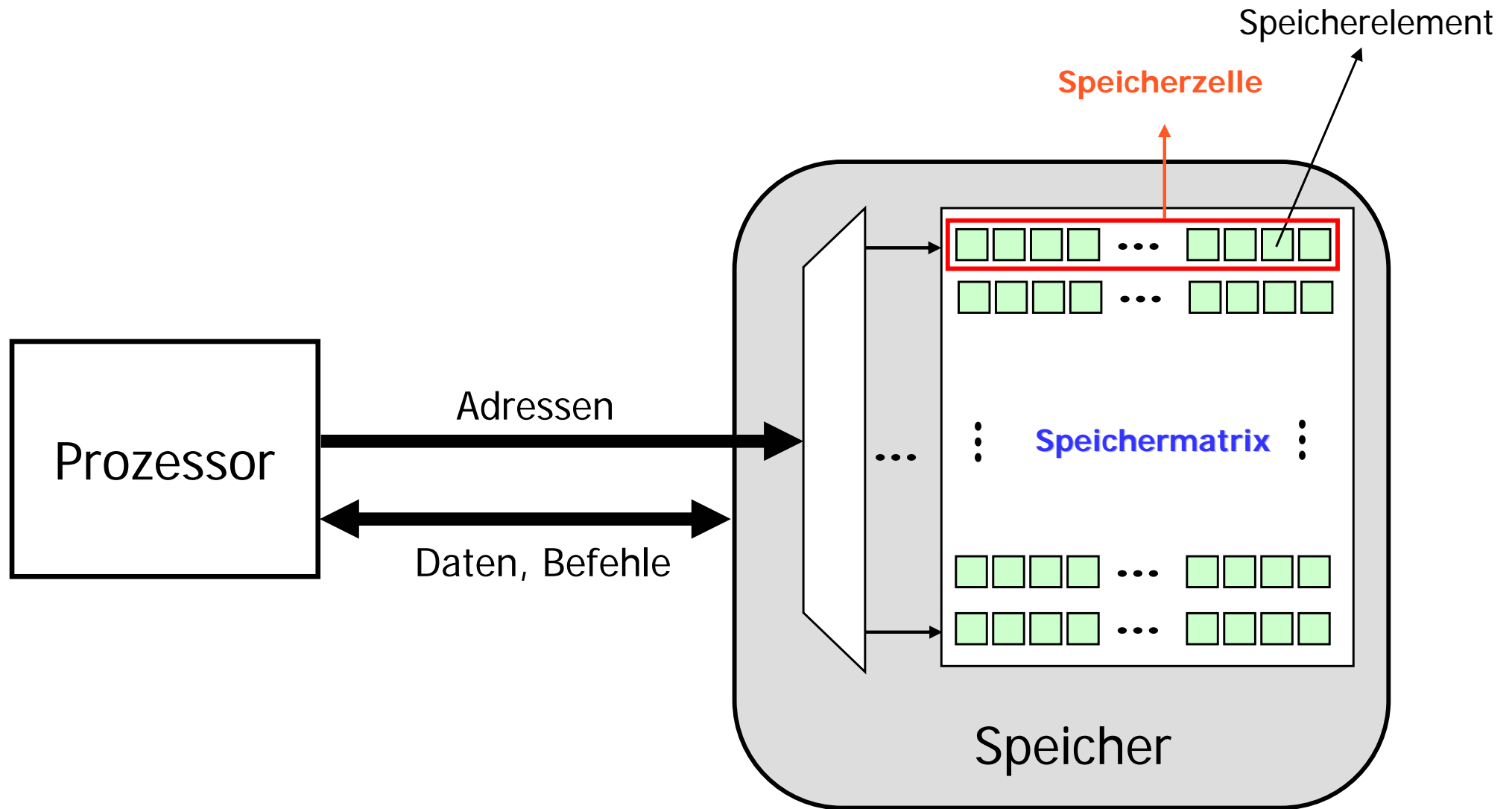


Kapitel 6

Speicherwerk

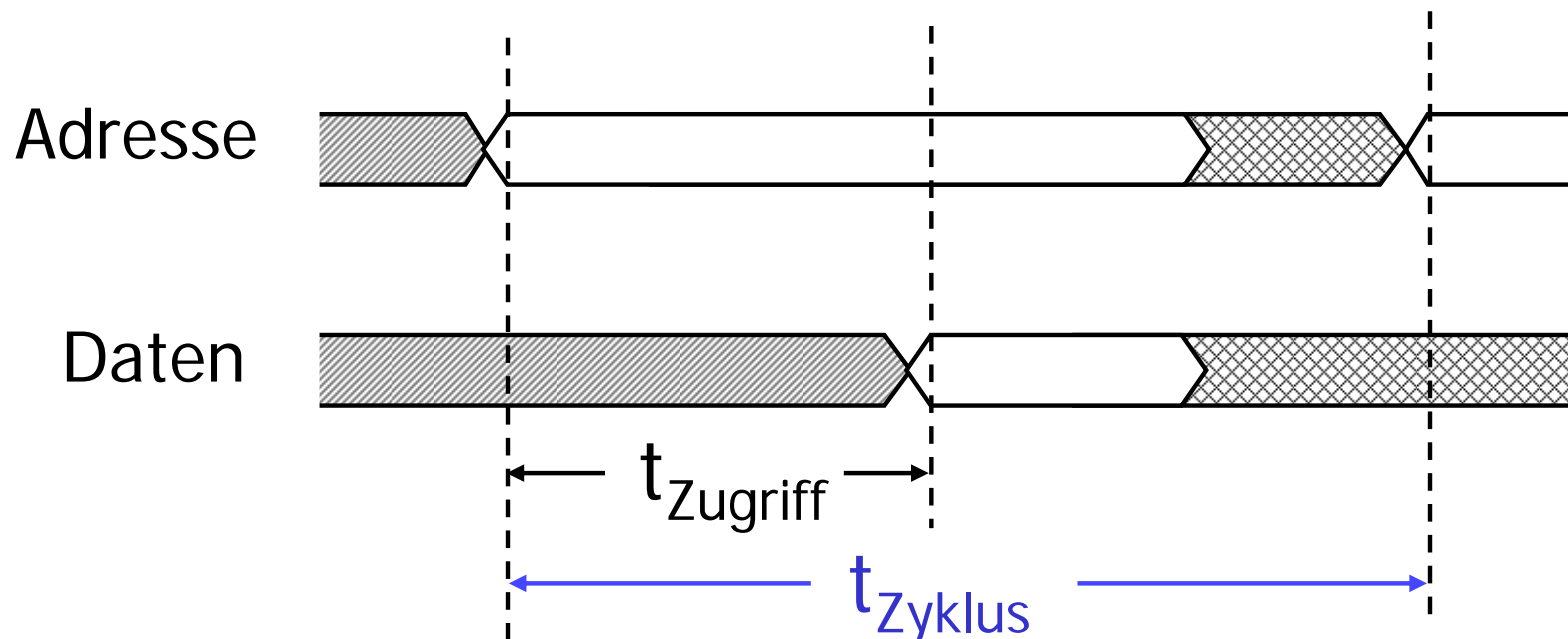
- Halbleiterspeicher
- Aufbau und Organisation
- Techniken zur Zugriffsbeschleunigung

Allgemeine Struktur: Matrixanordnung



Zugriffszeit / Zykluszeit

- **Zugriffszeit (*access time*)**
maximale Zeitdauer, die vom Anlegen einer Adresse an den Speicher bis zur Ausgabe der gewünschten Daten vergeht
- **Zykluszeit (*cycle time*):**
minimale Zeitdauer, die zwischen zwei hintereinander folgenden Aufschaltungen von Adressen an den Speicher vergehen muss.



Zugriffszeit / Zykluszeit

Die Zykluszeit kann erheblich länger als die Zugriffszeit sein!

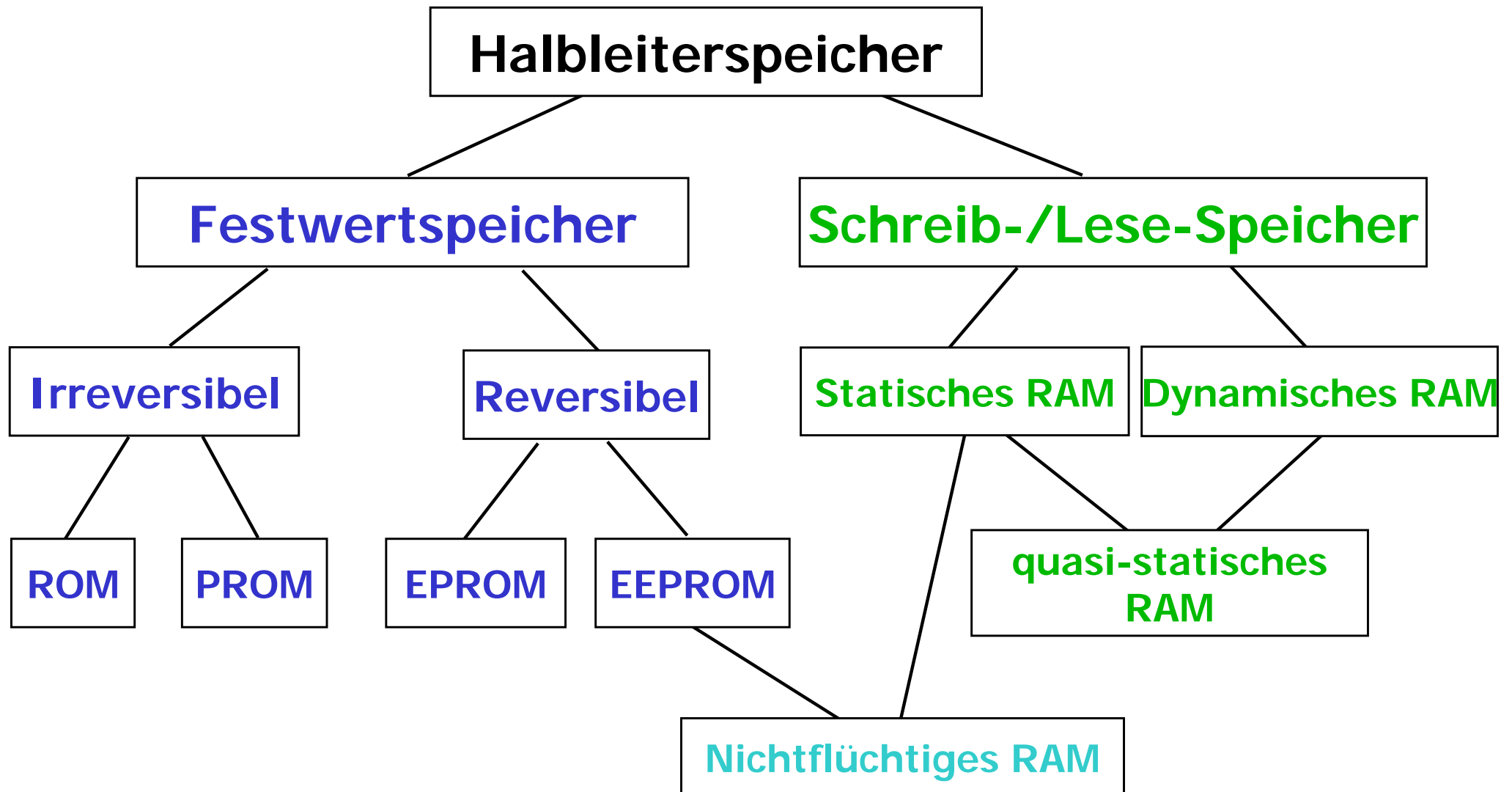
Gründe:

