

# Solid State Drive?

Die neue „Festplatte“  
(**Keine** HDD!)

Dr. Reiner Kupferschmidt

# Gliederung

- Begriff, Eigenschaften
- Technische Daten

# Begriff und Eigenschaften

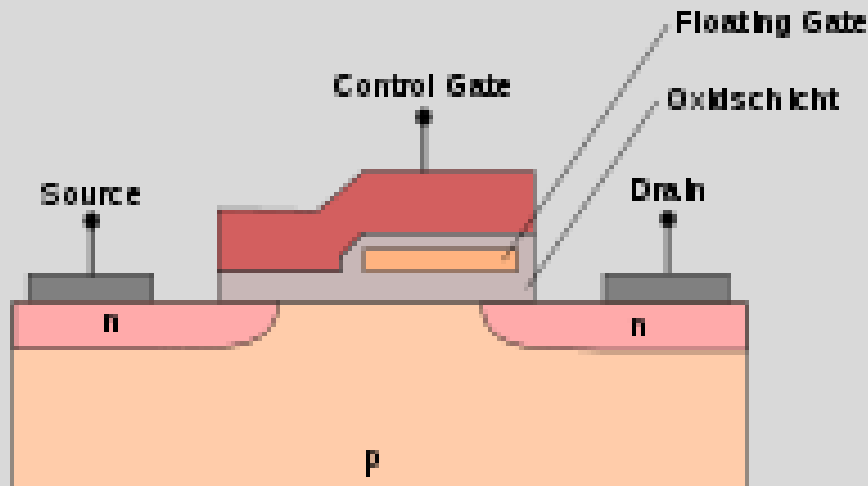
- Solid State Drive/Disk
- Kein mechanisches Laufwerk oder Scheibe vorhanden
- Sind
  - Lautlos
  - Stoßfest
  - Energiearm (Kapazitätsabhängig)
  - Kühl
  - Schnell (hohe Zugriffsgeschwindigkeiten Lesen/Schreiben)
  - Größenlimit
- Flash-Speicher

# Technische Daten

	HDD	SSD
Datenrate	0,6 Gbit/s	128 Gbit/s
Zugriffszeit	3,5 ms	0,1 ms

# Speicherprinzip

- Flash-Speicher (-EEPROM)
  - Nicht flüchtig (*non volatile memory*, kurz *NVM*)
  - Elektrisch blockweise löschbar
- MOSFET – Floating-Gate Transistor



# Speicherkapazität

- Hoher Platzbedarf > feinere Halbleiterstrukturen mehr Speicherkapazität bei gleicher Chipgröße
- Nachteil – Verkleinerung macht Floatinggate empfindlicher und stärkerer Stromverbrauch
- Änderung d. Ladezustandes (Schreiben) belastet Floating-Gate und umgebende Sperrschichten
- >Speicherichte pro Speicherzelle erhöhen – mehr Zustände pro Zelle (SLC > MLC > TLC)
- Anzahl der Schreibvorgänge sinkt, Verschleiß
- Fehlerkorrekturmaßnahmen notwendig
- Aktuelle Kapazität (03/2018): 100 TB,  
**Nimbus Data ExaDrive DC100, 10.000 €**

# Lesen und Schreiben

- Lesen
  - 2 – 4 kByte aus Flash-Zellen lesen und in Pufferspeicher schreiben – sehr schnell
- Schreiben
  - Flash-Speicher ist blockweise organisiert (128 bis 512 kByte)
  - Bei Änderungen muss der betreffende Block komplett neu geschrieben werden, davor muss er gelöscht werden
  - Hohe Löschspannung – alle Zellen verlieren den Inhalt
  - Daten aus Puffer in Flash-Zellen einlesen
  - NCQ (Native Command Queuing - fängt die Schreibzugriffe ab, sortiert sie um, intelligente Verteilung auf die einzelnen Speicherblöcke)

# Schreiben

- 1 Block = 256 Pages mit je 8 KiB
- Zum Schreiben und Löschen werden Blöcke nochmalig zusammengefasst
- Erst wenn alle zusammengefassten Blöcke nicht Aktuell sind, werden sie gelöscht

