

Universität Hamburg
Fachbereich Informatik
Seminar: Speicher und Dateisysteme
Leitung: Dr. Michael Kuhn

**SSD – Eine Einführung in die Funktionsweise
und Technik moderner Festplatten**
26.09.2016

Oskar Dörfler (6. Semester)
Email: 3doerfle@informatik.uni-hamburg.de
Matrikelnummer: 6523882

Inhaltsverzeichnis

1. Einleitung.....	3
2. Flash Speicher.....	3
2.1 P/N-Dotierung.....	4
2.2 Der Tunneleffekt.....	6
2.3 Die Speicherzelle – Aufbau.....	7
2.4 Die Speicherzelle – Lesen.....	8
2.5 Die Speicherzelle – Schreiben.....	9
2.6 Die Speicherzelle – Löschen.....	10
2.7 SLC, MLC, TLC.....	10
2.8 Speicherzellen Anordnung.....	12
3. Software und Datenorganisation in SSD's.....	14
3.1 Wear Leveling.....	14
3.2 Fehlerquellen Minimierung.....	15
3.3 Bad Block Management.....	16
3.4 Garbage Collection.....	16
4. Probleme heutiger SSD's.....	17
4.1 Write Amplification.....	17
4.2 Write Cliff.....	18
5. Zusammenfassung.....	20

1. Einleitung

Datenspeicherung ist in unserer heutigen Welt eines der am wichtigsten Themen. Ob es um die Einkaufsliste auf dem Handy geht oder um riesige Datenansammlungen auf Servern, wie die von Facebook, Datenspeicherung ist immer ein Thema über das sich Gedanken gemacht werden muss.

Zu Beginn für die Speicherung von großen Datenmengen standen die HDD's (Hard Disk Drive), welche auf magnetisch manipulierbaren Scheiben ihre Daten ablegten. Der größte Nachteil dieser Magnet Scheiben ist, dass sie mechanisch bewegliche Komponenten beinhalten darunter der Schreib-/Lesekopf und ein Motor für die Drehung der Platten.

Dieser Punkt macht HDD's träge und anfällig für physikalische Einwirkungen, wie Erschütterungen.

Durch das ständige bewegen des Schreib-/Lesekopfes vergeht zudem viel Zeit, was die Zugriffszeiten auf Daten stark erhöht.

Nun kamen jedoch immer neuere Technologien auf den Markt, unter anderem große Mainframes und die ersten Digitalkameras. Diese Technologien konnten sich jedoch nicht auf so lange Antwortzeiten, wie es bei HDD's der Fall ist verlassen, sondern brauchten wesentlich schnelleren Speicher.

So entstanden die ersten Speicher auf elektronischer Basis, welche keine mechanischen Bauteile mehr beinhalten und es ermöglichten Daten schnell und sicher abzulegen und wieder abzurufen.

Heutzutage ist die bekannteste Form solcher elektronischen Speichermedien die SSD (Solid State Drive), deren Datenspeicherung auf NAND verknüpften Flash Speicher basiert.

Um diese Technologie genau zu verstehen müssen zuerst die einzelnen physikalischen Funktionen der inneren Komponenten einer SSD betrachtet, sowie im Anschluss die auf diesen Komponenten laufende Software verstanden werden.

2. Flash Speicher

Wie in der Einleitung erwähnt basieren heutige SSD's in der Regel auf NAND verknüpften Flash Speicher. Was das genau bedeutet wird im folgendem Abschnitt erklärt.

Flash Speicher ist eine geordnete Struktur von einzelnen elektronischen Speicherzellen. In diesen Speicherzellen werden die Daten in Form von nicht flüchtigen Ladungen in einem sogenannten Floating Gate gespeichert.

Diese Speicherzellen werden auch MOSFET's genannt was für metal-oxide-semiconductor field-effect transistor steht.

Bevor jedoch auf den Aufbau und die Funktionsweise dieser näher eingegangen wird müssen vorerst zwei physikalische Prinzipien geklärt werden um die

Datenspeicherung zu verstehen.

Zum einen ist dies der Vorgang der Dotierung, welcher eine Verunreinigung eines Trägermaterials darstellt um die Leitfähigkeit dieses zu manipulieren.

Zum anderen ist es der Tunneleffekt welcher es Elektronen ermöglicht scheinbar unüberwindbare Hindernisse doch zu überwinden.

2.1 P/N-Dotierung

Jeder weiß, dass in der Halbleiterindustrie Silizium eine extrem große Rolle spielt, denn in jedem Computerchip steckt es.

Silizium ist ein chemisches Element der 4. Hauptgruppe damit hat es 4 Valenzelektronen (Außenelektronen, welche Bindungen mit benachbarten Atomen eingehen können). Liegt nun Silizium als festes Material bei Raumtemperatur vor sind die einzelnen Siliziumatome mit je 4 anderen Siliziumatomen in einer Kristallinen Struktur verbunden. Das kommt daher, dass jedes Atom das Bestreben hat in einem sehr stabilen Zustand zu sein und dies ist bei 8 Außenelektronen erreicht. Daher teilen die Siliziumatome mit ihren Nachbarn ihre 4 Außenelektronen, gehen also 4 kovalente Bindungen (Elektronenpaarbindung) ein, damit jedes Atom 8 Stück auf der äußeren Schale besitzt. Bei höheren Temperaturen lösen sich jedoch einige dieser Doppelelektronenbindungen und es werden Ladungsträger frei welche es ermöglichen einen Strom fließen zu lassen.

Da man aber sein PC nicht jedes mal vor dem benutzen auf die Herdplatte stellen will braucht es einen Weg welcher es ermöglicht diese Halbleiter auch bei niedrigeren Temperaturen leiten zu lassen. An dieser Stelle kommt die Dotierung ins Spiel. Dotierung beschreibt nämlich den Vorgang des Verunreinigens eines Halbleiters mit einem anderen Stoff.

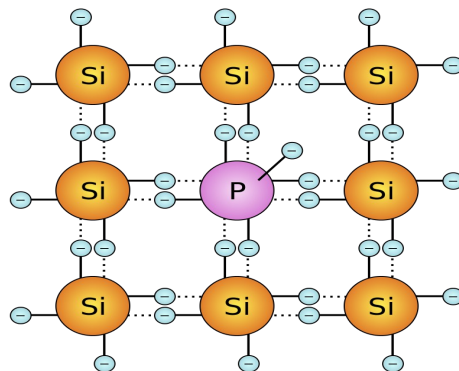


Abbildung 1: n-Dotierung im Siliziumkristallgitter mit Phosphor
[<https://de.wikipedia.org/wiki/Dotierung>]

In Abbildung 1 sieht man die Verunreinigung von Siliziumatomen (Si) mit Phosphoratomen (P). Da Phosphor aus der 5. Hauptgruppe kommt besitzt es 5 Außenelektronen doch hat nur 4 Partner um sich herum mit denen es Bindungen eingehen möchte. Nun gehen also 4 der 5 Außenelektronen

Doppelelektronbindungen mit den Außenelektronen benachbarter Siliziumatome ein. Daraus folgt jedoch, dass um das Phosphoratom nun 9 Elektronen existieren und wie oben beschrieben streben Atome immer zu 8 Außenelektronen, somit hat das 9. Außenelektron eine sehr schwache Bindung zu dem Phosphoratom und kann als freier Ladungsträger fungieren und somit Strom leiten.

Diese Art von Dotierung mit Elementen höherer Hauptgruppen nennt man auch n-Dotierung. Festzuhalten ist, dass bei der n-Dotierung die „überflüssigen“ Elektronen den Strom leiten, diese werden auch Minoritätsladungsträger genannt.

Die Zweite Form der Dotierung ist die Dotierung eines Halbleiters mit einem Element einer niedrigeren Hauptgruppe, diese Form nennt man die p-Dotierung.

