

Speichergeräte und RAID Systeme

Alexander Günther

6803826

Universität Hamburg

6guenther@informatik.uni-hamburg.de

27. Februar 2019

Inhaltsverzeichnis

1	Einleitung	3
2	Geschichte der Speichergeräte	3
2.1	Lochkarten	3
2.2	Disketten	4
2.3	Compact Disc	4
2.4	Hard Disc Drive	4
3	Speicherhierarchie in modernen Computern	5
4	RAID Speichersysteme	7
4.1	Grundlagen	7
4.1.1	Striping	8
4.1.2	Mirroring	8
4.1.3	Paritäten	8
4.2	RAID 0	8
4.3	RAID 1	9
4.4	RAID 2	10
4.5	RAID 3	10
4.6	RAID 4	12
4.7	RAID 5	13
4.8	RAID 6	14
4.9	RAID 01	15
4.10	RAID 10	15
5	Zusammenfassung	16

1 Einleitung

Speichersysteme haben sich, seit der Erfindung der Lochkarte im Jahr 1928, beständig weiterentwickelt. Heutzutage gibt es viele verschiedene Speicherarten, welche sich in ihrer Funktion in Rechnersystemen stark unterscheiden. Von den Registern im Prozessor bis hin zu riesigen Tapespeicherarchiven an hochmodernen Rechenzentren sind sie heute nicht mehr wegzudenken. Mit der starken Zunahme der zu speichernden Datenmengen/Datengrößen nahm auch die Notwendigkeit zu, sich vor einem Versagen einzelner Speicher abzusichern. Dieses Paper beschäftigt sich kurz mit der Geschichte von Speichergeräten und der heutzutage vorherrschenden Speicherhierarchie in Rechnersystemen, bevor es die verschiedenen Arten von RAID Speichersystemen 4 und ihre Vor- und Nachteile vorstellt.

2 Geschichte der Speichergeräte

Zu Beginn des Papers möchte ich eine kurze historische Einleitung bezüglich der Entwicklung von Speichersystemen geben. Insbesondere soll der Fokus hierbei auf der Speicherkapazität liegen.

2.1 Lochkarten

Bei der Lochkarte handelt es sich um eine Kartonkarte auf der Daten durch Einstanzen permanent gespeichert werden. Die ersten Lochkarten wurden 1890 für die Volkszählung in den USA eingeführt. Diese Lochkarten besaßen eine Speicherkapazität von 288 Bit (24 Spalten und 12 Zeilen [14]), es existieren auch Lochkartenformate mit bis zu 576 Bit (IBM-System/3 Lochkarten mit 96 Spalten mit je 6 bit [16]).

2.2 Disketten

Bei Disketten handelt es sich um rotierende magnetische Speicher. Disketten kommen in unterschiedlichen Größen vor: 3.5", 5.25" und 8". Ihre Speicherkapazität variiert von 79.75 kB bis zu 240.75 MB [15].

2.3 Compact Disc

Bei der **C**ompact **D**isc handelt es sich um einen optischen Speicher, welcher seit 1982 auf dem Markt ist. Die Speicherkapazität beläuft sich zwischen 650 und 90 MB.

2.4 Hard Disc Drive

Beim **H**ard **D**isc **D**rive handelt es sich um einen rotierenden magnetischen Speicher, welcher erstmals 1956 vorgestellt wurde [10]. Die Speicherkapazität wuchs mit den Jahren, von 3.75 MB im Jahre 1957 (IBM Ramic 350) auf über 15 TB. Wie auf Abbildung 1 erkennbar ist dieser Anstieg logarithmisch.