- Speicherzelle muss sich nach einem Zugriff "erholen"
- Bei einigen Speicherarten wird die Information durch das Auslesen zerstört und muss erst wieder eingeschrieben werden (refresh)

Idealfall: $\text{Zykluszeit} = \text{Zugriffszeit}$

Realität: meist $\text{Zykluszeit} > \text{Zugriffszeit}$

6.2 Klassifizierung von Halbleiterspeichern



Festwertspeicher (ROM, *Read Only Memory*)

Inhalt ist während des Normalbetriebs nur lesbar. Inhalt ist nicht flüchtig (*non volatile*), d. h. er geht bei Abschaltung der Versorgungsspannung nicht verloren

❑ **Irreversible Festwertspeicher:**

Das Einschreiben einer Information kann nicht wieder rückgängig gemacht werden

- maskenprogrammierten Festwertspeicher (ROM):
Programmierung bei der Herstellung, nur bei großer Stückzahl
- programmierbare Festwertspeicher (PROM):
einmaliges Programmieren durch den Anwender mit Programmiergerät (Durchbrennen von Verbindungen, fusable links)

Festwertspeicher (ROM, *Read Only Memory*)

□ **Reversible Festwertspeicher:**

Das Einschreiben der Informationen kann wieder rückgängig gemacht werden.

- UV-löschbaren Festwertspeicher (EPROM):
Löschen durch UV-Licht.
- Elektrisch löschbare Festwertspeicher (EEPROM):
können im μ P-System selbst elektrisch gelöscht werden.
Dieser Vorgang ist jedoch sehr langsam und nur begrenzt oft möglich.
- Flash-ROM:
Werden genauso wie EEPROMs elektrisch programmiert. Die Daten werden nicht byteweise sondern blockweise ausgelesen und geschrieben.

Schreib/Lese-Speicher (RAM, *Random Access Memory*)

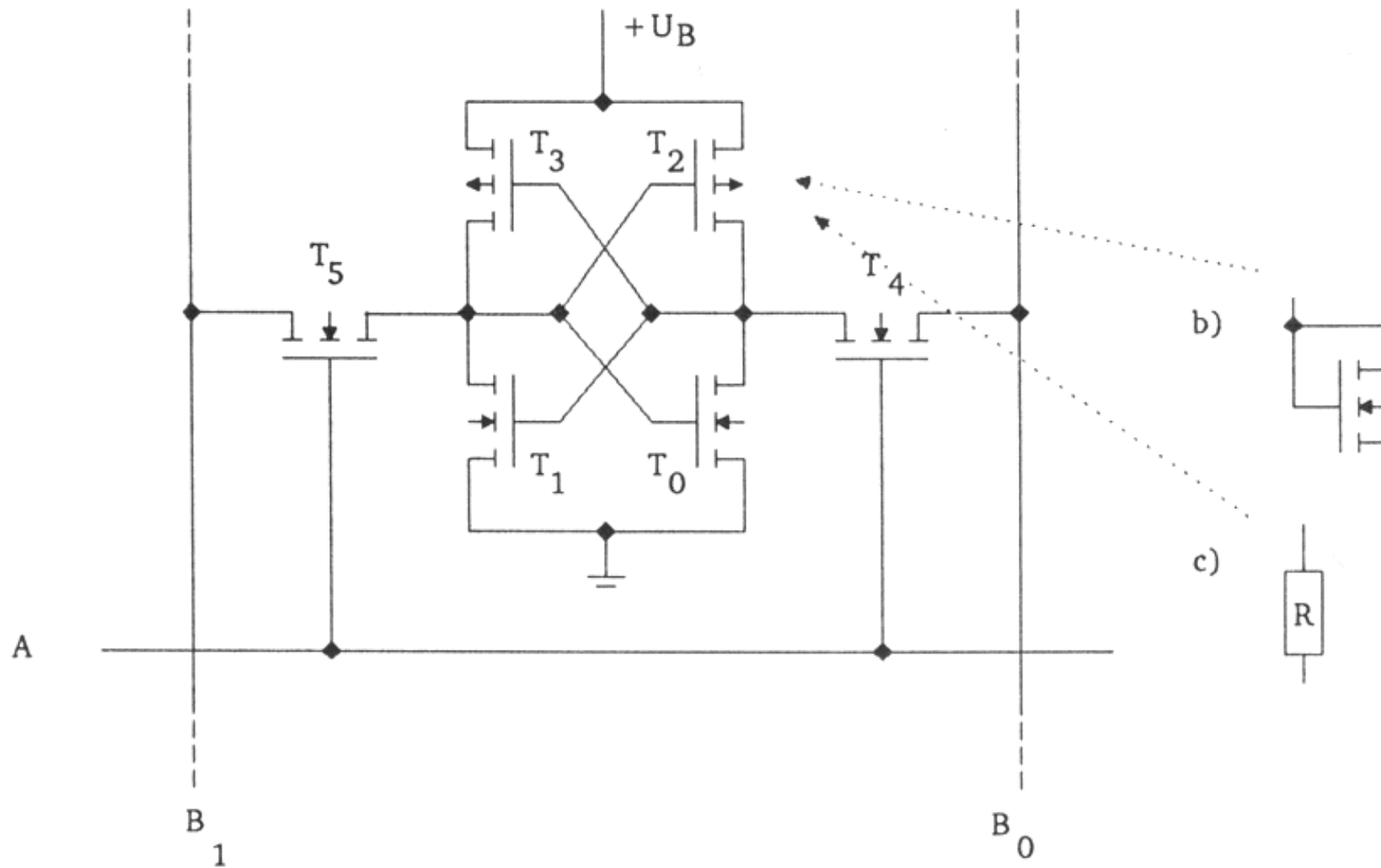
- ❑ Random Access Memory (RAM): Zugriffszeit unabhängig vom Ort (Adresse) im Speicher
 - ❑ Inhalt ist jederzeit lesbar und schreibbar.
 - ❑ Inhalt ist flüchtig (*volatile*), d. h. er geht bei Abschalten der Versorgungsspannung verloren.
-
- **Statische Schreib/Lese-Speicher (SRAM)**
 - **Dynamische Schreib/Lese-Speicher (DRAM)**

Schreib/Lese-Speicher (RAM, *Random Access Memory*)

□ **Statische Schreib/Lese-Speicher (SRAM)**

- Speichern Daten in Flipflops.
- Inhalt stabil, solange Versorgungsspannung vorhanden ist.
- Versionen:
 - aus Bipolar-Transistoren
(schneller, aber großer Flächenbedarf)
 - aus MOS-Transistoren
(langsamer, aber kleinerer Flächenbedarf)

Statische CMOS-Speicherzellen



Statische CMOS-Speicherzellen

- Eine Statische CMOS-Speicherzelle besteht aus
 - 2 kreuzweise rückgekoppelten Invertern (T_0, T_2 bzw. T_1, T_3).
 - 2 Transistoren T_4 und T_5 zur Ankopplung an die Bitleitungen

→ 6 Transistor Zelle

Vorteil der CMOS-Zelle:

nur zum Umschaltzeitpunkt fließt Strom

NMOS-Zelle:

