiedlichen Anforderungskategorien.

Die Statistikdaten zeigen eindeutig, dass bei den großen Datenmengen in professionellen Systemen das Fehlerrisiko sehr hoch ist. Deshalb müssen entsprechend bessere Vorkehrungen in Bezug auf Zuverlässigkeit und Lebensdauererwartung eingeplant werden. Die Daten in solchen Systemen sind auch meist unternehmensentscheidend. Die wertvollsten Daten minderwertigen Speichersystemen anzuvertrauen ist also definitiv keine gute Wahl. Die anfängliche Einsparung von einigen Euro für Produkte mit schlechter Qualität bei niedrigem Preis resultiert nach kurzer Zeit in unverhältnismäßig höheren Kosten. Es zahlt sich definitiv aus, gute Produkte von Herstellern mit langjähriger Erfahrung, wie Seagate, für die Speicherung von unternehmenswichtigen Daten einzusetzen.

Es geht nur mit Sicherungscodierung

Unabhängig von der eingesetzten Flash-Technik benötigt ein professionelles Speichersystem ein durchgängiges Sicherheitssystem. Die Daten erreichen das Flash-Speichersystem zusammen mit IOEDC- und IECC-Sicherungscodierung. Hier wird IOECC entfernt und die Daten werden in Einheiten



Durchgängige Fehlerentdeckung und die Korrektur

von Seitengröße organisiert und verschlüsselt. Diese Daten erhalten noch zusätzlich eine ECC-Codierung, bevor sie dann gespeichert werden.

Der schon vor der Übertragung eingefügte DIF-Code (ANSI-Standard »data integrity field«) ist in den Datenpaketen noch enthalten. Er wird beim Lesen der Daten zusammen mit den anderen Sicherungscodes zurückübertragen. Werden keine

Fehler gefunden, dann werden die Zusatzcodes an den dafür vorgesehenen Stellen entfernt. Werden Fehler erkannt, dann können diese nach Möglichkeit mit Hilfe der zusätzlichen Codes korrigiert werden, oder es muss eine Wiederholung der Datenübertragung angefordert werden.

Flash-Speicher benötigen intern viele komplexe Funktionen, die – über entsprechende Algorithmen – die geordneten

Schreib-, Speicherungs- und Lesevorgänge von Daten ermöglichen. Vieles kann schief gehen, wenn der Entwickler diese Vorgänge nicht versteht und entsprechend einplant. Ein Flash-Speichergerät (SSD) enthält hunderte von Flash-Chips. Alle müssen vor unberechtigtem Beschreiben oder vor Nichtbeschreiben bei Stromausfall geschützt werden. Der Inhalt von Zwischenspeichern muss möglicherweise mehrmals

geschrieben werden, um die Daten sicher zu speichern. Das kann einige Millisekunden dauern. Deshalb haben professionelle Systeme interne Supercaps oder Polycaps (Energiespeicher), um genügend Energie vorrätig zu halten, bis solche Abläufe vollständig abgearbeitet sind.

4 Flash-Leistungsmerkmale nutzen:

Aus Flash-Chips aufgebaute Geräte und Systeme haben gewisse Vorteile gegenüber den herkömmlichen Speicherelementen. Die sehr hohe Lesegeschwindigkeit ist besonders zu erwähnen. Deshalb sind lese-lastige Anwendungen besonders gut für Flash-Betrieb geeignet. Die Schreibgeschwindigkeit ist ebenfalls höher als bei herkömmlichen Produkten. Allerdings müssen Schreibvorgänge sorgfältig geplant, minimiert und in großen zusammenhängenden Blöcken durchgeführt werden.

Der Einsatz von Flash ist in der Praxis deutlich komplexer, als es die Theorie erwarten lässt. Flash-Zellen und Blöcke verhalten sich sehr individuell, wie alle natürlichen Elemente. In der Theorie reicht es, einfach nur alle Schreibvorgänge sequenziell auszuführen, um ein Maximum an Nutzungsdauer zu erreichen. Sequenziell angeordnete Blöcke oder Zellen zeigen nicht immer physikalisch identische Eigenschaften.