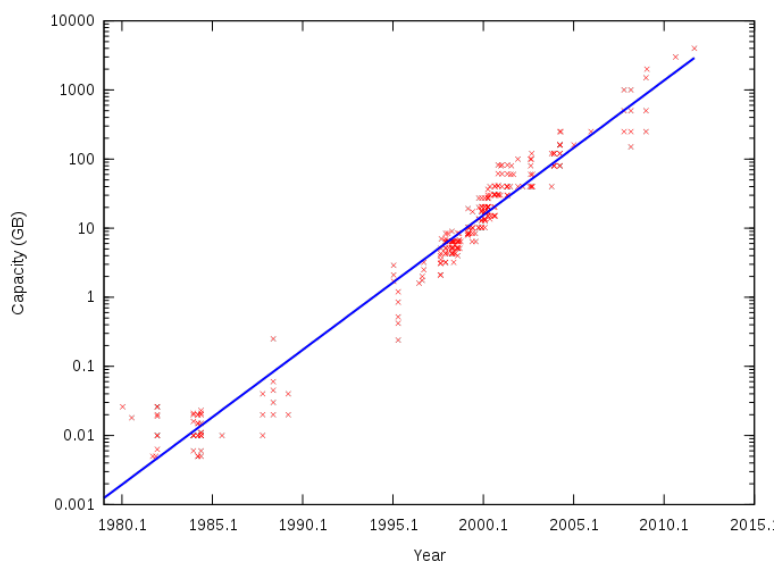


Abbildung 1: HDD Kapazitäten im Laufe der Zeit [10]

Wenn nun die Kapazitäten und Lese-/Schreibgeschwindigkeiten von HDDs verglichen werden ist festzustellen, dass die Kapazitäten einen stärkeren Anstieg verzeichnen (vergleiche Abb.1 und 2).

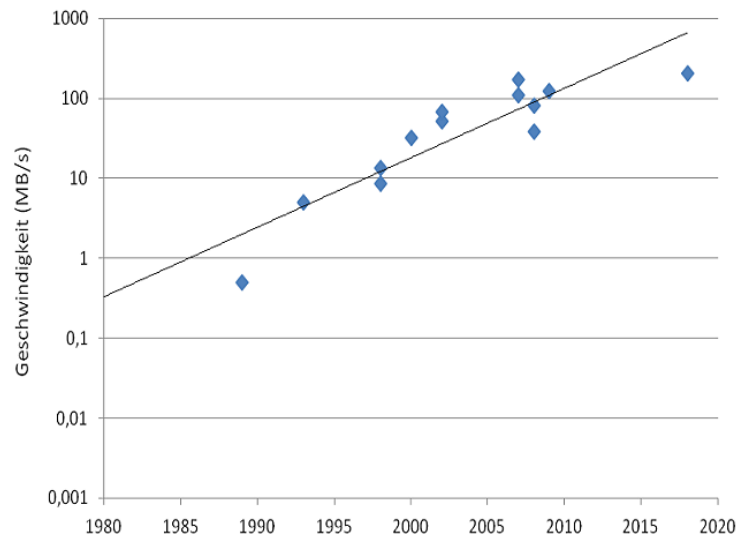


Abbildung 2: HDD Lese-/Schreibgeschwindigkeiten (Daten aufbereitet aus [10] sowie aktuellen Datenpunkten einer Handelsüblichen Enterprise HDD [9])

3 Speicherhierarchie in modernen Computern

In modernen Computern sind eine Vielzahl verschiedener Speicherarten, welche sich sowohl in ihrer Speicherkapazität, als auch Lese- und Schreibgeschwindigkeit stark unterscheiden, zu finden. Desweiteren sind einige Speicherarten *flüchtig* (eng. *non-volatile*). Das heißt sie behalten gespeicherte Daten nur solange Strom an ihnen anliegt. Bei Speichern, die auch ohne anliegenden Strom ihre Daten behalten, spricht man von *nicht flüchtigen* Speichern(eng. *volatile*). Flüchtige Speicher sind in der Abbildung 3 rot, nicht flüchtige in orange dargestellt.

Prozessor Register Auf diesen Registern werden die Berechnungen des Prozessors ausgeführt. Aufgrund der Unabhängigkeit zum BUS-System sind Lese- und Schreibgeschwindigkeit besonders hoch (bis zu 1000 GB/s)[8].