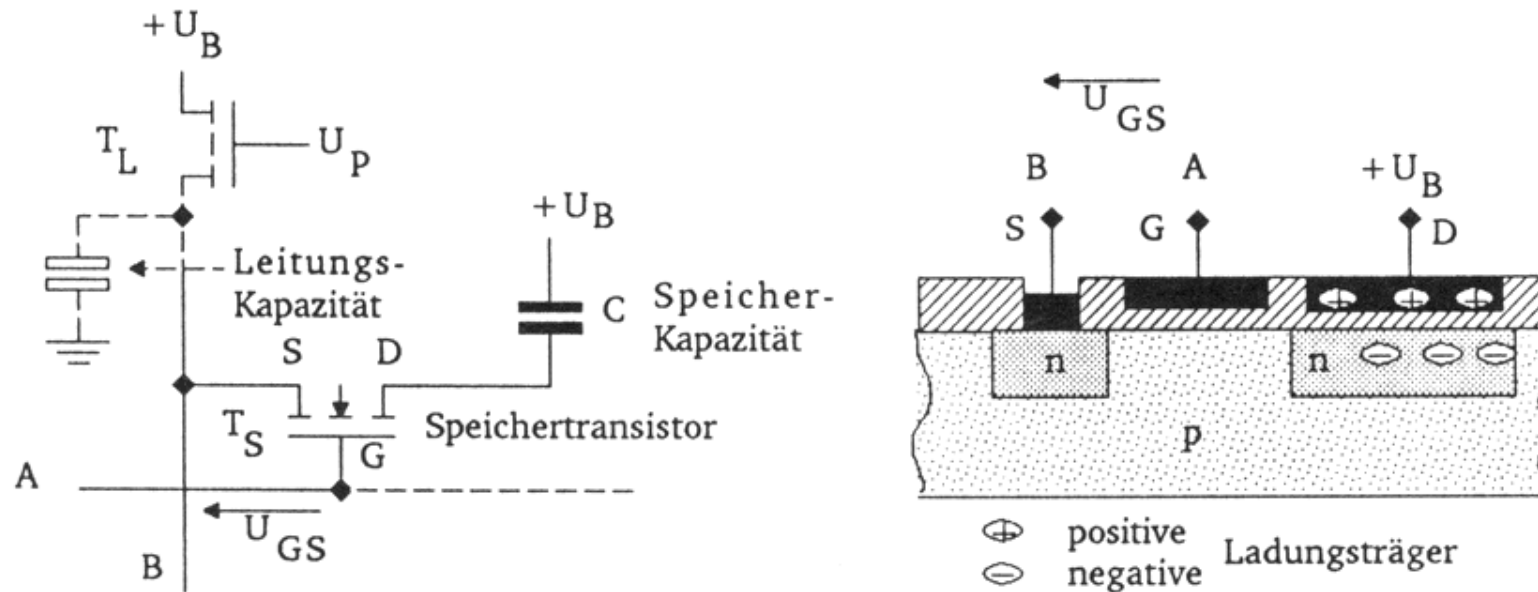
Inverter aus n-Kanal-Transistor und Widerstand
wird benutzt (Varianten b, c)

Schreib/Lese-Speicher (RAM, *Random Access Memory*)

□ **Dynamische Schreib/Lese-Speicher (DRAM)**

- Speichern Daten als elektrische Ladung in einem Kondensator
- Das Lesen bewirkt eine Entladung (*destructive read*)
➔ nach dem Lesen muss wieder eingeschrieben werden
- Die Ladung geht nach einiger Zeit auch durch Leckströme verloren ➔ periodische Auffrischung erforderlich (*refresh*)
- Integrationsdichte höher als SRAM (ca. 4 mal)

Dynamische MOS-Speicherzellen



1 Transistorzelle, kleinster Aufwand von allen betrachteten Zellen
(1/4 Platzbedarf einer SRAM-Zelle)

Die Information wird in einem Kondensator gespeichert.

Dieser Kondensator wird durch eine vergrößerte Drain-Zone gebildet, die durch eine dünne Isolierschicht vom Drain-Kontakt getrennt ist

Kapazität ca. 0,1 - 0,5 pF \rightarrow speichert 100 000 - 150 000 Elektronen

Dynamische MOS-Speicherzellen

Lesen:

Problem: Speicherkapazität hat ungefähr die gleiche Größe wie parasitäre Leitungskapazität der Bitleitung

Das Lesen erfolgt durch Ladungsvergleich zwischen Leitungskapazität und Speicherkapazität

Ablauf:

- Zunächst wird die Leitungskapazität vorgeladen (*precharge*), indem die Bitleitung kurz über T_L mit $+U_B$ verbunden wird
- Zum Lesen wird dann über A eine positive Spannung an das Gate des Speichertransistors gelegt

Ist die Speicherkapazität geladen, so findet ein Ausgleich mit den Ladungsträgern der Bitleitung statt

Leseverstärker am Ende der Bitleitung mißt diesen Strom

Dynamische MOS-Speicherzellen

□ Schreiben:

Durch Anlegen einer positiven Spannung U_{GS} wird der Speichertransistor leitend.

Liegt nun die Bitleitung B auf Masse

➔ Elektronen werden auf die Drain-Zone aufgebracht, der Speicherkondensator geladen

Liegt die Bitleitung B auf U_B

➔ Elektronen werden von der Drain-Zone abgesaugt, der Speicherkondensator entladen

(Zuordnung log. 0 / log. 1 zu Ladung / keine Ladung ist rein willkürlich)

Schreib/Lese-Speicher (RAM, *Random Access Memory*)

- **Quasi-statische Schreib/Lese-Speicher (iRAM, integrated RAM)**
 - dynamische Speicher, welche die Schaltung für das Wiederauffrischen mit auf dem Chip haben
 - ➔ sie sehen nach außen wie statische Speicher aus, sind jedoch im inneren dynamisch

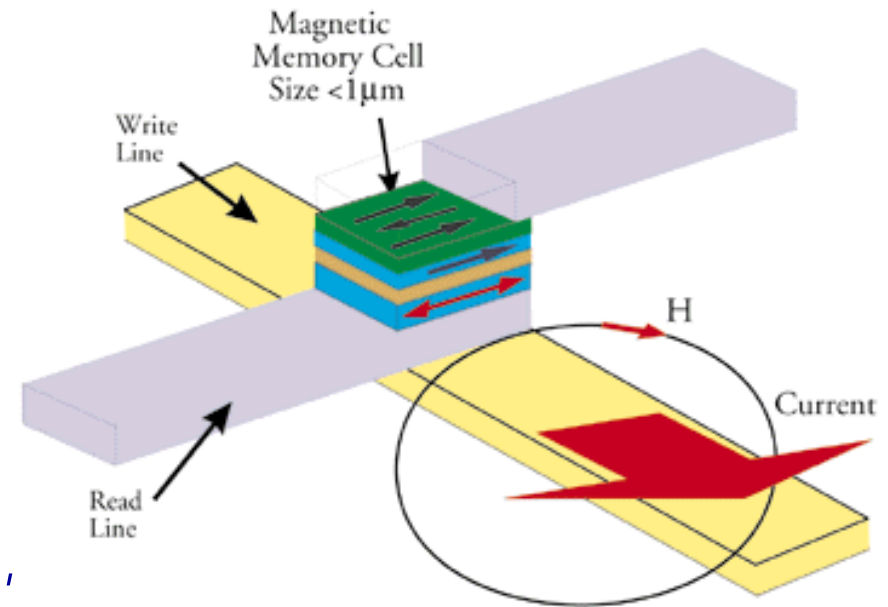
Nicht-flüchtige RAM's (NVRAM, *non volatile RAM*)

- ❑ Kombination aus Festwertspeicher und Schreib/Lese-Speicher auf einem Baustein.
- ❑ Jede Speicherzelle ist **doppelt** ausgelegt, einmal als **statische RAM-Zelle** und einmal als **EEPROM-Zelle**.
- ❑ Im Normalbetrieb werden die RAM-Zellen benutzt.
- ❑ Ein **spezielle Steuerschaltung** erlaubt jedoch das Kopieren des RAM-Inhalts in das EEPROM.

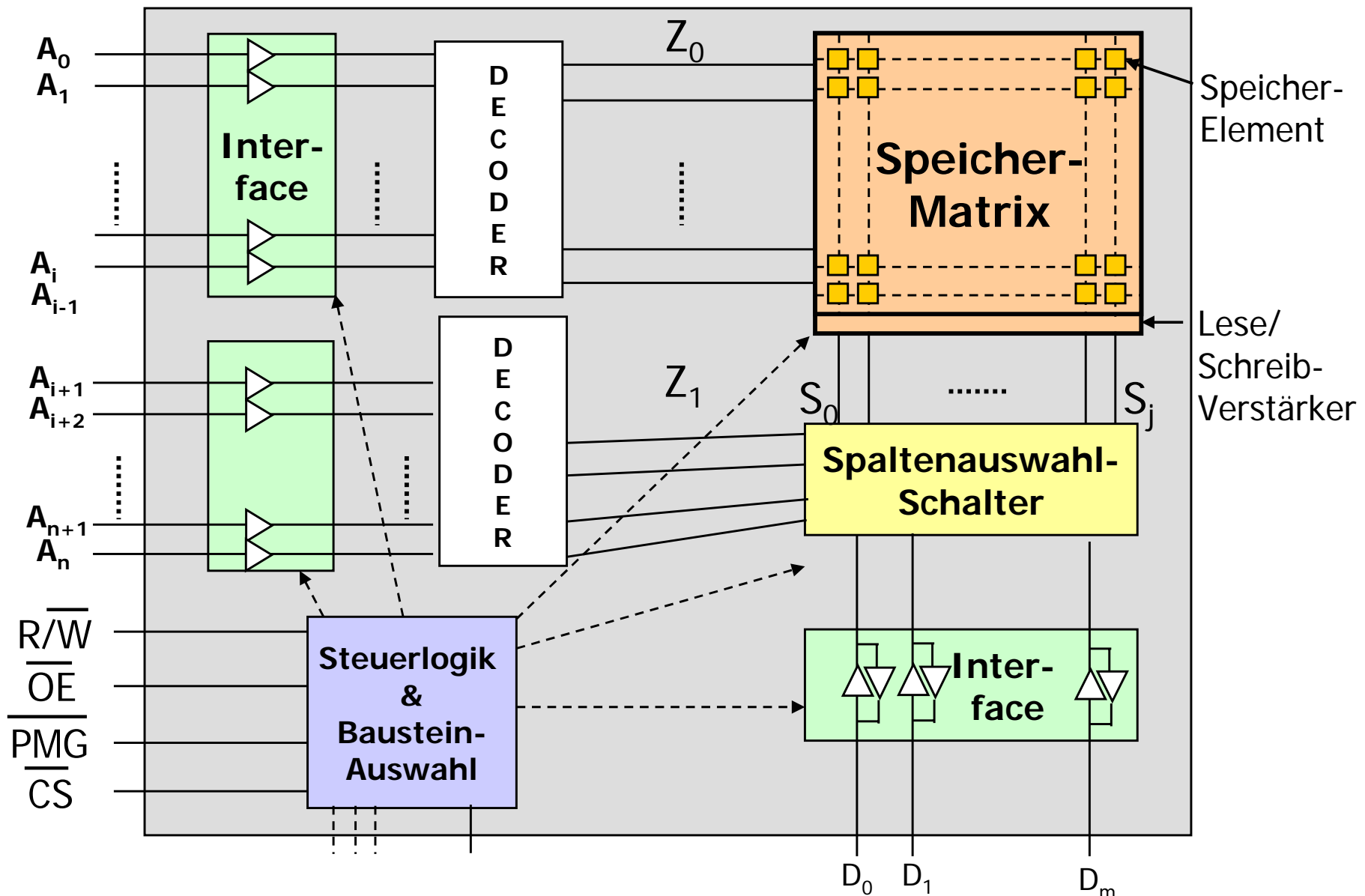
VLSI Symposium in Kyoto, Japan, 10 -14. Juni 2003

IBM und Infineon mit neuer Speichertechnik: **MRAM**
(*Magnetoresistive Random Access Memory*)
soll die Vorteile der verschiedenen RAM-Technologien vereinen:

- hohe Geschwindigkeit von SRAMs,
 - Dichte und die Kostenvorteile von DRAMs,
 - nichtflüchtige Speicherung von Flash-Speichern
-
- Magnetische statt elektronischer Ladungselemente für die Speicherung
 - Gefertigter Chip mit 128 Kbit in 0.18 μm -Technologie und 1,4 μm^2 Fläche
 - Mit der MRAMs kann z. B. der Bootvorgang bei PCs und Mobilgeräten entfallen.
 - Ab 2005 soll es mit Massenproduktion nach einer gemeinsamen Erklärung von IBM und Infineon soweit sein.



