

1.5 Halbleiter

- In den 1960er Jahren hat es einige Computer aus diskreten Bauteilen gegeben.
- In 1970 Jahren wurden die ersten 4-Bit-Mikroprozessoren von Intel (4004) entwickelt.
- Massenproduktion von Mikroprozessoren begann in den 1980er Jahren, mit dem 8080 von Intel und dem 6800 von Motorola.

2.2 Klassifizierung von Betriebssystemen

2.2.1 Bitbreite

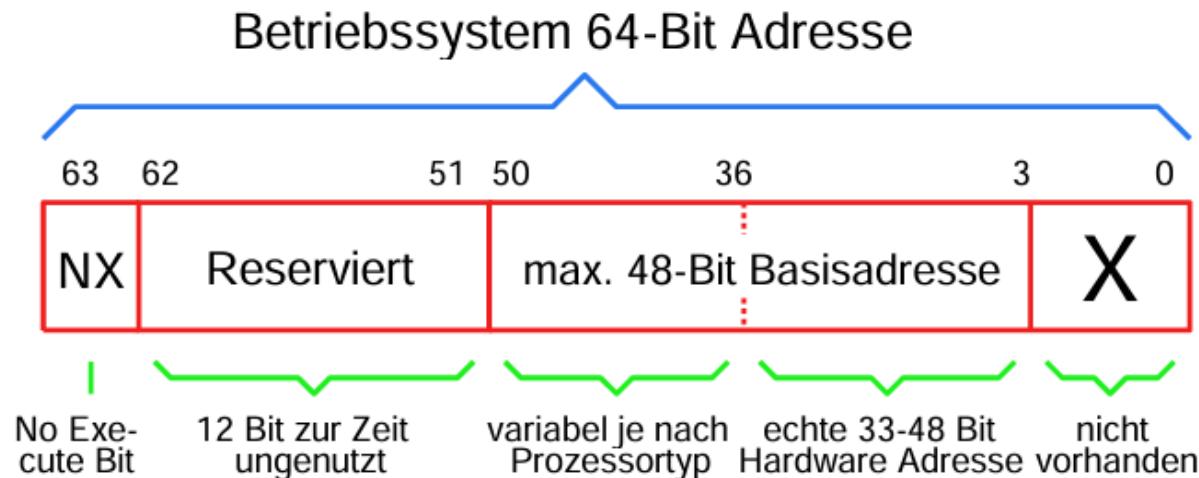
- Bitbreite von Betriebssystemen wird unterschieden durch die Menge des adressierbaren Speichers.
(manchmal wird die Speicherbandbreite künstlich limitiert)
- Heutzutage sind folgende Systeme auf dem Markt:
 - 16-Bit-Systeme (max. 64KB)
 - 20-Bit-Systeme (16Bit + Segmentierung bis 1MB)
 - 32-Bit-Systeme (max. 4GB)
 - 64-Bit-Systeme (max. 16EB = 18 Millionen TB)

2.2.2 64-Bit Verwechslung

- Adressierung eines Datums dauert bei 64-Bit doppelt so lange wie bei 32-Bit, da die Adresse über den Datenbus übertragen werden muss.
- Hype auf 64-Bit-Betriebssysteme entstand durch Verwechslung der Bitbreite des Betriebssystems mit der Systembreite der CPU. Bei einem 64-Bit-Betriebssystem muss jede Speicheradresse 64 Bit breit sein.
- Vorteile von 64-Bit-Betriebssystemen bestehen nur bei Nutzung von mehr als 4 Gigabyte Speicher und wenn eine Anwendung dies benötigt.
- Aktuell gibt es keine Prozessoren mit mehr als 45 echten Adressleitungen.
--> 64-Bit-Betriebssysteme sind in Wirklichkeit nur max. 45-Bit-Betriebssysteme und können nicht mehr als 256 Terabyte echten Speicher ansprechen.

2.2.3 NX-Bit

- Einige Prozessorhersteller haben begonnen, Sonderbedeutungen in diese Bits einzubauen.
- Beispiel: Höchstes Bit der Adresse (Bit Nr. 63) kennzeichnet, ob an der Adresse Daten oder ausführbarer Programmcode abgelegt ist.
- Zweck: Erschwerung für Schadprogramme, Schadcode in den Speicher einzuschleusen.
- No-eXecute-Bit (NX-Bit) dient dieser Sicherheitsfunktion.
- NX-Bit wurde 2003 von AMD mit dem Athlon64 eingeführt.
- Aktuell wird das NX-Bit von allen Betriebssystemen unterstützt.



2.3.3 Arbeitsspeicher/Cache

- Standard-Prozessoren wurden anfangs mit wenigen MHz betrieben.
- Dynamischer Speicher (DRAM) war als Arbeitsspeicher schnell genug.
- Bis in die 90er-Jahre wurde preiswertes dynamisches RAM verwendet.
- Mit Taktfrequenzen über 20 MHz wurde der langsame Speicher problematisch, da die CPU auf den Speicher warten musste.
- Lösung ab den ersten 32-Bit-Systemen: Ein schnellerer Zwischenspeicher (Cache) zwischen CPU und Speicher.
- Cache-Zwischenspeicher wurde als Statisches-RAM (SRAM) ausgeführt.
- Cache-Controller verwaltet, welche Daten im Cache stehen. (*ein eigener kleiner Computer für die Speicherverwaltung*)
- Ziel des Cache-Controllers: sicherstellen, dass benötigte Daten im Cache sind (Cache-Hit), bevor die CPU sie anfordert.
- Bei Cache-Miss muss die CPU auf den langsamen DRAM-Hauptspeicher warten.
- CPU-Hersteller haben ab steigenden Taktfrequenzen den Cache inklusive Cache-Controller in die CPU integriert (bei Intel ab i486).
- Anfangs waren nur kleine Cache-Größen (<16 kByte) möglich.
- Externer Cache wurde als Second-Level-Cache beibehalten und später integriert (ab ca. 2000).
- Neueste Entwicklungen (ab 2018) integrieren einen Third-Level-Cache (L3-Cache) bis über 20 MB in die CPU.

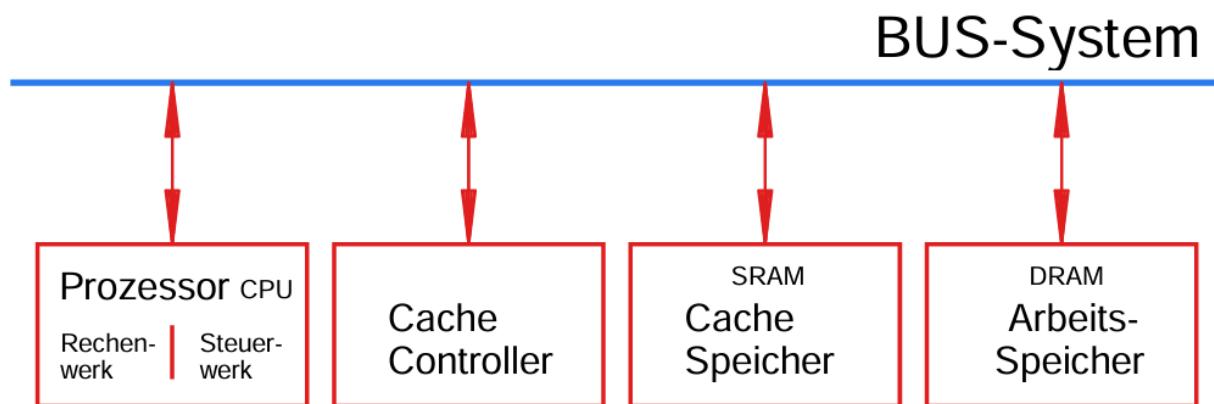


Abbildung 27 Cache

2.3.4 CPU

2.3.4.1 Generischer Aufbau einer CPU

- CPU (Central Processing Unit) ist das zentrale Element in einem Computer.
- Früher das einzige aktive Element, mittlerweile gibt es in modernen Computern mehrere, auch unterschiedliche CPUs.
- CPU bedient alle Busse: Adressbus, Datenbus, Steuerbus und setzt verschiedene Vorgänge in Gang.
- Interner Aufbau der CPU wird mit Registern dargestellt (1 Bit Speicher-Bausteine in Gruppen von 8-16-32 oder 64-Bit *jeweils Abhängig von der CPU-Architektur*)
- Frühe CPUs nutzten einen Akkumulator für arithmetische oder logische Funktionen.
- Moderne CPUs können Operationen mit jedem Register durchführen.

2.3.4.2 Aktueller CPU-Aufbau

- Langsamer Speicherzugriff auf DRAM-Hauptspeicher bleibt trotz 3 Cache-Ebenen ein Problem.
- Erste Lösungsversuche: 32-Bit-Prozessoren mit 64-Bit-Datenbus nach außen. --> Verdoppelung der Datenübertragungsgeschwindigkeit. **ABER** heutzutage nicht mehr möglich, da CPUs intern mit 64-Bit arbeiten.

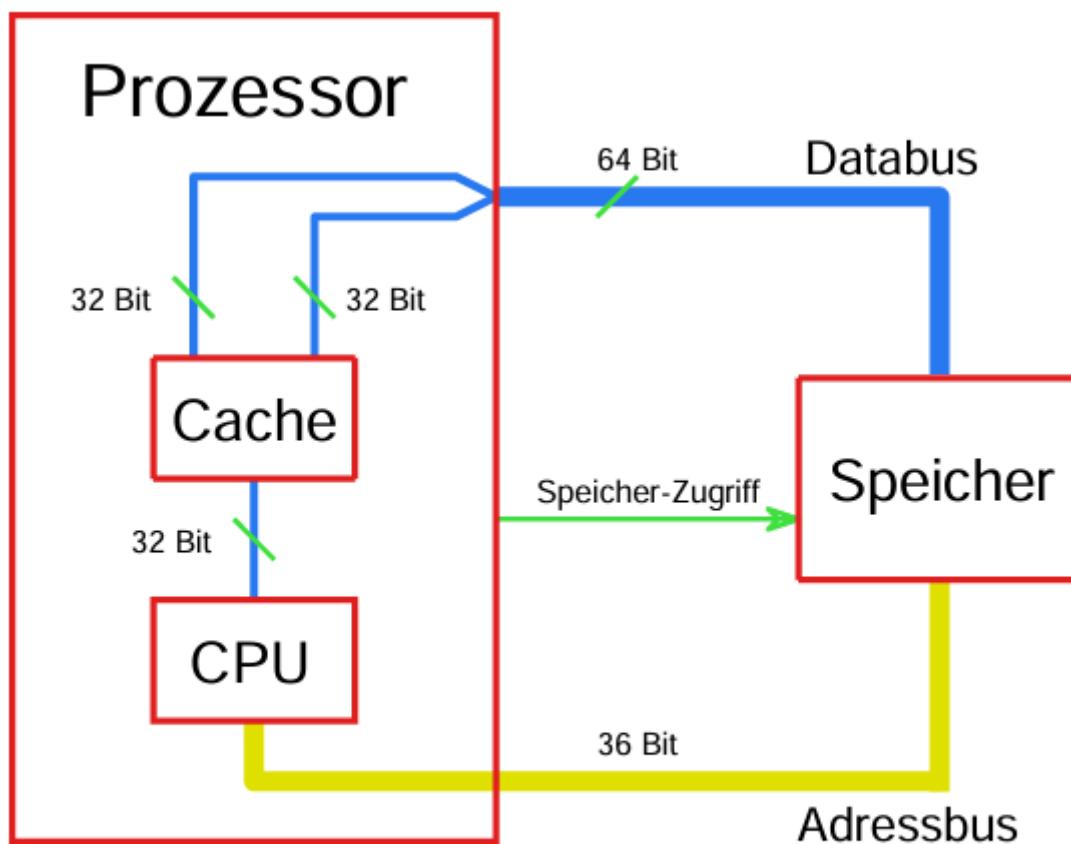


Abbildung 29 CPU 32/64 Bit Multiplexing

- Moderne Prozessoren (ab ca. 2010) haben keine klassische Busstruktur mehr außerhalb der CPU. (*Aber weiterhin innerhalb der CPU*)
- Speicherinterface wurde stark verändert, um die Datenübertragung zu beschleunigen.

- Mehrere Speicherkanäle und gemultiplexte Adressleitungen sind jetzt vorhanden. (DRAM-Speicher wird sowieso in 2 oder 4 Kanälen angesprochen, also Aufteilung der Speicheradressen)

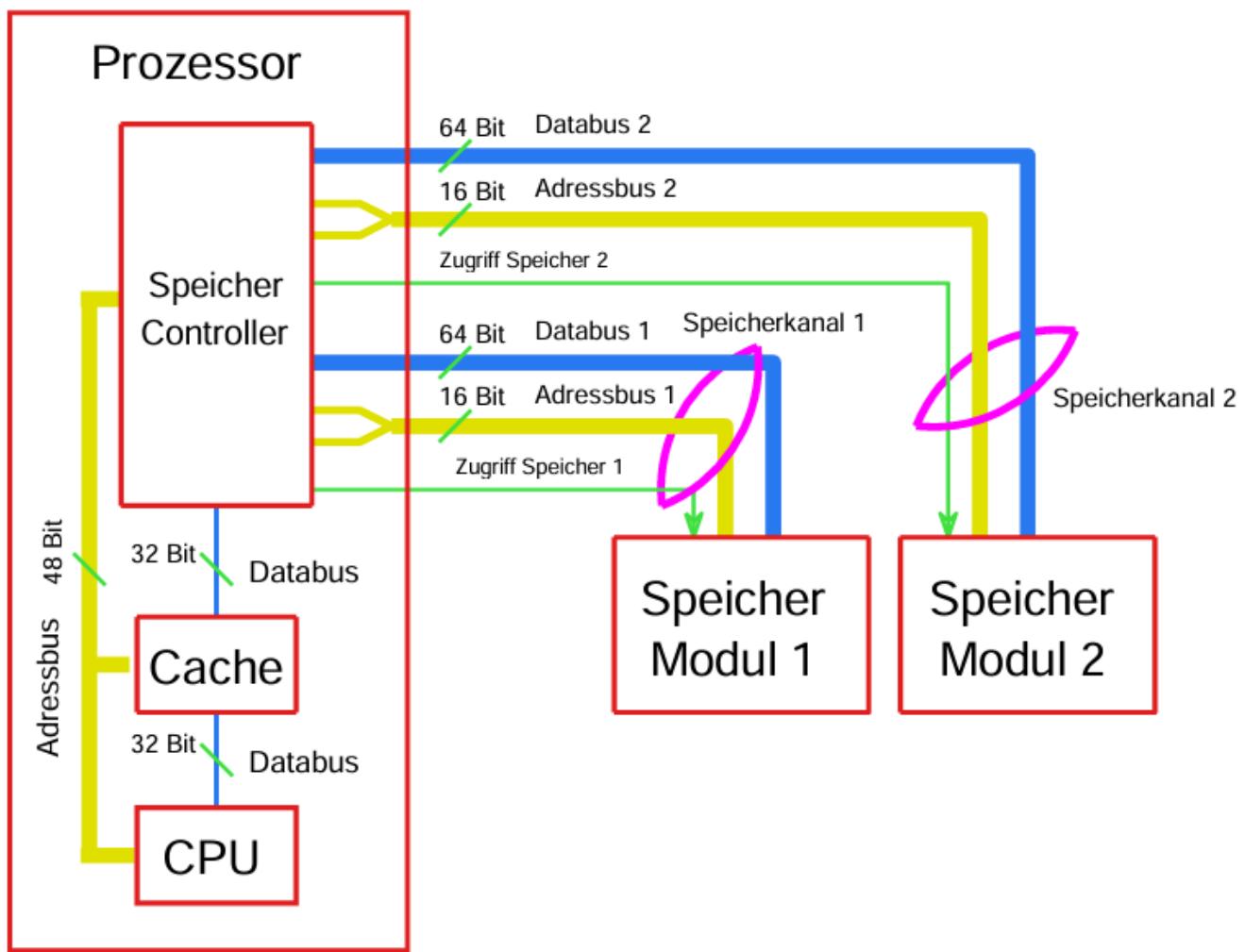


Abbildung 30 CPU mit 2 Speicherkanälen

2.3.4.3 Fehlende Adressleitungen (ab 32-Bit-Prozessoren)

- Speicheradressierung ist historisch bedingt bytewise organisiert.
- 32-Bit-Prozessoren übertragen immer 4 Bytes auf einmal, daher werden die beiden untersten Adressleitungen nicht mehr benötigt. ($da 2^2 = 4$)
 - Beispiel: Angenommen, ein 32-Bit-Prozessor möchte auf die Adresse 0x00000004 zugreifen. Da er immer 4 Bytes auf einmal liest, überspringt er die Adressen 0x00000001, 0x00000002 und 0x00000003, weil diese alle innerhalb des ersten 4-Byte-Wortes (Adresse 0x00000000) liegen würden.
- Einzelnes Byte wird durch zusätzliche Busleitungen "Byte Enable" BE0# - BE3# adressiert.
- Bei 64-Bit-Prozessoren wird dies mit acht Byte Enable Leitungen BE0# - BE7# weitergeführt. (Da hier immer 8 Bytes übertragen werden) ($2^3 = 8$)

2.3.4.4 Takt und Timing

- Erste Computer hatten genau einen Systemtakt von einem Taktoszillator (meistens Quarzoszillator).
- Ab den 1990ern wurde der Takt innerhalb der CPU erhöht, außerhalb entwickelte sich der Takt auf einige 100 MHz weiter.
- Hohe Taktfrequenzen auf Leiterplatten führen zu Signalintegritätsproblemen.

- Innerhalb der CPU wird mit einem höheren Takt gearbeitet, die Taktvervielfachung erfolgt mittels PLL (*Phase-Locked Loop: elektronische Schaltung zur Frequenzvervielfachung*)
- Speichercontroller in der CPU entkoppelt die unterschiedlich schnellen Busse.
- Maximale Taktfrequenz von 200-300 MHz wurde um das Jahr 2000 erreicht.
- Schnellere Datenübertragung durch 2 oder 4 Datenwörter pro Taktperiode.