Cache CPU Caches sind auf dem Prozessor befindliche Speichereinheiten. Ihre Aufgabe ist es Daten zwischenspeichern um die Notwendigkeit von ständigen Lese-/Schreibvorgängen auf dem RAM zu minimieren. Moderne CPUs verfügen über bis zu drei Level an Caches, welche eine Geschwindigkeit zwischen 250 und 500 GB/s und eine Kapazität zwischen einigen hundert kB und einigen mB besitzen. Je niedriger das Level ist, desto höher die Geschwindigkeit und kleiner die Kapazität [8].

RAM **R**andom **A**ccess **M**emory ist der Hauptspeicher aus dem Daten in die Caches der CPU geladen werden können. Die Geschwindigkeiten liegen bei 40 GB/s und handelsübliche Größen liegen zwischen 8 und 32 GB.

NV-RAM NV-RAM teilt die meisten Eigenschaften mit RAM, ist allerdings nicht flüchtig. Er findet bei modernen Rechensystemen noch keine reguläre Anwendung.

Flash Memory Flash Memory ist ein Halbleiterspeicher der bei Solid State Drives seine Anwendung findet. Lese-/Schreibgeschwindigkeiten sind kürzer als bei HDD's (500 MB/s), dafür ist die Lebenszeit verkürzt und die Kapazität geringer (von 256 GB bis zu einigen TB).

Hard Drives Bei HDDs sind Lese-/Schreibvorgänge langsamer als bei SSDs 2.4. Dies liegt unter anderem daran, dass gewartet werden muss, bis die richtige Stelle der Platte unter dem Schreibkopf liegt.

Tape Drives Bei Tapespeichern handelt es sich um ein magnetisches Speichermedium, bei dem die Informationen auf ein aufgerolltes Tape aufgespielt werden. Da bis zur richtigen Stelle des Tapes gespult werden muss um Da-

teilen zu lesen, sind die Zugriffszeiten besonders hoch. Dafür ist die mögliche Kapazität eines solchen Mediums mit bis zu 15 TB besonders hoch.

Moderne Rechnersysteme nutzen die meisten Speichersysteme um Kosten zu sparen. Um eine optimale Abwicklung von Prozessen zu ermöglichen, wird nicht unglaublich viel schnellen Speicher benötigt, wenn die benötigten Daten vorausschauend in den Cache oder RAM geladen werden. So kosten zum Beispiel eine 256 GB SSD in etwa das gleiche wie eine 1 TB HDD oder 8 GB RAM.

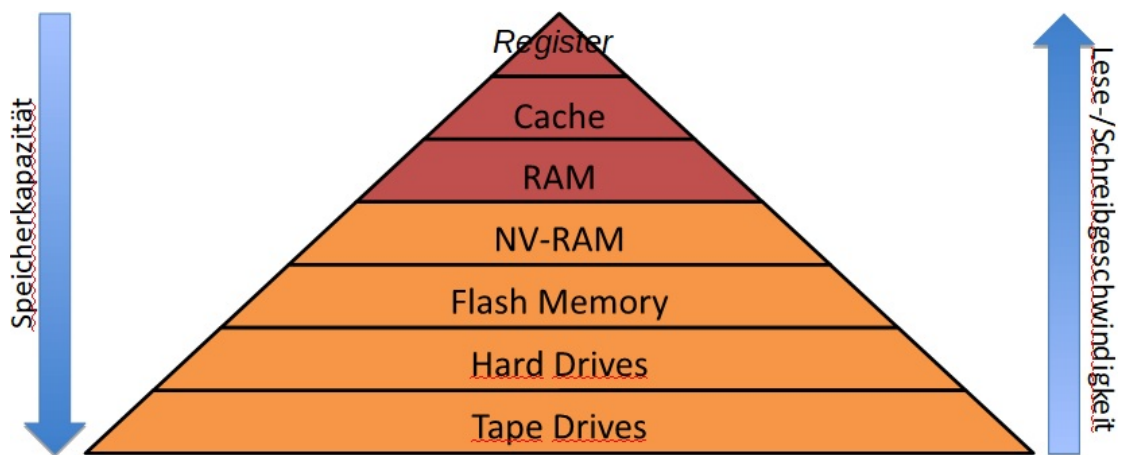


Abbildung 3: Speicherhierarchie in modernen Systemen (nach [8], Figure 2.2)