6.3 Organisation von Speicherbausteinen



6.3 Organisation von Speicherbausteinen

Speichermatrix:

Matrixförmige Anordnung der Speicherzellen

Auswahl einer Zelle durch

- Zeilenauswahlleitungen $Z_0 \dots Z_l$
- Spaltenauswahlleitungen $S_0 \dots S_j$

Minimierung der Anzahl Auswahlleitungen durch eine quadratische Matrix, d. h. $l \approx j$

6.3 Organisation von Speicherbausteinen

Steuerlogik & Bausteinauswahl:

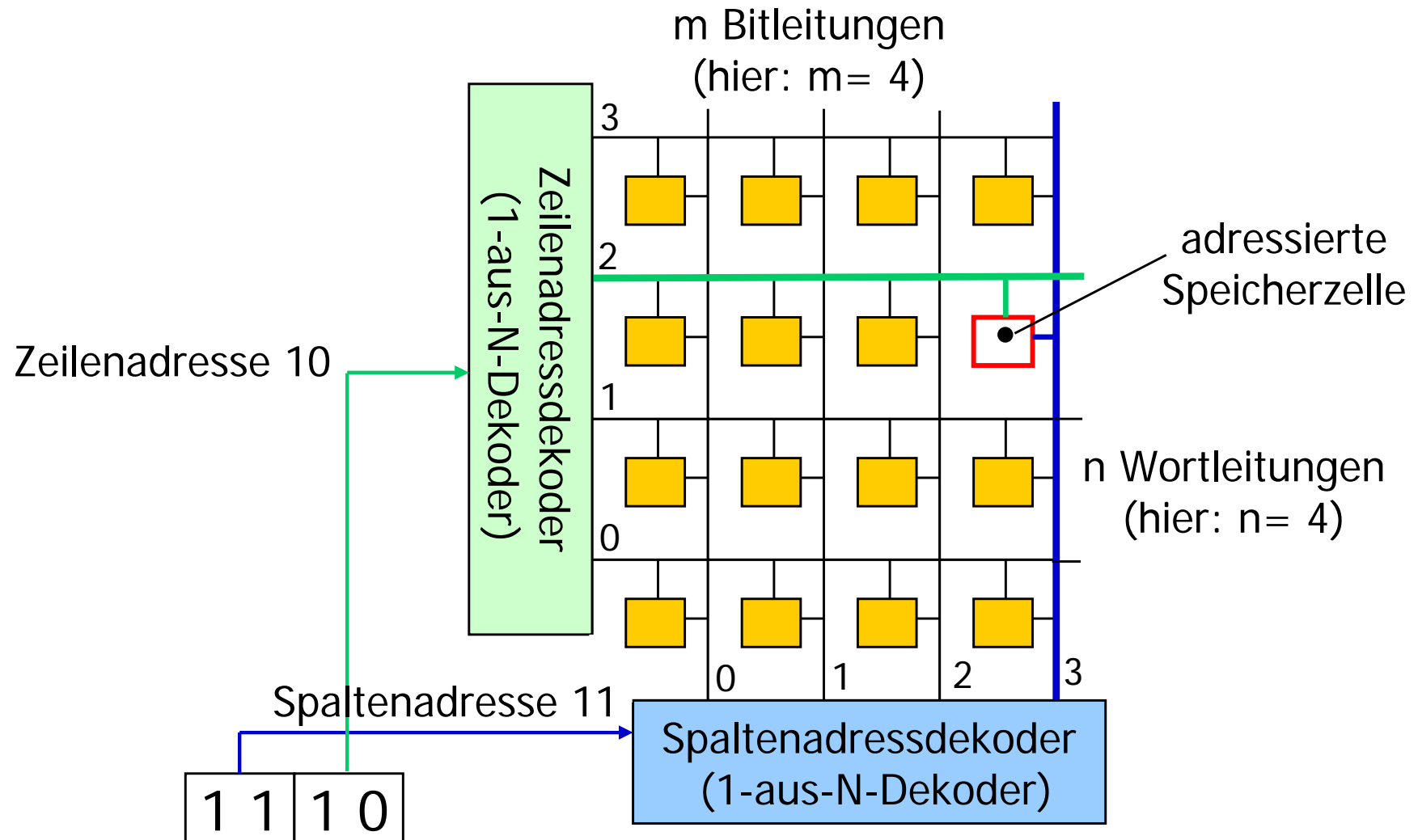
- ❑ $\overline{\text{CS}}$ (*chip select*) oder CE (*chip enable*): Bausteinauswahl, wird meist durch Adressdekoder aus den höchstwertigsten Adressbits erzeugt
- ❑ $\text{R}/\overline{\text{W}}$ (*read/write*): Auswahl Lesen oder Schreiben, nur bei RAMs, steuert die Richtung der Schreib/Leseverstärker und die Art der Zellenansteuerung
- ❑ $\overline{\text{OE}}$ (*output enable*): Aktiviert die Ausgangstreiber an der Datenbusschnittstelle (schaltet den Speicherbaustein als Quelle auf den Datenbus, Tristate Treiber)
- ❑ $\overline{\text{PGM}}$ (*program*): Neuprogrammieren des Bausteins, nur EPROMs, EEPROMs, NVRAMs, über U_p muß dann die Programmierspannung (12..21V) zugeführt werden

6.3 Organisation von Speicherbausteinen

Gewinnung der Zeilen- und Spaltenauswahlleitungen aus den Adressleitungen:

- Niederwertige Adressbits bilden über Interface-Treiber (Treiberleistung, Pegelanpassung, Adresslatch) und einen 1-aus-n Decoder die Zeilenauswahlleitungen
 - ➔ sie wählen eine Zeile der Speichermatrix aus
- Durch die Auswahl einer ganzen Zeile in einer möglichst quadratischen Matrix werden i. A. erheblich mehr Speicherzellen angesprochen als aktuell ausgelesen werden müssen.
 - ➔ Auswahl der anzusprechenden bzw. auszulesenden Spalten durch einen Spaltenauswahl-Schalter. Dieser wird über Treiber und Decoder von den höherwertigen Adressbits gesteuert

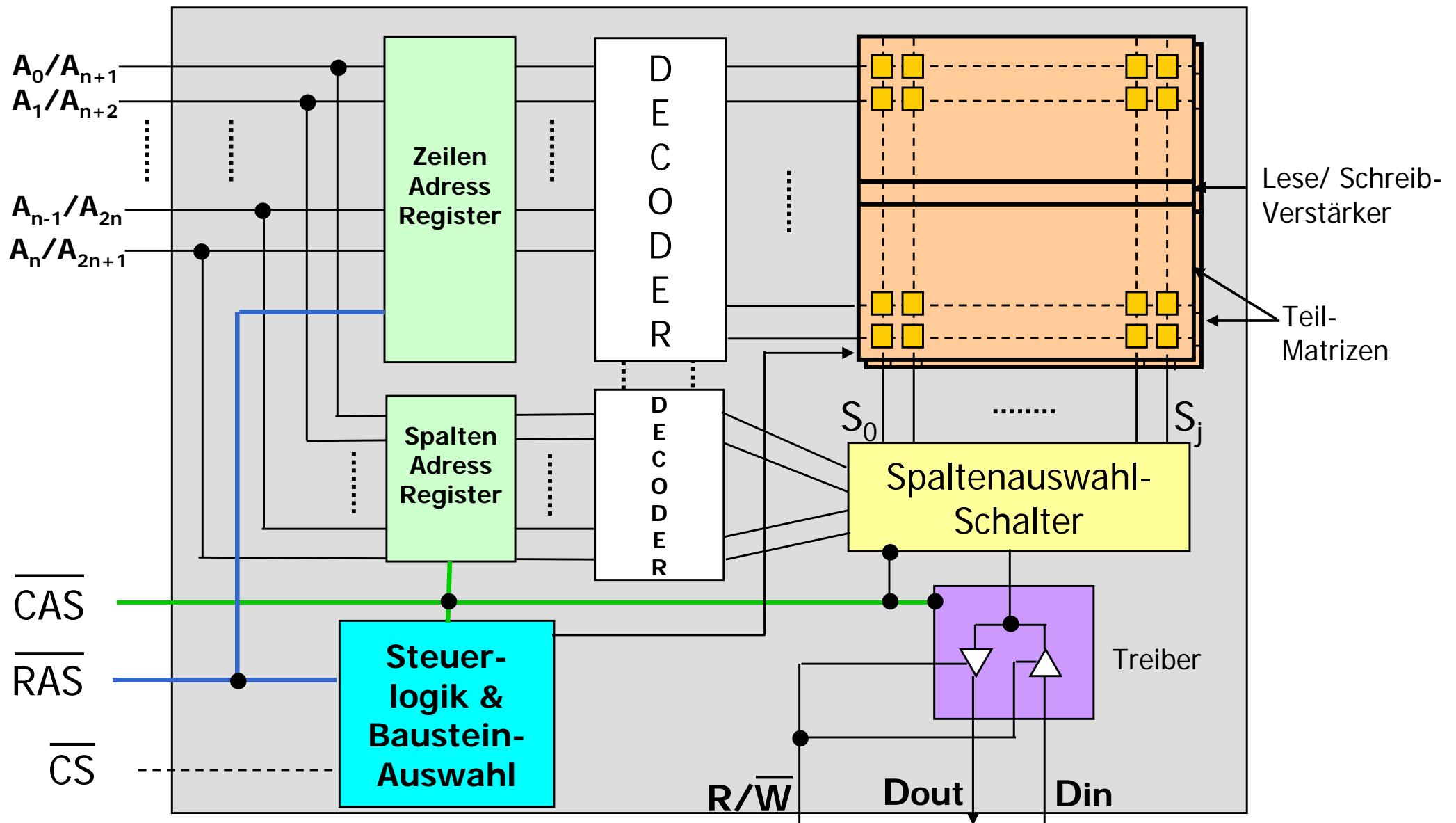
Beispiel: Selektieren einer Speicherzelle aufgrund der gegebenen Speicheradresse