Erfahrene Entwickler nutzen Routinen oder Algorithmen, mit denen sie das individuelle Verhalten in Bezug auf optimale Spannungspegel für Schreiben/Löschen und Lesen überwachen, oder die Zugriffszeiten zu den Signalen individuell anpassen, um den Ablauf zu optimieren.

Diese Parameter werden in besonders schnellen Zwischenspeichern abgelegt, um sie für einen optimalen Mix von Elementen mit ähnlichen Parametern auszuwerten. Diese Parameter ändern sich im Laufe der Zeit. Deshalb müssen für einen optimalen Betrieb entsprechende Anpassungen vorgenommen werden. Dazu sind offensichtlich fundiertes Fachwissen und langjährige Erfahrung für einen optimalen Flash-Speicherbetrieb nötig.

Es genügt also für optimierte Speichersysteme nicht einfach nur die Theorie unverändert in praktischen Produkten anzuwenden. Diese individuelle Behandlung der Flash-Elemente gibt es nur in professionellen Speichersystemen. Für Wegwerfprodukte oder für Heim- und manche Bürosysteme verschwendet niemand Zeit oder Betriebsmittel, um solche Systeme zu optimieren.

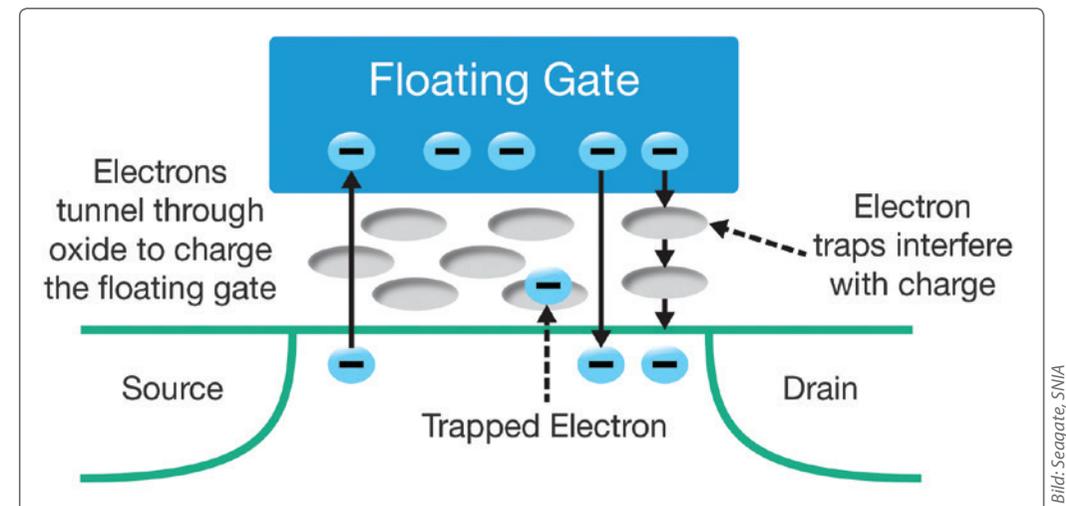
5 Vermeidung von Problemen: Die wichtigste Maßnahme beim Einsatz von Flash ist es die Anzahl der Schreib-Lösch-Vorgänge

(Programmier-Löschzyklen) zu reduzieren, da die Flash-Zellen abgenutzt werden, was ihre Nutzungsdauer verkürzt.