4 RAID Speichersysteme

4.1 Grundlagen

Bei **R**edundant **A**rray of **I**ndependent **D**isks Speichersystemen handelt es sich um eine Organisation von Massenspeichereinheiten mit dem Ziel die Ausfallsicherheit und/oder den Datendurchsatzes des Systems zu erhöhen. Es gibt verschiedene Arten von RAID-Systemen, welche unterschiedliche Vor- und Nachteile mit sich bringen.

4.1.1 Striping

Striping ist der Prozess des Aufteilens einer Datei in gleichgroße Blöcke, welche dann auf verschiedene Festplatten verteilt werden können.

4.1.2 Mirroring

Beim Mirroring werden Daten 1 zu 1 auf eine andere Festplatte kopiert.

4.1.3 Paritäten

Paritäten sind berechnete Redundanzinformationen, welche dazu benutzt werden können, verlorene Daten wieder herzustellen.

4.2 RAID 0

In RAID 0 Speichersystemen werden Dateien in gleich große Blöcke (sog. Chunks) geteilt und abwechselnd auf den Festplatten gespeichert [siehe Abb. 4]. Durch das Striping kann ein höherer Datendurchsatz erzielt werden, da auf beiden Festplatten gleichzeitig gelesen und geschrieben werden kann. Da keine Redundanz besteht bleibt die komplette Kapazität aller Festplatten bestehen. Sollte jedoch eine dieser Festplatten ausfallen, so ist es nicht mehr möglich die Datei wieder vollständig zusammenzufügen. Es werden mindestens zwei Festplatten für diesen RAID Level benötigt. Mit erhöhter Anzahl steigt jedoch auch die Wahrscheinlichkeit eines Systemausfalls. Die Größe der Chunks hat kaum Auswirkung auf die Performance des Festplattenverbandes [12].

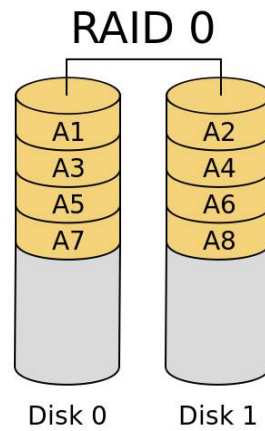


Abbildung 4: RAID 0 System stript Daten ohne Redundanzen zu erzeugen [1]

4.3 RAID 1

RAID 1 spiegelt die kompletten Informationen von einer Festplatte auf alle anderen. Der verfügbare Speicherplatz entspricht dem der kleinsten Festplattengröße im Verbund. Solange noch mindestens eine Festplatte funktionsfähig ist, besteht kein Datenverlust. Der Datendurchsatz ist jedoch geringer als bei RAID 0, da hier alle Schreibvorgänge von allen Festplatten ausgeführt werden müssen. Die Lesegeschwindigkeit ist relativ hoch, da die Software die Daten von der Festplatte liest, deren Lese-/Schreibkopf am nächsten an der Position der zu lesenden Datei ist [13].

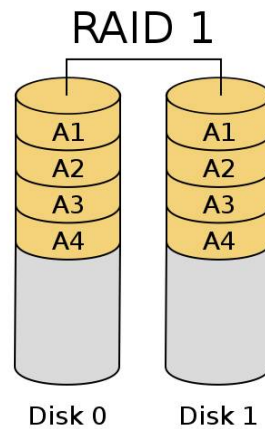


Abbildung 5: TODO

4.4 RAID 2

RAID 2 verwendet bitweises Stripen in Verbindung mit Paritätenberechnung mittels Hamming-Codes. Hierbei wird die gespeicherte Datei so auf die Festplatten verteilt, dass jedes aufeinanderfolgende Bit auf einer anderen Festplatte gespeichert wird. Die Parität berechnet sich durch den Hamming-Code und wird auf einer separaten Festplatte gesichert. Dieses Verfahren hat nur historische Bedeutung und wird heutzutage nicht mehr eingesetzt.