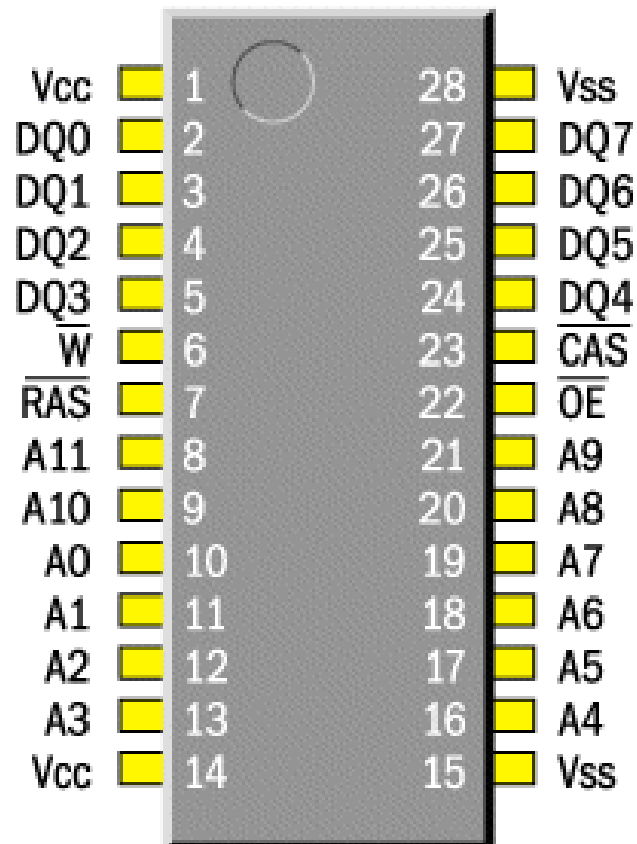
6.4. Dynamische RAM-Bausteine

- ❑ Größte Integrationsdichte aller Halbleiterspeicher
- ❑ Bitweise organisiert, getrennter Dateneingang und Datenausgang
- ❑ Speicheradressen gemultiplext (spart Anschlüsse) und im Speicherchip in Registern zwischengespeichert:
 - Auswahl der Zeilenadresse über das $\overline{\text{RAS}}$ -Signal (Row Address Select)
 - Auswahl der Spaltenadresse über das $\overline{\text{CAS}}$ -Signal (Column Address Select)

6.4. Dynamische RAM-Bausteine

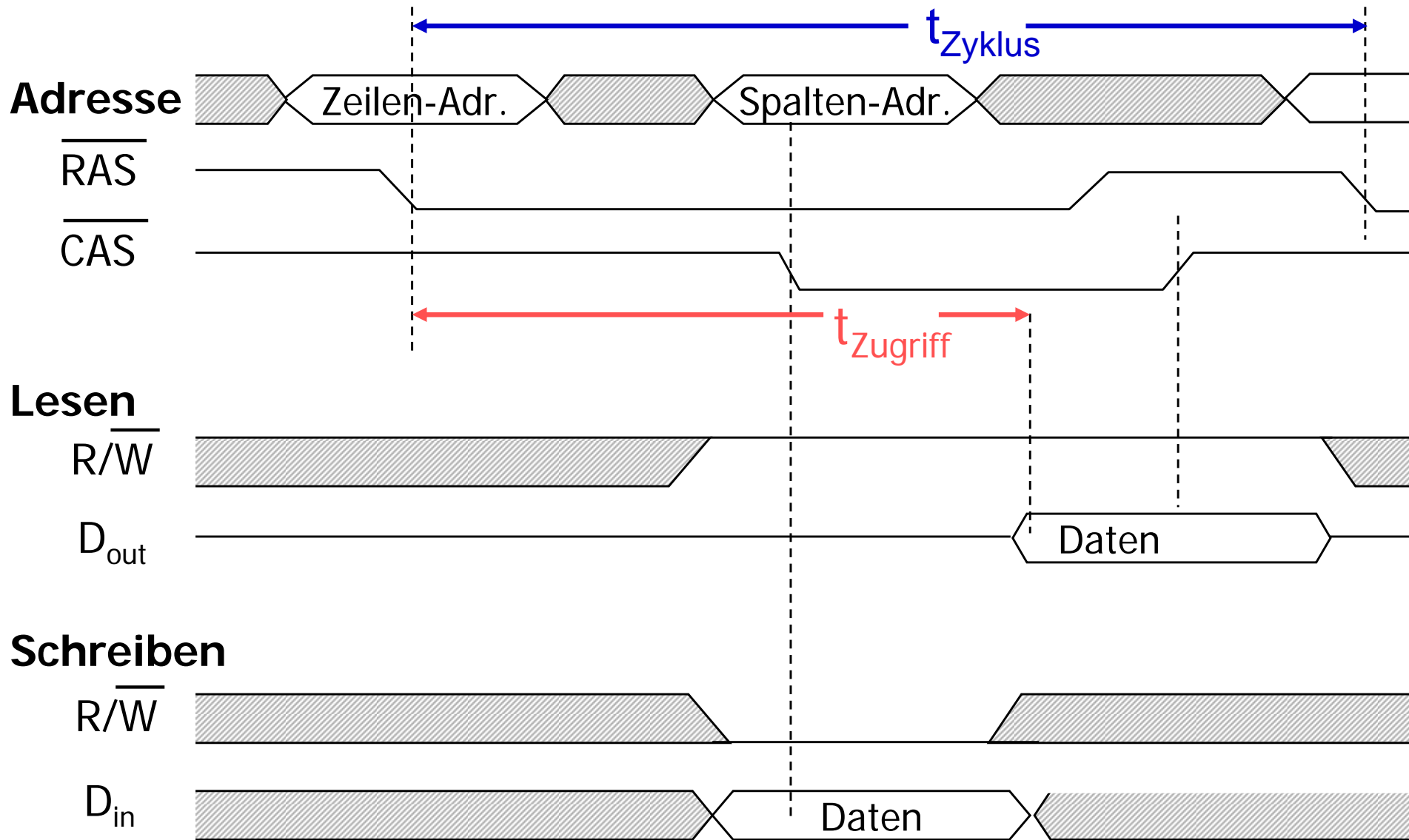


Pin-Belegung DRAM-ICs mit 2Mx8 Organisation



Pin Name	Pin Function
A0 - A11	Address Inputs
DQ0 - 7	Data In/Out
Vss	Ground
$\overline{\text{RAS}}$	Row Address Strobe
$\overline{\text{CAS}}$	Column Address Strobe
$\overline{\text{W}}$	Read/Write Input
$\overline{\text{OE}}$	Data Output Enable
Vcc	Power (+5V)

Adressierung eines dynamischen RAM-Bausteins



Adressierung eines dynamischen RAM-Bausteins

Adressieren:

negative RAS-Flanke übernimmt die Zeilenadresse ins Adressregister, negative CAS-Flanke die Spaltenadresse

Lesen:

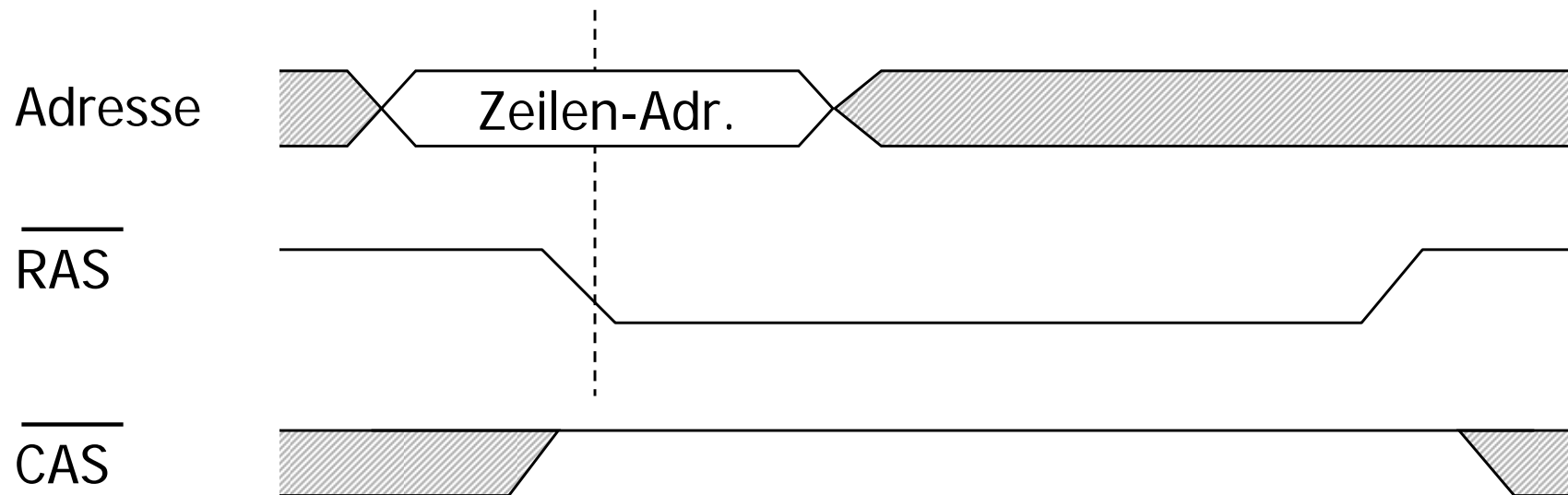
Daten erscheinen eine gewisse Zeit nach der negativen CAS-Flanke am Ausgang, Zyklus wird durch $CAS = 1$ wieder beendet