Es gibt viele Verfahren, mit denen, ohne Anwenderbeteiligung die Zahl von ungewollten Zyklen intern verringert werden können. Die Abläufe sind nur auf Seiten- (Schreiben) oder Blockebene (Löschen) möglich. Deshalb müssen individuelle Schreibdaten vorübergehend in Zwischenspeichern gehalten und erst später in Vielfachen von Blöcken geschrieben werden, weil das nötige Löschen nur auf Blockebene möglich ist. Zu löschende Daten innerhalb eines Blocks werden typischerweise

als gelöscht markiert, aber tatsächlich nicht physikalisch gelöscht. Dies geschieht erst, wenn der ganze, fast der ganze Block oder ein Vielfaches zum Löschen markiert ist. Noch enthaltene gültige Daten werden temporär zur späteren endgültigen Speicherung im Zwischenspeicher gelagert.

Der somit »leere« Block wird dann in einem Arbeitsgang gelöscht. Dieser leere Block kann nun unmittelbar beschrieben werden, wodurch die Schreibdauer verkürzt wird. Demgegenüber müssen Schreibdaten – wie oben beschrieben – für einen noch nicht gelöschten Block, zusammen mit den noch gültigen Daten in diesem Block, vor-



Elektronen können im Floating-Gate (FG) oder in der Isolationsschicht (Oxid) ungewollt eingelagert werden.

übergehend in einem Zwischenspeicher geparkt werden. Der Block muss dann erst gelöscht werden, bevor diese zwischengespeicherten Daten eingeschrieben werden können, idealerweise mit genügend Daten, um diesen Block gleich in einem Ablauf vollständig zu füllen. Diese Schreibverzögerung (write penalty) wird auch als Schreibklippe (write cliff) bezeichnet.

Schreibvorgänge werden zur Geschwindigkeitssteigerung oft gleichmäßig über mehrere NAND-Chips auf viele NAND-Busse innerhalb eines Speichergeräts (SSD) verteilt. In professionellen Speichersystemen werden zur Leistungssteigerung weitere ausgeklügelte Verfahren zur Vermeidung von Kollisionen auf Chip- oder Busebene angewandt.

Diese professionell ausgereiften Verfahren gibt es üblicherweise nicht für Wegwerfprodukte (USB-Sticks) und Verbrauchereingegeräte. Die Elemente innerhalb von Flash-Chips werden ständig im Zuge des technischen Fortschritts kleiner, um die Kapazitäten und Geschwindigkeiten zu erhöhen. Deshalb werden wegen der höheren Fehleranfälligkeit bei den kleineren Strukturen die bestmöglichen, gut durchdachten Fehlerkorrektur-Verfahren benötigt, um die wertvollen Daten auf der höchsten Qualitätsebene zu sichern.

So funktioniert Speicherbereinigung (garbage collection)

Die Funktion Speicherbereinigung (GC) gibt es schon sehr lange als automatisches Speicherungsverfahrens für APL, BASIC oder LISP. Außerdem wird GC beispielsweise in JAVA, .NET oder anderen Programmpaketen genutzt. In anderen Sprachen wie C oder C++ muss GC manuell durchgeführt werden. Das ist fehleranfällig. Die Speicherbereinigung ist ein sehr komplexes Verfahren mit ganz besonderen Auswirkungen für Flash. Die Grundidee ist die systematische Suche nach nicht mehr benötigtem Speicherplatz, um diesen wieder für andere Anwendungen zur Verfügung zu stellen.

Einige GC-Routinen führen auch in unbestimmten Zeitabständen eine Defragmentierung durch. Das erzeugt zusätzliche Schreib-Lösch-Zyklen, benötigt Rechnerzeit, ist zeitlich nicht vorhersehbar und daher nicht-deterministisch mit Problemen in Echtzeitanwendungen. In Mehrfachrechnersystemen oder in parallelläufigen Systemen muss die Speicherbereinigung üblicherweise als einziges, nicht-unterbrechbares Programm ablaufen. Dieser Ablauf verschwendet sehr viel Rechnerzeit.

Besondere Probleme gibt es bei der Speicherbereinigung in Flash-Systemen, weil die aufgesammelten Speicherbereiche

normalerweise nicht genau innerhalb von zu löschenden Blöcken anfallen, und diese Blöcke zuerst gelöscht werden müssen, bevor sie wiederverwendbar sind. Traditionelle Betriebs- oder Dateisysteme haben keine Verfahren, um diese blockweise Verarbeitung zu berücksichtigen.