2.3.6.2 Magnetischer Massenspeicher

Bandlaufwerke

- Anfänge: Magnetische Massenspeicher begannen mit Magnetbändern, wobei anfangs echte Audio-Tonbänder verwendet wurden. In den 1980er Jahren wurden normale Audiokassettenrekorder mit Kompaktkassetten genutzt.
- Datenaufzeichnung: Häufig erfolgte die Aufzeichnung frequenzmoduliert, was gut zur Audiotechnologie passte.
- Aktuelle Nutzung: Magnetbänder werden heute noch als Streamer für Backups verwendet. Es gibt verschiedene Formate, die digital auf Band aufzeichnen.
- Vor- und Nachteile: Der sequenzielle Zugriff ist sowohl ein Hauptvorteil als auch ein Nachteil. Die Zugriffszeit kann im Minutenbereich liegen, was den schnellen Zugriff auf Daten behindert, aber auch vor Schadprogrammen schützen kann, da nicht alle Daten direkt erreichbar sind und somit nicht sofort zerstört werden können.
- LTO-Ultrium-Bandlaufwerke: Die aktuellen Streamer-Techniken basieren auf LTO-Ultrium-Bandlaufwerken (Linear Tape Open), die verschiedene Generationen von LTO1 bis LTO8 umfassen, wobei einige rückwärtskompatibel sind.

<i>Laufwerkstyp</i>	<i>Speichergröße</i>
LTO-1	100 GB
LTO-2	200 GB
LTO-3	400 GB
LTO-4	800 GB
LTO-5	1.5 TB
LTO-6	2.5 TB
LTO-7	6 TB
LTO-8	12 TB

Diskettenlaufwerke (Floppy-Disk)

- Historie: Die Diskette war in den 1980er Jahren der mobile Massenspeicher für Computer. Sie fungierte als Zwischenlösung zwischen Magnetbandspeicher und Magnetplatte.

- Mechanisches Verhalten: Ähnlich dem Magnetband mit dem Vorteil des wahlfreien Zugriffs.
- Nachteile: Da es direkten Kontakt zwischen Schreiblesekopf und Medium gab, litten Disketten unter starkem Verschleiß und hatten eine geringe Lebensdauer.
- **ACHTUNG** Das Datenformat auf Disketten ist über verschiedene Computertypen hinweg nicht kompatibel.
- Entwicklung: Disketten wurden im Laufe der Jahre kleiner und wiesen eine höhere Speicherdichte auf.
- Ende der Nutzung: Mit dem Aufkommen von Flash-Speichern im USB-Stick-Format wurden Disketten ab ca. 2005 fast vollständig aufgegeben, abgesehen von einigen professionellen Anwendungen (z.B. Musiker-Equipment).

Übliche Größen	Einführt	Speicher-Kapazität
8 Zoll	1970er	80 kByte bis 256 kByte
5,25 Zoll	1980er	360 kByte bis 1,2 MByte
3,5 Zoll	1980er	720 kByte bis 1,4 MByte

Abbildung 39 Diskettengrößen

Festplattenlaufwerke (HDD)

- Funktionsweise: Magnetplatten funktionieren ähnlich wie Magnetbänder oder Disketten, wobei der Schreib-/Lesekopf über der Platte schwebt. Der Luftwirbel, der durch die Rotation der Platte entsteht, sorgt dafür, dass der Schreib-/Lesekopf die Platte im Datenbereich niemals berührt. Falls die Platte ausgeschaltet wird und die Umdrehung stoppt, fährt der Schreib-/Lesekopf auf eine Parkposition. (*Auf der Parkposition sind keine Daten gespeichert*)

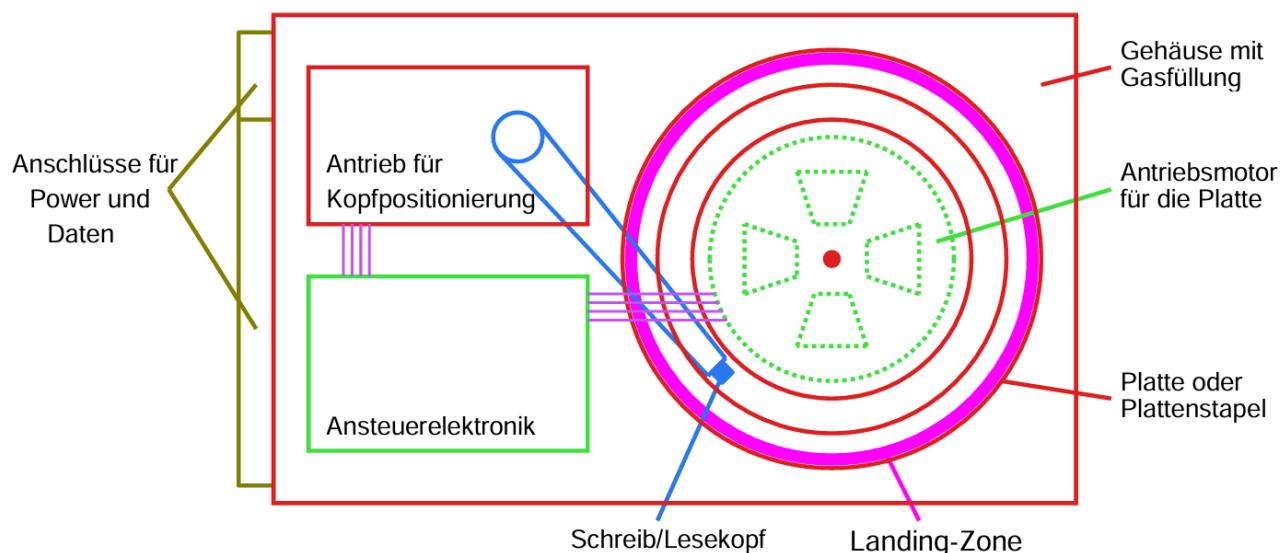


Abbildung 41 Aufbau Standard-Festplatte

- Vorteile: Durch die hohe Rotationsrate bieten Festplatten kurze Zugriffszeiten und hohe Übertragungsraten.
- Entwicklung: Festplattenkapazitäten haben sich von anfangs einigen Megabytes auf heute über 10 Terabyte pro Platte gesteigert. Die Schnittstellen haben ebenfalls eine Evolution durchlaufen, anfangs MFM, dann parallel (IDE/SCSI) und schlussendlich zu Serial ATA (SATA) und SAS.

- Giant Magneto Resistance: Bis zur Entdeckung des GMR-Effekts (Giant Magneto Resistance) war der Speicherzuwachs stagnierend. Mit dem GMR-Effekt konnten die Speicherdichten weiter erhöht werden.
- Shingled Magnetic Recording:
 - Prinzip:
 - Schreibkopfgröße: Der Schreibkopf ist physikalisch größer als der Lesekopf. Dies führt dazu, dass beim Schreiben die benachbarten Spuren teilweise überschrieben werden.
 - Überlappendes Schreiben: Bei SMR werden die Magnetzonen absichtlich überlappend geschrieben. Der nachfolgende Schreibvorgang überschreibt teilweise die vorherigen Magnetzonen, was Platz spart und die Datendichte erhöht.
 - Vorteile:
 - Erhöhte Speicherkapazität: Durch das Überlappen der Spuren können mehr Daten auf derselben Plattenfläche gespeichert werden, was zu einer höheren Speicherdichte führt.
 - Nachteile:
 - Komplexes Schreiben: Da benachbarte Spuren überschrieben werden, müssen beim Ändern von Daten größere Datenblöcke erneut geschrieben werden.
 - Verwaltungsaufwand: Es muss ein zusätzlicher Verwaltungsaufwand betrieben werden, um sicherzustellen, dass bei einer Änderung der Daten der gesamte betroffene Block neu geschrieben wird.
 - Technische Details:
 - Lücken: Um die vollständige Überschreibung der gesamten Platte zu vermeiden, werden in regelmäßigen Abständen Lücken freigelassen.
 - Schreibvorgänge: Bei einer Datenänderung muss der gesamte Block bis zur nächsten Lücke zuerst ausgelesen und danach wieder komplett neu geschrieben werden.

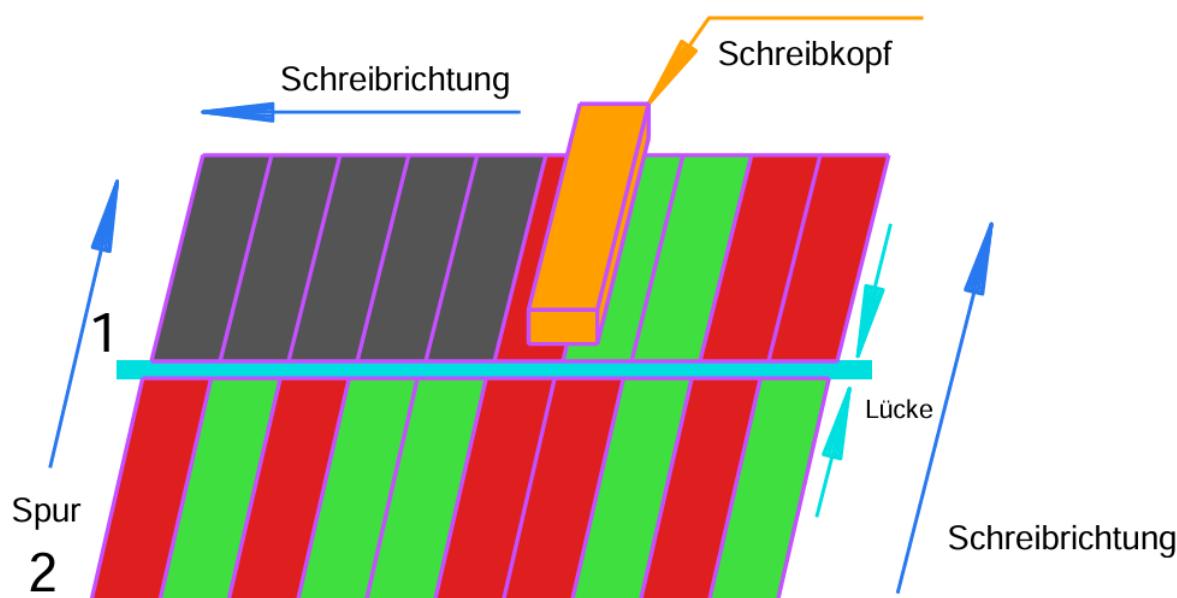


Abbildung 45 Herkömmliche Aufzeichnung

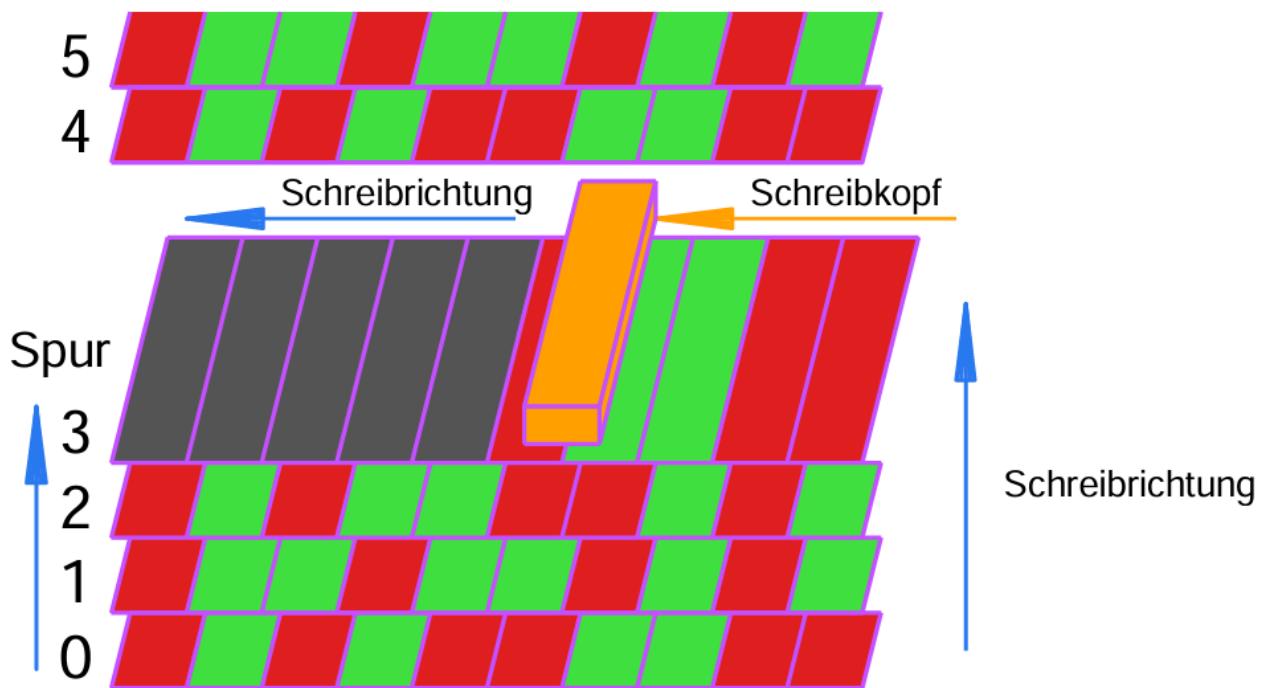


Abbildung 46 Shingled Magnetic Recording

- Ausblick:
 - Heat Assisted Magnetic Recording (HAMR): --> Mithilfe eines Lasers wird die Magnetschicht erhitzt, wodurch die notwendige magnetische Feldstärke reduziert wird und damit der Schreibkopf kleiner gemacht werden kann.
 - Microwave Assisted Magnetic Recording (MAMR): --> Mithilfe von elektromagnetischer Mikrowellenstrahlung wird die Magnetschicht "magnetisch aufgeweicht", wodurch die notwendige magnetische Feldstärke reduziert wird und damit der Schreibkopf kleiner gemacht werden kann.

2.4 Architekturen

2.4.1 Von-Neumann-Architektur

- Grundkonzept:
 - Ein einziges Bussystem: Es gibt nur ein Bussystem, das für die Übertragung aller Datenarten verwendet wird. Dies umfasst sowohl die Programmdaten als auch die Programmbefehle.
- Speicher:
 - Der gleiche Speicher wird für Daten und Programme genutzt, was bedeutet, dass der Speicher und der Bus abwechselnd für beide genutzt werden müssen.
- Vorteile:
 - Einheitlicher Speicher: Es wird nur eine "Art" von Arbeitsspeicher benötigt, der sowohl für Daten als auch für Programme zuständig ist. Dies vereinfacht die Speicherverwaltung und reduziert die Kosten.
- Nachteile:
 - Bus- und Speicherüberlastung: Da der Bus und der Speicher sowohl für Daten als auch für Programme verwendet werden, kann es zu Engpässen kommen, wenn beide gleichzeitig benötigt werden.

- Cache-Trashing: Mit der Einführung von Caches in den 90er-Jahren entstand das Problem, dass Daten und Programmbefehle sich gegenseitig aus dem Cache verdrängten. Dieses Problem wurde teilweise durch die Einführung von getrennten Caches für Daten und Programmbefehle gelöst, was eher einer Harvard-Architektur entspricht.
- Selbstmodifizierender Code: Ein großer Nachteil der Von-Neumann-Architektur ist die Möglichkeit, dass Software ihren eigenen Programmcode ändern kann. Dies kann von bösartiger Software ausgenutzt werden, um sich auszubreiten.

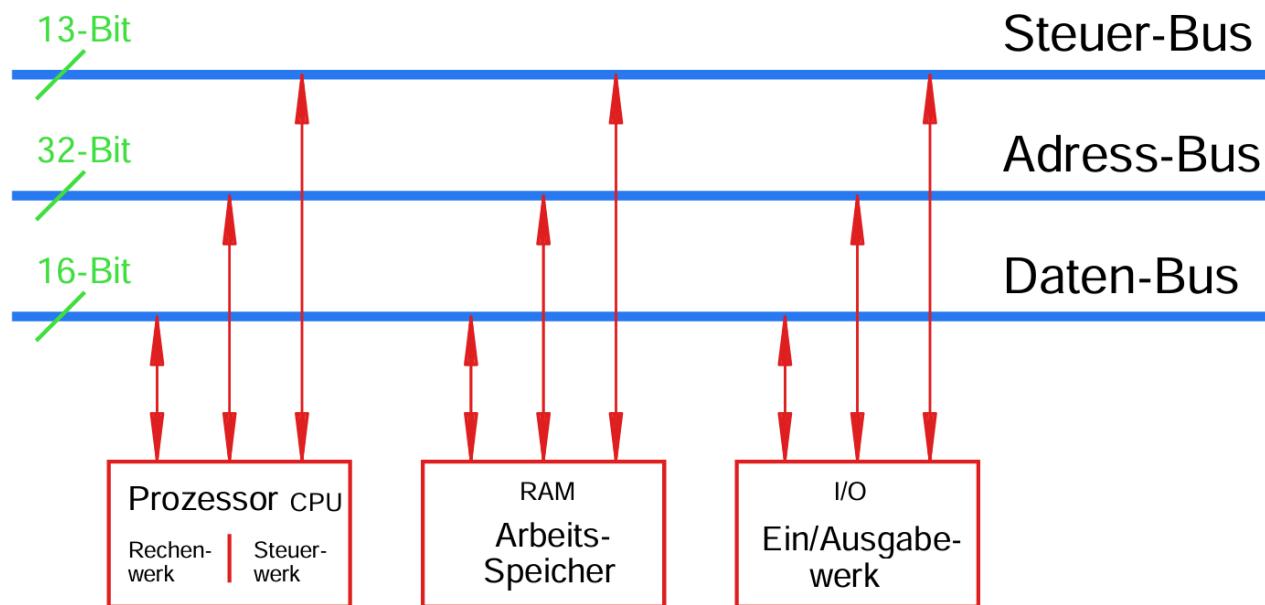


Abbildung 62 Von-Neumann

2.4.2 Harvard-Architektur

- Grundkonzept:
 - Getrennte Speicher und Busse: Die Harvard-Architektur zeichnet sich durch physikalisch getrennte Speicher- und Bussysteme für Programmcode und Daten aus.
- Vorteile:
 - Schneller Speicherzugriff: Durch die Trennung der Busse für Programmcode und Daten kann der Prozessor gleichzeitig auf beide Arten von Speicher zugreifen. Dies führt zu einer theoretischen Verdopplung der Zugriffsgeschwindigkeit im Vergleich zur Von-Neumann-Architektur.
 - Sicherheit: Ein laufendes Programm kann seinen eigenen Programmcode *fast* nicht überschreiben, was das Risiko von bösartigen Softwaremanipulationen reduziert. Der Programmspeicher ist während der normalen Programmausführung nur lesbar und oft in ROM (Read-Only Memory) gespeichert, insbesondere bei eingebetteten Systemen.