Schreibvorgang	1	2	3	4	5
1. Block	Bild.jpg	unaktuell	unaktuell	löschen	Bild.jpg
2. Block	Leer	Bild.jpg	Unaktuell	löschen	Leer
3. Block	Leer	Leer	Bild.jpg	löschen	Leer
4. Block	Leer	Leer	Leer	Leer	Leer
5. Block	Leer	Leer	Leer	Leer	Leer



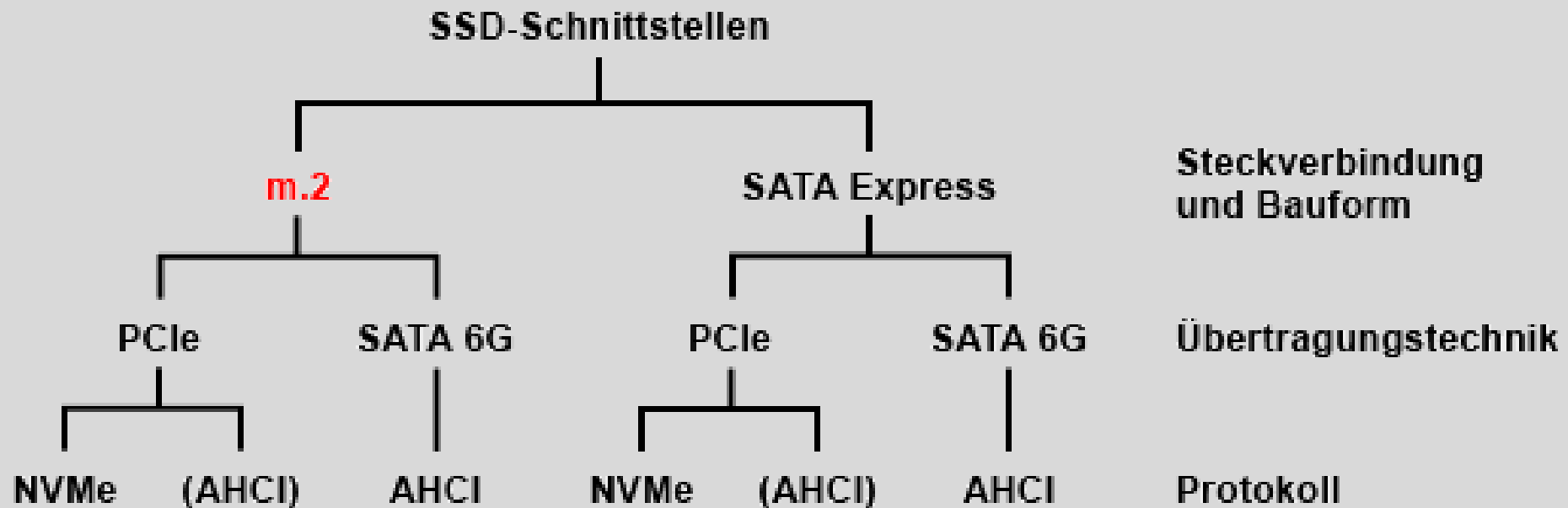
# Bauformen - Schnittstellen



SATA  
mSATA  
PCIe  
M.2  
SATAe



# Übersicht Schnittstellen



\*Advanced Host Controller Interface

\*\*Non-Volatile Memory Express

\*\*\*Peripheral Component Interconnect express

# Schnittstellen

- 2,5-Zoll-SSD mit SATA 6G
  - 6 GBit/s bzw. netto 600 MByte/s
  - Optimierung für NCQ
  - abwärtskompatibel zu SATA und SATA-II
- Speicherkarte mit M.2 (Nachfolger von Next Generation Form Factor (NGFF))
- 2,5-Zoll-SSD mit SATAe
  - 1 GByte/s (x1, x2, x4)
- PCIe-Erweiterungskarte mit PCIe (Peripheral Component Interconnect)
  - Server und Workstations (intern mit SATA oder SAS-Anbindung)
  - Ziel mit m.2 oder SATAe
  - 500 MByte/s pro Lane (V 2.0)
  - 1.000 MByte/s pro Lane (V 3.0)

# PCI(e)

Identifizieren der verschiedenen PCI-Steckplätze

Peripheral Component Interconnect oder PCI ist eine häufige Methode zum Anschließen von Zusatz-Controller-Karten und anderen Geräten an ein Computer-Motherboard. Diese Art Konnektor stammt aus den frühen 1990er Jahren und wird heute noch verwendet. Derzeit gibt es drei wesentliche PCI-Motherboard-Konnektoren (meist als „Slots“ oder Steckplätze bezeichnet.)