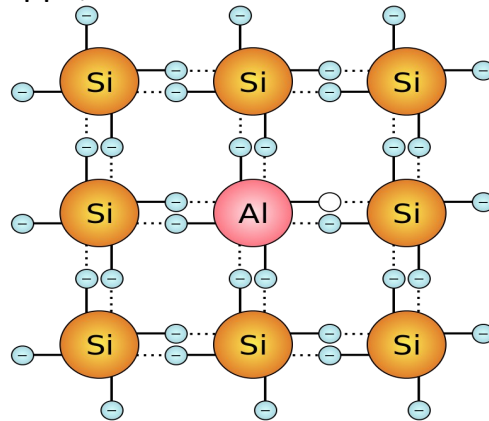


Abbildung 2: p-Dotierung im Siliziumkristallgitter mit Aluminium
[<https://de.wikipedia.org/wiki/Dotierung>]

In obiger Abbildung wurde Silizium mit Aluminium (Al) „verunreinigt“.

Alluminium steht in der 3. Hauptgruppe und besitzt damit 1 Valenzelektron weniger als Silizium. Daraus entstehen 3 Doppelelektronenbindung und eine Einzelelektronenbindung. Auch das Aluminium möchte jedoch gerne 8 Außenelektronen haben, deswegen können diese in der Einfachbindung enthaltenen „Löcher“ sozusagen positive Ladungsträger sein, da sie das Bestreben besitzen gefüllt zu werden. Diese Löcher, welche auch Majoritätsladungsträger genannt werden können somit wandern, was einen Strom darstellt. Dies geschieht so, dass wenn eine Spannung angelegt wird sich Elektronen aus ihren Bindungen lösen und in das freie Loch „springen“, doch nun sind an den Stellen wo diese sich herausgelöst haben wieder Löcher, welche wiederum gefüllt werden müssen und so weiter. Somit springen die Löcher immer wieder Richtung minus Pol. Diese Art Strom zu leiten wird auch Lochleitend genannt.

Daraus ergeben sich nun einige für uns wichtige Effekte die bei der Verwendung unterschiedlich dotierter Materialien auftreten. Denn es kann nur ein Strom (negative Ladungsträger) von einem n-dotierten Material zu einem p-dotierten Material fließen. Diese Eigenschaft wird später noch sehr wichtig wenn wir uns npn-dotierte Bereiche ansehen, diese bestehen aus zwei n-Dotierten Schichten, welche durch eine p-dotierte Schicht getrennt sind. Durch solch eine npn-Schicht fließt normalerweise kein Strom, doch wenn die mittlere p-Schicht durch ein elektrisches

Feld geschickt manipuliert wird kann diese auch n-leitend gemacht werden und ein Stromfluss ist möglich. Diese Technik wird in jedem Feld Effect Transistor eingesetzt. Um die p-Schicht n-leitend zu machen müssen nämlich die enthaltenen Löcher verdrängt bzw. gestopft werden. Dies wird gemacht indem oberhalb der p-Schicht eine Positive Spannung angelegt wird, dadurch werden Elektronen zu dieser Spannung hingezogen. Die Spannungsquelle (später das „Gate“ unserer Speicherzellen) muss zudem isoliert von der p-Schicht sein, damit kein Stromfluss stattfindet. Ist die Elektronen anziehende Spannung stark genug werden alle Löcher in der Nähe der Spannungsquelle geschlossen und überschüssige Elektronen dienen als Minoritätsladungsträger und somit wird ein kleiner Kanal nahe der Spannungsquelle n-leitend.

2.2 Der Tunneleffekt

In dem späteren Kapitel über den Aufbau und die Funktion der Speicherzelle wird immer wieder davon gesprochen wie Elektronen durch einen Isolator hindurchtunneln. Da wir nach der klassischen Physik sagen können, dass ein Strom nicht durch einen Nichtleiter fließen kann müssen wir vorerst klären wie so etwas überhaupt möglich ist. An dieser Stelle bedient sich die Flash Speicher Technologie dem Tunneleffekt, dieser ist ein Quantenmechanischer Effekt welcher nicht mit der klassischen Physik zu erklären ist.

Bei der Quantenmechanik handelt es sich um den Versuch der Beschreibung von unseren kleinsten bekannten Teilchen, wie zum Beispiel Quarks, Neutronen oder wie für uns wichtig den Elektronen. Entgegen der klassischen Physik kann man für ein Quanten Objekt nie eine genaue Position angeben, sondern nur die Wahrscheinlichkeit mit der es an einer bestimmten Position ist.

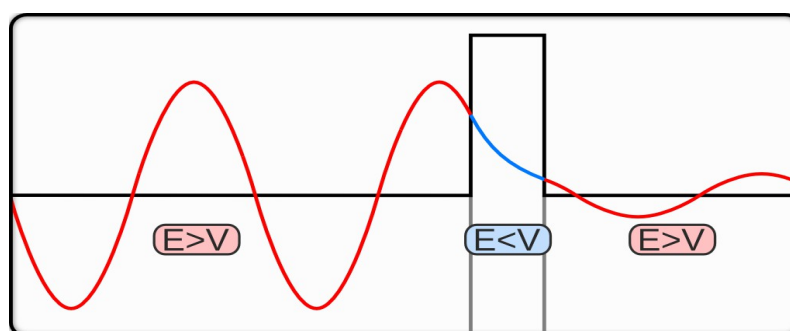


Abbildung 3: Schematische Darstellung des Tunneleffekts
[<https://de.wikipedia.org/wiki/Tunneleffekt>]

Auf Abbildung 3 ist eine solche Wahrscheinlichkeit für den Aufenthaltsort eines Elektrons eingetragen, diese wird durch die rote Kurve beschrieben. Umso höher die Amplitude desto höher ist die Wahrscheinlichkeit, dass sich dort das Teilchen aufhält.

Außerdem ist in der Grafik eine Wand eingetragen diese stellt eine Potentialbarriere

dar, welche bei der Speicherzelle mit dem Isolator gleichzusetzen ist.

Obwohl das Elektron eigentlich nicht die Energie besitzt die Potentialbarriere zu überwinden, was durch $E < V$ verdeutlicht wird, können wir anhand der kleineren Amplitude trotzdem sehen, dass es eine sehr geringe Wahrscheinlichkeit gibt, dass der Aufenthaltsort des Elektrons hinter der Barriere liegt.

Den bekanntesten Beweis für diesen Effekt kann man jeden Tag am Himmel beobachten, ohne den Tunneleffekt wäre nämlich die Kernfusion auf der Sonne nicht möglich. Zwei Teilchen mit gleicher Ladung stoßen sich ab, das dürfte jedem klar sein, doch wenn man zwei gleich geladene Teilchen nahe genug bringt wird die starke Wechselwirkung zwischen den Teilchen größer als die Abstoßung und sie ziehen sich an und verschmelzen. Die Grenze an der die Abstoßung und die Anziehung sich ausgleichen nennt man Coulombbarriere. Nun ist der Druck und die Temperatur auf der Sonne nicht ausreichend um ein Teilchen hinter die Coulombbarriere eines anderen zu bringen, doch auf Grund des Tunneleffekts gibt es eine Wahrscheinlichkeit, dass ein Teilchen, welches so schon nahe genug an einem Zweiten liegt hinter der Coulombbarriere ist und diese verschmelzen. So führt die Quantenmechanik dazu, dass wir Licht auf der Erde haben.

2.3 Die Speicherzelle – Aufbau

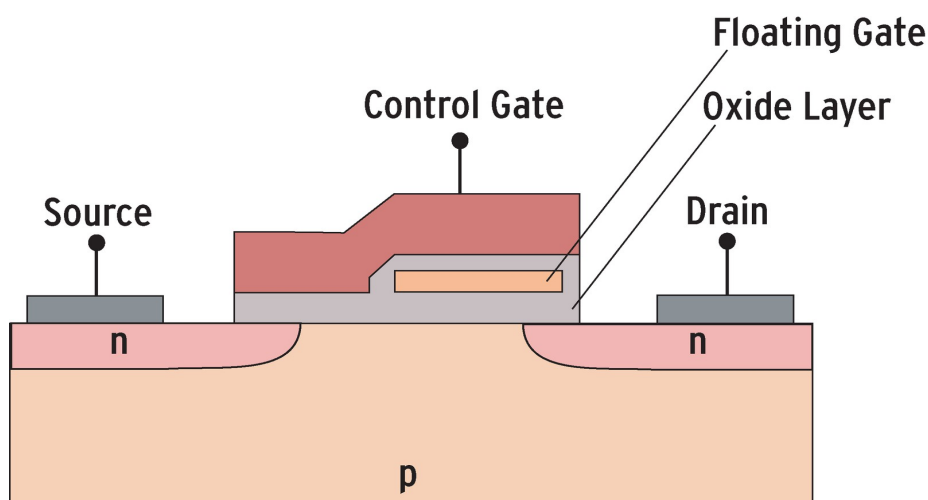


Abbildung 4: Struktureller Aufbau einer Speicherzelle
[<https://de.wikipedia.org/wiki/Flash-Speicher>]

Auf Abbildung 4 kann man die wichtigsten Komponenten einer Speicherzelle sehen. Die Funktionsweise basiert auf dem Prinzip des FET's (Feld Effekt Transistor), welches durch die P/N- Dotierung ein Stromfluss zwischen Source und Drain nur dann ermöglicht wenn an dem Control Gate eine Spannung angelegt wird. Als weitere Komponenten kommen zu dem FET ein Floating Gate hinzu welches die für die Datenspeicherung wichtigen Elektronen speichert und eine Oxid Schicht welche das Floating Gate vom Rest der Bauteile isoliert.