4.5 RAID 3

Bei RAID 3 wird die Datei byteweise gestript und dann auf mehrere Festplatten verteilt. Eine einzelne Festplatte ist bei diesem System jedoch dafür vorbehalten, die Paritätsinformationen zu speichern. Um die Paritäten zu berechnen wird eine XOR-Operation über die Blöcke mit gleichem Index auf den Speicherlaufwerken angewendet. Bei Abbildung 6 also jeweils über A1, A2, A3 sowie A4, A5 und A6. Für den XOR Operanden gilt:

$$A \oplus B = 1 \leftrightarrow A \neq B$$

$$A \oplus B = X \rightarrow A \oplus X = B$$

Ein einfaches Beispiel:

Werte			Ergebnis
0	0	0	0
1	0	0	1
1	0	1	0
1	1	1	1

Dies wird nun mit jeder Stelle der bytegroßen Stripes gemacht, um die Parität zu berechnen:

Index	A1	A2	A3	Ergebnis auf A_p
1	1	1	0	0
2	1	1	1	1
3	1	0	0	1
4	1	0	1	0
5	0	1	0	1
...

Somit würde die Parität $A_p(1 - 3)$ in unserem Beispiel mit 01101 anfangen.

Sollte nun eine beliebige, einzelne Festplatte ausfallen, sodass uns ihre Daten nicht mehr zur Verfügung stünden, gingen keine Daten verloren. Wir könnten über die verbleibenden Werte einfach wieder eine XOR-Operation durchführen und die so verlorenen Daten wiederherstellen.

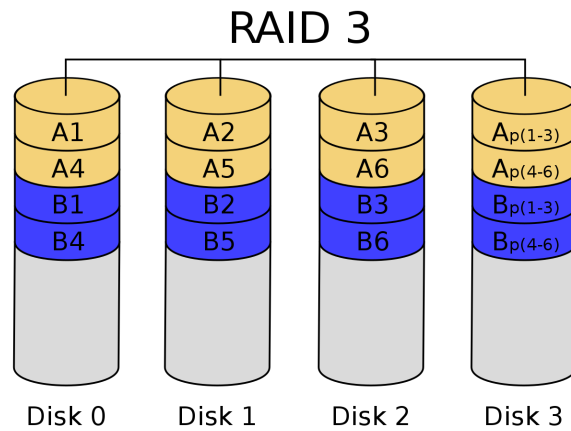


Abbildung 6: byteweises Striping mit Paritäten auf getrennter Festplatte [2]

4.6 RAID 4

RAID 4 funktioniert im Wesentlichen wie RAID 3, jedoch wird nicht byteweise sondern in Blocks (den sogenannten Chunks) gestriped. Diese größeren Blöcke führen zu einer höheren Geschwindigkeit im Vergleich zu RAID 3. Ein großer Nachteil, den sich RAID 3 und RAID 4 teilen ist, dass bei jeder Änderung der Daten auf einem der Datenspeicher auch die Paritätsfestplatte einen Schreibvorgang tätigen muss. Somit müssen Schreibvorgänge auf mehreren Festplatten immer darauf warten, dass der vorherige Schreibvorgang beendet ist. Deswegen ist die Geschwindigkeit der Paritätsfestplatte immer die begrenzende maximal mögliche Geschwindigkeit. Ein Vorteil gegenüber RAID 1 ist eine erhöhte Kapazität.