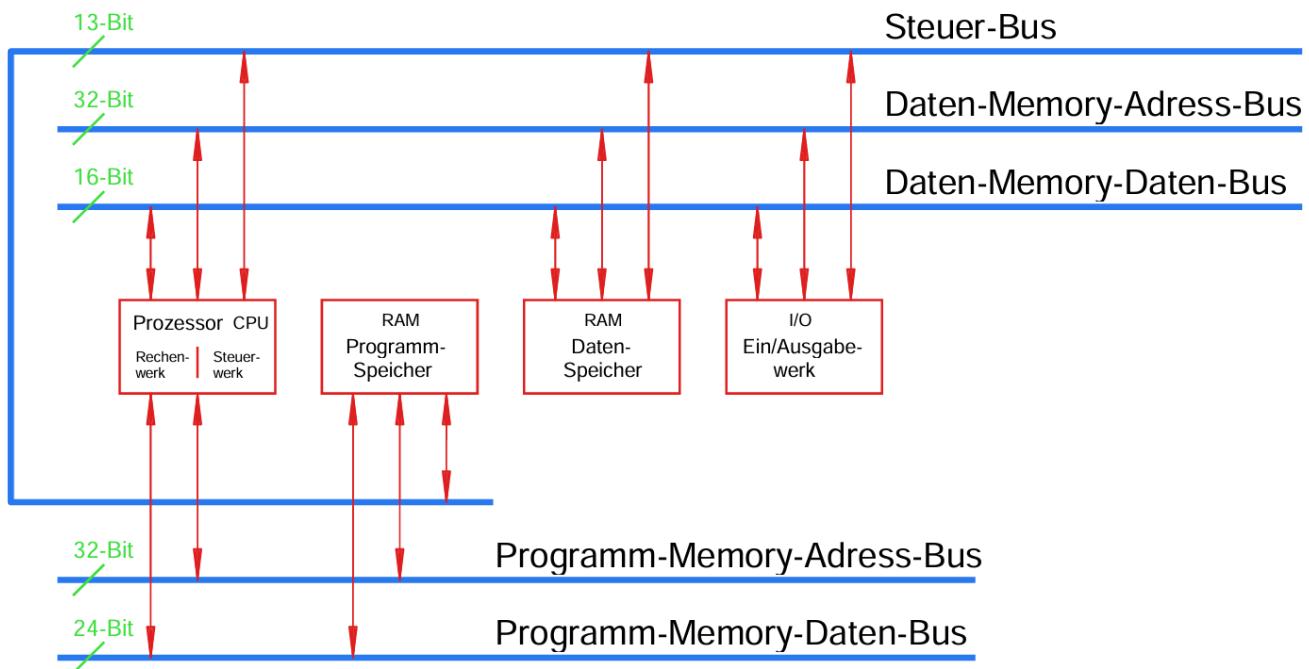
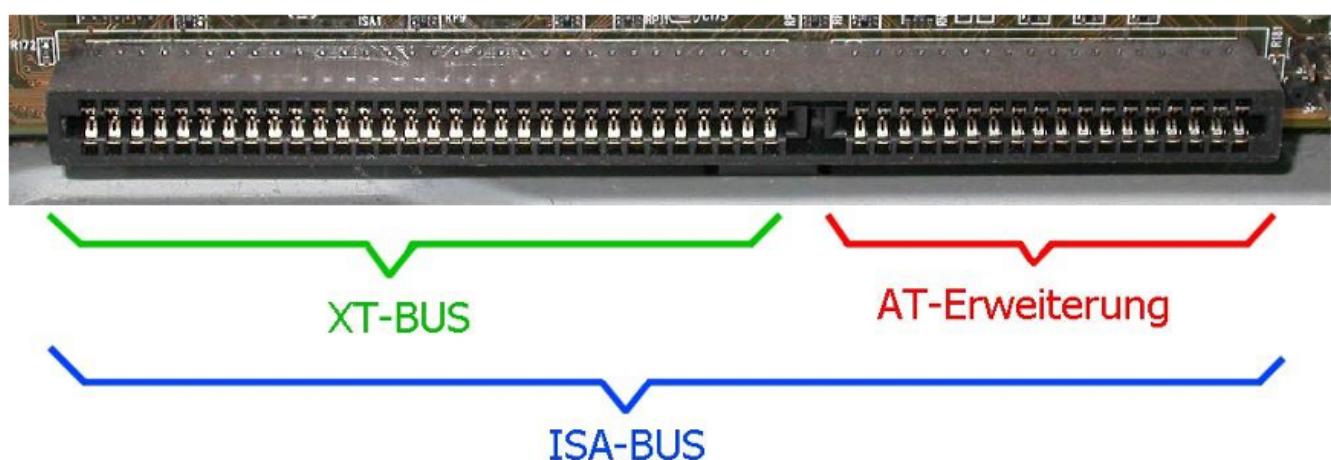


Abbildung 63 Full Harvard

2.5 PC-Bussysteme

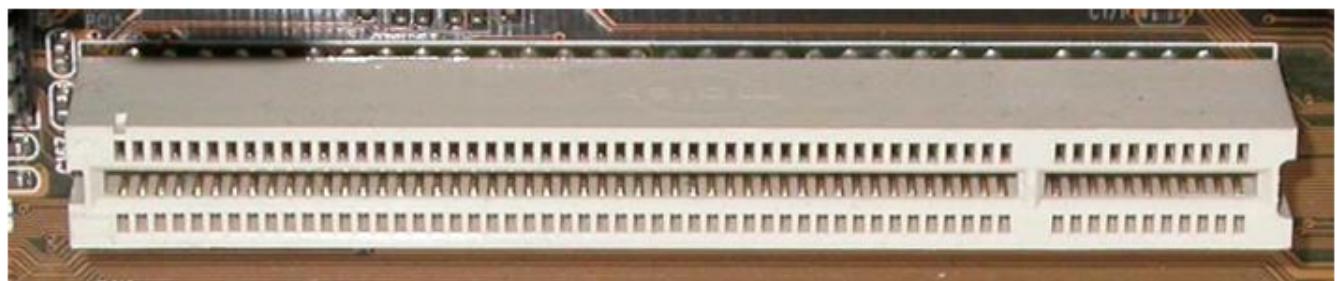
2.5.1 ISA-Bus (*Industry Standard Architecture*)

- Einführung und Entwicklung:
 - Entwickelt von IBM in den 1980er-Jahren.
 - Ursprünglich als XT-BUS mit 8-Bit-Datenbusbreite und 4,7 MHz.
 - Später erweitert auf 16 Bit mit einer Taktfrequenz von 8,33 MHz (Übertragungsrate 16,6 MB/s).
- Eigenschaften:
 - Der gesamte Systembus des Computers ist direkt auf den ISA-Stecker geführt.
 - Ein Entwurf zur Erweiterung auf 32 Bit (EISA) wurde nicht erfolgreich umgesetzt.
- Verwendung:
 - Auch heute noch in Industrieanwendungen verwendet, oft mit zusätzlicher ISA-Hardware-Emulation, um moderne Systeme nicht mit niedrigen Geschwindigkeiten zu betreiben.



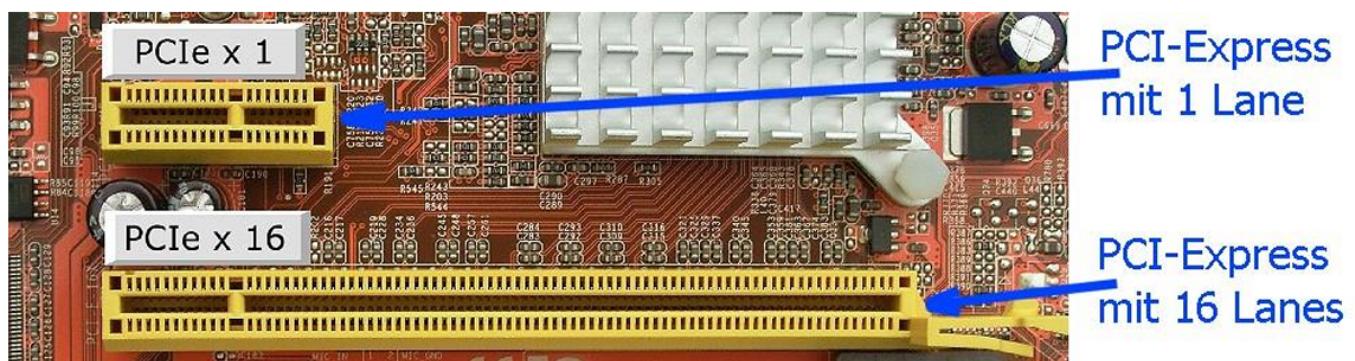
2.5.2 PCI-Bus (*Peripheral Component Interconnect*)

- Einführung und Entwicklung:
 - Mitte der 1990er-Jahre von einem Konsortium mehrerer Firmen entwickelt.
 - **Keine** direkte Verbindung mehr zur CPU, sondern über den Chipsatz des Prozessors.
- Eigenschaften:
 - Daten und Adressleitungen werden gemultiplexed.
 - Daten- und Adressbreite: 32 Bit.
 - Taktfrequenz: 33 MHz.
 - Übertragungsrate: Maximal 133 MB/s im Burst-Modus.
- Erweiterungen:
 - 64-Bit-Variante (PCI-X) für Serverboards, die Taktfrequenz wurde erhöht und QDR eingeführt, wodurch Übertragungsraten bis zu 4,2 GB/s erreicht wurden.



2.5.3 PCIe-Bus (PCI Express)

- Problematik paralleler Bussysteme:
 - Signale müssen auf allen Leitungen gleich lang unterwegs sein, was durch Mäander auf Leiterplatten erreicht wurde.
 - Physikalische Grenzen bei parallelen Konzepten wurden erreicht.
- Eigenschaften:
 - Serielle Punkt-zu-Punkt-Verbindung (Lane) ersetzt den eigentlichen BUS.
 - Taktfrequenz: ca. 2,5 GHz (PCIe-1), was eine Verdopplung der Übertragungsrate von 32-Bit-PCI ergibt.
 - Serielle Taktübertragung erfolgt mittels PLL (Phase-Locked Loop).
 - Mehrere Lanes können zusammengeschaltet werden (bis zu 16 Lanes), ohne eine parallele Übertragung zu bilden.
- Vorteile:
 - Störsichere Übertragung durch differenzielle Paare.
 - Für Software unsichtbar, da die parallel-seriell-parallel-Wandlung von der Hardware übernommen wird.



3.3.1 Rechenwerk Addition

3.3.1.1 Halbaddierer

Addition von einstelligen Dualzahlen

Rechenregeln:

Für die erste Stelle

$$0+0=0$$

$$0+1=1$$

$$1+0=1$$

$$1+1=0$$

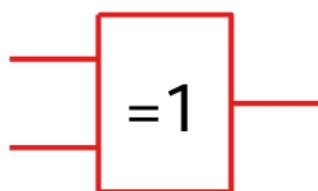


Abbildung 72 Exklusiv ODER

Für die zweite Stelle

$$0+0=0$$

$$0+1=0$$

$$1+0=0$$

$$1+1=1$$

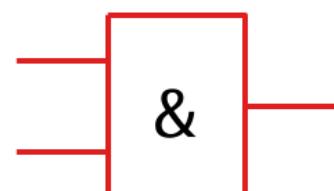


Abbildung 73 UND

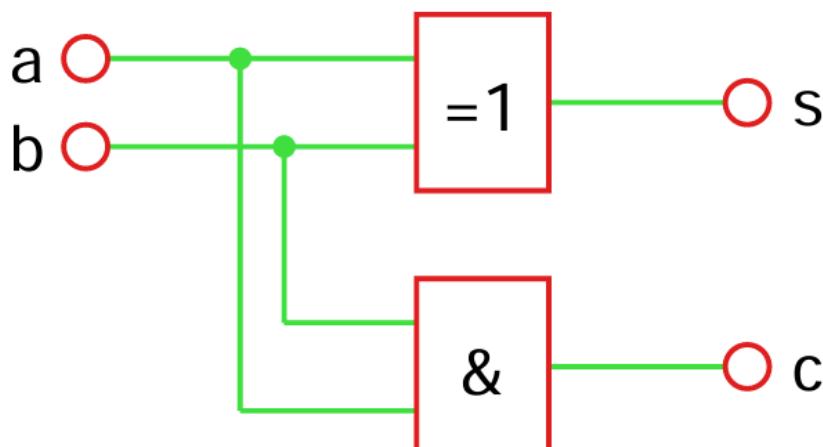


Abbildung 74 Halbaddierer

Wahrheitstabelle:

a	b	c	s
0	0	0	0
0	1	0	1
1	0	0	1
1	1	1	0

Rechenregeln:
Für die erste Stelle

$$\begin{aligned}0+0 &= 0 \\0+1 &= 1 \\1+0 &= 1 \\1+1 &= 0\end{aligned}$$

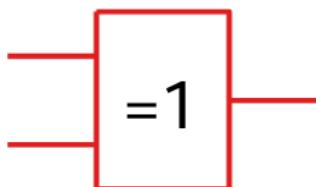


Abbildung 72 Exklusiv ODER

Für die zweite Stelle

$$\begin{aligned}0+0 &= 0 \\0+1 &= 0 \\1+0 &= 0 \\1+1 &= 1\end{aligned}$$

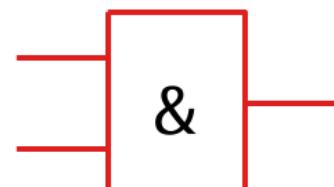


Abbildung 73 UND

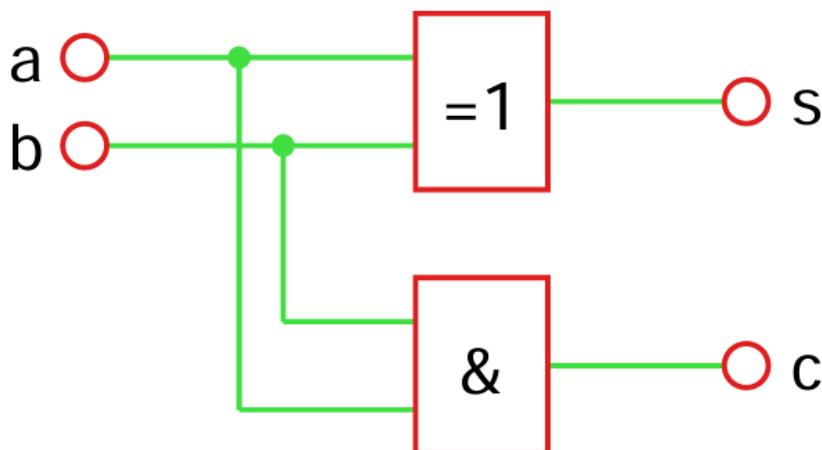


Abbildung 74 Halbaddierer

Wahrheitstabelle:

a	b	c	s
0	0	0	0
0	1	0	1
1	0	0	1
1	1	1	0

3.3.1.2 Volladdierer

Addition von Dualzahlen mit Übertrag

Addition von zwei einstelligen Dualzahlen mit einem Übertrag

Wahrheitstabelle:

c_{i-1}	a_i	b_i	c_i	s_i
0	0	0	0	0
0	0	1	0	1
0	1	0	0	1
0	1	1	1	0
1	0	0	0	1
1	0	1	1	0
1	1	0	1	0
1	1	1	1	1

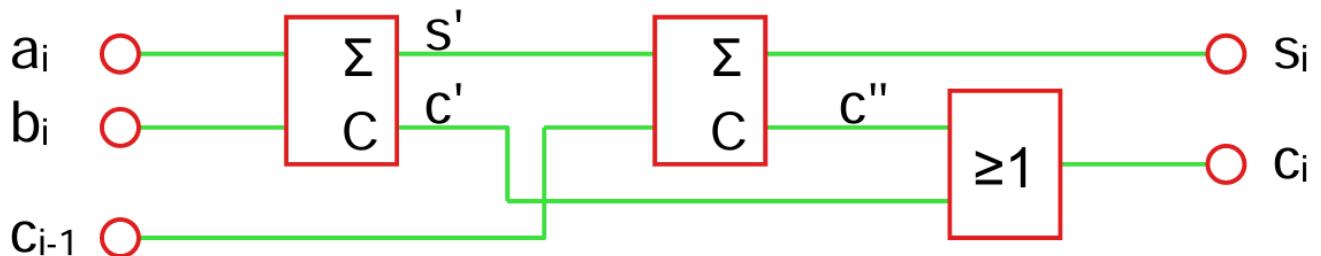


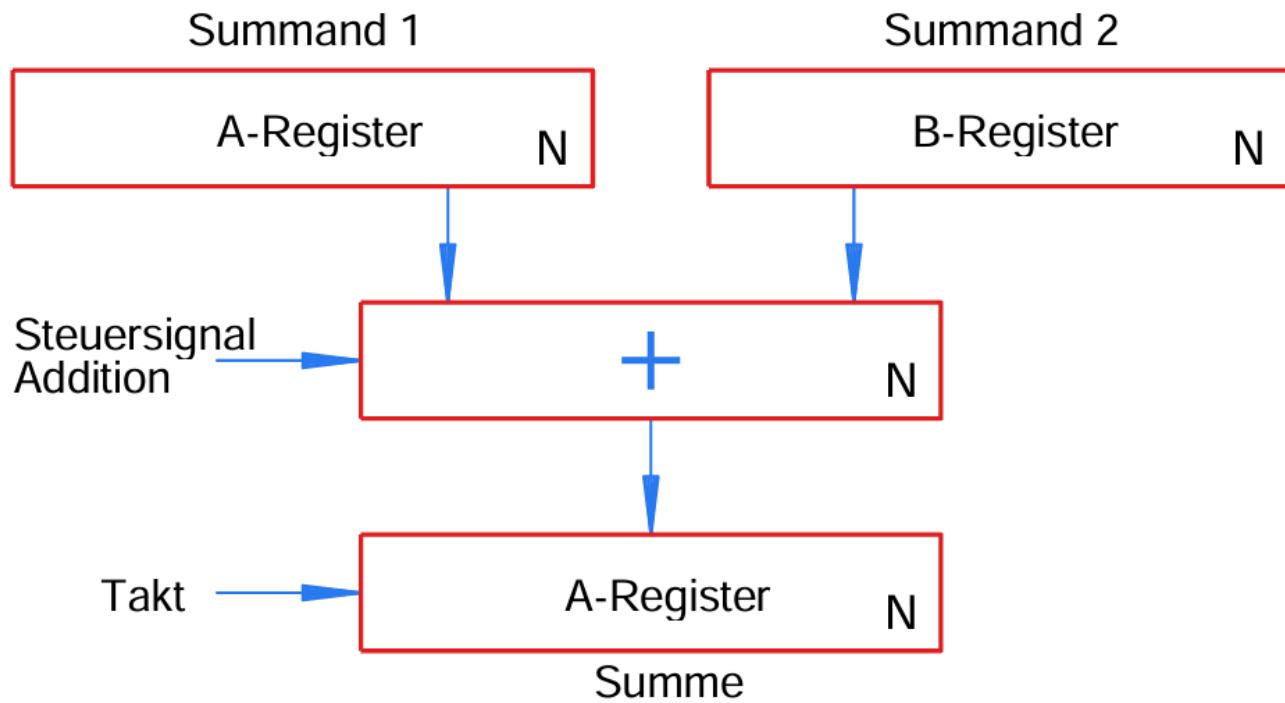
Abbildung 77 Volladdierer aus Halbaddierern

3.3.1.3 Paralleladdierwerk

Addition von Dualzahlen mit mehreren Stellen

Ablauf: Um eine komplette Addition zweier Register durchzuführen:

1. Summanden stehen in zwei Registern mit Bitbreite N.
2. Steuersignal "Addition" aktiviert das Paralleladdierwerk.
3. Mit der nächsten Taktflanke wird das Ergebnis in das A-Register übernommen.