2.4 Die Speicherzelle – Lesen

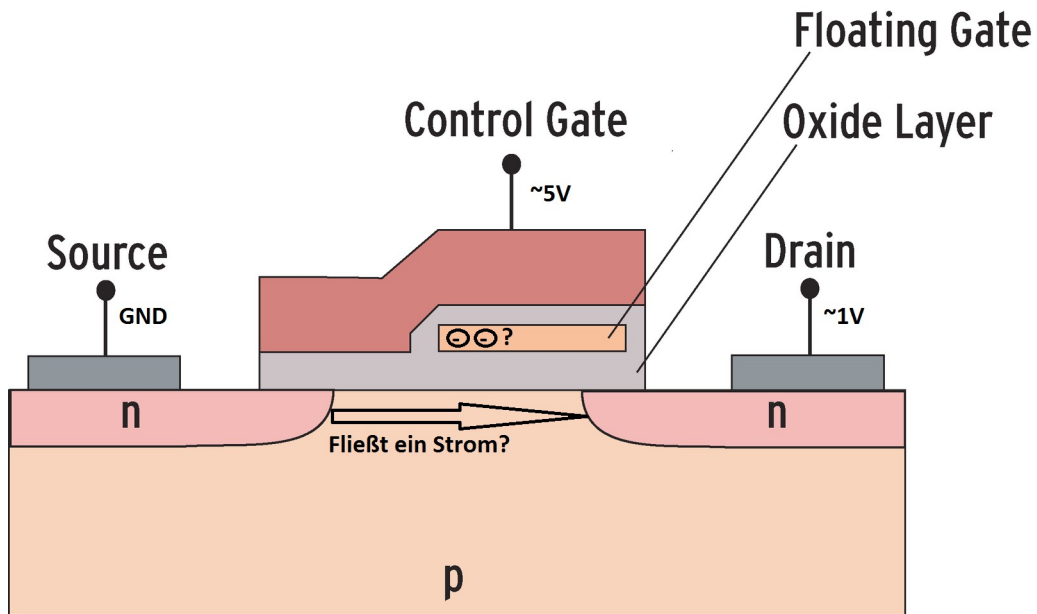


Abbildung 5: Struktureller Aufbau einer Speicherzelle - lesen
[<https://de.wikipedia.org/wiki/Flash-Speicher>]

Um nun aus den Elektronen welche eventuell im Floating Gate gespeichert wurden Informationen zu gewinnen muss diese Ausgelesen werden.

Nun liegt es nahe zwischen Source und Drain eine Spannung anzulegen und zu schauen ob ein Strom fließt oder nicht und diesen nun als 0 oder 1 zu Interpretieren. Und fast so einfach ist es auch, denn normalerweise fließt kein Strom durch eine npn-dotierte Schicht, dies muss erst durch das Gate ermöglicht werden. Wenn an das Control Gate eine Spannung angelegt wird zieht diese Elektronen zu sich welche, wie im letzten Abschnitt des Kapitels P/N-Dotierung beschrieben einen kleinen Kanal des p-Substrates n-leitend machen.

Die Spannungsschwelle ab der alle Löcher im Kanal „gestopft“ sind und es freie Minoritätsträger im p-Substrat nahe des Gates gibt ist in diesem Fall unsere Grenzwertspannung. Wird dieser Spannungswert vom Gate überschritten kann also ein Strom von Source nach Drain fließen.

Genau auf diesen Vorgang des Verdrängens der Majoritätsträger aus dem P-Dotierten Bereich nimmt das Floating Gate Einfluss. Zum auslesen wird nämlich an das Gate eine Spannung angelegt, welche etwas über der Grenzwert Spannung liegt.

Befinden sich nun Ladungsträger im Floating Gate wirken diese dem Elektrischen Feld durch die Gate-Spannung entgegen und verhindern eine Verdrängung der Majoritätsträger aus dem p-dotiertem Substrat, daraus folgt, dass kein Stromfluss möglich ist und dieser Zustand als 0 interpretiert werden kann.

Ist das Floating Gate jedoch ungeladen findet, wie in einem normalen Feld Effekt Transistor auch hier ein Stromfluss statt, dieser kann als 1 interpretiert werden.

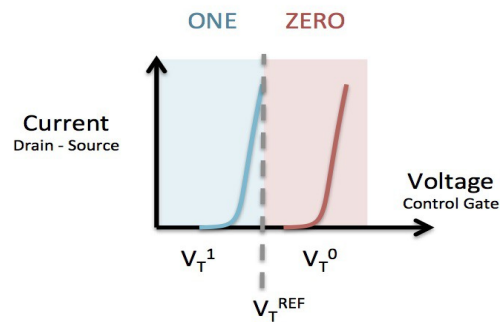


Abbildung 6: Stromfluss beim Auslesen einer Speicherzelle
[\[https://flashdba.com/2015/01/09/understanding-flash-floating-gates-and-wear/\]](https://flashdba.com/2015/01/09/understanding-flash-floating-gates-and-wear/)

Diese Abbildung verdeutlicht noch einmal den Vorgang des Auslesens.

In der Mitte erkennt man die zum Auslesen genutzte Spannung V_T^{REF} , welche größer als die eben beschriebene Grenzwertspannung ist. Die blaue Kurve gibt an, wie hoch der Stromfluss zwischen Drain und Source für eine angelegte Spannung ist, wenn das Floating Gate ungeladen ist, und die rote für die geladene Variante des Floating Gates.

Zwischen diesen 2 Zuständen kann es jedoch noch mehr geben, wenn die in dem Floating Gate gespeicherte Ladung variiert wird, somit kann in einer einzelnen Speicherzelle noch mehr Information als nur 0 und 1 gespeichert werden. Doch dazu im späteren Kapitel „SLC, MLC, TLC“ mehr.

2.5 Die Speicherzelle – Schreiben

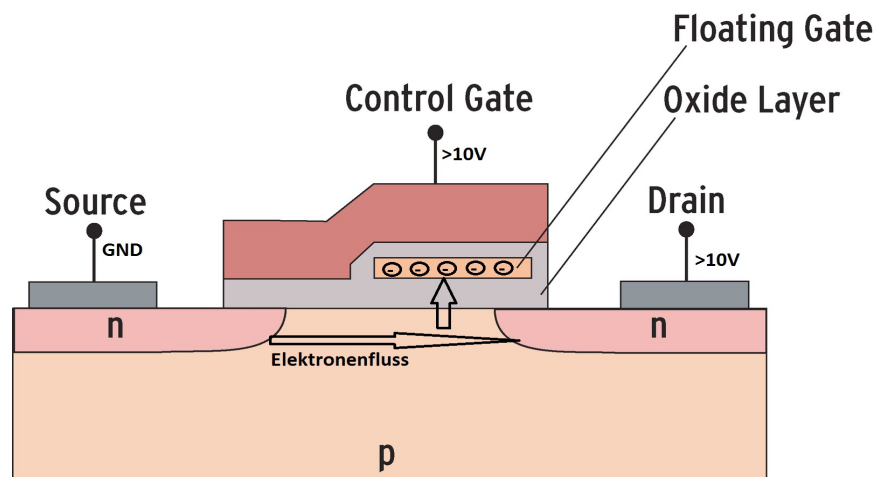


Abbildung 7: Struktureller Aufbau einer Speicherzelle - schreiben
[\[https://de.wikipedia.org/wiki/Flash-Speicher\]](https://de.wikipedia.org/wiki/Flash-Speicher)

Um eine Speicherzelle mit Informationen zu füllen, ist es vorerst notwendig, dass das Floating Gate dieser Speicherzelle keine Elektronen enthält, also noch nicht beschrieben ist, da sonst die Menge an Elektronen im Floating Gate nicht eingeschätzt werden kann und es beim Auslesen zu falschen Resultaten kommen

könnte.

