

Link: <https://www.channelpartner.de/a/unter-der-ssd-haube,2555854>

Solid State Discs: Die Technik

Unter der SSD-Haube

Datum: 17.03.2014
Autor(en): Klaus Manhart

SSDs erscheinen auf den ersten Blick simpel und nicht sehr anders zu arbeiten als herkömmliche Festplatten. Doch das ist ein Trugschluss: Damit die Flash-Speicherlaufwerke langfristig schnell und zuverlässig funktionieren, betreiben sie einen enormen Aufwand bei der Datenorganisation.

SSDs funktionieren ganz anders als HDDs: Im Gegensatz zu Festplatten mit ihren beweglichen Teilen sind die internen Komponenten von SSDs nicht beweglich und enthalten keine anfälligen mechanischen Bauteile. In erster Linie bestehen SSDs aus einem Controller und mehreren NAND-Flash-Bausteinen als physischem Speicher.

Der Controller als Steuereinheit kontrolliert die Kommunikation zwischen den SSD-Speicherzellen und dem Computer. Er kann die Flash-Bausteine parallel ansprechen, vervielfacht damit den Datendurchsatz und bildet so einen zentralen Knotenpunkt, den alle Daten zwangsweise passieren müssen.

Flash-Speicherbausteine, die zweite wesentliche SSD-Komponente, sind nichtflüchtig. Das heißt, die in ihren Zellen gespeicherten Informationen gehen nicht verloren, wenn die Stromzufuhr abgebrochen wird. Jedes einzelne Bit wird in Form einer elektrischen Ladung auf einem "Floating-Gate-Transistor" gespeichert.

Alle gängigen SSDs basieren auf NAND-Speicher-Chips. Der Ausdruck "NAND-Technik" bezieht sich auf die serielle Anordnung der einzelnen Speicherzellen. Grundsätzlich unterscheidet man dabei zwischen SLC- oder MLC-NANDs. SLC steht für "Single Level Cell" und bedeutet, dass eine Flash-Zelle auch nur ein Bit speichert. Dieses definiert sich durch einen festgelegten Spannungslevel. Beim Flash-Typ "Multi Level Cell" (MLC) lassen sich in einer Zelle aktuell zwei oder drei Bit-Zustände speichern.

Je nach Kapazität sind heute bis zu 64 Speicherbausteine üblich, diese sind teilweise auch noch in Sandwich-Bauweise auf dem Datenträger verbaut. Gerade die Kombination und die Qualität der Controller und Speicherbausteine haben enormen Einfluss auf die Geschwindigkeit und Zuverlässigkeit der SSD.

Aufbau von Flash-Speicher: Pages und Blocks

Damit die SSD langfristig schnell und zuverlässig funktioniert, führt der SSD-Controller ständig komplexe Operationen und teils aufwendige Berechnungen durch. Sie optimieren die Performance und Langlebigkeit der SSD. Um zu verstehen, was bei SSDs unter der Haube abläuft und warum bestimmte Funktionen und Algorithmen so wichtig sind, muss man sich mit den Charakteristika von NAND Flash-Speicher vertraut machen.

NAND-Flashspeicher arbeitet grundsätzlich page- und blockorientiert. Die Page ist die kleinste Speichereinheit, in der mehrere Speicherzellen zusammengefasst sind. Sie hat bei den meisten SSDs eine Größe von 4 KByte. Die Pages selbst werden ebenfalls wieder gebündelt und zu einem Block gruppiert. Ein solcher Block umfasst bei größeren Speicherbausteinen 128 Pages.

Die Pages lassen sich einzeln beschreiben. Das ist allerdings nur möglich, wenn die Page physisch leer ist. Enthält eine Page bereits Daten, müssen diese Daten zuerst gelöscht werden. Aufgrund der Gruppierung ist ein Löschen einer Page nur über das Löschen des Blocks möglich, in dem sie liegt. Eine SSD kann deshalb nur ganze Blocks löschen. Somit werden immer 128 Pages auf einen Schlag gelöscht.

Generell verteilt der Flash-Controller die Schreibzugriffe gleichmäßig über alle Zellen. Die Daten in Zellen, die mit selten veränderten Daten wie Betriebssystem und Programmen belegt sind, werden ab und zu umgeschichtet, um so wieder weniger stark genutzte Zellen zu bekommen.