Ripple-Carry-Addition: Die Addition erfolgt bitweise von LSB (Least Significant Bit) zu MSB (Most Significant Bit). Der Übertrag wird von einem Bit zum nächsten weitergereicht.

--> Hierbei wird pro Dualzahl ein Volladdierer benötigt.

ACHTUNG Im ungünstigsten Fall muss ein Übertrag durch alle Bits gereicht werden, was zu einer langen Verzögerung führt.

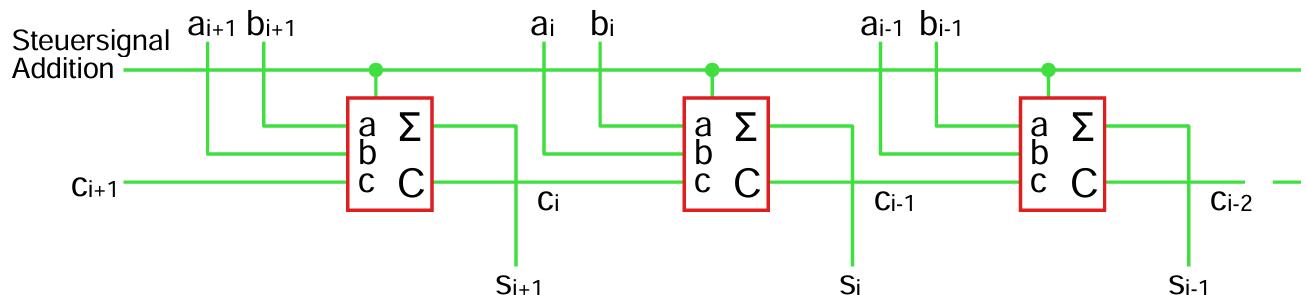


Abbildung 80 Paralleladdierer mit Volladdierern

Carry-Lookahead-Addition: Die Überträge werden parallel berechnet, was zu einer schnelleren Addition führt.

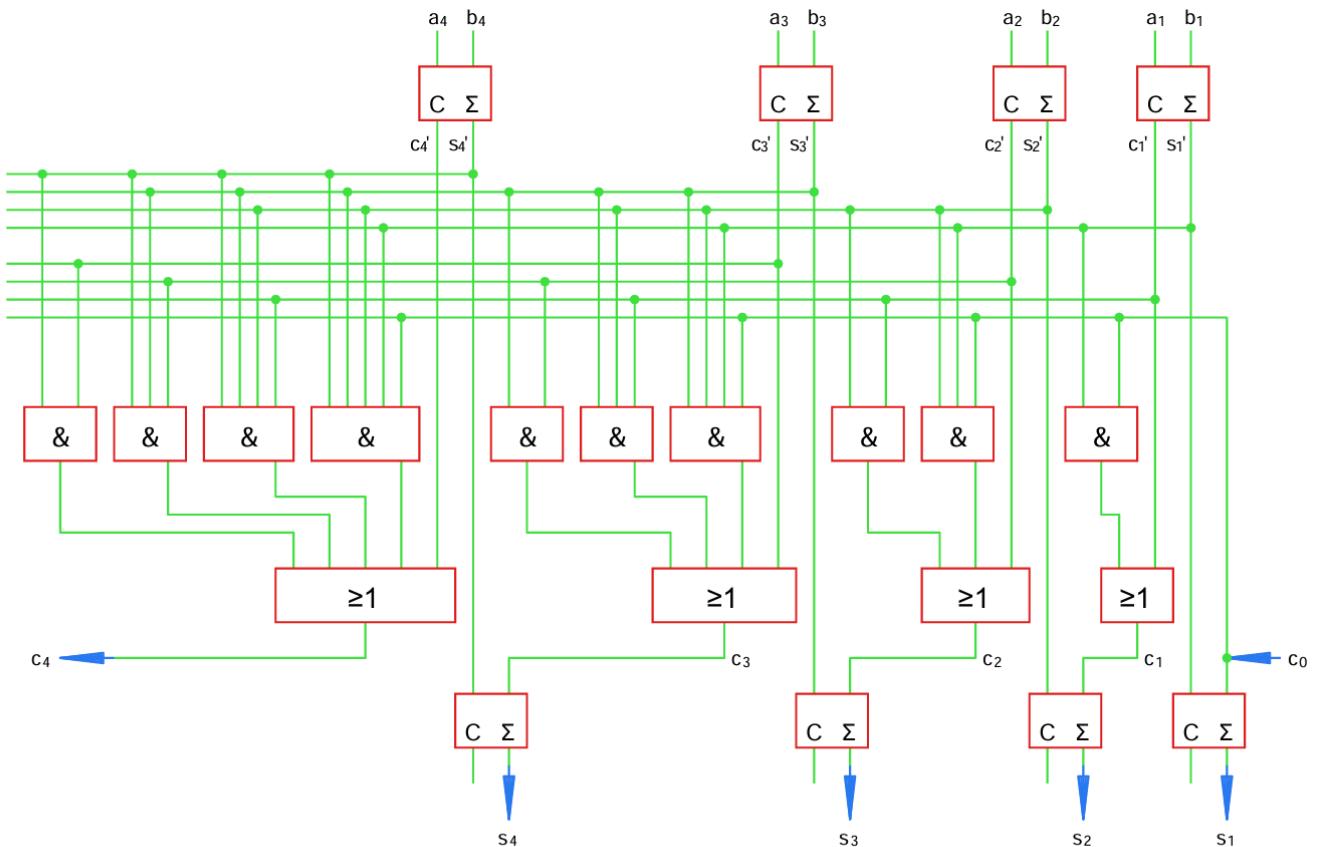


Abbildung 82 4-Bit Carry Look Ahead Addierer

3.3.1.4 Inkrement

Beschreibung: Inkrementieren ist eine häufig benötigte arithmetische Funktion, z.B. bei Schleifenberechnungen.

Optimierung:

- Addition mit 1, wobei fast alle Stellen des einen Summanden null sind.
- Pro Binärstelle (außer LSB) kann ein Volladdierer durch einen Halbaddierer ersetzt werden.

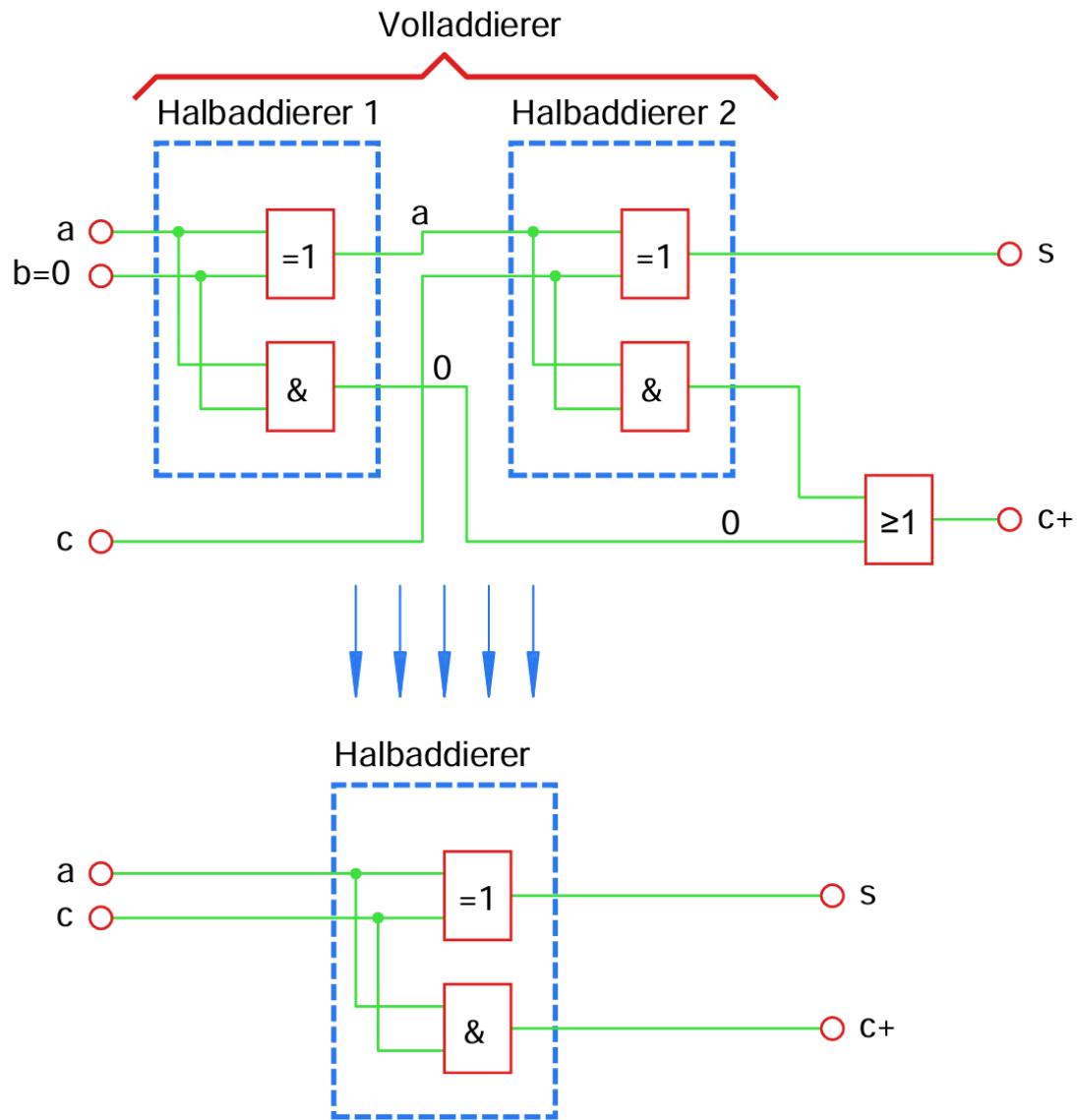


Abbildung 83 Inkrementator

3.3.7 Faktor 256 hoch x

Binäre Schiebeoperation Die Multiplikation mit 256 in einer binären Zahl entspricht einer Schiebeoperation (Shift-Operation) um 8 Bit nach links.

$$\rightarrow 256 = 2^8$$

- Optimierung durch Compiler: Der Compiler erkennt die Shift-Operation um 1-Byte und führt keine Multiplikation durch, sondern greift direkt auf die Speicheradresse um 1-Byte weiter zu. \rightarrow Dadurch wird die Zahl ohne Rechenoperation der ALU um Faktor 256 verkleinert. Alle Faktoren um 256 hoch x können ersetzt werden durch eine Verschiebung um x Byte.

3.3.9 SISD/SIMD (*Single Instruction Single Data / Single Instruction Multiple Data*)

Traditionelle SISD-Architektur:

- Eine Instruktion wird auf ein ganzes Register angewendet. **Problem:** heutige Register sind 64-Bit breit, was zu einer Verschwendungen von Rechenleistung führt, wenn nur 8-Bit Daten verarbeitet werden.
- **Lösung:** SIMD-Architektur:
- Aufteilung des Registers in mehrere Teile, die gleichzeitig verarbeitet werden können. (*bis zu 8x 8-Bit Daten*)

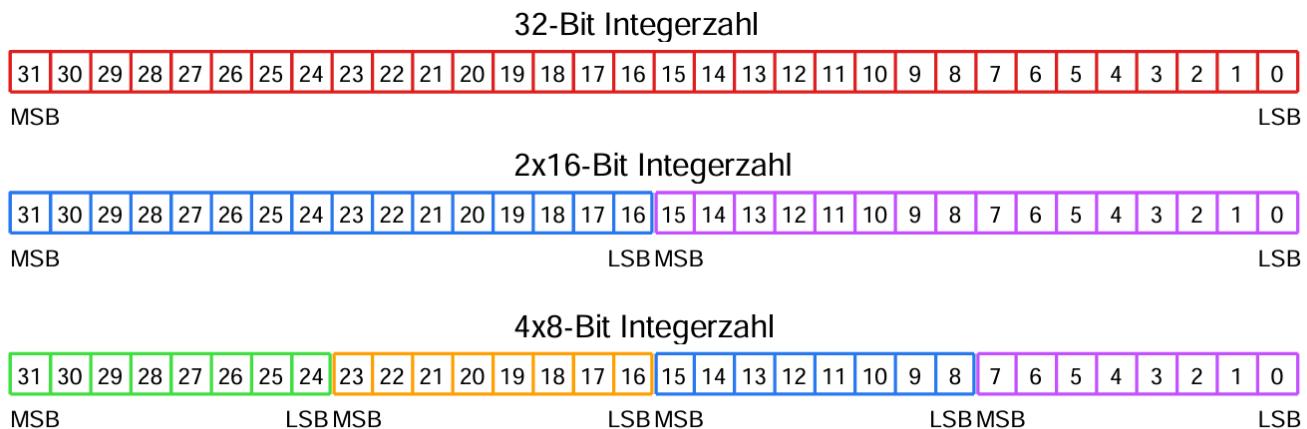
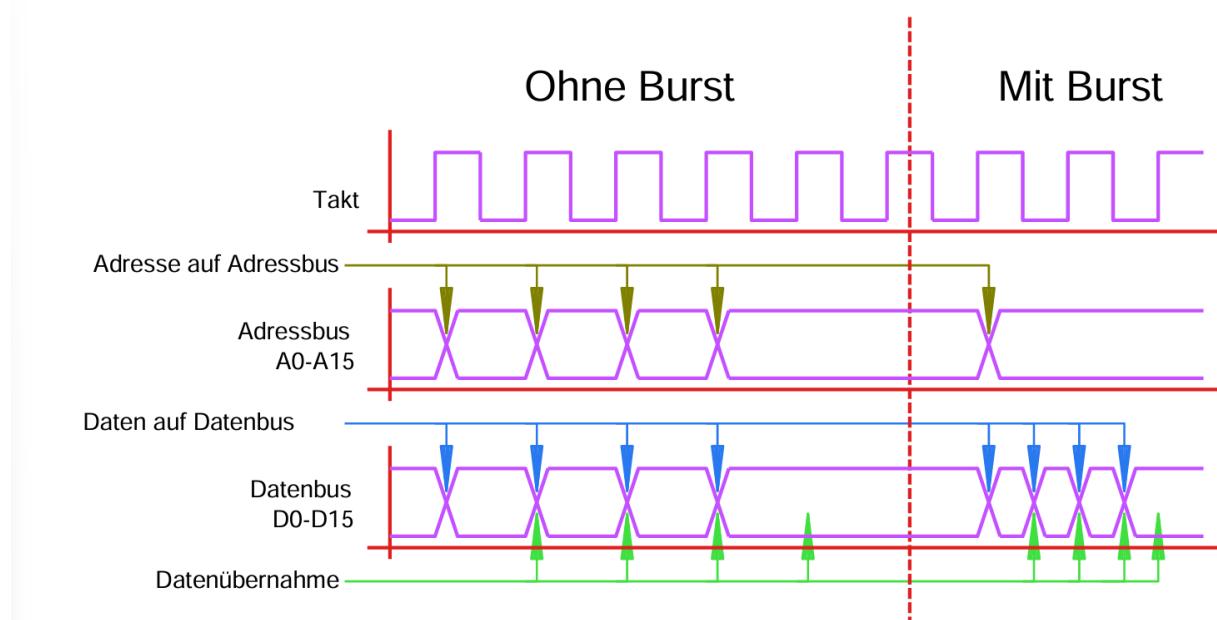


Abbildung 97 SIMD Beispiele bei 32-Bit-Registern

4.4 Burst-Mode

Motivation: Durch aufwendige statistische Erforschung von Speicherzugriffen wurde in den 1990er Jahren bemerkt, dass eine Software oft auch die Daten der Nachfolgeadressen des gerade angeforderten Datums benötigt.

Burst-Mode: Ein Burst ist eine Übertragung von gleich mehreren Datenwörtern direkt hintereinander, ohne jedes Mal die Adresse mit zu übertragen.