Um eine Speicherzelle mit der Information 0 zu beschreiben müssen also Elektronen in das Floating Gate gelangen. Dafür wird an das Control Gate so wie an den Drain eine sehr hohe Spannung angelegt was zu einem starken Stromfluss zwischen Source und Drain führt und gleichzeitig die Elektronen stark in Richtung Control Gate zieht. Doch zwischen dem p-Substrat und dem Floating Gate liegt ein Isolator und genau an dieser Stelle kommt der Tunneleffekt zum Einsatz, denn nur dieser Quantenmechanische Effekt ermöglicht es den Elektronen den Isolator zu überwinden und im Floating Gate zu landen, sie sind nun im Floating Gate „gefangen“.

Je nach Dauer und Stärke der angelegten Spannungen kann beim schreiben variiert werden wie stark die Ladung des Floating Gates ist.

2.6 Die Speicherzelle – Löschen

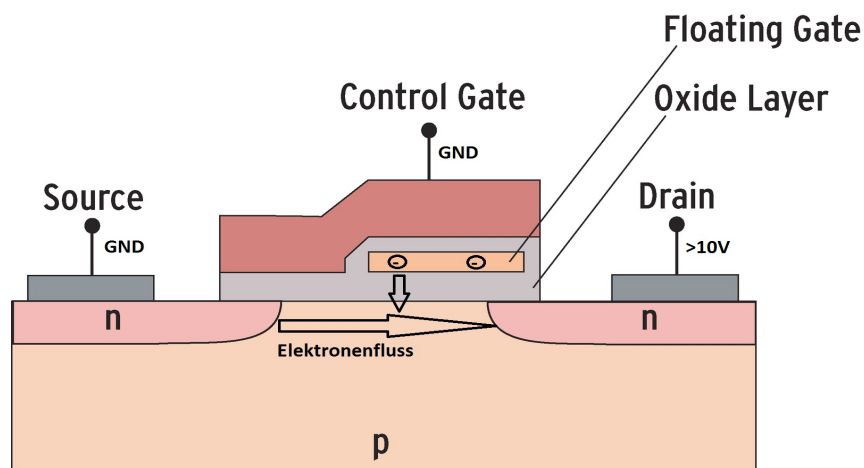


Abbildung 8: Struktureller Aufbau einer Speicherzelle - löschen
[<https://de.wikipedia.org/wiki/Flash-Speicher>]

Da die Speicherzellen zum erneuten Beschreiben leer sein müssen, sollte vorher gelöscht werden. Dieser Vorgang ist sehr ähnlich dem beschreiben einer Speicherzelle jedoch wird hierbei keine Spannung oder eine negative, welche die Elektronen abstößt an das Control Gate angelegt. Durch die hohe Spannung am Drain Anschluss Tunneln die Elektronen aus dem Floating Gate zurück in das p-Substrat und, da ein Stromfluss von p nach n möglich ist fließen die überschüssigen Elektronen im Drain ab.

2.7 SLC, MLC, TLC

Eine einfache Speicherzelle welche nur die Unterscheidung zwischen 0 und 1 bietet wird Single Level Cell genannt, also SLC.

Wie in vorangegangenen Kapiteln schon erwähnt gibt es zudem die Möglichkeit mehr als nur 0 und 1 in einer Speicherzelle zu speichern.

Diese Speicherzellen werden Multi Level Cell's oder Triple Level Cell's genannt.

Die Funktionsweise der Zellen die mehr als 1 bit Speichern können werde ich an folgender Abbildung erklären.

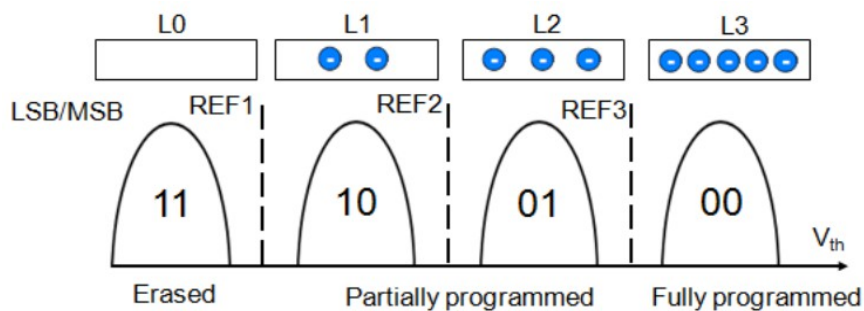


Abbildung 9: Mehrfache statusabhängige Grenzwertspannungen einer MLC Speicherzelle [<https://janzhou.org/2015/error-patterns-in-mlc-nand-flash-memory.html>]

Zu sehen ist eine MLC, wobei der Begriff MLC eigentlich ausschließlich für Zellen eingesetzt wird welche genau zwischen 4 Zuständen unterscheiden und demzufolge 2 Level besitzen. Diese MLC bietet statt nur der Unterscheidung zwischen 0 und 1 also die Unterscheidung zwischen 11, 10, 01 und 00. Die Kästen oberhalb der Zustände geben die Füllung des Floating Gates an, also ist im Gegensatz zu einer SLC der komplett gelöschte Zustand als 11 und nicht als 1 anzusehen und der voll programmierte als 00 und nicht nur als 0. Die Zustände 10 und 01 werden beim schreiben erreicht indem der Schreibvorgang kürzer oder unter einer geringeren Spannung stattfindet. Doch hier entpuppt sich schon das erste Problem, das Programmieren, denn das Beschreiben einer Zelle muss sehr präzise ablaufen und hat eine viel kleinere Fehlertoleranz als das Beschreiben von SLC Zellen.

Doch auch beim auslesen einer MLC im Vergleich zu einer SLC gibt es Unterschiede. Bei einer SLC müsste in obiger Abbildung nur die Spannung REF2 angelegt werden um zu sehen wie die Zelle Beschrieben ist. Doch bei MLC's muss hier eine Fallunterscheidung vorgenommen werden. Zuerst wird auch Spannung REF2 an das Control Gate angelegt fließt nun ein Strom muss immer noch geschaut werden ob das Ergebnis 11 oder 10 lautet, also muss als nächstes Spannung REF1 angelegt werden. Diese Fallunterscheidung muss äquivalent dazu für den Fall zwischen 01 und 00 vorgenommen, falls nach anlegen von REF2 kein Strom fließt.

Damit ist der Lesevorgang mit viel größerem Aufwand und Risiko einer Fehlinterpretation verbunden als bei SLC's.

Des weiteren war noch von TLC also Tripple Level Cell's die Rede, diese unterscheiden nicht nur zwischen 2 oder 4 Zuständen sondern zwischen 8, also 111, 110, 101, 100, 011, 010, 001 und 000 die Funktionsweise entspricht der gleichen wie der der MLC's, jedoch mit noch höherer Fehleranfälligkeit beim beschreiben und beim Auslesen. Bei dem Auslesevorgang von TLC Zellen müssen ganze 3 mal verschiedene Spannungen angelegt werden, dies strapaziert die Speicherzelle enorm und ist zudem wiederum sehr Fehleranfällig.

Die Vor und Nachteile dieser 3 Varianten liegen auf der Hand, zusammengefasst kann man sagen je weniger Level desto Langlebiger ist die Zelle und desto mehr

Level desto mehr Speicherplatz hat die Zelle.

In folgender Tabelle stehen dazu einige Daten von gemittelten Werten welche die Unterschiede zwischen SLC, MLC und TLC noch einmal veranschaulichen.