Reserve zur Datenverwaltung: Over Provisioning

Mit zunehmenden Lösch- und Wiederbeschreib-Vorgängen benötigt eine SSD immer länger für die Datenspeicherung. Der Grund: Während bei magnetischen Festplatten die gelöschten Daten einfach überschrieben werden, braucht eine SSD dafür zwei Schritte. Wenn eine Datei auf einer SSD über das Betriebssystem gelöscht wird, dann wird auf dem SSD-Laufwerk nur der Eintrag dieser Datei im Inhaltsverzeichnis des Speichers gelöscht - die eigentlichen Daten befinden sich nach wie vor in den Speicherzellen der SSD.

Mit der Zeit sind so immer weitere Bereiche auf der SSD mit eigentlich gelöschten Daten gefüllt. Beim erneuten Schreiben muss eine SSD diese Zellen deshalb erst leeren, was die Schreibgeschwindigkeit etwas verlangsamt.

Was passiert, wenn die SSD fast voll beschrieben ist? Wenn kaum mehr freier Speicherplatz vorhanden ist, so gibt es natürlich weniger freie Blöcke. Damit wird die SSD gezwungen, freie Blöcke zur Verfügung zu stellen. Nur dann kann das Laufwerk hohe Performance sicherstellen. Dieser Vorgang, also das ständige Optimieren der Datenhaltung, benötigt Zeit und freien Speicherplatz. Das wird gegebenenfalls in Form geringerer Performance spürbar.

Aus diesem Grund kommt eine Strategie namens Over Provisioning zum Einsatz: Dabei wird eine bestimmte Menge an Speicherplatz speziell für SSD-interne Zwecke reserviert und kann vom Anwender nicht genutzt werden. In diesem reservierten Speichersegment werden die "Garbage Collection" und andere Wartungsfunktionen durchgeführt. Die Garbage Collection selbst wird auf dem ganzen verfügbaren Speicherbereich durchgeführt. Die SSD behandelt den Over-Provisioning-Teil als großen Pool. Diese Mechanismen halten die Maximalleistung einer SSD aufrecht.

Datenmüll beseitigen: Garbage Collection

Da bei SSDs jede Zelle vor dem nächsten Überschreiben zunächst gelöscht werden muss, wurden die Verfahren TRIM und Garbage Collection entwickelt. Garbage Collection sorgt für die automatische Speicherbereinigung. Der Mechanismus löscht nicht mehr benötigte Daten bereits vor einer neuen Beschreibung aus den Speicherzellen. Damit steht dieser Speicherbereich von Haus aus für zukünftige Schreibvorgänge bereit.

Technisch betrachtet sammelt Garbage Collection die in Blöcken verbliebenen Pages und fasst diese blockweise zusammen. Die als ungültig (invalid) markierten Pages sollen dabei möglichst viele neue freie Blöcke ergeben. Diese Bereinigung entspricht also bildlich einer Müllabfuhr, welche für eine Minimierung des Speicheranspruchs eines Prozesses steht. Damit lassen sich die ineffiziente Nutzung des bereits zugeteilten Speichers sowie eine Verlangsamung der Datentransfers vermeiden.

Es gibt zwei Methoden der Garbage Collection: Entweder diese arbeitet passiv im Hintergrund, wenn die SSD beziehungsweise der Controller gerade wenig oder gar nichts zu tun hat. Oder aber die Garbage Collection findet aktiv statt, wenn die SSD gerade Schreibdaten bearbeitet. Samsung SSDs implementieren die letztere, aktive Variante mit Abwandlungen. Das erlaubt den Laufwerken, aktiv Daten zu verwalten und freie Blöcke "on the fly" zu generieren.

Zuarbeiter für den Müllschlucker: TRIM

Ein zweiter Mechanismus, der die Leistungsfähigkeit von SSDs sichert, ist TRIM. TRIM beugt dem Performance-Verlust bei SSDs entgegen. Diese Betriebssystemfunktion benachrichtigt den SSD-Controller umgehend, wenn Daten gelöscht wurden oder nicht mehr gebraucht werden. Mit anderen Worten: TRIM arbeitet der Garbage Collection direkt zu.

Die Funktion hat ihren Ursprung in der Tatsache, dass Daten bislang auf Festplatten abgelegt wurden. Dabei handelt es sich um rotierende magnetische Speicher, die die Beschränkungen hinsichtlich Schreiben und Löschen nicht haben. Gelöschte Daten werden schlicht als gelöscht markiert, und die betreffenden Pages können jederzeit neu beschrieben werden. Damit war es bei Festplatten auch nicht notwendig, dass das Betriebssystem gelöschte Daten an das Laufwerk kommuniziert.