- 64-Bit-PCI
- 32-Bit-PCI und PCI-X
- PCI-Express (PCI-E)

Jeder PCI-Steckplatz sieht anders aus und akzeptiert andere Geräte. Wird eine PCI-Karte an den falschen Steckplatz gesteckt, wird die Karte beschädigt und eventuell wird der gesamte Computer zerstört.

64-Bit-PCI:

64\_bit\_pci

LaCie stellt keine Produkte mehr her, die mit diesem Steckplatz kompatibel sind. Für PCs wurden sie meist nicht verwendet, aber für alle G4 und G3 Macintosh-Computer. Dieser Steckplatz kann anhand der drei Segmente identifiziert werden, wobei sich das kürzeste Segment in der Mitte befindet.

32-Bit-PCI und PCI-X:

pci\_x

LaCie stellt verschiedene Erweiterungskarten her, die diesen Konnektor verwenden. Der Unterschied zwischen dem normalen 32-Bit-PCI-Steckplatz und dem PCI-X ist der mittellange Steckplatz auf der linken Seite. Nur PCI-X hat dieses letzte Segment. Zusätzlich funktionieren 32-Bit-PCI-Karten an einem PCI-X-Steckplatz, PCI-X-Karten funktionieren jedoch nicht an einem Standard-32-Bit-PCI-Steckplatz.

Nahezu alle PC-Motherboards haben mindestens den 32-Bit-PCI-Steckplatz. Macintosh G5 verwendete PCI-X bis zur Änderung auf flüssigkeitsgekühlte Hardware-Versionen. Apple kann eine G5-Seriennummer nachschlagen, um zu bestimmen, ob PCI-X-Steckplätze zur Verfügung stehen oder ob stattdessen die neue PCI-E verwendet werden muss.

Diese Steckplätze können vom 64-Bit-PCI anhand der Anordnung der Segmente unterschieden werden. Das kleine Segment kommt an erster Stelle statt in der Mitte.

PCI-Express (PCI-E):

Die meisten Computer, die nach 2005 hergestellt wurden (einschließlich Macintosh), verfügen über PCI-E-Steckplätze. Diese sind manchmal schwer zu identifizieren, da die Länge des Steckplatzes unterschiedlich sein kann. Diese Varianten werden „Lanes“ genannt und sind meist mit einer Nummer, gefolgt von einem x (1x 8x 16x usw.), bezeichnet. In der Abbildung ist ein 16x-PCI-Steckplatz dargestellt. Ein 1x-Steckplatz beginnt mit demselben kleinen Segment, gefolgt von einem zweiten kleinen Segment. Mehrere Lanes bedeuten im Allgemeinen eine größere Schnittstellengeschwindigkeit. Die meisten PCIx-Karten sind 1x oder 4x (mit Ausnahme der Video-Controller, die fast immer 16x-Karten sind.)

Eine 1x-PCI-E-Erweiterungskarte passt in einen 16x-Steckplatz. Aufgrund der physischen Größe bleibt der Rest des Steckplatzes unbelegt, dies ist jedoch normal. Dieser Steckplatz kann von anderen (insbesondere 32-Bit-PCI) durch seine physikalische Größe unterschieden werden. Alle Konnektoren an einem PCI-E-Steckplatz sind deutlich kleiner und er wird weiter in das Motherboard eingesteckt als andere PCI-Steckplätze.

# M.2

## Nachfolger von Next Generation Form Factor (NGFF)

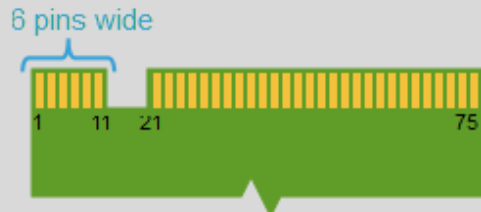
### Ablösung von mSATA (mini SATA)