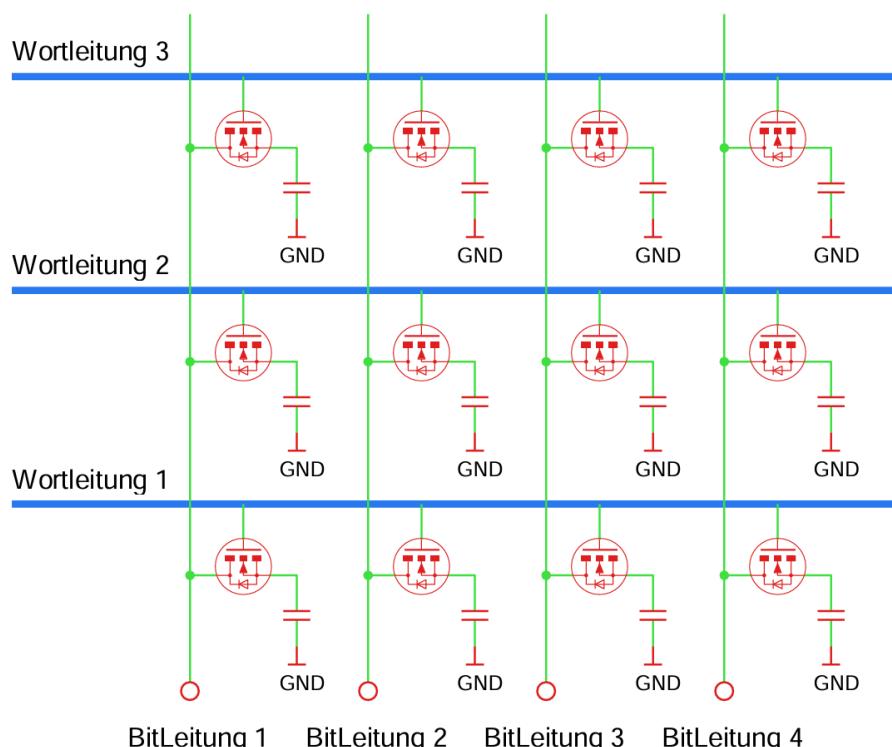


- Die Datenbusbreiten der ersten CPUs mit Burstübertragung lagen bei 32 Bit (i486). Die Größe einer internen Cacheline lag damals bei 16 Bytes. Die ersten Burstlängen lagen deshalb bei 4 32-Bit-Zugriffen direkt hintereinander. Aktuelle Prozessoren mit 64-Bit-Bus haben interne Cachelines von 64 Bytes und übertragen Bursts jetzt mit 8 Zugriffen direkt hintereinander.

- **RAM: Großer** Vorteil beim Zugriff auf DRAM-Speicher, da die verschachtelte RAS-CAS-Addressierung eine viel kürzere Zugriffszeit nur auf direkt aufeinanderfolgende Adressen ermöglicht.
- **PCI-Bus:** Kann auch im Burst-Mode betrieben werden, was hier ein starker Vorteil ist, da Daten- und Adressbus gemultiplexed sind. Die Burstlänge ist nicht begrenzt. (*Man kann mehrere Megabyte direkt hintereinander übertragen*)

5.2.2.1 Standard-DRAM

- **Grundprinzip**
 - DRAM (Dynamic Random Access Memory) basiert auf einem kleinen Kondensator, der den Wert eines Bits in Form einer elektrischen Ladung speichert.
 - Im Vergleich zu SRAM benötigt DRAM nur einen Transistor, um ein Bit zu speichern.
- **Abbildung: DRAM-MOSFET-Zelle**
 - Die Größe des Kondensators einer DRAM-Zelle liegt heute im Attofarad-Bereich.
 - Der Kondensator entlädt sich von selbst durch Leckströme innerhalb von Millisekunden.
- **Refresh-Zyklus**
 - Um den Speicherinhalt zu erhalten, muss bei jedem Bit, bevor die Ladung verloren geht, der Inhalt ausgelesen, zwischengespeichert und zurückgeschrieben werden.
 - Dieser Vorgang wird Refresh genannt und zyklisch für den gesamten Speicher durchgeführt.
- **Organisation des Speicherarrays**
 - DRAM-Speicher werden in einem zweidimensionalen Speicherfeld organisiert, wobei die Wortleitung und die Bitleitungen vom Adressdecoder in zwei Teile geteilt werden.

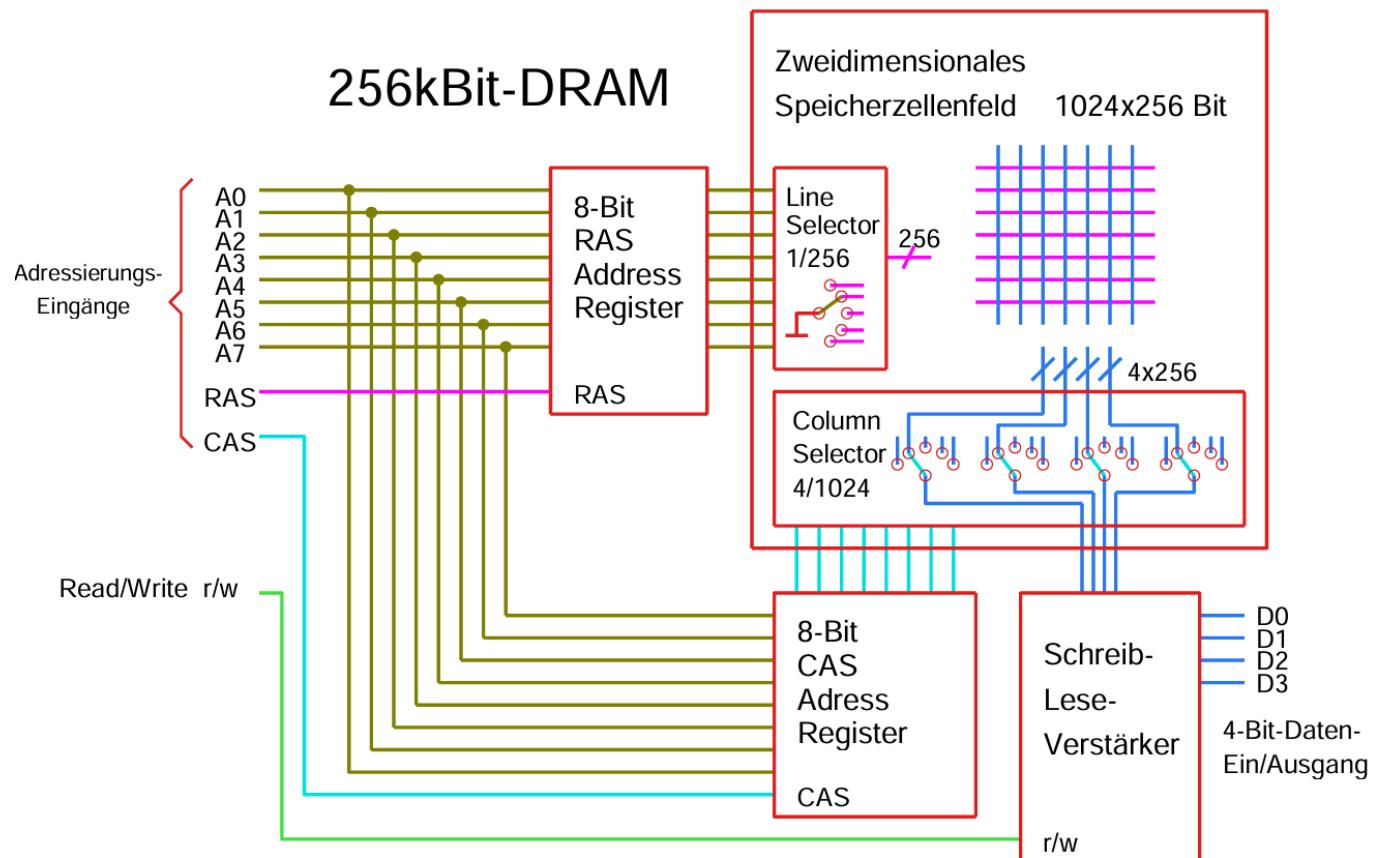


DRAM-Speicher werden gewöhnlich in einem zweidimensionalen Speicherfeld organisiert. Dabei werden die Wortleitung und die Bitleitungen vom Adressdecoder in zwei Teile geteilt.

Abbildung 126 DRAM-Array

- **Adressierung**
 - Die Adressierung des Speicherarrays erfolgt in zwei Schritten: RAS (Row-Address-Select, Zeilenadresse) und CAS (Column Address-Select, Spaltenadresse).

- Die wahlfreie Zugriffszeit beträgt ca. 40-60 ns, was einer maximalen Zugriffstaktfrequenz von 16-25 MHz entspricht.
- RAS-CAS-Adressierung
 - Bei einem Adressierungsvorgang, bei dem nur eine der beiden Adressteile verändert wird, erfolgt die Adressierung etwa doppelt so schnell (20-30 ns).



• Parallelschalten der Adressierungseingänge

- Je nach Anzahl der nach außen geführten Datenleitungen werden nur ein Teil des selektierten Datenworts vom CAS-Adressdecoder ausgewählt.
- Früher wurden oft nur ein einzelnes Bit selektiert, während moderne DRAM-Bausteine normalerweise 4-Bit-Daten herausführen.

• Buffered/Registered RAM

- Beim Parallelschalten der Adressierungseingänge kann es zu einer Überlastung des Adressbusses kommen. Um dies zu verhindern, werden zusätzliche Buffer vor die Adressleitungen der Bausteine geschaltet.
- Wenn ein Zwischenspeicher (Latch) für die Adresse vorhanden ist, spricht man vom Registered-RAM.

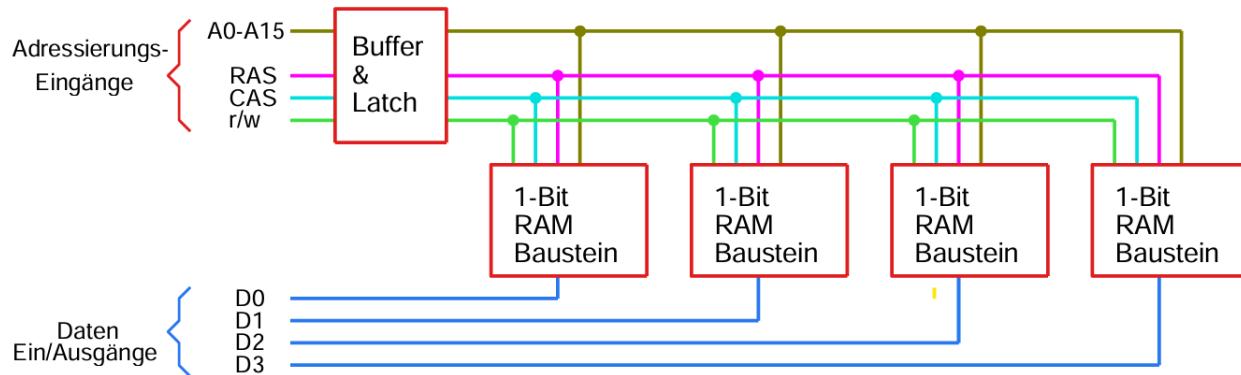


Abbildung 128 Buffered/Registered RAM

5.2.2.3 SDRAM (*Synchronous Dynamic Random Access Memory*)

- **Einführung eines gemeinsamen Taktsignals**

- Vor dem Jahr 2000: Speicherzugriffe erfolgten asynchron.
- Ab SDRAM: Alle Steuersignale beziehen sich auf ein gemeinsames Taktsignal. Dies synchronisiert den gesamten RAM-Baustein.

- **Vorteile durch Burst-Modus**

- Moderne Prozessoren nutzen den Burst-Modus für Speicherzugriffe, bei dem mindestens 16 Bytes direkt hintereinander gelesen oder geschrieben werden.

- **Optimierungen innerhalb von SDRAM:** --> führt zu ca. einer Verdoppelung der Datenrate

- Pipelining: Mehrere Befehle werden gleichzeitig verarbeitet, indem sie in verschiedene Phasen zerlegt werden.
- Vervielfachung der Speicherbänke: Erhöhung der Anzahl der unabhängigen Speicherbänke, die gleichzeitig arbeiten können.
- Mehrere Buffer: Zwischenspeicher innerhalb des RAM-Bausteins verbessern die Zugriffsgeschwindigkeit.

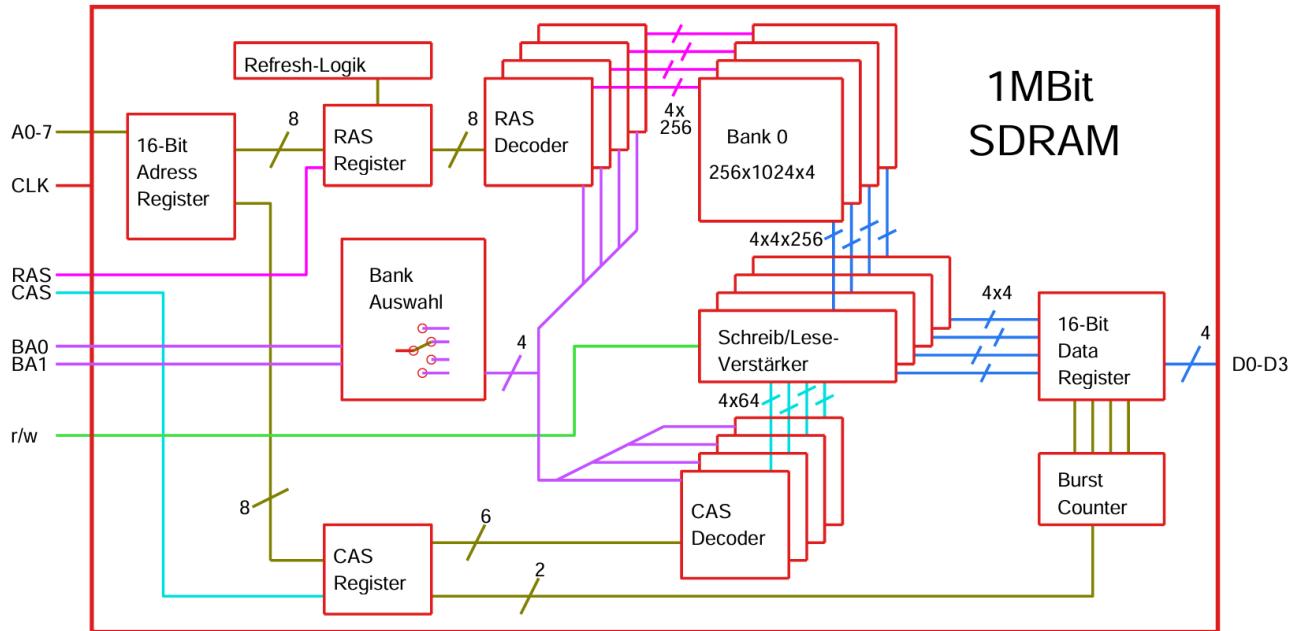


Abbildung 131 SDRAM

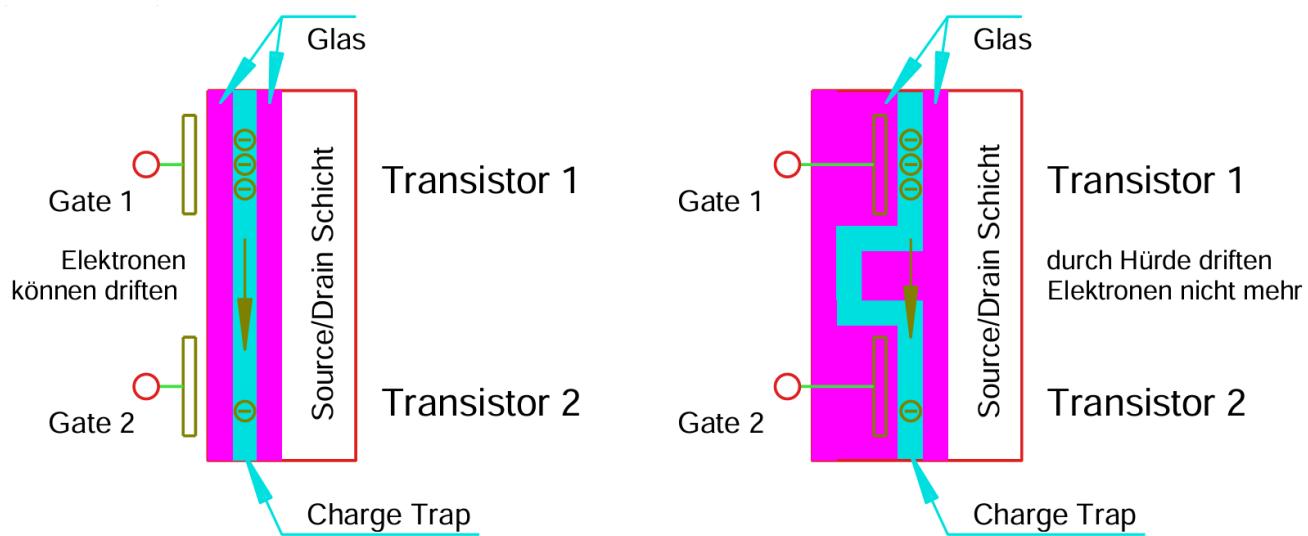
- **Struktur von SDRAM**
- *Bankauswahl und Zugriffsoptimierung*
 - Bank-Select Leitungen: Zwei zusätzliche Leitungen (BA0, BA1) ermöglichen die direkte Auswahl von vier separaten Speicherbereichen.
 - Diese Struktur ermöglicht schnelle Zugriffe, ohne die RAS- oder CAS-Adresse ändern zu müssen.
- *CAS-Latency (CL)*
 - **Definition:** Die Zeitverzögerung in Taktzyklen zwischen dem Anfordern und Bereitstellen von Daten.
 - *Beispiel:*
 - Übertragungsrate: 1333 MByte/s bei 64 Bit Busbreite.
 - Taktfrequenz: $1333/8 = 166 \text{ MHz}$ (8 Byte werden gleichzeitig übertragen).
 - CL = 10: Der Speichercontroller muss 10 Takte warten, bis das RAM-Modul bereit ist.
 - Max. wahlfreie Zugriffsrate: ca. 16,6 MHz.

5.3.4.6 3D-NAND-Flash

- **Problem:**
 - Die Menge der unterscheidbaren Elektronen kann nicht weiter verringert werden, ohne mehr Fehler zu verursachen. Dies bedeutet, dass die Integrationsdichte nicht weiter gesteigert werden kann.
--> Die Speicherdichte von Flash-Speichern stößt an physikalische Grenzen.
- **Lösung: (Eintreten in die 3.te Dimension)**
 - Die Speicherzellen werden nicht mehr nur in einer Ebene angeordnet, sondern auch in die Höhe gestapelt.
 - Herstellung des Floating-Gate-Transistors aus Siliziumnitrid, das als Isolator dient. (*wird auch als Charge Trap bezeichnet*)

- **Anmerkung:**

- Der Übergang vom Floating-Gate zum Charge Trap wurde schon bei 2D-NAND-Flash durchgeführt.
--> vereinfacht die Litographie bei der Fertigung stark.



6.3.2.4 RS485 - Schnittstelle

- **Grundsätzliches**

- RS485 = Weiterentwicklung von RS422.
- Halbduplex-Betrieb: Daten können in beide Richtungen übertragen werden, aber nicht gleichzeitig.
- Bis zu 32 Teilnehmer an einem Bus. (*Mit speziellen Treibern bis zu 256 Teilnehmer*)

- **Halbduplex-Betrieb**

- Sendeverstärker jedes Gerätes muss abschaltbar sein, um Kollisionen zu vermeiden.
--> Abschaltung erfolgt mittels Kommunikationsprotokolls.