TRIM-Befehl hilft viel, aber nicht bei allem

Eine gewisse Möglichkeit bietet der TRIM-Befehl in neueren Linux- und Windows-Systemen. Dieser Befehl berücksichtigt automatisch die Besonderheiten von Schreib-Lösch-Zyklen in Flash-Systemen. Die Speicherbereinigung in Flash-Systemen unterscheidet sich von ähnlichen Verfahren für Festplattensysteme (HDDs). TRIM kann nicht alle Flash-Speicherbereinigungsprobleme lösen. Wird eine Verschlüsselung auf Softwarebasis eingesetzt, dann werden durch TRIM die zu bearbeitenden Blöcke aufgedeckt. TRIM kann gemäß dem entsprechenden ANSI-Standard nicht in eine Warteschlange gestellt werden. Damit wird der TRIM-Befehl bei nicht-sachgemäßer Nutzung zu einer massiven Leistungsbremse.

Für SATA 3.1 wurde ein Warteschlangentauglicher TRIM-Befehl definiert. Je nach Betriebssystem, Chipsatz oder RAID-Verfah-

ren gibt es weitere Einschränkungen. In Speichersystemen auf Basis von SCSI gibt es ähnliche Befehle unter den Bezeichnungen UNMAP oder WRITE SAME. RAID- und Speicherbereinigungsverfahren, die aus herkömmlichen Systemen übernommen werden, erzeugen unverhältnismäßig viele Schreib-Lösch-Zyklen in Flash-Systemen. Professionelle Experten, wie beispielsweise bei Seagate, wählen sorgfältig erprobte Verfahren passend für Flash-Systeme aus, und optimieren diese für Langzeit-Nutzung und optimale Amortisierung.

So funktioniert der Nutzungsausgleich (wear-leveling)

Mit den Nutzungsausgleichsverfahren werden die Schreib-Lösch-Zugriffe gleichmäßig auf den Flash-Speicher verteilt. So wird die Nutzungsdauer des Flash-Speichers verlängert. Bei diesem Verfahren wird die Anzahl der Schreib-Lösch-Zyklen für jeden Block berücksichtigt, und die Zuordnung von logischen zu physikalischen Blöcken mit speziellen Algorithmen dynamisch ausgeglichen, damit die Abnutzung gleichmäßig über den ganzen Speicherbereich verteilt wird. Durch die Bereitstellung von mehr Speicher als nominell vorhanden (over-provisioning), ist es möglich, ausgefallene Speicherzellen im laufenden Betrieb auto-

matisch zu ersetzen und somit die Nutzungsdauer zu verlängern.

Der Nutzungsausgleich kann dynamisch oder statisch erfolgen. Bei der dynamischen Methode werden die logischen Blockadressen (LBAs) vom Betriebssystem den physikalischen Blöcken im Flash zugeordnet. Jedes Mal, wenn das Betriebssystem Daten überschreibt, wird der ursprüngliche Block als ungültig markiert, und ein neuer Block genutzt. Allerdings werden Blöcke, die niemals überschrieben werden, auch nicht abgenutzt. Damit gibt es noch ungenutzte gute Blöcke, auch wenn das Laufwerk schon ausgefallen ist.

Die statische Methode nutzt ebenfalls die Zuordnung von logischen Blockadressen (LBAs) zu physikalischen Blöcken. Zusätzlich werden »ruhende« Blöcke, die nur ganz selten überschrieben werden, auch überwacht. Diese werden dann an eine andere Stelle kopiert, damit diese besonders wenig genutzten Blöcke für andere Daten nutzbar werden. Mit dieser Rotationsmethode bleiben alle Speicherblöcke gleichmäßig bis zum Ende der Nutzungsdauer einsatzfähig.