	SLC (Single Level Cell)	MLC (Multi Level Cell)	TLC (Triple Level Cell)
Bit pro Zelle	1 Bit	2 Bit	3 Bit
Speicherbare Zustände	2 (2^1)	4 (2^2)	8 (2^3)
Lebensdauer	100.000 Schreibvorgänge	3.000 Schreibvorgänge	ca. 1.000 Schreibvorgänge
Fehlerrate	sehr niedrig	mittel	hoch
Geschwindigkeit	sehr hoch	niedrig	niedrig
Stromverbrauch	sehr niedrig	hoch	hoch

Abbildung 10: Tabellenvergleich zwischen SLC, MLC und TLC [<http://www.elektronik-kompodium.de/sites/com/1906241.htm>]

2.8 Speicherzellen Anordnung

In den vorangegangenen Kapiteln wurde besprochen wie einzelne Speicherzellen funktionieren. Um solche Speicherzellen jedoch sinnvoll zu benutzen benötigt man tausende oder gar Millionen von ihnen. Diese vielen Speicherzellen müssen jedoch in irgendeiner Struktur verbunden werden um sie dicht zu packen und einfach beschreiben und auslesen zu können.

Die zwei verbreitetsten Arten der Speicherzellen Anordnung sind die NOR- und die NAND-Struktur.

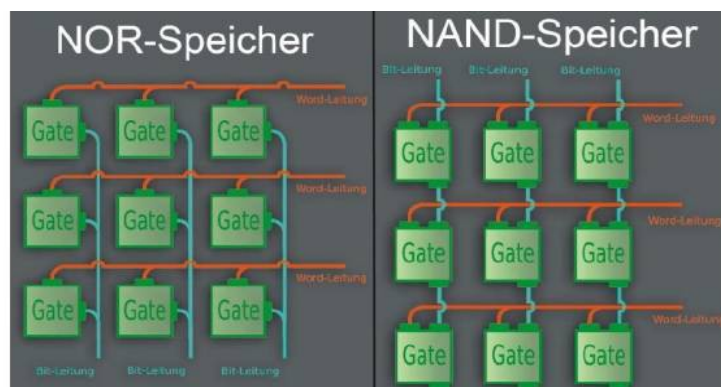


Abbildung 11: Vergleich NOR vs. NAND (schematisch)
[http://www.pcwelt.de/ratgeber/Die_Technik_hinter_Solid_State_Drives__SSDs__Flash-Speicher-8387351.html]

In obiger Abbildung sind diese beiden Anordnungs-Strukturen gegenüber gestellt. Bei NOR-Speicher werden die einzelnen Speicherzellen parallel Verknüpft und haben jede ihre eigene Verbindung zur Bitleitung. Der Vorteil der sich daraus ergibt ist, dass die Speicherzellen einzeln ansprechbar sind und somit die Latenz beim Auslesen bis zu 5 mal geringer ist als bei NAND-Speicher. Der offensichtliche Nachteil ist jedoch, dass man viel mehr Datenleitungen braucht um jede Zelle anzuschließen und somit NOR-Speicher bei gleicher Datenmenge in etwa 2.5 mal soviel Platz benötigt wie NAND-Speicher.

In NAND-Speicher sind alle Zellen einer Bitline in Reihe geschaltet, dies bedeutet eine viel geringere Anzahl an Datenleitungen und eine bis zu 10 mal längere Haltbarkeit als NOR-Speicher. Aufgrund der genannten Vorteile und der günstigeren Herstellungskosten des NAND-Speichers wird dieser in so ziemlich allen SSD's eingesetzt, weswegen sich in diesem Kapitel ausführlicher mit NAND auseinandergesetzt wird. Im Gegensatz dazu wird NOR-Speicher häufig z.B. in Eingebetteten Systemen verwendet, da dort meistens nur ein kleinerer Speicher von Nöten ist und selten Daten geschrieben werden, aber häufig und schnell ausgelesen werden müssen.

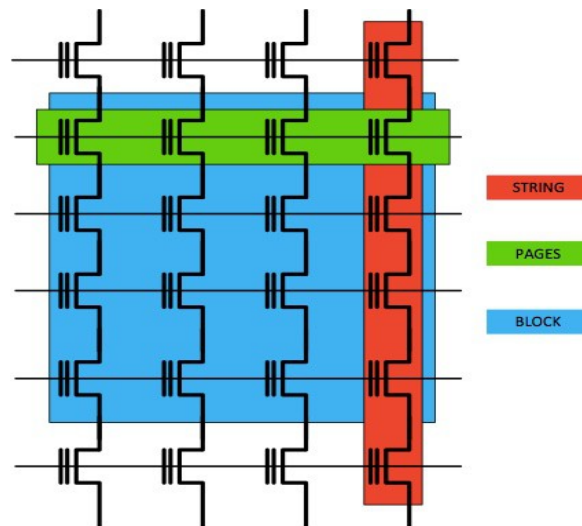


Abbildung 12: Anordnung von Speicherzellen in NAND-Flash
[\[http://www.flashstorage.com/flash-storage-technical-economic-primer/\]](http://www.flashstorage.com/flash-storage-technical-economic-primer/)

Diese Abbildung veranschaulicht einmal den logischen Aufbau von NAND-Speicher. NAND-Speicher können immer nur Page weise ausgelesen und Blockweise gelöscht werden, wieso das so ist wird im folgenden erklärt.

In der Horizontalen liegen die Wordlines und in der Senkrechten die Bitlines.

Möchte man nun eine Page, welche meistens eine Größe zwischen 512Byte und 8kByte hat auslesen wird auf der Wordline die für das auslesen benötigte Spannung angelegt. Die Wordline ist wiederum mit den Floating Gates aller Zellen verbunden welche auf der gleichen Page liegen. Außerdem muss auf jeder anderen Zelle welche im Block aber nicht auf der gleichen Page liegt der Transistor auf durchlass geschaltet werden um nur die eine in der Reihe liegende Zelle auslesen zu können. Dies geschieht indem auf allen anderen Wordlines des selben Blocks eine Spannung angelegt wird, welche selbst wenn Elektronen im Floating Gate sind stark genug ist, dass ein Strom durch die npn-Schicht fließen kann.

Da alle anderen Zellen auf durchlass geschaltet sind kann man auf der Bitleitung erkennen ob die eine Zelle geladen oder ungeladen ist und da dies für jede Zelle auf der gleichen Wordline gilt, liest man gleich die gesamte Page aus.

3. Software und Datenorganisation in SSD's

Da nun die Funktionsweise des in der SSD verbautem Speichers klar ist muss sich nun angesehen werden wie diese Daten in der SSD abgelegt und organisiert werden. Diese Aufgaben übernimmt heutzutage in jeder SSD der Controller. Dieser ist ein Computerchip welcher entscheidet wo Daten abgelegt werden, der überprüft ob die Speicherzellen noch richtig funktionieren, defekte gegebenenfalls ersetzt und vieles mehr.

3.1 Wear Leveling

Eine der wichtigsten Aufgaben des Controllers ist das Wear Leveling.

In Abbildung 10 ist in der Spalte Lebensdauer zu erkennen, dass eine Speicherzelle nicht unbegrenzt funktioniert. Würden nun immer wieder die gleichen Zellen der Festplatte beschrieben würden zunächst diese kaputt gehen und der verfügbare Speicher der SSD immer weiter schrumpfen. Um dieses Phänomen zu verhindern gibt es das sogenannten Wear Leveling.

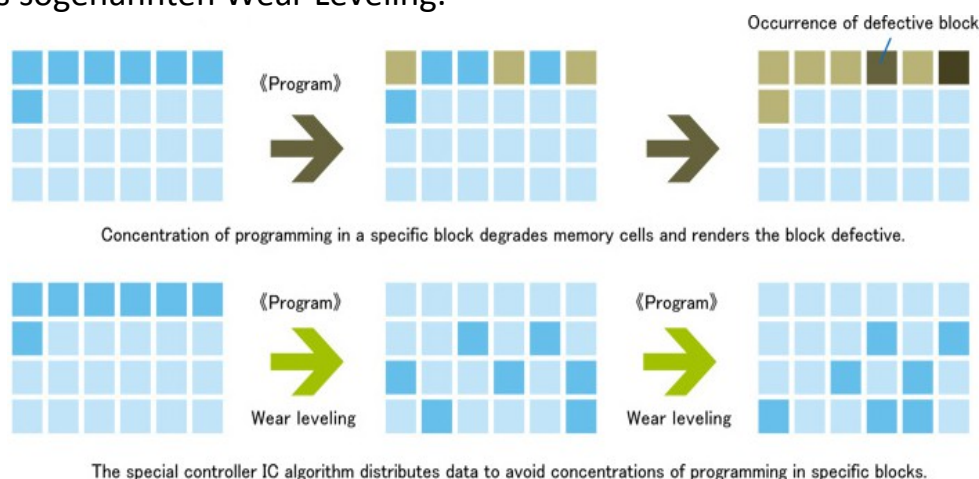


Abbildung 13: Datenverteilung mittels Wear Leveling

[https://product.tdk.com/info/en/techlibrary/archives/techjournal/vol14_sdg3b/contents03.html]

Dabei wird wie im unteren Teil der Abbildung 13 zu sehen beim neu beschreiben nicht immer nur der Anfang der Platte überschrieben sondern immer andere Zellen um die Auslastung der Platte über alle Speicherzellen zu verteilen.