<https://de.wikipedia.org/wiki/M.2>

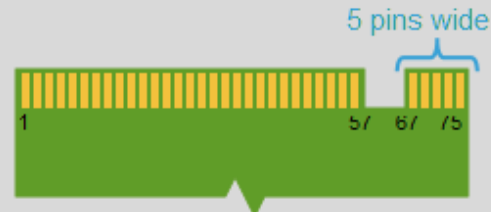


Socket for "B key" edge connector

Socket for "M key" edge connector



"B key" edge connector

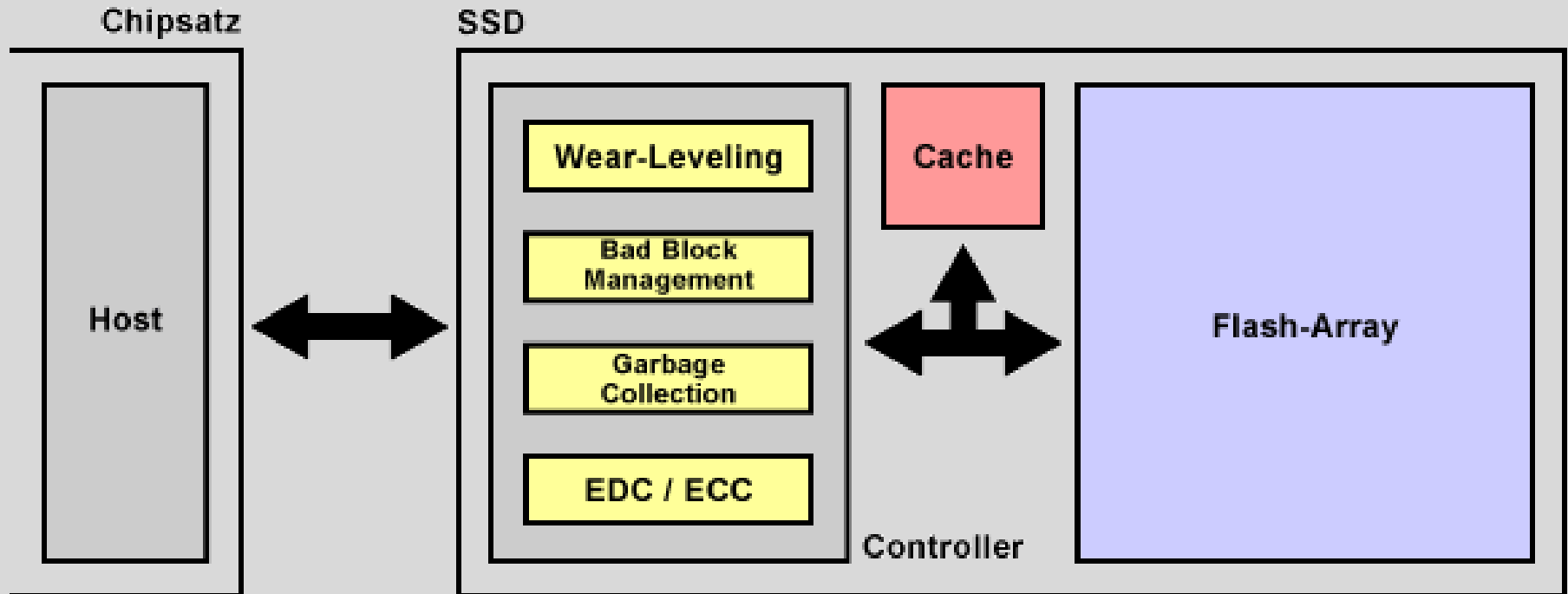


"M key" edge connector



"B & M key" edge connector

# Aufbau



\*ECC – Error Correction Code

\*EDC – Error Detection Code

# Wear-Leveling

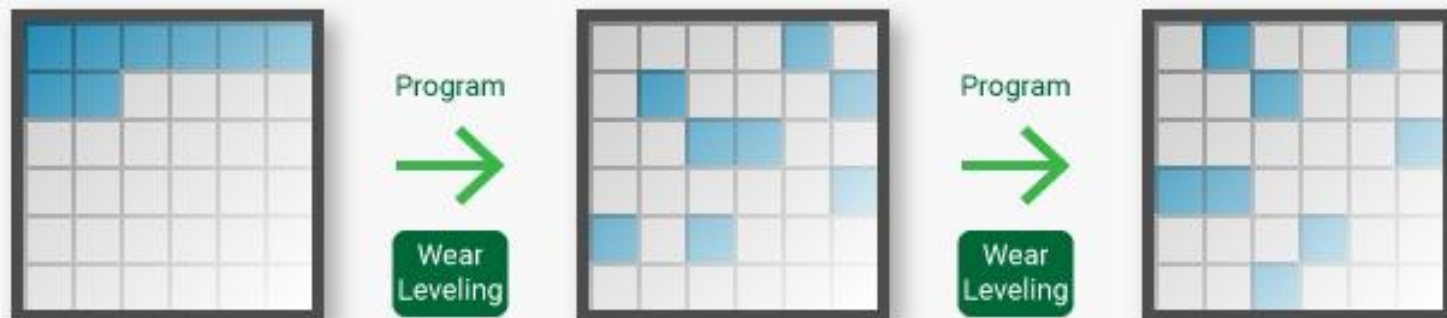
- Begrenzte Lebensdauer, 1.000 bis 1.000.000 Schreibvorgänge
- Ständig schreibende Anwendungen ungeeignet
- Wear-Leveling
  - Gleichmäßige Verteilung der Schreiblast über die Zellen durch den Controller
  - Statisches WL
    - Relativ konstante Daten werden in abgenutzte Bereiche verschoben
    - Erhöhung der Lebensdauer des Speichers
    - Geht zu Lasten der Performance
    - Problem gelöschte Dateien – **TRIM** – BS teilt Flashcontroller „gelöschte“ Bereiche mit
  - Dynamisches WL
    - Schreibzugriffe werden gleichmäßig über freie oder frei werdende Blöcke verteilt
    - Häufig genutzte Bereiche nutzen sich stärker ab und fallen aus
- Defragmentierung, Dateiindexierung und Prefetching des Betriebssystems abschalten

# Wear-Leveling

Device without Wear Leveling



Device with Wear Leveling





# Overprovisioning

- Überprovisionierung - übliches Verfahren zur Steigerung der Lebensdauer und der Leistungsfähigkeit eines SSD-Speichers
- Ein bestimmter Anteil des vorhandenen Speicherplatzes wird für den Controller und das Datenmanagement reserviert
- Anwender oder Applikationen können auf diesen Speicherplatz nicht zugreifen
- Überprovisionierung hält einen konstanten Bereich der SSD für den Controller und die temporären Verschiebevorgänge frei