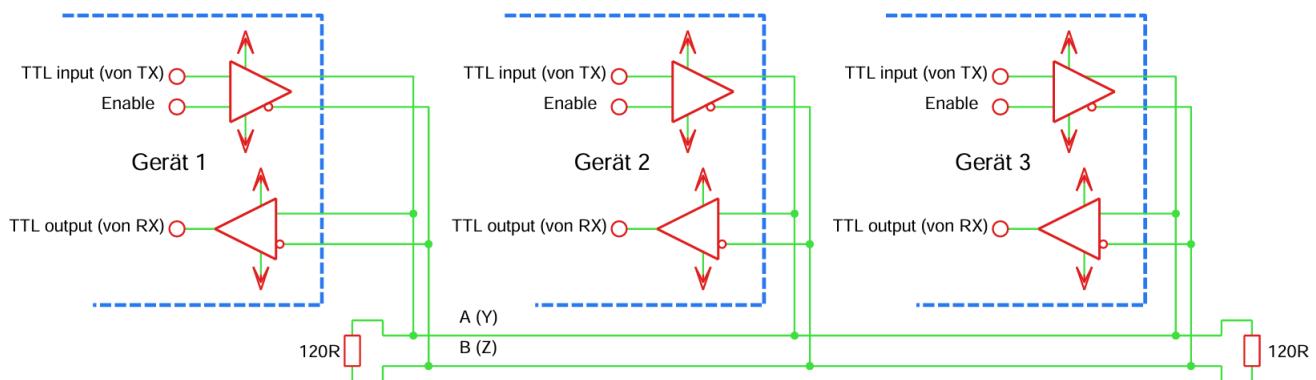


Abbildung 172 RS485 Verschaltung

- **Bezugsleitung**

- Theoretisch kein Stromfluss über GND-Bezugsleitung aufgrund gegenphasiger Signale

- Bezugsleitung könnte weggelassen werden (*Bezugspotenzial kann über Erdpotenzial realisiert werden*)
- Maximaler Gleichtaktfehler: -7 bis +12 V
- Bei großen Leitungslängen wird eine Bezugsleitung zur Vermeidung von Störungen empfohlen.
- **Gleichtaktfehler**
 - Differenzbildung der Signale verhindert Störungsanzeigen beim Empfänger

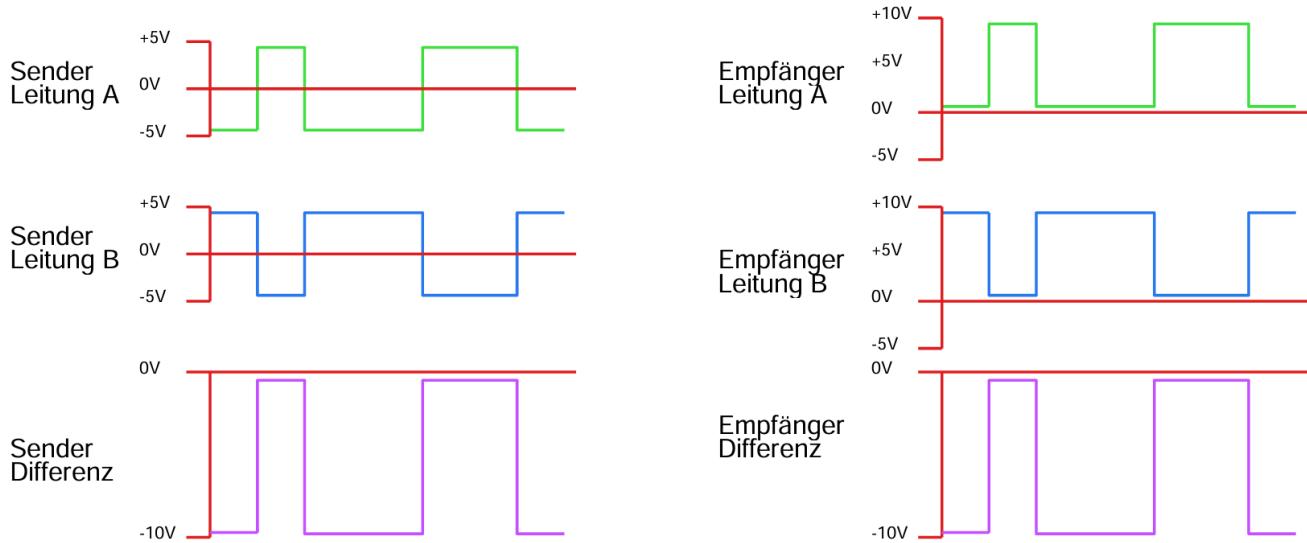


Abbildung 173 RS485 Gleichtaktfehler

- **Busabschluss**
 - Mögliche Situation: kein aktiver Sender am Bus
 - Einfacher Busabschluss (120 Ω-Widerstand): Leitung fällt auf undefinierten 0 V Pegel, wenn kein Sender aktiv ist
 - **Lösung:** Failsafe-Busabschluss für definierten Pegel ohne aktive Sender



Abbildung 174 RS485 Beispiele Busabschluss

- **Neuere Versionen**
 - Absolutpegel spielt fast keine Rolle für Störabstand, nur Signaldifferenz relevant
 - Moderne RS485-Bausteine verzichten auf negative Spannungskomponenten
--> *Signale bewegen sich zwischen 0 V und +5 V*
 - Vereinfachtes elektrisches Design, keine negativen Hilfsspannungen nötig
 - Nur geringe Verschlechterung des Störabstands

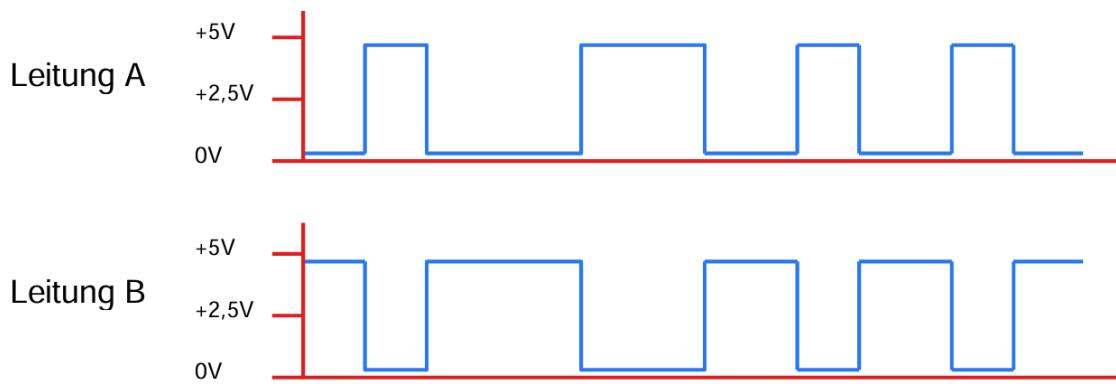


Abbildung 175 RS485 positive Pegel

6.3.8.4 USB 3 (*Universal Serial Bus*)

- Einführung: 2008
- Änderungen zu USB 2:
 - Neue Steckertyp mit zusätzlichen Daten-Leitungen
 - Rückwärtskompatibel und Vorwärtskompatibel
 - 5 zusätzliche Leitungen als 2 differenzielle Full-Duplex-Datenleitungen und eine Masseleitung
 - Datenübertragungsraten von bis zu 5 GBit/s (*auch gennant Super-Speed-USB*)
 - TR/VIA: Blaue Farbe für den Steckers
 - Stromversorgung: Erhöhung von 0,5 A auf 0,9 A --> **Maximale Leistung jetzt 4,5 W**
 - Maximale Kabellänge: 3 m

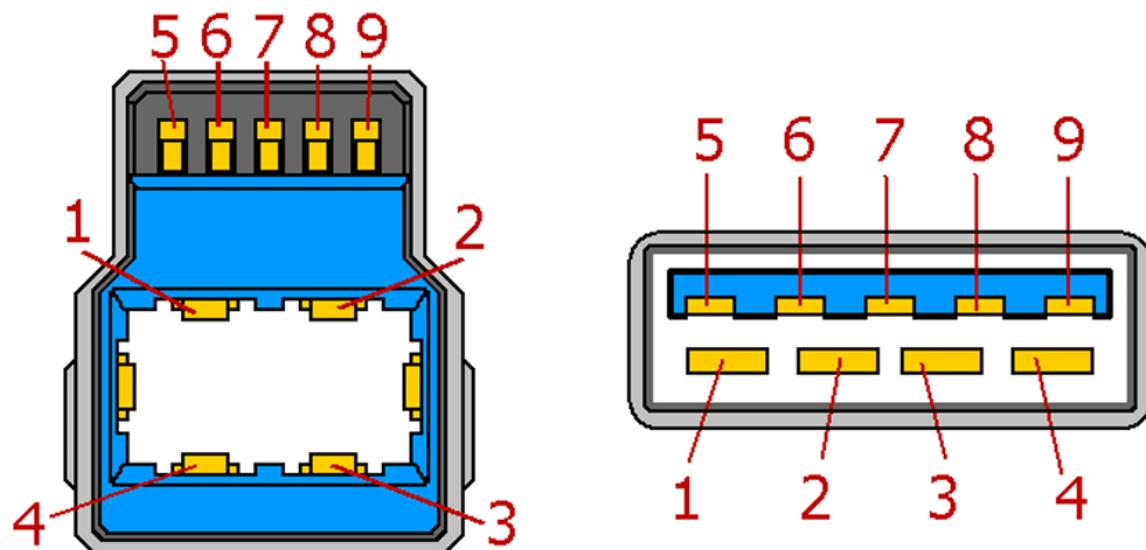


Abbildung 193 USB 3 Stecker

Pin Nr.	Name	Beschreibung
1	VBUS	+5V
2	D+	differentielles Paar 1+
3	D-	differentielles Paar 1-
4	Masse	0V
5	SSTX- / SSRX-	differentielles Paar 2- TxRx
6	SSTX+ / SSRX+	differentielles Paar 2+ TxRx
7	Masse	Masse für differentielles Paar 2
8	SSRX- / SSTX-	differentielles Paar 2- TxRx
9	SSRX+ / SSTX+	differentielles Paar 2+ TxRx

Abbildung 194 Belegung USB3-Stecker

- **USB 3.1**

- Neuer symmetrischer Steckertyp-C
- Host und Client mit demselben Steckertyp
- Kabel mit 24 Leitungen
- Effizientere Codierung vorher 8b10b-Code jetzt 128b132b-Code
- Höhere Übertragungsrate von bis zu 10 GBit/s
- Verschiedene Spannungen möglich
- Mehr elektrische Leistung möglich
- **Verringerung** der maximalen Kabellänge: 1 m

- **USB 3.2**

- Definition von USB 3.2 im Jahr 2017
- Erhöhung der Übertragungsrate auf 20 GBit/s
- Verwendung von 2 Leitungen für 2x 10 GBit/s
- Kompatibilität mit USB 3.0 und USB 3.1

6.3.9.1 Drahtgebundene Netzwerke

- **Allgemein**

- Serielle Verbindungen: Physikalisch bedingt; parallele Verbindungen haben Laufzeit-Probleme.
- Historie: Verschiedene Netzwerke durch Ethernet ersetzt; industrielle Netzwerke wie RS485 bestehen noch.
- Galvanische Trennung: Notwendig bei langen Leitungen; Nutzung von Übertragern für gleichanteilsfreie Bitmuster.

- **Koaxial**

- Verwendung: In den 1990er-Jahren; alle Teilnehmer eines Netzwerks über ein Koaxial-Kabel verbunden.
- **Notwendige** Abschlusswiderstände: 50Ω an beiden Enden des Netzwerks.
- **Probleme:** Fehleranfällig, Kommunikation bei Problemen unterbrochen. --> Single Point of Failure (einzelnes Kabelbruch führt zum Ausfall des gesamten Netzwerks)

- Übertragungsrate: 10 MBit/s, Halbduplex, maximale Länge 185 m.
- Signalpegel: 0 und -2.2 V mit Manchester-Kodierung; Kollisionserkennung durch tiefere Pegel.
- Frequenzbandbreite: Ca. 10 MHz

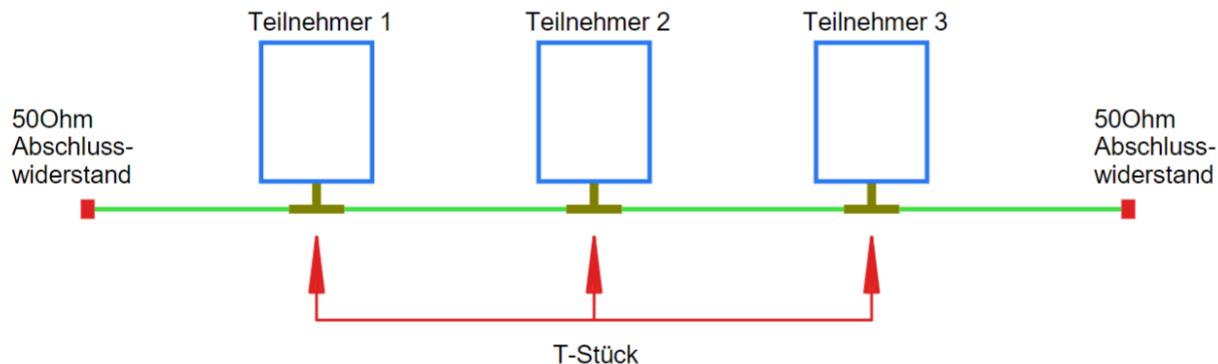


Abbildung 197 Koax Ethernet

- **Twisted Pair 2x2**

- Einführung: CAT3 Kabel mit RJ45 Steckern, sternförmige Verbindungen.
- Übertragung: ± 2.5 V, symmetrisch, Fullduplex, 2 Adernpaare.
- Übertrager: An Kabelenden, oft in Netzwerkbuchsen integriert.
- Umstieg auf 100 MBit/s: Durch CAT5 Kabel, verbesserte Codierung, 3 Spannungspegel (PAM-3) (+1V/ 0V/ -1V), Frequenzbandbreite 31,25 MHz. (*Immer noch 2 Adernpaare*)

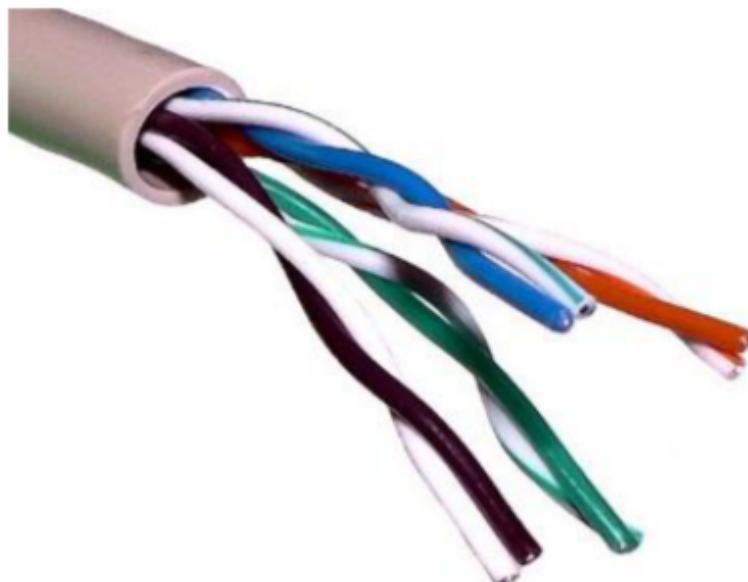


Abbildung 199 Twisted Pair Kabel



Abbildung 198 Ethernet Buchse mit Übertrager

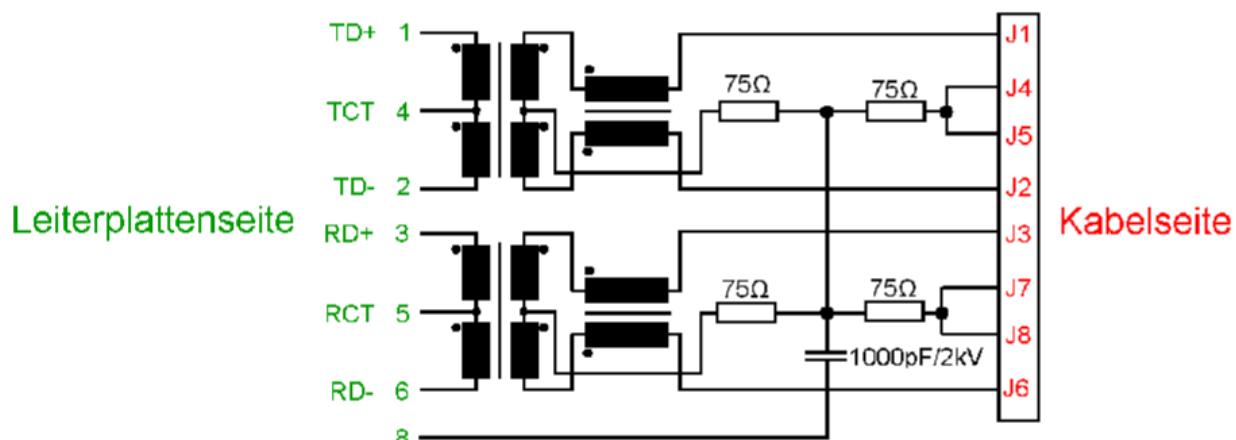


Abbildung 200 RJ45 Übertrager

- **Twisted Pair 4x2**
 - 1 GBit/s Übertragung: Nutzung aller 4 Adernpaare, Fullduplex auf allen Aderpaaren.
 - Echokompensation: An beiden Enden des Kabels.
 - Signalpegel: 5 Pegel (PAM-5), Frequenzbandbreite 62,5 MHz. (+1 V/+0,5 V/0 V/-0,5 V/-1 V)
 - Höhere Geschwindigkeiten: 2,5/5/10 Gbit/s mit 16 Pegeln (PAM-16).