Um zu entscheiden welche Speicherzelle als nächstes beschrieben wird gibt es zwei verschiedene Verfahren, das statische und das dynamische Wear Leveling.

Beim Statischen Wear Leveling werden alle Speicherzellen der gesamten SSD betrachtet. Jede Speicherzelle hat einen sogenannten Write-Counter welcher zählt wie oft die jeweilige Zelle schon beschrieben wurde, ist dieser sehr hoch neigt sich die Lebensdauer der Zelle wahrscheinlich dem Ende entgegen. Nun werden also von allen Speicherzellen die Write-Counter verglichen und die beschrieben welche den geringsten hat um so einen Lastausgleich zu schaffen. Stehen in der zu beschreibenden Zelle schon Daten müssen diese jedoch erst verschoben werden

und je nach Controller entscheidet dieser ob dies sinnvoll ist, wenn die dort geschriebenen Daten z.B. selten angefasst werden oder ob er nach einem anderen Ort sucht. Dies braucht hohen Rechenaufwand und dauert seine Zeit, doch erhöht die Lebensdauer der SSD so enorm, dass sich Wear Leveling immer lohnt.

Beim dynamischen Wear Leveling betrachtet der Controller nur den unbeschriebenen Teil der Platte und auch hier werden die Write-Counter der Zellen verglichen und die gewählt, welche den niedrigsten hat.

Hier ist natürlich der Zeitaufwand geringer, da keine Daten erst verschoben werden müssen, doch verringern Daten die nie benutzt werden wieder die Lebensdauer der SSD, da die Zellen die damit beschrieben sind nie optimal ausgenutzt werden können.

3.2 Fehlerquellen Minimierung

Die meisten haben schon von ECC also Error Correction Code gehört. Beispiele um Fehler zu erkennen sind die Huffman-Kodierung oder der Hamming-Code. Auch in SSD's werden solche Fehlercodes eingesetzt, doch hat der Controller noch andere Möglichkeiten um die Fehleranfälligkeit der Zellen zu reduzieren.

Eine Möglichkeit des Controllers ist das randomisierte ablegen von Daten welche häufig den gleichen Wert besitzen.

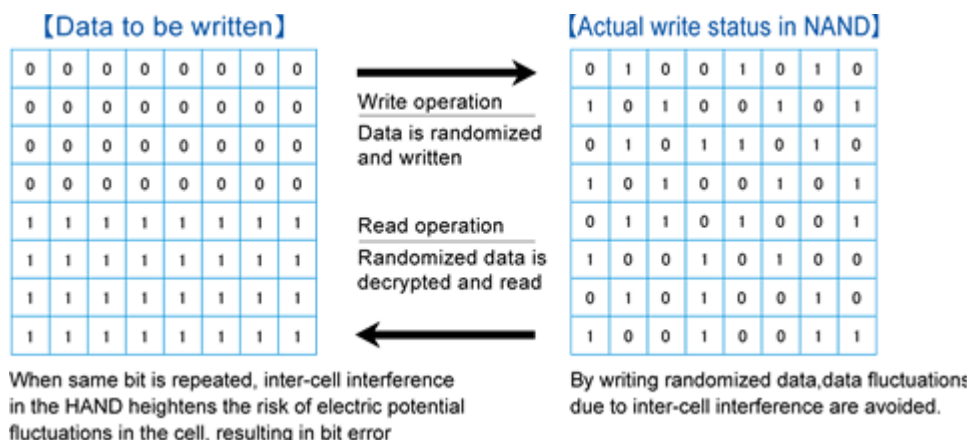


Abbildung 14: Datenrandomisierung

[https://product.tdk.com/info/en/techlibrary/archives/techjournal/vol14_sdg3b/contents05.html]

Nehmen wir an die Daten wären so abgelegt wie im linken Teil der Abbildung 14, beim auslesen der Daten einer Page werden zwar alle anderen Zellen durch eine sehr hohe Spannung auf durchlas geschaltet, trotzdem kann es den ausgelesenen Wert minimal verändern wenn alle umliegenden Zellen den gleichen Wert besitzen. Außerdem wird durch die Randomisierung die Interferenz zwischen den Speicherzellen minimiert, welche zu bitflips führen könnte.

Eine zweite Möglichkeit des Controllers ist der Auto-refresh. Wenn in einem Datenblock ein Fehler auftritt ist es nämlich sehr Wahrscheinlich, dass Zellen im selben Block auch bald einen Fehler erzeugen könnten. Beim Auto-refresh werden,

sobald ein Fehler auftritt, alle Zellen des selben Blocks ausgelesen und mit diesen Daten neu beschrieben. Dies führt dazu, dass Ladungen, welche mit der Zeit etwas schwächer wurden nun wieder aufgefrischt sind. Dieser Vorgang kann vom Controller zu beliebigen Zeitpunkten ausgeführt werden, wobei dafür meist der Idle-State genutzt wird.

Eine dritte Möglichkeit ist das verändern der Auslese Spannung einer Zelle. Mit der Zeit kann es nämlich passieren, dass sich aufgrund von Abnutzung die in obigen Kapiteln erwähnte Grenzwertspannung etwas verschiebt. Wenn der Controller nun die Fähigkeit besitzt diese korrekt anzupassen muss eine solche etwas abgenutzte Zelle nicht gleich als Defekt markiert werden, sondern kann mit der neuen Grenzwertspannung immer noch weiter benutzt werden.

3.3 Bad Block Management

Durch Fehler bei der Herstellung, fehlerhafte Schreibvorgänge oder durch zu häufiges Beschreiben und Löschen von Zellen können diese beschädigt oder sogar unbrauchbar werden.

Ist dies der Fall hat der Controller die Möglichkeit defekte Zellen zu markieren damit diese nicht weiter benutzt werden. Um zu erkennen, dass eine Zelle kaputt ist werden vom Controller mehrere Werte überprüft. Zu diesen Werten zählen die Anzahl der P/E-Cycles (Programm/Erase – Cycles), also wie oft eine Zelle schon neu beschrieben wurde, der ECC Counter, also wie oft in diesem Datenblock Gebrauch vom Error Correction Code gemacht wurde und das eventuelle Auslesen von Daten außerhalb des erwarteten Bereiches, wie z.B. ein viel zu hoher Stromfluss beim Auslesen. Für diese Fälle besitzt jede SSD einen reservierten Bereich von Speicherzellen welche dazu dienen solche fehlerhaften Zellen zu ersetzen und nicht sofort die Speicherkapazität der SSD zu verringern. Diese Art der Reservierung von Zellen nennt man Overprovisioning. Dieser Zellenbereich wird des Weiteren dazu genutzt um das oben besprochene Wear Leveling zu optimieren und um Effekte wie die Write-Cliff abzuschwächen, auf welche später noch eingegangen wird.

3.4 Garbage Collection

Die Garbage Collection also die Müllabfuhr ist eine weitere Aufgabe des Controllers. Da immer nur ganze Blöcke gelöscht werden können kann es auch schon bei geringer Speicherplatz Auslastung zu Leistungseinbrüchen kommen. Dies geschieht wenn alle Blöcke mindestens ein bisschen beschrieben sind, da nun vor jedem Schreibvorgang erst einmal gelöscht und erst dann neu beschrieben werden kann. Um diesen Effekt zu verhindern gibt es die Garbage Collection. Der Controller führt dabei mehrere Blöcke die nur teilweise beschrieben sind zu einem Block zusammen um bei einem Schreibvorgang zu garantieren, dass es freie Blöcke gibt und nicht erst gelöscht

werden muss. Außerdem ist ein weiterer Teil der Garbage Collection das Blöcke, welche Daten enthalten die nicht mehr gebraucht werden, für das spätere neu beschreiben gelöscht werden.