- Sollten Blöcke des regulär genutzten Speichers aufgrund von Alterungsprozessen oder anderen Problemen nicht mehr nutzbar sein, können Blöcke aus dem reservierten Bereich die Aufgabe der defekten Speicherblöcke übernehmen. Der überdimensionierte Speicher wird zwar kleiner, aber die Lebensdauer der SSD steigt
- Grad der Überprovisionierung wird in als Prozentzahl angegeben
- Quotienten aus reservierter Speicherkapazität und nutzbarer Speicherkapazität
- Consumer-SSDs 7 % Überprovisionierung übliche
- Professionelle Anwendungen - schreibintensive Vorgängen - Over-Provisioning > 20 Prozent bis 50 %

# Overprovisioning

**QNAP**

## SSD Internal Over-provisioning

- SSD as Flash storage, its capacity should be presented with Gbabyte (GiB)  
 $2^{30} = 1,073,741,824$  Bytes °
- As storage device, SSD actually report its capacity back with Gigabyte (GB)  
 $10^9 = 1,000,000,000$  Bytes °
- Every GB in SSD contain 73,741,824 Bytes (7.37%) for the Garbage Collection operation, and enterprise SSD may reserved more!

SSD actual size (GiB)	SSD present size (GB)	Vendor designed OP %	Actual internal OP %	Size show in the NAS (GiB)
256 GB	256 GB	0%	7%	238 GB
256 GB	240 GB	7%	15%	223 GB
256 GB	200 GB	28%	37%	186 GB

In the NAS system, all capacity are actually be calculated and showed in the GiB format  
<https://www.seagate.com/tw/zh/tech-insights/ssd-over-provisioning-benefits-master-ti/>

# Bad Block Management / Defect Management

- Überwacht die Speicherzellen
- Defekte Zellen werden markiert und aus Reserve ersetzt (Overprovisioning)
- SLC-SSDs Reserve 2 % - geringe Abnutzung
- MLC-SSDs Reserve 7 % - höhere Abnutzung
- Kein Verlust von Speicherzellen
- Erhalt von Lebensdauer, Zuverlässigkeit und Speicherkapazität