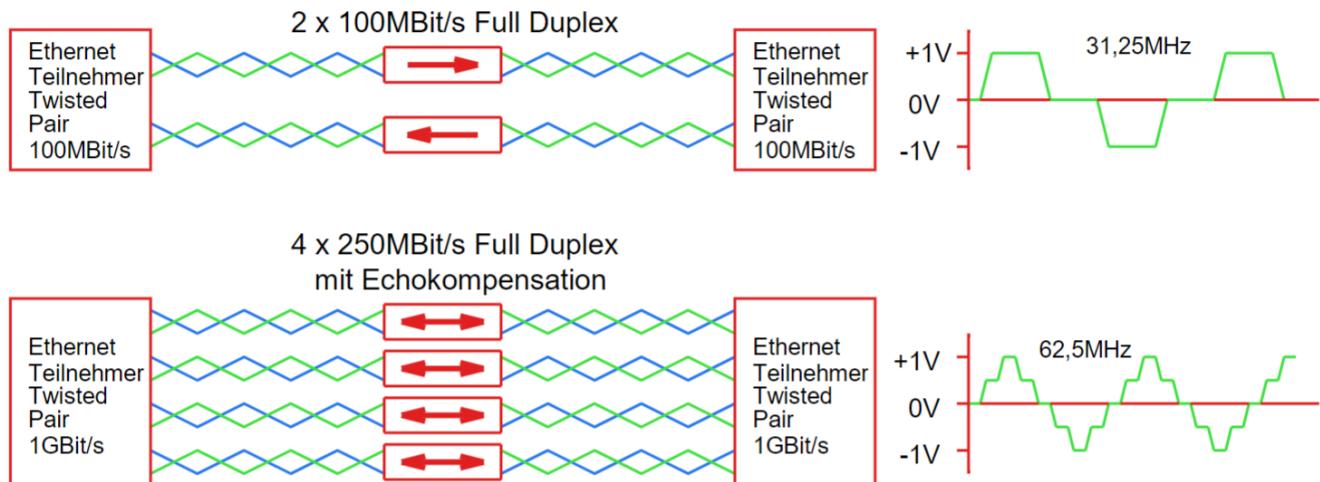


Abbildung 202 Vergleich Ethernet 100 Mbit/s und 1 Gbit/s



Abbildung 203 Ethernet Industriestecker



Abbildung 204 Vergleich RJ45 <-> ix

- **ix-Industrial**

- Neue Steckverbinder: Einführung 2018, kompakter als RJ45, für industrielle Anwendungen.
- PoE (Power over Ethernet): Versorgung von Clients mit Energie über Datenleitungen



Abbildung 203 Ethernet Industriestecker



Abbildung 204 Vergleich RJ45 <-> ix

- **Twisted Pair 1x2**

- Single Pair Ethernet (SPE): Verwendung eines Adernpaars, CAT7 Kabel, bis zu 1 Gbit/s, maximale Länge 15-40 m.
- PoDL (Power over Data Line): Gleichzeitige Energieversorgung, Spannungen 12-48 V, Leistung 0,5-50 W.

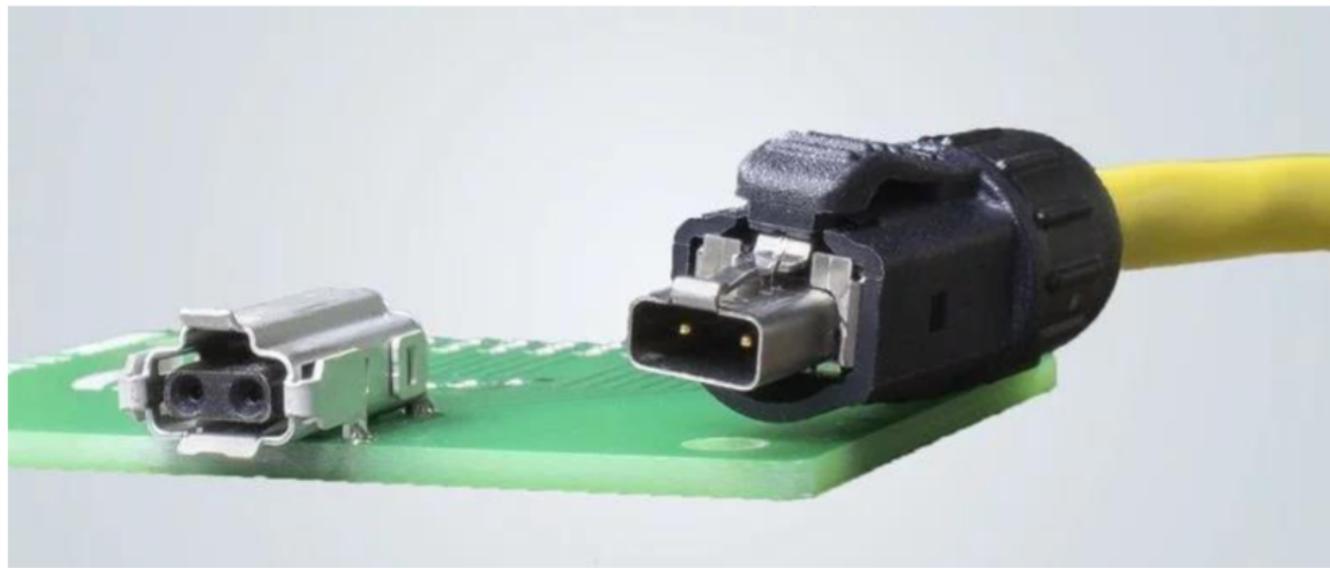
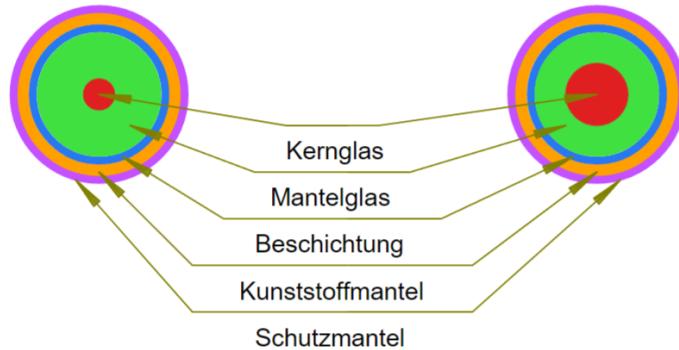


Abbildung 205 Single Pair Ethernet

- **Glasfaser**

- Parallel zu Kupferkabeln: Standards für optische Übertragung ab 1992.
- Glasfasertypen: Monomode, Multimode, Gradientenindex.
- Anwendungen: Multimode für kurze Distanzen, Monomode für längere Distanzen.
- Lichtwellenlängen: 850 nm-1550 nm, meist Infrarot.
- Vorteile: Nahezu perfekte Signalübertragung bei Monomode, komplexere Technik.
- Medienkonverter: Einfacher Wechsel von 1 GBit/s-Kupferleitungen auf Glasfaserleitungen.

Monomode-Glasfaser



Multimode-Glasfaser

Der Kern einer Monomodefaser ist nur einige μm dick. Die Verbindungstechnik für Monomodeglasfasern ist entsprechend mechanisch sehr aufwendig und teuer. Im Netzwerkbereich werden deshalb fast ausschließlich Multimodefasern eingesetzt.

Abbildung 207 Glasfaser Typen

- **TOSLINK**

- Consumer-Audiotechnik: Optische S/PDIF Verbindung, rotes Licht (650 nm).
- Übertragungsrate: Deutlich unter 10 MBit/s.
- Vorteile: Galvanische Trennung, geringe elektromagnetische Störbeeinflussung.
- **Keine Glasfaserübertragung**

7.2.1 Leuchtdioden

7.2.1.1 Einzelne Leuchtdioden

- Ansteuerung: Ein einfacher Port-Pin eines Mikrocontrollers ist ausreichend.

- High-Aktiv: Leuchtdiode leuchtet bei High-Signal.
- Low-Aktiv: Leuchtdiode leuchtet bei Low-Signal; meist besser geeignet, da Digitalausgänge bei Low-Pegel mehr Strom aufnehmen können.
- Durchlassspannung: Unterschiedlich je nach Farbe der LED.
 - Blaue LEDs: Oft über 3 V Durchlassspannung, problematisch bei Mikrocontrollern mit 3 V Versorgung.
 - Weiße LEDs: Enthalten blaue LED-Chips.
 - True-Green LEDs: Verhalten sich wie blaue LEDs.

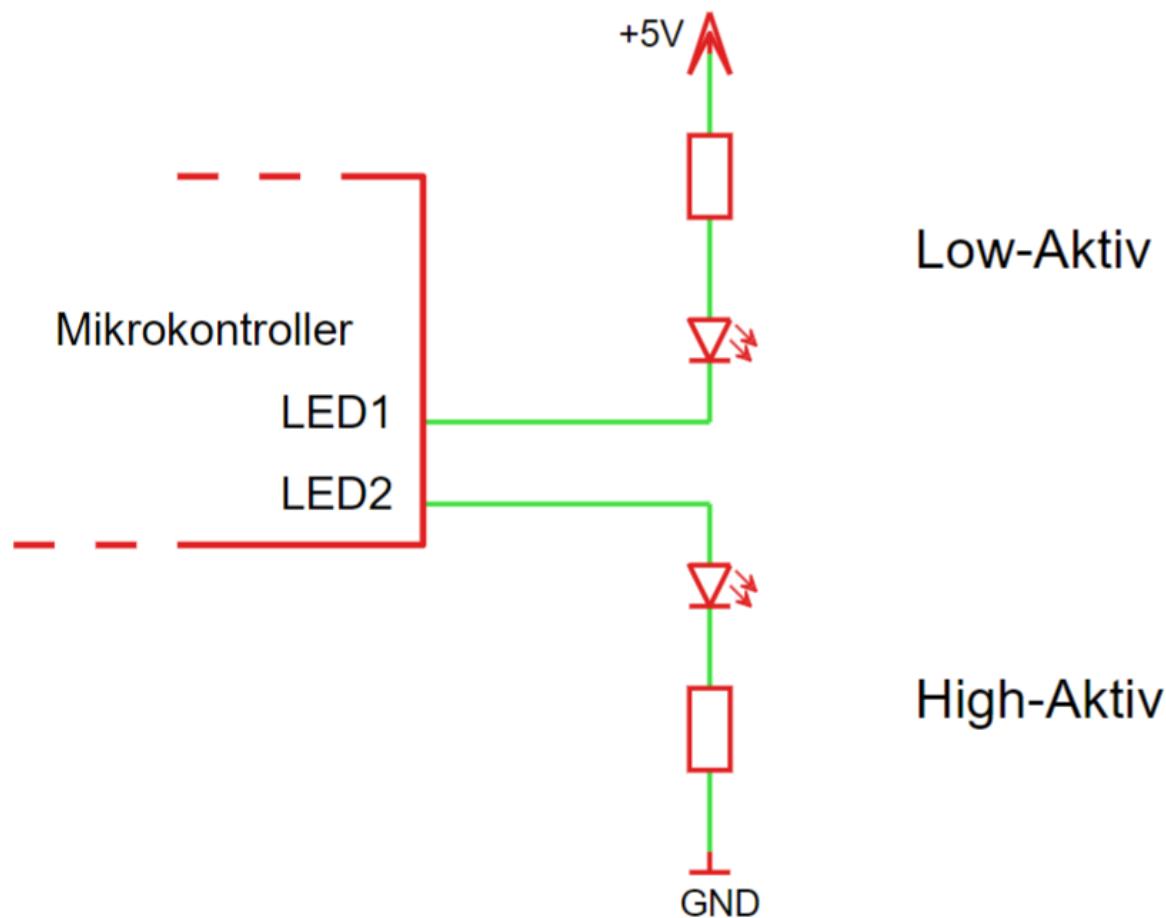


Abbildung 214 LED-Ansteuerung

7.2.1.2 Gruppierte Leuchtdioden

- **Einfache Matrix:** LEDs in einer gemultiplexten Matrix ansteuern.
 - Multiplexing: LEDs leuchten nur kurzzeitig --> Strom muss ausreichend hoch sein, um eine ausreichende Helligkeit zu erzeugen.
 - Kurzzeitige Ströme: Manche Matrix-LEDs vertragen Ströme von einigen Ampere.

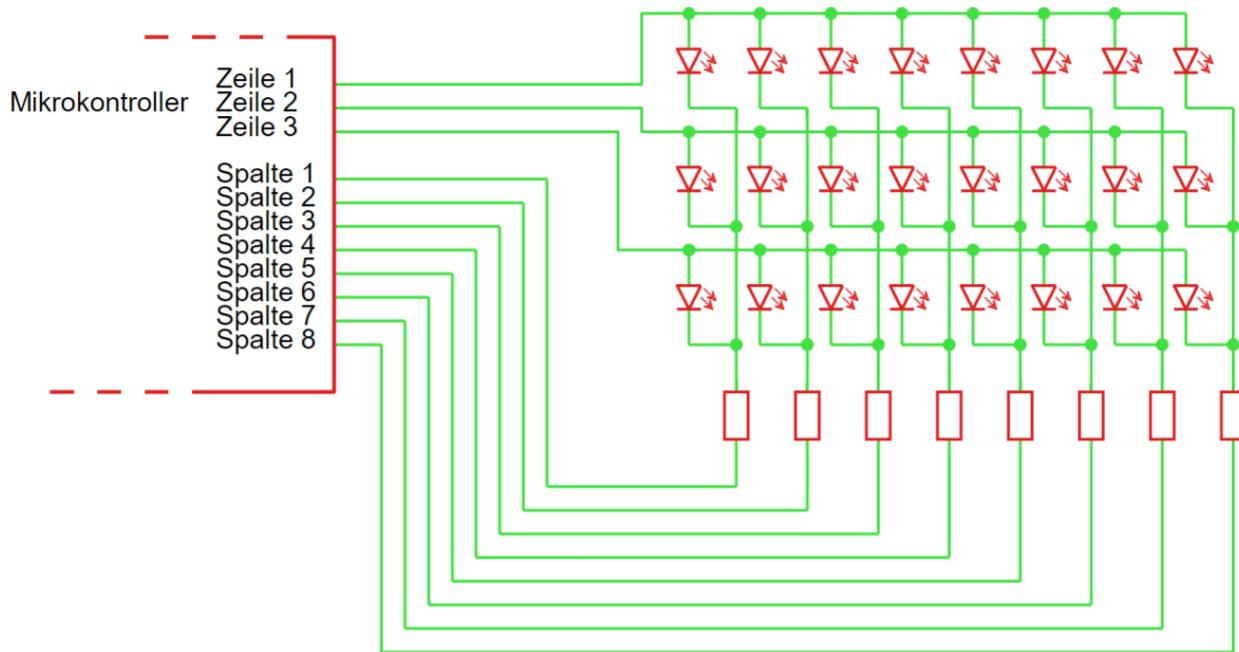


Abbildung 215 LED-Multiplexing

- **7-Segment-Anzeigen:** Häufige Anordnung von gruppierten LEDs; zwei Ausführungen (*gemeinsame Kathode/Anode*):
 - Gemeinsame Kathode: Kathoden aller Segmente sind verbunden.
 - Gemeinsame Anode: Anoden aller Segmente sind verbunden.
 - Ansteuerung: Transistoren müssen die Ausgangspins des Mikrocontrollers verstärken.
 - Versorgung: Mikrocontroller und LEDs müssen die gleiche Versorgungsspannung haben, wenn Variante mit gemeinsamer Anode verwendet wird.

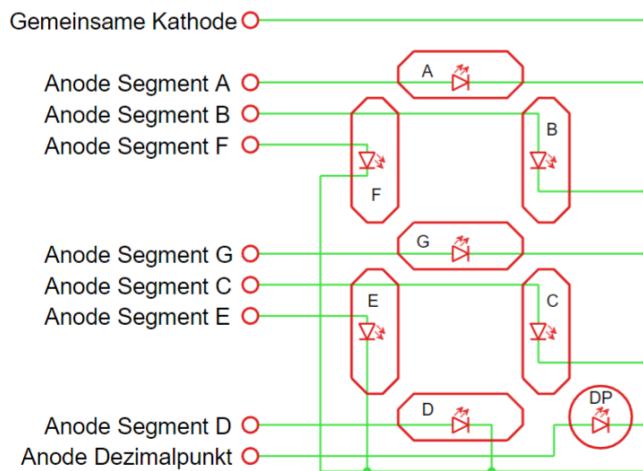


Abbildung 216 Gemeinsame Kathode

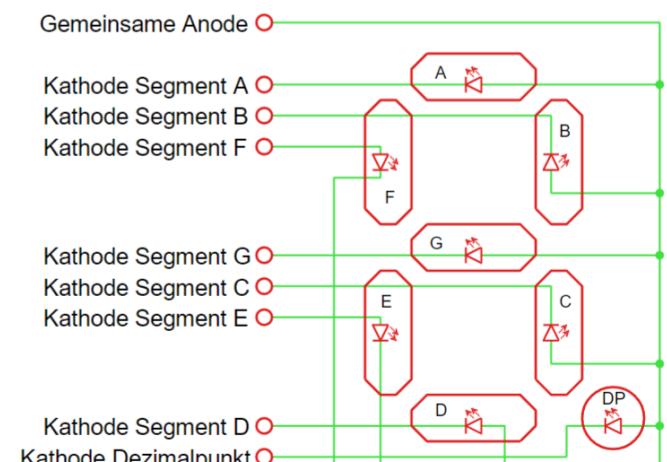


Abbildung 217 Gemeinsame Anode

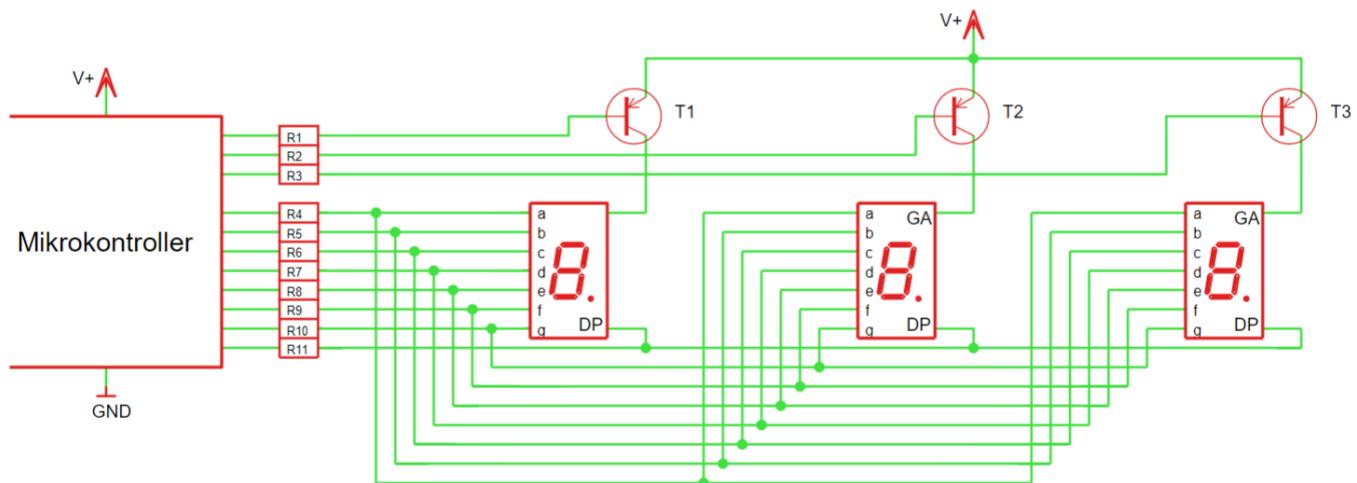


Abbildung 218 7-Segment mehrstellig

7.2.2 LC-Anzeigen

7.2.2.1 Grundprinzip

- **Flüssigkristall-Anzeigen (LCD):** Flüssigkristalle beeinflussen die Polarisationsrichtung von Licht.
 - TN-Zelle (Twisted Nematic): Grundprinzip der LCDs.
 - Polfilter: Licht wird durch Polfilter polarisiert. (*50% Helligkeitsverlust*)
 - Flüssigkristalle: Zwischen zwei Glasscheiben; drehen die Polarisationsebene des Lichts.
 - Spannungsanwendung: Zerstört die schraubenförmige Anordnung der Kristalle, blockiert Licht.
 - Hintergrundbeleuchtung: Früher Kaltkathodenleuchtstofflampen, jetzt weiße LEDs.
 - Farbige Darstellung: Farbfilter reduzieren Wirkungsgrad, neuere Technologien verbessern Farbdarstellung.
 - QLED: Nutzung von Quantenpunkten für bessere Farbdarstellung und Wirkungsgrad.