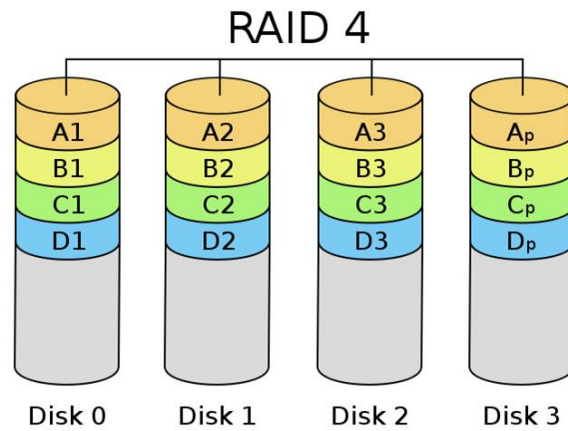


Abbildung 7: blockweises Striping mit Paritäten auf getrennter Festplatte [3]

4.7 RAID 5

Auch bei RAID 5 werden die Paritäten mittels XOR berechnet. Diese werden jedoch, im Gegensatz zu RAID 3 und 4, auf alle Festplatten gleichermaßen verteilt. Dabei bleibt die Sicherheit gegenüber Datenverlust dieselbe. Es ist nun aber möglich Daten auf mehreren Festplatten gleichzeitig zu schreiben. Wenn zum Beispiel A_1 und B_2 überschrieben werden soll, dann muss hierfür auch A_p und B_p beschreiben werden. Dieses kann gleichzeitig durchgeführt werden, da Disk 1 und 2 vom Schreibvorgang von A_1 auf Disk 0 und 3 nicht betroffen sind. Somit erhält man einen höheren Datendurchsatz im Vergleich mit RAID 3 und RAID 4. Desweiteren wird keine Festplatte übermäßig vielen Lese/Schreibvorgängen ausgesetzt, wie das bei der Paritätsfestplatte bei RAID 4 der Fall ist.

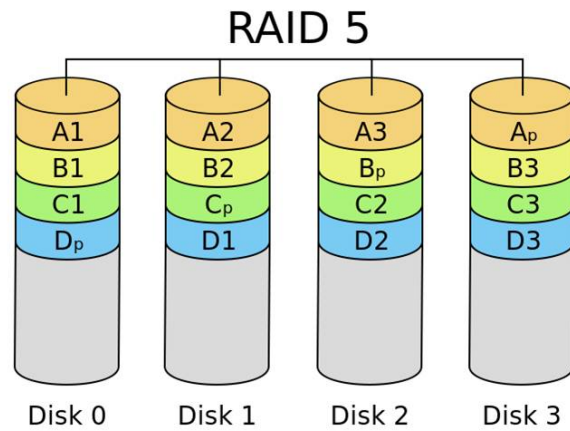


Abbildung 8: blockweises Striping mit verteilten Paritäten [4]

4.8 RAID 6

Bei RAID 6 Systemen ist die Vorgehensweise identisch zu RAID 5 mit dem entscheidenden Unterschied, dass eine zweite Parität über die Daten berechnet wird. Der Vorteil bei diesem System ist, dass es möglich ist alle Daten zu lesen, selbst wenn zwei Datenträger ausgefallen sind. Die Lesegeschwindigkeit ist fast identisch zu RAID 5. Die Schreibgeschwindigkeit hingegen ist schlechter [11].

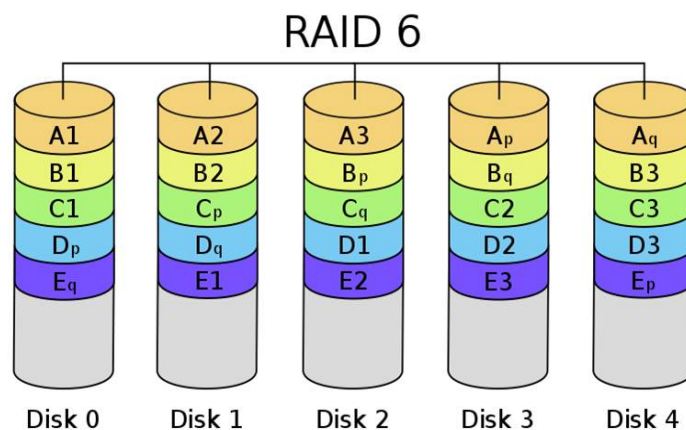


Abbildung 9: blockweises Striping mit doppelter, verteilter Parität [5]

4.9 RAID 01

RAID 01 spiegelt mehrere RAID 0 Verbände. Der Datendurchsatz ist bei RAID 01 äquivalent zu dem von RAID 0 und die Sicherheit zu dem von RAID 1.