Auch diese Aufgabe kann der Controller zu einem beliebigen Zeitpunkt ausführen und sie wird deshalb meist auf den Idle-State verschoben.

4. Probleme heutiger SSD's

Zum Abschluss werden in diesem Kapitel zwei Probleme beschrieben mit denen heutige SSD's kämpfen. Zum einen ist dies die Write Amplification welche dafür sorgen kann, dass die SSD bei einem kleinen Schreibvorgang viel mehr schreibt und liest als eigentlich notwendig ist und zum anderen ist das die sogenannte Write Cliff welche die SSD's von Herstellerseite aus sehr schnell erscheinen lässt, sie im Betrieb jedoch langsamer ist als erwartet.

4.1 Write Amplification

Wie schon erwähnt ist eine Restriktion bei der Verwendung von NAND-Speicher, dass Zellen nur als gerammter Block gelöscht werden können. Durch die Garbage Collection werden die Auswirkungen davon zwar verringert, aber welche Auswirkung es haben kann wenn ein Block vor dem schreiben gelöscht werden muss wird im folgenden Beispiel aufgezeigt.

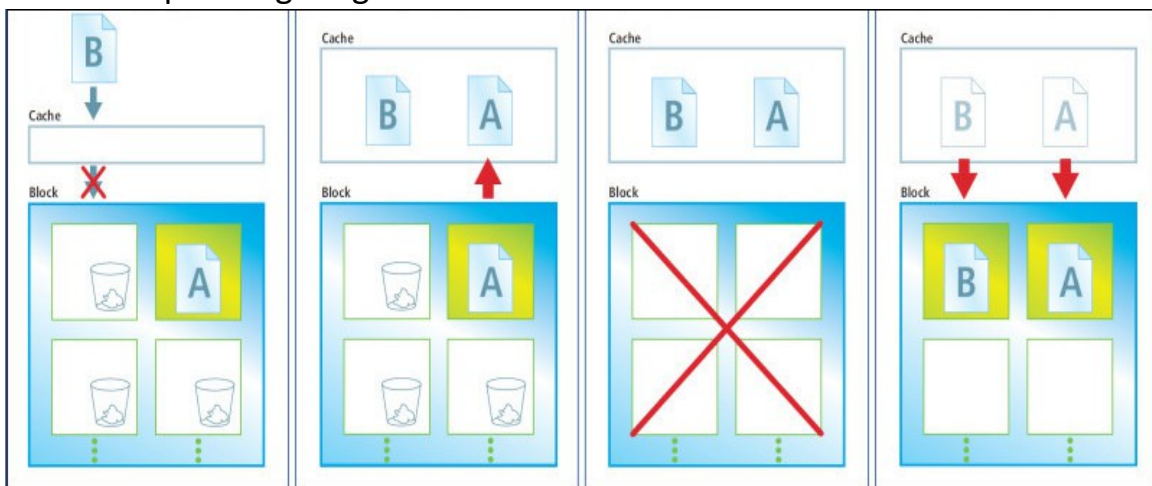


Abbildung 15: Read-Modify-Write-Zyklus [http://www.com-magazin.de/praxis/ssd/solid-state-drives-61179.html?page=4_write-amplification-und-read-modify-write-zyklus]

In Abbildung 15 soll eine Datei B geschrieben werden. In unserem Beispiel hat diese Datei eine Größe von 4kB. Die Datei B soll dabei in den gezeigten Block geschrieben werden, welcher jedoch voll mit Müll und einer wichtigen Datei ist. Also wird Datei B vorerst in den Cache geladen und wir haben somit unsere ersten 4kB geschrieben. Anschließend wird die Datei A, welche ebenfalls 4kB groß ist eingelesen und ebenfalls in den Cache geschrieben. Somit haben wir bis jetzt einen Schreibaufwand von 8kB (Cache) und 4kB gelesen (Flash-Speicher). Da der Cache sehr schnell ist ist

dieser Aufwand bis jetzt jedoch sehr gering, nur das Lesen hat uns einige Mikrosekunden an Zeit gekostet. Jetzt muss jedoch der Block gelöscht werden und das kann zwischen 1.5 und 3 **Millisekunden** dauern, was 10 bis 100 mal so lange ist, als ein Schreibvorgang benötigt. Zum Schluss können wir die Dateien A und B wieder in unseren Block reinschreiben.

Summa summarum haben wir also insgesamt einen Schreibaufwand von 16kB (8kB Cache und 8kB Flash-Speicher) für eine nur 4kB große Datei.

Diesen Effekt des „mehr Schreiben als benötigt“ nennt man Write Amplification. Offensichtlicher Weise tritt diese natürlich desto häufiger auf, je weniger freie Blöcke es auf der Platte gibt und je schlechter der Controller in Hinsicht auf die Garbage Collection arbeitet. Dieser Effekt ist deshalb einer der Gründe dafür, weshalb sehr volle SSD's wesentlich langsamer arbeiten als kaum beschriebene.

4.2 Write Cliff

Um noch eine letzte Problematik vorzustellen möchte ich an dieser Stelle die Write Cliff erwähnen. SSD's werden oft mit ihrer immensen Geschwindigkeit angepriesen, doch diese auf dem Datenblatt angezeigten Werte täuschen aufgrund der Write Cliff oft die Kunden.

Diese ausgeschriebenen Werte werden nämlich oft nur bei kleinen Daten erreicht, welche in den Cache passen und nicht wenn direkt auf die Platte geschrieben wird. Somit kann ein Schreibvorgang, welcher mehr Daten beinhaltet als in den SSD-Cache passen in etwa wie folgt aussehen.

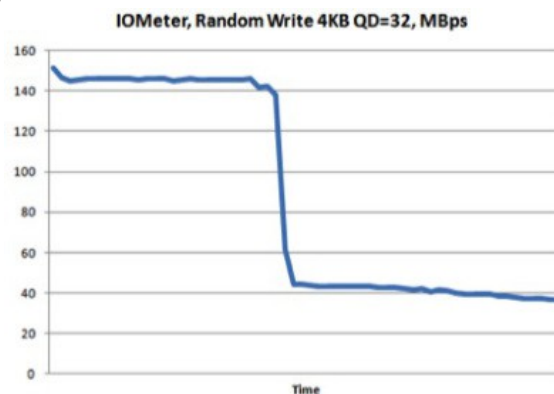


Abbildung 16: Write Cliff

[<https://flashstorageguy.wordpress.com/tag/write-cliff/>]

Meist tritt dieser Effekt verstärkt auf je billiger die vorliegende SSD ist. Das liegt daran, dass in billigen SSD's häufig MLC oder sogar TLC-Speicherzellen verbaut sind, welche wie oben erwähnt beim Schreib- und Lesevorgang viel langsamer sind als SLC-Speicher. Außerdem besitzen billige SSD's häufig einen kleineren Cache und eine schlechtere Speicherverwaltung als ihre teureren Pendanten.

Um diesen extremen Leistungseinbruch zu verringern können zum Beispiel die internen Flash-Chips über mehr Flash-Channels angebunden werden, was die interne Datenübertragungsrate verbessern würde, aber wiederum teurer ist und mehr Platz

braucht.

Als Beispiele aus der Praxis wurden 2 SSD's mit ähnlicher Speicherkapazität aber stark unterschiedlichen Preis verglichen.