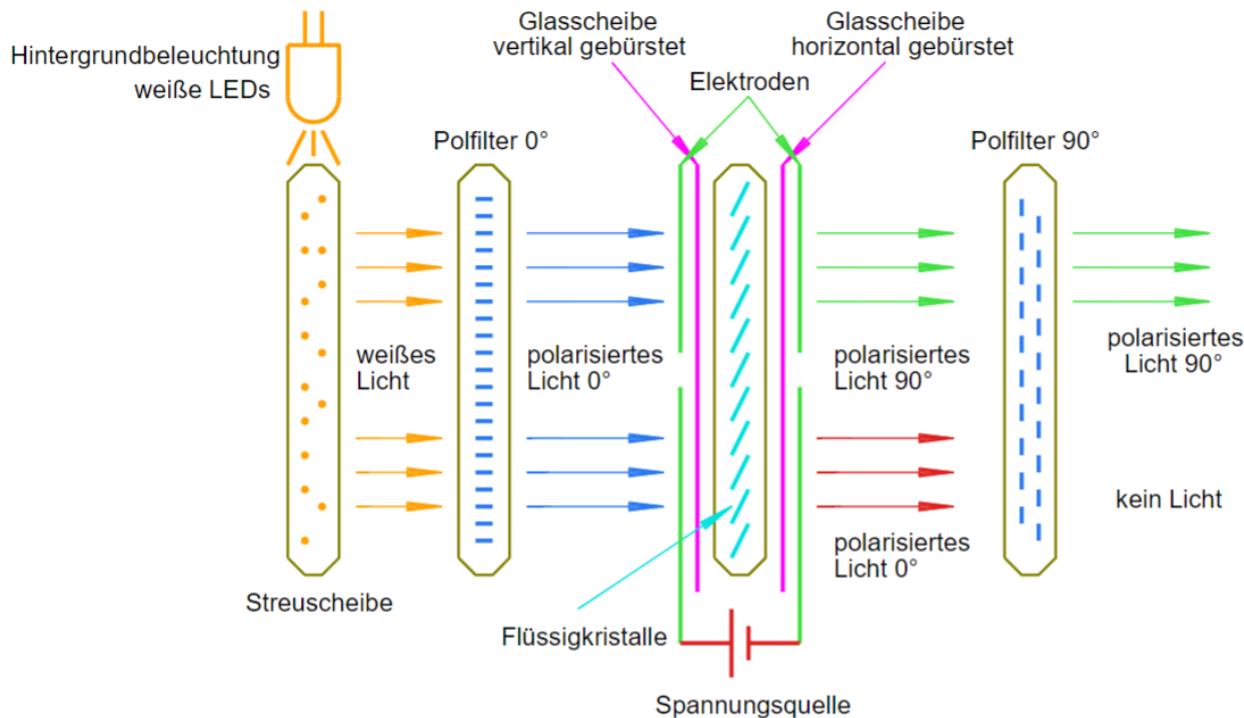


Abbildung 219 LCD-TN-Zelle

7.2.2.2 Passive LCDs

- Einfach: Keine farbige Darstellung, nur Ein/Aus Zustände.
 - Ansteuerung: Jeder Pin führt zu einem Segment, direkt gegenüber einem Backplanesignal.
 - Backplanesignale: Mehrere für große Segmentgruppen.
 - Gleichspannung: Nicht dauerhaft anlegen, da es zu irreversiblen elektrochemischen Prozessen führt.
 - AC-Signal: Pulsierende Gleichspannung, max. 1 kHz, erzeugt eine echte Wechselspannung für die Segmente.

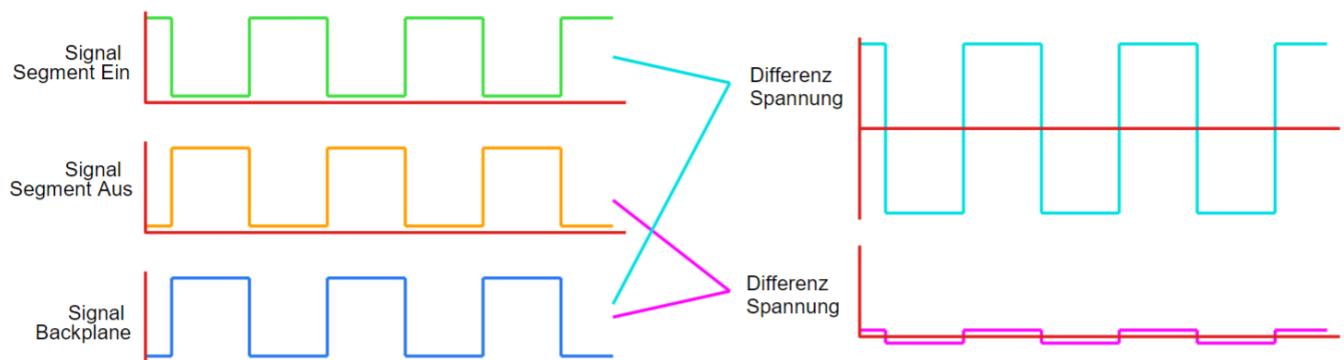


Abbildung 221 LCD-AC-Signal

7.2.2.3 LCDs mit integrierter Ansteuerung

- **Aktive LCDs:** Viele Typen; HD44780-Interface (*von Hitachi*) weit verbreitet.
 - Controller: 4-Bit-Modus benötigt 7 Leitungen.
 - Kontrast: Einstellbare negative Spannung für die Kontrastleitung, oft durch eine Ladungspumpe erzeugt.

- Grafische LCD-Module: Serielle SPI-Schnittstelle hat sich durchgesetzt.

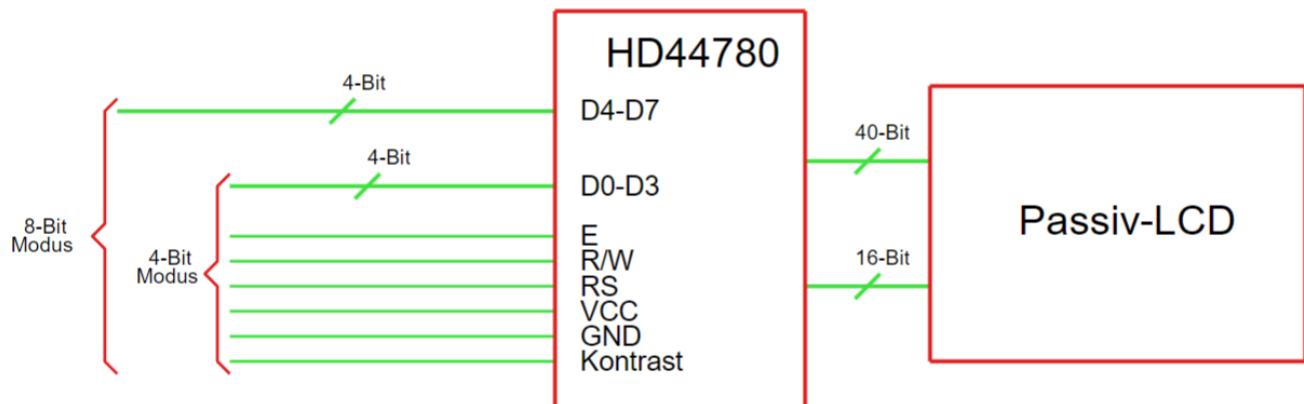


Abbildung 222 LCD-Controller

7.3.3 DVI (*Digital Visual Interface*)

- **Motivation für DVI (Digital Visual Interface):**

- Ende der 1990er Jahre kamen Flachbildschirme auf dem Markt --> Ansteuerung mit VGA-Signalen umständlich.
- Elektronik in Flachbildschirmen benötigte digitale Bilddaten.
- VGA-Signale mussten digitalisiert werden, was zu Qualitätsverlusten führte.

- **Eigenschaften von DVI:**

- Synchrone, symmetrische Übertragung mittels 6 differenziellen Signalen (TDMS-Links). (*Transition Differential Minimized Signaling*)
- Übertragung bis zu 10m.
- Synchraones Clock-Signal wird symmetrisch mitübertragen.
- Single-Link (3 TDMS-Signale) und Dual-Link-Verbindungen (6 TDMS-Signale) für höhere Auflösungen. (*bis zu 7,44Gbit/s*)
- Rückwärtskompatibilität: Übertragung von 4 analogen Signalen (RGB+H-Sync) möglich.

- DVI-D nur die digitalen Leitungen werden genutzt
- DVI-A nur die analogen Leitungen werden genutzt
- DVI-I analoge und digitale Leitungen werden genutzt

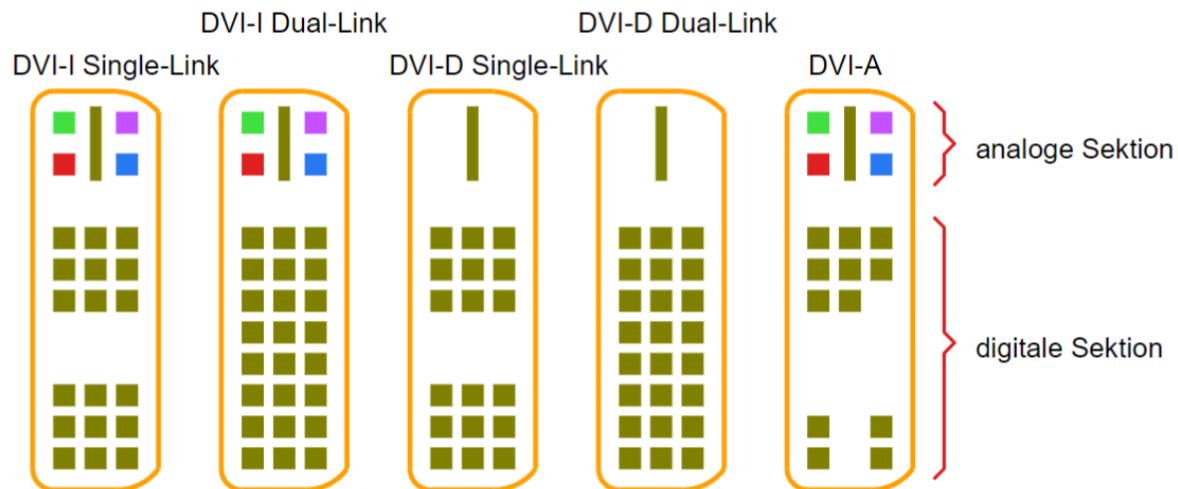


Abbildung 227 DVI-Ausführungen

Wichtig

Die digitale TDMS-Kodierung ist ein 8b10b-Code. Es werden aus dem 10-Bit-Vorrat nur Bitmuster verwendet, die gleich viele Anteile von Einsen und Nullen enthalten, damit die Übertragung immer gleichanteilsfrei bleibt.

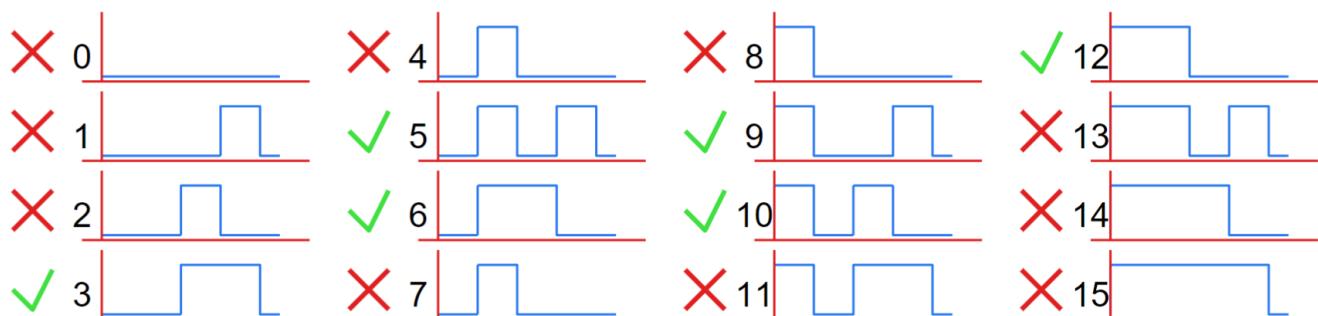


Abbildung 228 Beispiel gleichanteilsfreier 4-Bit-Code

7.4.2.4 GPU in CPU integriert

- **Frühe Integration von Grafik in die CPU:**
 - Versuch, DRAM-Speicher für Grafikausgabe einzusparen.
 - Ein Teil des Hauptspeichers wurde als Bildspeicher verwendet.
 - Einige Rechnerentwürfe in den 1990er Jahren verwendeten externe Grafikkontroller.
 - Pixeldaten wurden über den Systembus geleitet, was die Systeme stark ausbremste.

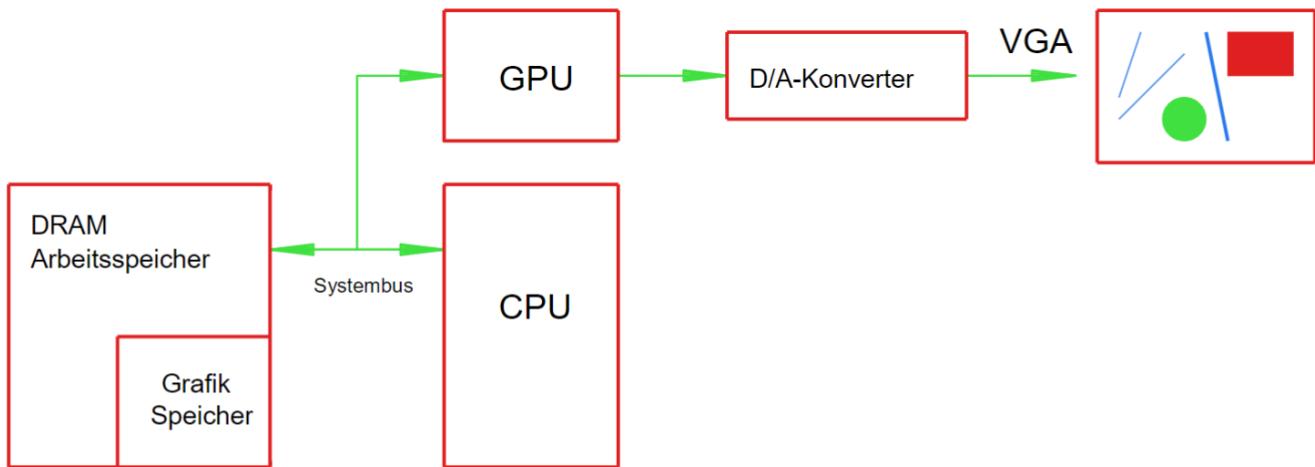


Abbildung 237 Pixel im Hauptspeicher

- **Entwicklung bis zur modernen Integration:**

- Mit zunehmender Integration von Cache-Speichern in CPUs und mehrkanaligen DRAM-Speicherinterfaces ist die separate Nutzung von Speicher für Grafikzwecke obsolet geworden.

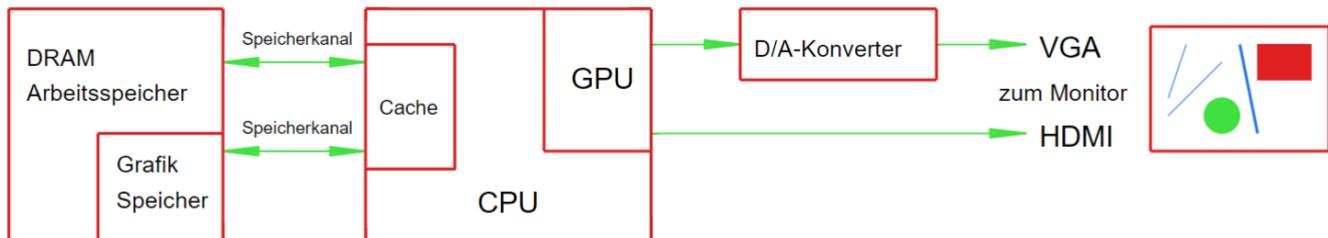


Abbildung 238 CPU GPU RAM

- **Integration von Grafikbeschleunigungsfunktionen:**

- Ab etwa 2010 wurden alle benötigten Beschleunigungsfunktionen in den Chip der Haupt-CPU integriert (z. B. Intel HD Graphics, Intel Graphics Media Accelerator).
- Die Leistung der integrierten 3D-Funktionen ist im Vergleich zu externen Grafiksystemen mit eigenem Speicher weniger stark.