# Garbage Collection

- Wird vom Betriebssystem angestoßen
- Beim Löschen wird nur der Name gelöscht, Inhalt der Speicherzelle bleibt erhalten
- Nicht geleerte Speicherzellen prüfen und leeren
- > Steigerung der Schreibgeschwindigkeit

# Haltbarkeit und Zuverlässigkeit

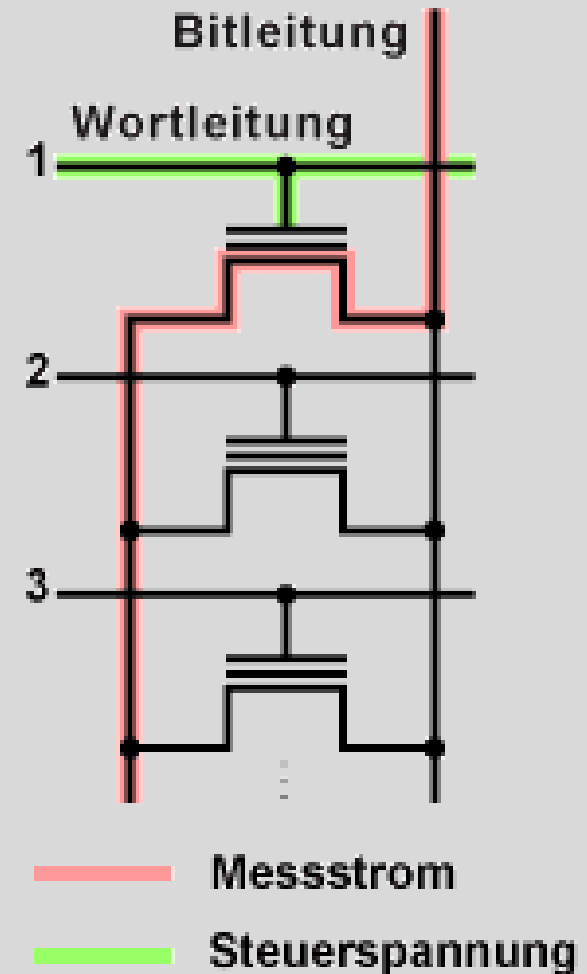
- 700 TByte und 1 PByte – maximale Datenmenge abhängig
- Garantiebeispiel 80 TByte in 10 Jahren (nicht für Server) > 20 GByte pro Tag
  - (Auslagerungsdateien
  - Start des BS
  - Herunterfahren des BS)
- Kein Schreiben mehr möglich – keine Ansprechbarkeit oftmals ohne Vorwarnung

# Endurance (Ausdauer, Haltbarkeit)

- maximal zulässige Anzahl an Lösch- bzw. Speicherzyklen (NVRAM)
- hohen Spannungen (10 bis 18 V) schädigen die Struktur der Zelle
- Möglichst dünne Isolation – negative Auswirkung auf Datenhaltung (Retention)
- Schreiben mit 10 ... 18 V, Oxidschicht nimmt Schaden (wird dünner)
- Schreibvorgang schreibt 16 bis 128 kByte gleichzeitig – geringe Änderungen des Speicherinhaltes erfordert Neuschreiben vieler Speicherzellen