Schreiben:

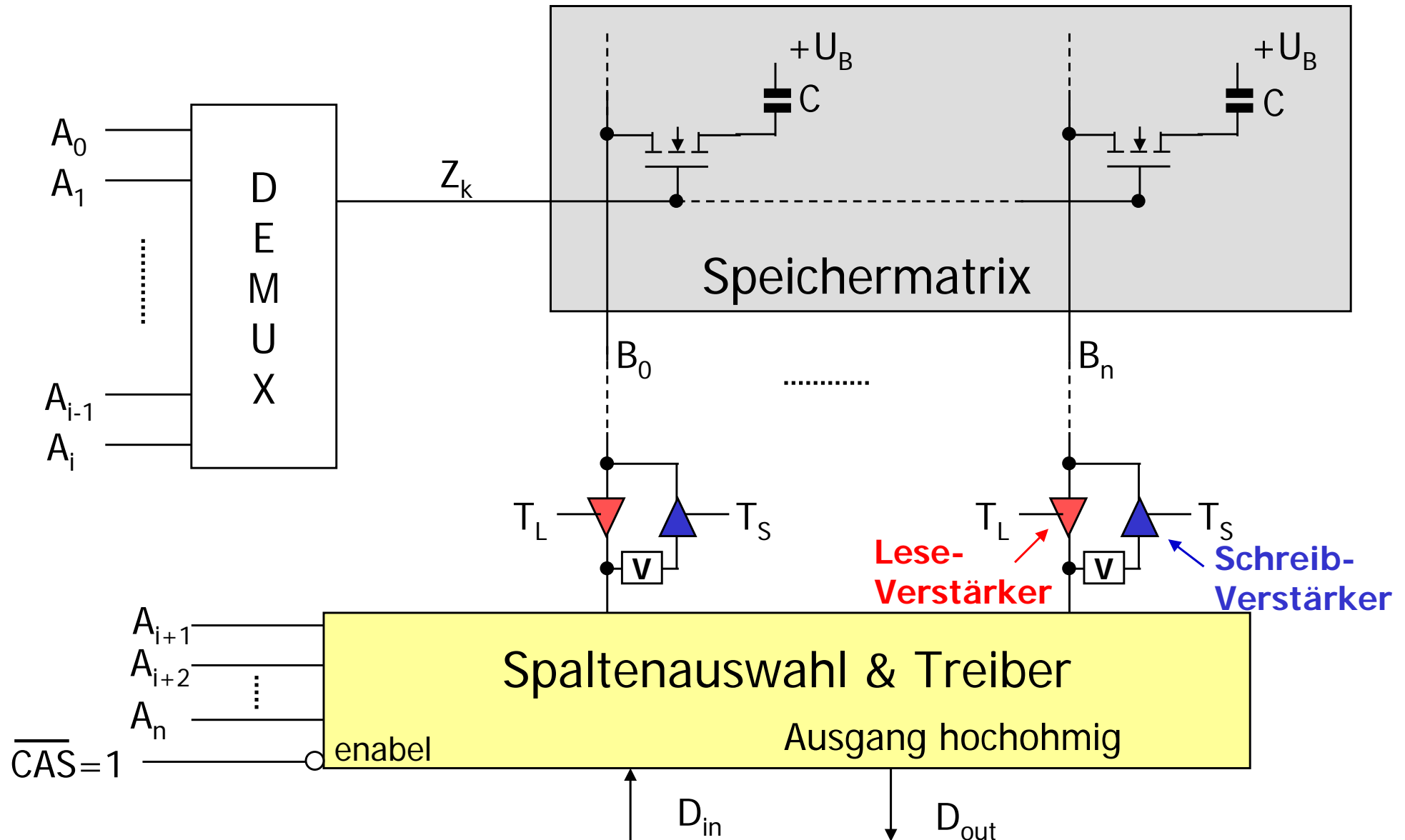
Das zu schreibende Datum muss gleichzeitig mit Spaltenadresse am Dateneingang anliegen.
Datenübernahme geschieht mit negativer CAS-Flanke

Auffrischen dynamischer RAMs

- Geschieht zeilenweise, jede Zeile muss aufgefrischt werden (alle 2 msec)
- Nur die Zeilenadresse wird an den Baustein angelegt, $\overline{\text{RAS}} = 0$ und $\overline{\text{CAS}}$ konstant auf 1



Aufbau der Auffrischlogik



Auffrischen dynamischer RAMs

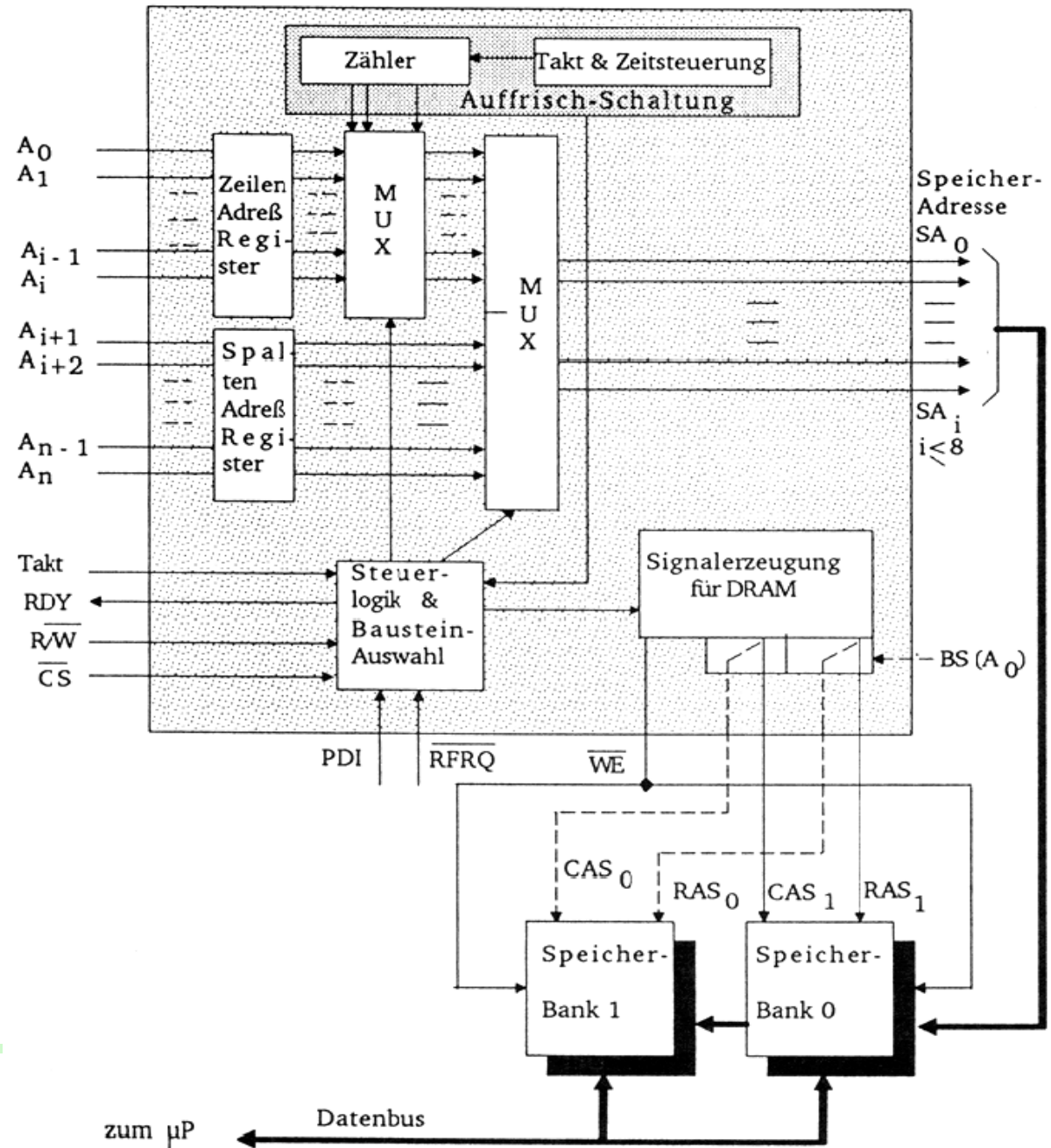
- Der Zeileninhalt wird zunächst, gesteuert vom Takt T_L über die Leseverstärker ausgelesen
- Über eine kleine Verzögerung V wird dieser Inhalt dann mittels der Schreibverstärker gesteuert vom Takt T_S zurückgeschrieben

Um die Leitungskapazitäten besser zu verteilen, werden bei hochintegrierten Bausteinen die Schreib-/Lese-Verstärker auch oft in die Mitte der Spaltenleitungen verlegt (statt am Ende wie im Bild dargestellt)

DRAM-Controller

□ Aufgabe:

- Ansteuerung der DRAM-Bausteine
- Erzeugen der RAS- und CAS-Signale
- Multiplexen der Adressen
- Speicher-Refresh



DRAM-Controller

Bis zu 4 getrennte $\overline{\text{RAS}}$ - / $\overline{\text{CAS}}$ -Ausgänge bei DRAM-Controllern → Unterstützung von bis zu 4 Speicherbänken

Vorteil:

durch Zyklusüberlappung (*interleaving*) kann der Zugriff verkürzt werden.