6 Optimierung der Betriebsmittel: Die wichtigsten Vorteile der Flash-Technik sind Geschwindigkeit, sehr kleine Latenzzeiten

und niedriger Energieverbrauch. Daraus ergibt sich, dass der sehr teure Flash-Speicher eigentlich nur dort eingesetzt wird, wo diese Vorteile zum Tragen kommen. Das setzt eine ausgewogene Speicherarchitektur voraus. Es wäre eine Verschwendung von Betriebsmitteln, Flash dort einzusetzen, wo es gegenüber langzeiterprobten herkömmlichen Festplatten (HDDs) keine Vorteile bringt.

Für große, professionelle Speichersysteme bedeutet dies die Nutzung von Hybrid-Systemen mit einer Mischung aus Flash-Geräten mit hoher Leistung (hohe Geschwindigkeit, niedrige Latenzzeit) und zur Archivierung optimierten Festplatten für wenig genutzte Daten sowie weiteren Gerätekategorien dazwischen. Oft werden bei den Festplatten SAS-Geräte für schnelle Datenübertragung und SATA-Geräte an SATA-Schnittstellen oder an SAS-Anschlüssen mit SATA-Funktionalität im gleichen System eingesetzt. So können die Systeme bedarfsgerecht erweitert, aufgerüstet oder die Daten umgelagert werden, ohne zu viel oder zu wenig an Speicherplatz und Extrakosten zu haben.

Wie wir erfahren, verkürzt sich die Nutzungsdauer von Flash durch Schreibzyklen. Es ist deshalb vorteilhaft, das Lesen-zu-Schreiben-Verhältnis in jeder größeren Anwen-

dung zu ermitteln. Typische Datenanwendungen haben oft ein Verhältnis von 30 Prozent Schreiben zu 70 Prozent Lesen. Das kann optimiert werden, wenn schreibintensive Anwendungen (VDI etc.) auf SLC-Flash und leseintensive Anwendungen auf eMLC- oder MLC-Flash geleitet werden. Das setzt voraus, dass diese Abläufe bekannt sind.

Bei IOPS sind Flash-Speicher um Längen besser

Wie bereits erwähnt, sind die Geschwindigkeitsunterschiede in Flash-Systemen zwischen Lesen und Schreiben sehr groß. Oft ist die Schreibgeschwindigkeit nur halb so hoch wie die Lesegeschwindigkeit. Das sollte berücksichtigt werden, wenn Funktionen, Aufgaben oder Speicherbereiche auf Speichergeräte verteilt werden. Ähnliche Verhältnisse sind für die Ein-/Ausgabevorgänge pro Sekunde (IOPS) zu berücksichtigen.

Bei den Ein-/Ausgabevorgängen ist der Unterschied zwischen Flash und magnetischer Speicherung besonders groß. Die Flash-Technik (SSD) ist üblicherweise um mehrere Größenordnungen schneller als die magnetische Technik (HDD). So ist es sinnvoll, Datenzugriffe mit großen Mengen von zufällig verteilten Ein-/Ausgabeabläufen auf Flash-Systemen zu speichern, und

solche mit sequenziellen Datenzugriffen auf Festplatten. Die Schwierigkeit bei der Verteilung der Datenströme besteht darin, die Verteilung der Abläufe in einer bestimmten Anwendung zu kennen.

Viele IT-Experten kennen die für eine Entscheidung benötigten Parameter nicht. Es ist dringend erforderlich, die Verteilung der Zugriffsarten und der Datenströme sorgfältig zu ermitteln. Auswahlkriterien sind Durchlaufverzögerung, Ein-/Ausgabeabläufe je Sekunde, Schreib-/Leseverhältnis, Blockgröße und das Verhältnis zwischen wahlfreien und sequenziellen Zugriffen. Das Ergebnis dieser Analyse sind besonders optimierte Systeme zu erheblich niedrigeren Kosten.