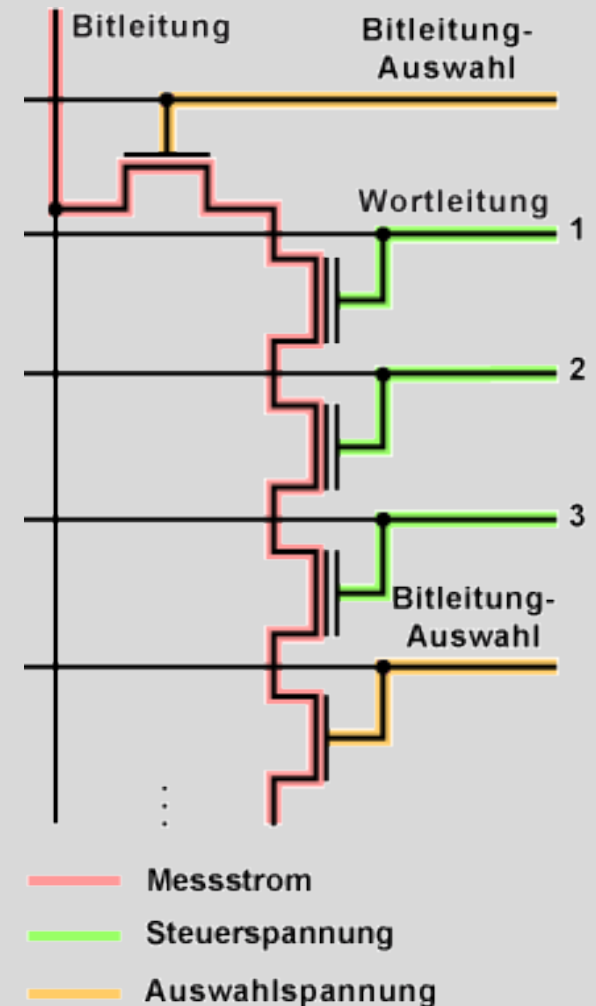
# NOR-Flash

- Unterscheidung in Speicherdichte und Zugriffsgeschwindigkeit
- Speicherzellen sind parallel verschaltet
- Zugriff erfolgt wahlfrei und direkt
- Kurze Zugriffszeiten
- Geringer Widerstand zwischen Stromquelle und Auswerteschaltung
- Programmspeicher von MC
- BIOS



# NAND-Flash

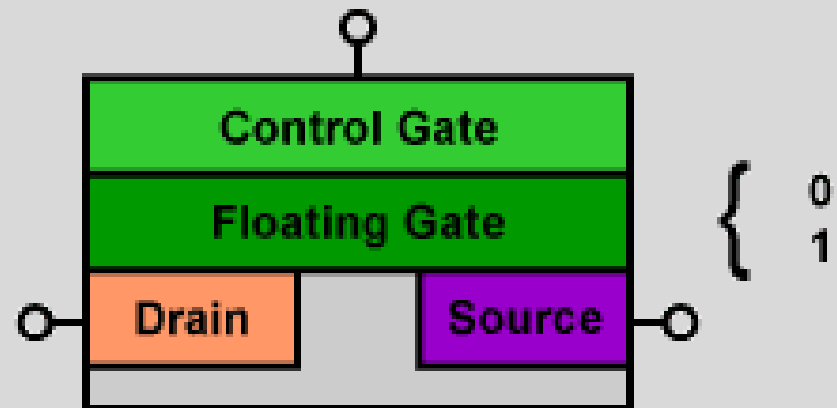
- Serielle Verschaltung – Lesen und Schreiben nur in Blöcken möglich
- Geringe Anzahl an Datenleitungen – weniger Platz
- Daten werden blockweise gelesen und geschrieben – Festplattenersatz
- Hohe Speicherdichte, geringe Kosten, schnell. Geringer Stromverbrauch
- Controller – Geschwindigkeit





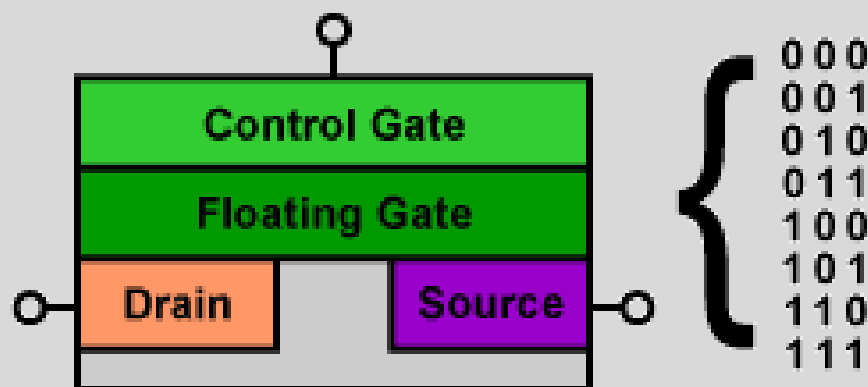
# SLC-Flash (Single Level Cell)

- ein Bit pro Speicherzelle
- zuverlässiger Flash-Speicher für SSDs – 100.000 Schreibzyklen
- Teuer