Die Auswahl der Speicherbank (BS) wird z. B. mit A_0 beschaltet → fortlaufende Speicheradressen liegen in unterschiedlichen Speicherbänken

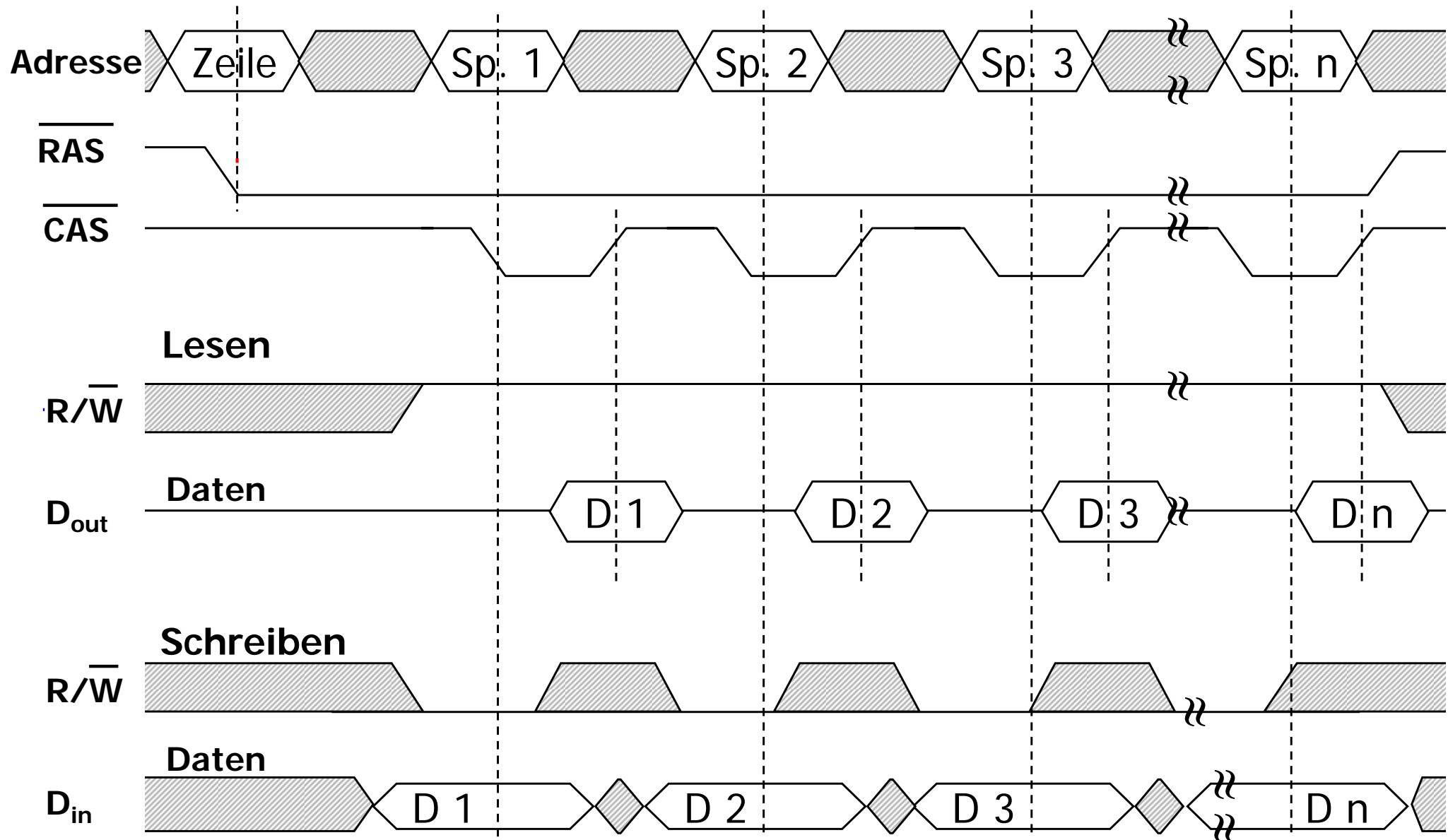
6.5 Techniken zur Zugriffsbeschleunigung

Ausgangspunkt:

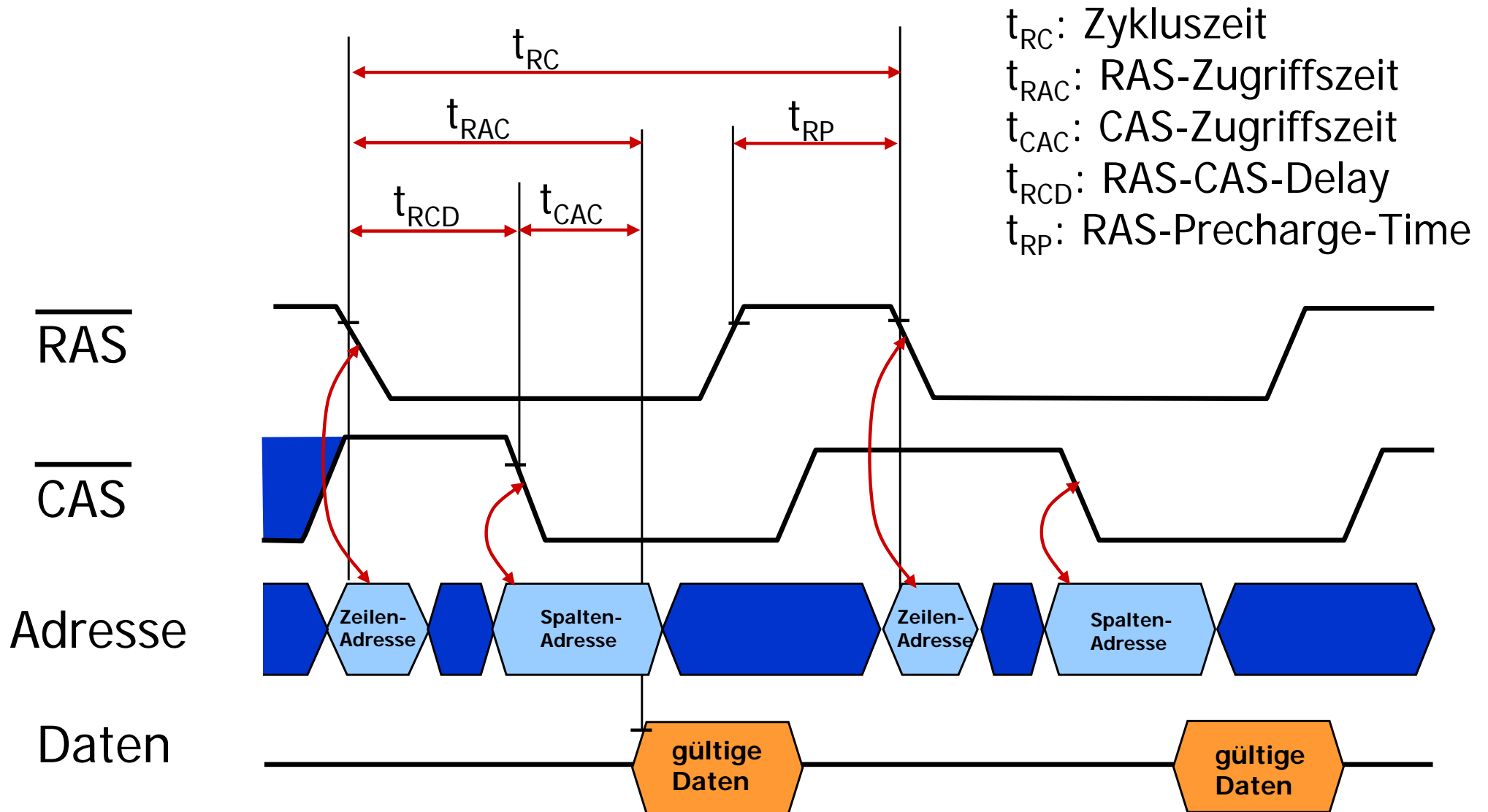
Zwischen Prozessor und Speicher oder Speicher und Speicher werden keine einzelnen Bytes übertragen, sondern benachbarte Gruppen von Bytes (Blöcke)

- Beschleunigter Zugriff auf den Speicher-Baustein, wenn alle zu lesenden oder schreibenden Speicherzellen in einer Zeile (Seite) liegen.
- Die Zeilenadresse wird bei einem wiederholten Zugriff auf diese Zeile (auch *page* genannt) nur einmal angelegt (wird im Register gespeichert). Dann werden in schneller Folge die Spaltenadressen angelegt (*fast page mode: FPM-DRAM*)
➔ erheblich beschleunigter Zugriff