Mit der Einführung von SSDs wurde es allerdings notwendig, dass sich das System und die SSD darüber austauschen können, welche Daten valide sind und welche nicht. Durch TRIM wird die SSD darüber informiert, wenn Daten "invalid" sind, so dass die Garbage Collection gelöschte Pages als solche markiert und solche invalide Daten nicht unnötigerweise herumschieben muss. Sobald ein kompletter Block nur noch aus invaliden Pages besteht, wird dieser als frei markiert und die SSD kann diesen Block neu vergeben.

Moderne Betriebssysteme wie Windows 7 oder 8 senden den TRIM-Befehl automatisch an den SSD-Controller, und zwar immer, wenn eine Datei gelöscht wird. Auf älteren Windows-Systemen wird TRIM meist nicht unterstützt beziehungsweise kann nur manuell mit Software-Tools ausgelöst werden.

Fehler korrigieren: Bad Block Management

Fehlerkorrekturverfahren (Error Correcting Code - ECC) dienen dazu, Defekte bei der Speicherung und Übertragung von Daten zu erkennen und wenn möglich, zu korrigieren. Sie kompensieren den Ausfall von Einzelbits im täglichen Betrieb und vermeiden so Datenverlust. Bei SSDs nimmt die Anzahl der Fehler zu, wenn eine Flash-Zelle in die Nähe ihrer maximalen Schreib-/Löschzyklen kommt.

Sobald der Controller einen Lesefehler registriert, wird er den ECC-Mechanismus einbinden und die Daten rekonstruieren. Wenn die Wiederherstellung nicht mehr funktionieren sollte, wird das "Bad Block Management", also die Verwaltung fehlerhafter Sektoren, den gesamten Block aus dem Speicherbereich entfernen. Dateien aus fehlerhaften Zellen werden in funktionierende Zellen verschoben. Die defekte Zelle wird dann für die künftige Speicherung von Daten ausgeschlossen, und eine neue nimmt dafür ihren Platz ein.

Bad Blocks können während des Lesens, Programmierens oder Löschens von NAND Flash festgestellt werden. Sie werden aktiv verwaltet, um zu jeder Zeit die erwartete Leistungsfähigkeit und langfristige Zuverlässigkeit von SSDs sicherzustellen. So werden Flashzellen bereits dann als fehlerhaft markiert, wenn sie stark abgenutzt sind.

Mit der Betriebsdauer einer SSD wird damit zwar theoretisch der nutzbare Speicherplatz weniger, Datensicherheit und Datentransfergeschwindigkeit bleiben aber auf gleichem Niveau. Üblicherweise werden Samsung SSDs ohne jegliche Formatierung verkauft. Der erwähnte Reservebereich befindet sich im Bereich des Over Provisionings, den jede SSD hat.

Schreibvorgänge verteilen: Wear Leveling

Als ein Hauptproblem von SSDs gilt die begrenzte Lebensdauer. Je nach Flash-Typ geht eine Speicherzelle nach rund 1.000 bis 100.000 Speichervorgängen kaputt. Zwar wird nicht der gesamte Speicher zerstört. Doch es machen sich Verschleißerscheinungen bemerkbar, die auch zu Datenverlusten führen können.

Um dieses Problem in den Griff zu bekommen, implementieren alle SSD-Controller so genannte Wear-Leveling-Mechanismen. Diese stellen sicher, dass alle Schreibvorgänge auf alle vorhandenen NAND-Zellen gleichmäßig verteilt werden. Damit werden einzelne Zellen nicht über die Maßen belastet, so dass sie frühzeitig ausfallen können. Würden zu viele Zellen versagen, dann würde der SSD-Controller den gesamten Block abschalten (wie oben beschrieben). Die Anzahl der Reserveblöcke ist jedoch begrenzt, so dass dem Wear Leveling eine wichtige Bedeutung zukommt.

Alle der oben genannten Funktionen, mit Ausnahme von TRIM auf älteren Windows-Systemen, arbeiten transparent und erfordern keine Eingriffe seitens des Anwenders. Während die Implementierung der Funktionen von Hersteller zu Hersteller variiert, verfügen praktisch alle modernen SSDs über solche Funktionen. Wear Leveling und ECC, um die Lebenserwartung zu maximieren und die Datenintegrität zu gewährleisten, und TRIM sowie Garbage Collection, um hohe Performance jederzeit abrufbar zu machen, sorgen für Qualität und eine gute "User Experience".