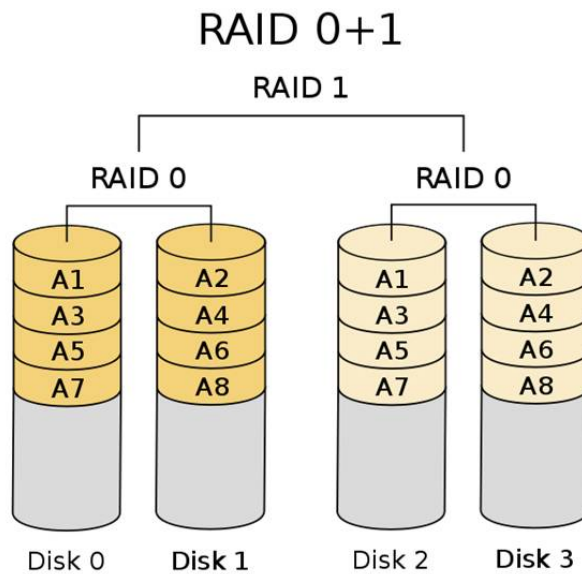


Abbildung 10: RAID 1 über mehrere RAID 0 [6]

4.10 RAID 10

RAID 10 stript Daten auf mehrere RAID 1 Verbände. Der Datendurchsatz bei RAID 10 ist äquivalent zu dem von RAID 0. Die Ausfallsicherheit ist höher als bei RAID 01. Dieses hat den Grund, dass ein RAID 0-System unbrauchbar wird, sobald eine Platte ausfällt. Sowohl RAID 01, als auch RAID 10 überstehen den Verlust einer Festplatte ohne Probleme. Beim Verlust der zweiten Festplatte besteht bei RAID 01 am Beispiel der Abbildung: 10 die Chance von $2/3$, dass es sich beim nächsten Ausfall um eine Festplatte des anderen RAID 0-Systems handelt. Ist dieses der Fall, so fällt das ganze System aus. RAID 10 fällt nur dann aus, zwei Platten ausfallen, welche identische Informationen tragen. Die Wahrscheinlichkeit in Abbildung 11 beträgt also

Beispielsweise nur 1/3.

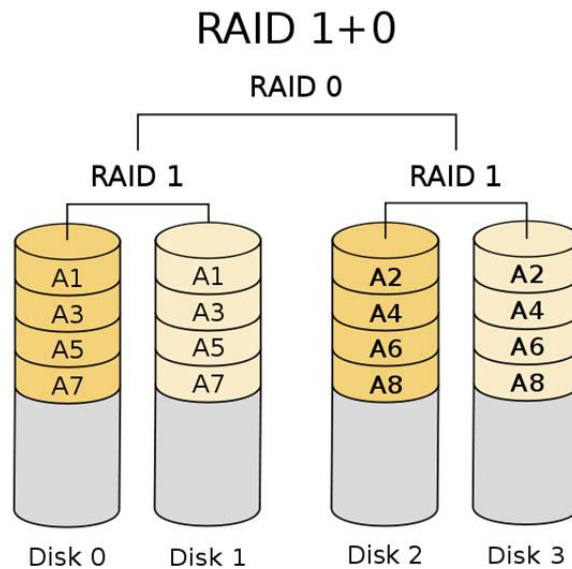


Abbildung 11: RAID 0 über mehrere RAID 1 [7]

5 Zusammenfassung

Zusammenfassend ist zu erkennen, dass Anstieg von Kapazität und Datendurchsatz von Festplatten nicht im gleichen Maße ansteigt. Daher müssen Lösungen gefunden werden, welche diese Lücke zu schließen. Desweiteren herrscht in modernen Rechnersystemen eine Speicherhierarchie vor um Kosten einzusparen und trotzdem die Berechnungen nicht zu verlangsamen. Zuletzt ist festzustellen, dass RAID Speichersysteme verwendet werden um Kapazität, Zuverlässigkeit oder Datendurchsatz gegenüber einzelnen Festplatten zu erhöhen. Für verschiedene Anwendungsfälle gibt es verschiedene RAIDs.