Seitenzugriff bei DRAMs



DRAM Timing-Parameter



DRAM Timing-Parameter

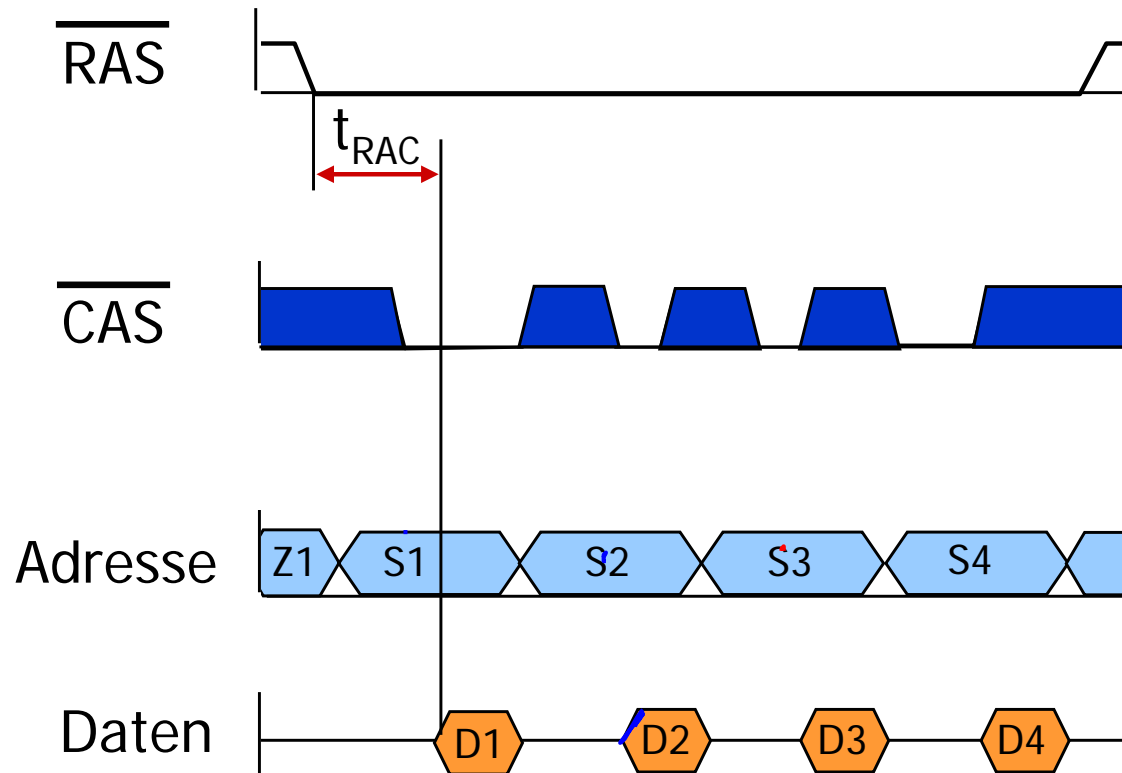
- **t_{RAC}** : minimale Zeitdauer, die zwischen der fallenden Flanke von RAS bis zur Ausgabe der gewünschten Daten vergeht.
 - $t_{\text{RAC}} = 60 \text{ ns}$ (bei einem 4 MB DRAM)
- **t_{RC}** : Minimale Zeitdauer von Beginn eines Zeilenzugriff bis zum nächsten (Zykluszeit).
 - $t_{\text{RC}} = 110 \text{ ns}$ bei einem 4Mbit DRAM mit t_{RAC} von 60 ns
- **t_{CAC}** : minimale Zeitdauer, die zwischen der fallenden Flanke von CAS bis zur Ausgabe der gewünschten Daten vergeht.
 - $t_{\text{CAC}} = 15 \text{ ns}$ bei einem 4Mbit DRAM mit t_{RAC} von 60 ns.
- **t_{RC}** : Minimale Zeitdauer vom Beginn eines Spaltenzugriff bis zum nächsten.
 - $t_{\text{RC}} = 35 \text{ ns}$ bei einem 4Mbit DRAM mit t_{RAC} von 60 ns

Fast Page Mode DRAM (FPM-DRAM)

Oft liegen aufeinander folgende Speicherzugriffe in der gleichen Zeile des DRAMs. Das wird bei FPM-DRAMs ausgenutzt:

- ❑ Initialisieren des ersten Lesezugriffs läuft wie bei normalem DRAM.
- ❑ Nach dem ersten Lesezyklus lässt die Speichersteuerung das RAS-Signal einfach auf aktiv → Zeile (page) bleibt aktive.
- ❑ Bei den folgenden Lesezugriffen übergibt die Speichersteuerung nur noch jeweils eine neue Spaltenadresse an das DRAM.
- ❑ RAS-Precharge-Zeit (t_{RP}) und die RAS-CAS-Delay (t_{RCD}) fallen bei den Folgezugriffen weg.

Timing-Diagramm eines FPM-DRAM



Zugriffszeit reduziert sich erheblich. Erst wenn der Speicherzugriff des Prozessors auf eine andere Zeile erfolgt, muss der Chipsatz einen normalen Lesevorgang initialisieren.

FPM-DRAM

Typische FPM-DRAM-Zugriffszeiten: 60 oder 70 ns

CAS-Zugriffszeit (t_{CAC}) bei einem 60 ns Baustein im FPM-Mode beträgt ca. 40 ns → Daten können im Abstand von 40 ns gelesen werden.

PC mit 66 MHz Bustakt (Taktperiode = 15 ns)

→ Prozessor kann im FPM-Mode nur bei jedem dritten Takt auf eine Zeile im Speicher zugreifen.

Initialisierung eines Lesevorgangs mit dem Anlegen von Zeilen- und Spaltenadresse dauert dagegen ganze fünf Takten 75 ns.

Maximale Datentransferrate: 200 Mbyte/sec

Prozessoren ab Intel Pentium führen durch ihren 64-Bit breiten Datenbus Speicherzugriffe mit 8 Byte durch. Bei einer CAS-Zykluszeit von 40 ns lassen sich 64 Bit in diesem Zeitraum übertragen.

EDO-RAM

EDO-RAM (*Extended Data Output*):

Die Datenausgabe wird beim Lesen vom $\overline{\text{CAS}}$ -Signal mittels interner Pufferung entkoppelt.

- die Daten stehen länger am Ausgang zur Verfügung
- bessere Verschachtelungsmöglichkeiten bei Lesezugriffen

Der Prozessor kann Daten auslesen, während die Speichersteuerung eine neue Spaltenadresse an das DRAM übergibt.

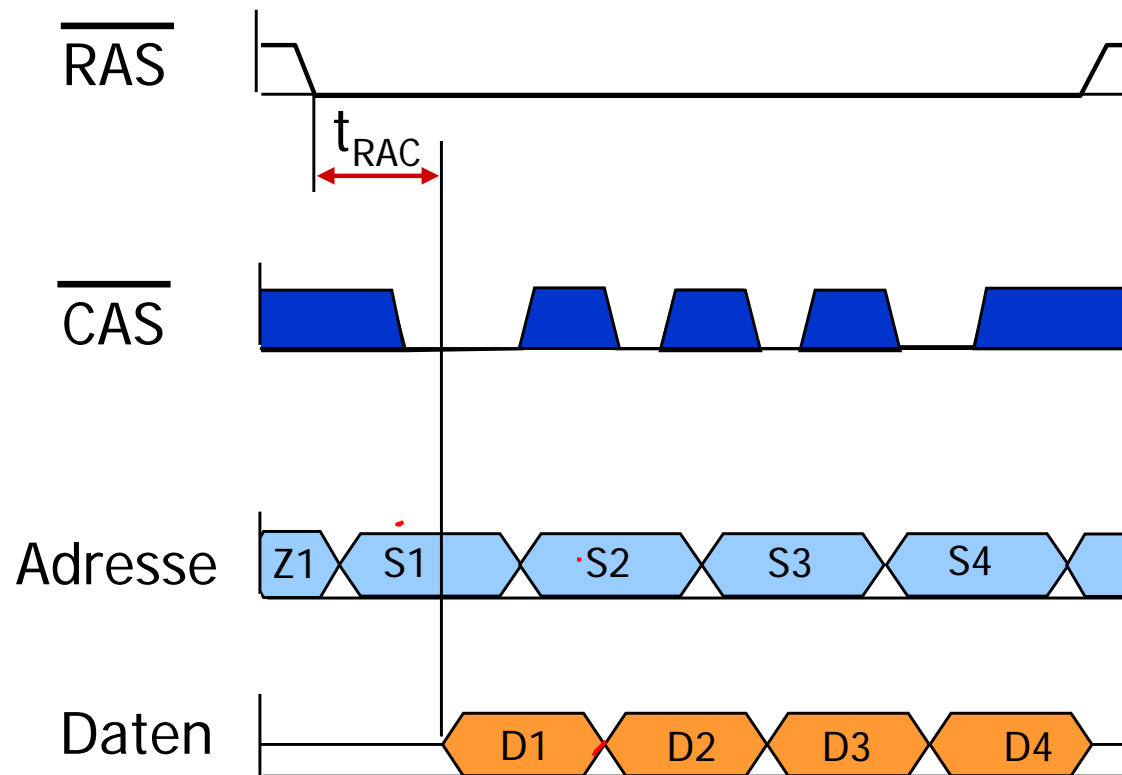
EDO-RAM

Weiterentwicklung der FPM-DRAMs. Durch eine einfache Modifikation der Steuerung wird eine Geschwindigkeitssteigerung erreicht.

- ❑ EDO-DRAMs sind um einen Latch-Speicher am Ausgang erweitert.
- ❑ Die gelesenen Daten bleiben bis zum nächsten Aktivieren des CAS-Signal gültig.
- ❑ Der Prozessor kann Daten auslesen, während die Speichersteuerung eine neue Spaltenadresse an das DRAM übergibt.
- ❑ Durch dieses „Pipelining“ verkürzt sich die Wartezeit zwischen zwei aufeinander folgenden CAS-Impulsen → höherer Datendurchsatz

Schreibzugriffe bleiben wie bei FPM-Speicher unbeschleunigt und entsprechen von der Performance her den normalen Standard-DRAM.

Timinig-Diagramm eines EDO-DRAM



Eine neuer Lesevorgang kann beginnen, bevor der alte abgeschlossen ist.

EDO-RAM

Typische EDO-DRAM-Zugriffszeiten: 50 bis 70 ns CAS-Zugriffszeit (t_{CAC}) bei einem 60 ns EDO-DRAM-Baustein auf 25 ns (gegenüber 40 ns bei FPM-DRAM)

PC mit 66 MHz Bustakt (Taktperiode = 15 ns)

➔ Prozessor kann im EDO-Mode bei **jedem zweiten Takt** auf eine Zeile im Speicher zugreifen.

Initialisierung eines Lesevorgangs mit dem Anlegen von Zeilen- und Spaltenadresse dauert, wie bei FPM-DRAM fünf Takte.

Maximale Datentransferrate: 300 MByte/sec (bei 64-Bit Datenbus)

Performance-Steigerung von 50 %. In der Praxis fallen die Geschwindigkeitsgewinne wesentlich geringer aus (wenige Prozente).

Zusammenfassung: FPM, EDO

- ❑ Sie arbeiten asynchrone zum Systembus
- ❑ Für eine Datenübertragung ist ein Handshaking-Verfahren notwendig. Ein Lesevorgang läuft wie folgt:
 - Prozessor signalisiert der Speichersteuerung, dass eine Adresse anliegt.
 - Wenn die Daten am Ausgang des DRAMs bereitliegen, teilt die Speichersteuerung dem Prozessor dies mit (BRDY-Signal). Erst dann liest der Prozessor die Daten ein.
 - Dazwischen ist die CPU im Leerlauf und führt Wartezyklen aus.
- ❑ Varianten von EDO-DRAM (BEDO-DRAM) können die Daten ohne Wartezyklen liefern, aber nur bis zu einem Bustakt von 66 MHz.