Komprimierung oder Deduplikation gefährdet Flash-Leistungskonstanz

In Flash-Systemen benötigen alle Zugriffe die gleiche Zeit, weshalb solche Systeme eine konstante Leistung (deterministisch, gleichmäßig) zur Verfügung stellen. Allerdings wird diese Leistungskonstanz durch Verfahren wie Komprimierung oder Deduplikation gefährdet, wenn diese nicht unter Berücksichtigung der Anwendungsart sorgfältig geplant werden.

Ein weiteres Kriterium ist auch der Ort, wo diese Abläufe im Datenkanal durchge-

führt werden. Diese Verfahren sollten während der Übertragung oder zumindest vor der Speicherung im Flash erfolgen. Das ist nötig, um die Anzahl der Schreibzyklen für eine lange Nutzungsdauer zu minimieren.

In Flash-Systemen wird üblicherweise mehr Speicher zur Verfügung gestellt als nominell verfügbar ist. In diesem für den Anwender nicht zugänglichen Bereich werden bereits gelöschte Blöcke zum sofortigen Be-

schreiben in einem virtuellen Bereich vorgehalten. Dieser Bereich bildet auch eine Reserve, falls defekte Blöcke im Anwenderbereich ersetzt werden müssen. Weitere Parameter sind die Granularität dieser Abläufe, ihr temporärer Speicherbedarf und die Verarbeitungszeit (Latenz) für ihre Ausführung. Ein sehr großer Teil des Geschwindigkeitsvorteils von Flash-Speichersystemen könnte durch die vorgenannten Verfahren

zunichte gemacht werden, wenn die Experten nicht sorgfältig bei der Entwicklung vorgehen.

Fazit: Flash sollte gut geplant eingesetzt werden

Für eine lange Zeit wird Flash deutlich (einige Arten erheblich) teurer sein als magnetische Speicherung. So ist es sinnvoll, jede dieser Speichertechnologien so einzu-

setzen, dass sich der bestmögliche Nutzen in einem optimalen System ergibt. Das muss von professionellen Experten unter Berücksichtigung von individuellen Anforderungen sorgfältig geplant werden. Das Resultat einer entsprechenden Planung ist ein über lange Zeit zuverlässiges Speichersystem bei optimalem Kapitaleinsatz bei niedrigen laufenden Kosten.

Hermann Strass

Sonderdruck

Veröffentlichung vom 05.02.2015

So lassen sich die Vorteile von Flash richtig nutzen

Powered by:

Seagate Technology GmbH



Messerschmittstr. 4

D-80992 München

Tel. +49 (0) 89/14 30 50 00

Fax +49 (0) 89/14 30 51 00

E-Mail: DiscSupport@Seagate.com

Web: www.seagate.de

speicherguide.de GbR

Karl Fröhlich, Ulrike Haak, Engelbert

Hörmannsdorfer

Bahnhofstr. 8, D-83727 Schliersee

Tel. +49 (0) 80 26/928 89 96

E-Mail: redaktion@speicherguide.de

Chefredaktion:

Karl Fröhlich (verantwortlich für den redaktionellen Inhalt)

E-Mail: redaktion@speicherguide.de

Projektleitung:

Engelbert Hörmannsdorfer

Redaktion:

Hermann Strass

Layout/Grafik:

Uwe Klenner, Layout und Gestaltung,

Rittsteiger Str. 104, 94036 Passau,

Tel. 08 51-9 86 24 15

www.layout-und-gestaltung.de

Titelbild: © iStockphoto.com/teekid

Mediaberatung:

Claudia Hesse,

Tel. +41 (0) 41 - 780 04 86

E-Mail: media@speicherguide.de

Urheberrecht:

Alle in diesem Sonderdruck erschienenen Beiträge sind urheberrechtlich geschützt. Alle Rechte

(Übersetzung, Zweitverwertung) vorbehalten.

Reproduktion, gleich welcher Art, sowie elektronische Auswertungen nur mit schriftlicher Genehmigung der Redaktion.

Aus der Veröffentlichung kann nicht geschlossen werden, dass die verwendeten Bezeichnungen frei von gewerblichen Schutzrechten sind.