Literatur

- [1] Colin M.L. Burnett. [Online; accessed 10-February-2019]. URL: https://commons.wikimedia.org/wiki/Redundant_array_of_independent_disks#/media/File:RAID_0.svg.
- [2] Colin M.L. Burnett. [Online; accessed 10-February-2019]. URL: https://commons.wikimedia.org/wiki/Redundant_array_of_independent_disks#/media/File:RAID_3.svg.
- [3] Colin M.L. Burnett. [Online; accessed 10-February-2019]. URL: https://commons.wikimedia.org/wiki/Redundant_array_of_independent_disks#/media/File:RAID_4.svg.
- [4] Colin M.L. Burnett. [Online; accessed 10-February-2019]. URL: https://commons.wikimedia.org/wiki/File:RAID_5.svg.
- [5] Colin M.L. Burnett. [Online; accessed 10-February-2019]. URL: https://commons.wikimedia.org/wiki/Redundant_array_of_independent_disks#/media/File:RAID_6.svg.
- [6] Colin M.L. Burnett. [Online; accessed 10-February-2019]. URL: https://commons.wikimedia.org/wiki/Redundant_array_of_independent_disks#/media/File:RAID_01.svg.
- [7] Colin M.L. Burnett. [Online; accessed 10-February-2019]. URL: https://commons.wikimedia.org/wiki/Redundant_array_of_independent_disks#/media/File:RAID_10_01.svg.
- [8] Kira Isabel Duwe. „Simulation of Storage Tiering and Data Migration“. Masterthesis. Universität Hamburg. URL: edoc.sub.uni-hamburg.de/informatik/volltexte/2017/233/pdf/master_duwe.pdf.
- [9] Seagate Technology LLC. *Exos X12 Enterprise Festplatte Datenblatt*. [Online; Stand 10. Februar 2019]. URL: https://www.seagate.com/www-content/datasheets/pdfs/exos-x-12-DS1946-2-1712DE-de_DE.pdf.

- [10] Han Kwang Nienhuys. *Hard drive capacity over time*. URL: https://de.wikipedia.org/wiki/Datei:Hard_drive_capacity_over_time.svg.
- [11] Linux Kernel RAID wiki Team. *Introduction*. URL: <https://raid.wiki.kernel.org/index.php/Introduction>.
- [12] Linux Kernel RAID wiki Team. *Linux Kernel RAID Performance*. URL: <https://raid.wiki.kernel.org/index.php/Performance>.
- [13] Linux Kernel RAID wiki Team. *What is RAID and why should you want it?* URL: https://raid.wiki.kernel.org/index.php/What_is_RAID_and_why_should_you_want_it%3F.
- [14] Wikipedia. *Lochkarte* — *Wikipedia, Die freie Enzyklopädie*. [Online; Stand 10. Februar 2019]. 2019. URL: <https://de.wikipedia.org/w/index.php?title=Lochkarte&oldid=184927068>.
- [15] Wikipedia contributors. *Floppy disk* — *Wikipedia, The Free Encyclopedia*. [Online; accessed 10-February-2019]. 2019. URL: https://en.wikipedia.org/w/index.php?title=Floppy_disk&oldid=879673370.
- [16] Wikipedia contributors. *IBM System/3* — *Wikipedia, The Free Encyclopedia*. [Online; accessed 10-February-2019]. 2019. URL: https://en.wikipedia.org/w/index.php?title=IBM_System/3&oldid=879382508.