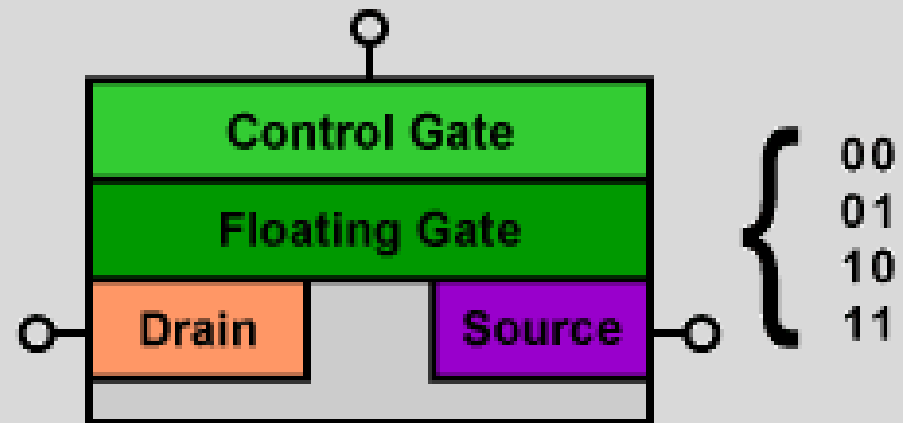
# TLC-Flash (Triple Level Cell)

- Drei Bit pro Speicherzelle
- 1.000 Schreibvorgänge
- USB-Sticks, Speicherkarten, extrem billige SSD
- Nicht für Server



# MLC-Flash (Multi Level Cell)

- Zwei Bit pro Speicherzelle
- Günstigere Fertigung, Massenmarkt
- Langsamer als SLC
- Häufige Lesefehler Fehlerkorrektur
- 3.000 Schreibzyklen
- USB-Sticks



# SLC-, MLC- und TLC-Flash

	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

# Vorteile von Flash-Speicher

- Die gespeicherten Daten bleiben auch bei fehlender Versorgungsspannung erhalten. Auf eine Erhaltungsladung kann verzichtet werden. Somit ist auch der Energieverbrauch und die Wärmeentwicklung geringer.
- Wegen fehlender beweglicher Teile ist Flash geräuschlos, unempfindlich gegen Erschütterungen und magnetische Felder.
- Im Vergleich zu Festplatten haben Flash-Speicher eine sehr kurze Zugriffszeit. Lese- und Schreibgeschwindigkeit sind über den gesamten Speicherbereich weitestgehend konstant.
- Die erreichbare Speichergröße ist durch die einfache und platzsparende Anordnung der Speicherzellen nach oben offen.

# Nachteile von Flash-Speicher

- begrenzte Schreib- bzw. Löschvorgänge
- „begrenzte“ Speicherkapazität
- „hoher“ Preis
- „Datenwiederherstellbarkeit“
- „Löschbarkeit“

# Weiterentwicklung von Flash-Speicher

- Längere Haltbarkeit
- Kleinere Zugriffszeiten und hohe Geschwindigkeit
- Senkung der Schreib- und Lesefehler der MLC in Bereich der SLC
- Größere Kapazitäten und preiswerter
- 3D-V-NAND, iSLC oder eMLC – mehrere MLC Lagen übereinander – vertikale Strukturen (20 Lagen)
- Einsparung von Siliziumfläche – Zellen können größer sein – höhere Speicherkapazität, Lebensdauer, Geschwindigkeit, Datenhaltung

# Fragen

- Wie arbeiten SSDs?
- Was ist der Unterschied zwischen "NAND"- und "NOR"-SSD-Modellen?
- Welche Vorteile bieten SSDs gegenüber anderen Speichermedien?
- Welches sind die Nachteile von SSDs?
- Welche Formfaktoren und welche maximalen Speicherkapazitäten bieten die SSD-Hersteller zurzeit an?
- SSDs kommen bisher vor allem in Notebooks und Netbooks zum Einsatz. Welchen Einfluss haben SSD-Komponenten auf Server- und Storage-Systeme?
- Was verstehen Sie unter Wear-Leveling?
- Was verstehen Sie unter TRIM?
- Was verstehen Sie unter Garbage-Collection?
- Was verstehen Sie unter Overprovisioning?



# Ende

Recherche über Daten und Preise derzeit vertriebener SSD.....