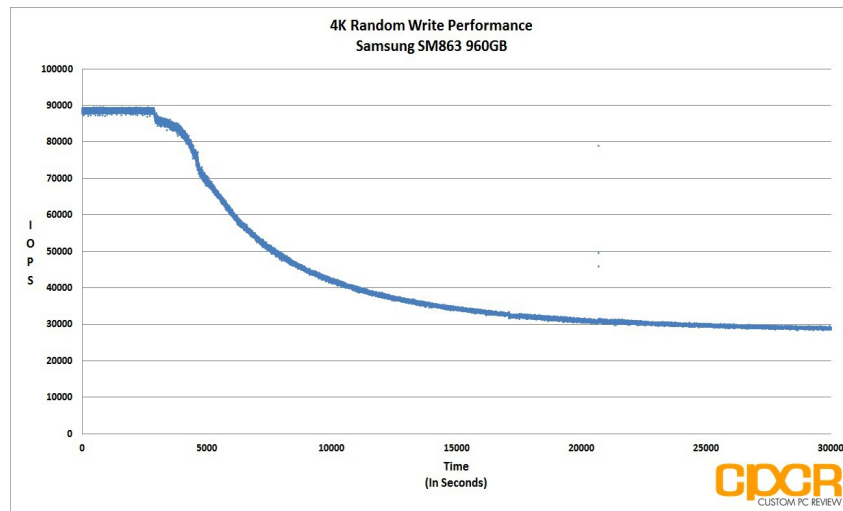


Abbildung 17: Write Performance SM863

[www.custompcreview.com/reviews/review-samsung-pm863-sm863-960gb-enterprise-ssd/26068/5/]

Als erstes liegt die SM863 von Samsung vor, welche bei 480GB einen Preis von 330\$ hat. In der Abbildung 17 ist zwar das 960GB Modell zu sehen, was aber keine großen Auswirkungen auf die zu erwartende Kurve hat. Dieses Modell ist für Server konzipiert und deshalb sehr teuer. Gut erkennbar ist, dass der Geschwindigkeitsabfall erst nach etwa 4000 Sekunden beginnt und erst nach etwa 20000 Sekunden seinen Tiefpunkt bei etwa 1/3 der Maximalgeschwindigkeit erreicht hat.

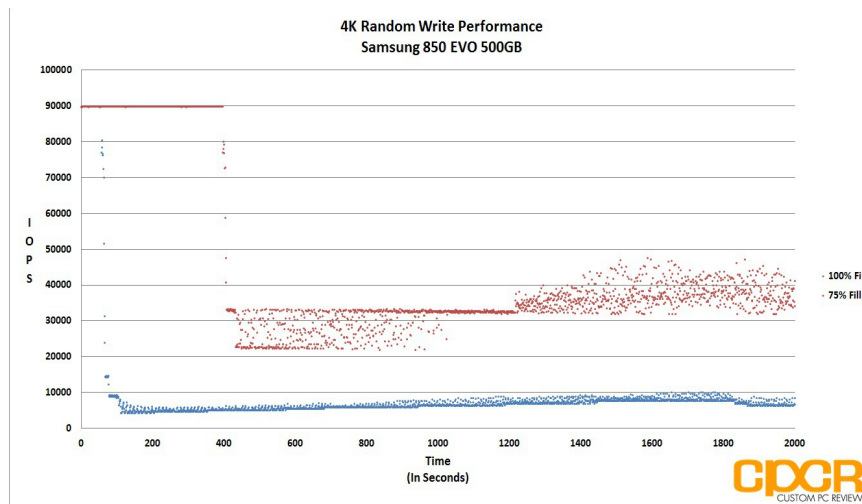


Abbildung 18: Write Performance 850 evo

[www.custompcreview.com/reviews/samsung-850-evo-500gb-ssd-review/22307/8/]

Abbildung 18 zeigt die 850 Evo, welches eine 500GB Consumer SSD von Samsung zum Preis von 170\$ ist. Deutlich zu erkennen ist, dass hier bereits nach 400 Sekunden, falls die Platte noch nicht voll ist und schon nach etwa 70 Sekunden, falls die Platte nahezu komplett belegt ist die Geschwindigkeit auf 1/10 bis 1/3 der

Leistung abfällt, also viel schneller als bei dem Enterprise Modell.

Der immense Preisunterschied zwischen diesen beiden Modellen hat nicht nur als Vorteil, dass der Write Cliff Effekt erst viel später eintritt, sondern das bessere Modell hat noch viele weitere Eigenschaften welche sie von der billig SSD abhebt. Dazu zählt zum Beispiel die längere Lebensdauer der Festplatte, wo bei der billig SSD nur Garantie auf bis zu 150 geschriebenen Terabytes besteht sind es bei dem teuren Modell 3080 Terabyte.

5. Zusammenfassung

Diese Arbeit sollte einen kleinen Einblick in die Funktionsweise der heutigen SSD's verschaffen und hat zum Ziel, dass Leser beim Kauf einer neuen SSD nicht nur blind auf den Preis/GB achten sondern noch etwas weiter denken sollten.

Welche SSD für jemanden die richtige ist kann kein anderer sagen, denn es kommt immer auf den Einsatzzweck an und natürlich auf den Portemonnaie Inhalt.

Jedoch werden SSD's immer billiger und einige Prognosen behaupten sogar, dass Flash-Speicher an sich ca. 2023 anfängt billiger zu werden als ihr härtester Konkurrent die HDD. Andere Quellen geben auch an, dass jetzt schon die Kosten, welche Anschaffung und Betrieb über mindestens 4 Jahre mit einbeziehen für gewerbliche Datenspeicherung unter Verwendung von SSD's billiger sind als wenn man HDD's für den gleichen Zweck einsetzt. Diese Behauptungen sollten jedoch mit Vorsicht genossen werden und treffen offensichtlich nicht auf jeden zu. Dennoch wird die Flash Technologie so schnell vorangetrieben, dass wir uns auf rosige Zeiten mit billigen und extrem schnellen Festplatten freuen können.

Abbildungsverzeichnis

Abbildung 1: n-Dotierung im Siliciumkristallgitter mit Phosphor [https://de.wikipedia.org/wiki/Dotierung]	4
Abbildung 2: p-Dotierung im Siliciumkristallgitter mit Aluminium [https://de.wikipedia.org/wiki/Dotierung]	5
Abbildung 3: Schematische Darstellung des Tunneleffekts [https://de.wikipedia.org/wiki/Tunneleffekt]	6
Abbildung 4: Struktureller Aufbau einer Speicherzelle [https://de.wikipedia.org/wiki/Flash-Speicher]	7
Abbildung 5: Struktureller Aufbau einer Speicherzelle - lesen [https://de.wikipedia.org/wiki/Flash-Speicher]	8
Abbildung 6: Stromfluss beim auslesen einer Speicherzelle [https://flashdba.com/2015/01/09/understanding-flash-floating-gates-and-wear/]	9
Abbildung 7: Struktureller Aufbau einer Speicherzelle - schreiben [https://de.wikipedia.org/wiki/Flash-Speicher]	9
Abbildung 8: Struktureller Aufbau einer Speicherzelle - löschen [https://de.wikipedia.org/wiki/Flash-Speicher]	10
Abbildung 9: Mehrfache statusabhängige Grenzwertspannungen einer MLC Speicherzelle [https://janzhou.org/2015/error-patterns-in-mlc-nand-flash-memory.html]	11
Abbildung 10: Tabellenvergleich zwischen SLC, MLC und TLC [http://www.elektronik-kompodium.de/sites/com/1906241.htm]	12
Abbildung 11: Vergleich NOR vs. NAND (schematisch) [http://www.pcwelt.de/ratgeber/Die_Technik_hinter_Solid_State_Drives__SSDs_-Flash-Speicher-8387351.html]	12
Abbildung 12: Anordnung von Speicherzellen in NAND-Flash [http://www.flashstorage.com/flash-storage-technical-economic-primer/]	13
Abbildung 13: Datenverteilung mittels Wear Leveling [https://product.tdk.com/info/en/techlibrary/archives/techjournal/vol14_sdg3b/contents03.html]	14
Abbildung 14: Datenrandomisierung [https://product.tdk.com/info/en/techlibrary/archives/techjournal/vol14_sdg3b/contents05.html]	15
Abbildung 15: Read-Modify-Write-Zyklus [http://www.com-magazin.de/praxis/ssd/solid-state-drives-61179.html?page=4_write-amplification-und-read-modify-write-zyklus]	17
Abbildung 16: Write Cliff [https://flashstorageguy.wordpress.com/tag/write-cliff/]	18
Abbildung 17: Write Performance SM863 [www.custompcreview.com/reviews/review-samsung-pm863-sm863-960gb-enterprise-ssd/26068/5/]	19
Abbildung 18: Write Performance 850 evo [www.custompcreview.com/reviews/samsung-850-evo-500gb-ssd-review/22307/8/]	19