

# **Paměti**

## **1 Základní údaje**

1.1 Rozdělení pamětí

1.2 Důležité parametry pamětí

1.3 Paměti paralelní a sériové

1.4 Vnitřní uspořádání paměti s adresovým výběrem

## **2 Paměti RAM**

2.1 Statické paměti RAM

2.2 Dynamické paměti RAM

2.3 Synchronní dynamické paměti RAM

## **3 Permanentní paměti**

3.1 Elektricky programovatelné paměti

3.2 Paměť NOR FLASH

3.3 Paměť NAND FLASH

## **4 Dvoubránová paměť**

## **5 Paměť fronty**

# Paměti

Účelem paměti je uložení velkého množství dat. Existuje mnoho fyzikálních jevů, na jejichž základě lze (někdy zatím jen teoreticky) zkonstruovat paměť. Významného rozšíření se však zatím dočkaly jen paměti polovodičové v technologii **CMOS** a paměti magnetické. V tomto textu budou rozebrány různé typy polovodičových pamětí. Vývoj pamětí vždy probíhal velmi intenzivně, takže všechny jejich podstatné parametry, jako je vybavovací doba a kapacita, se rychle mění. Aktuální hodnoty je třeba hledat ve firemních materiálech.

## 1. Rozdělení pamětí

Paměti lze rozdělit podle různých hledisek:

Podle způsobu adresace:

- Paměť **s adresovým výběrem** - data jsou uložena na stabilních adresách a dodání adresy je nutné pro operaci čtení i zápisu.
- Paměti **bezadresové** - zásobníková paměť LIFO (Last In - First Out), nebo paměť fronty FIFO (First In - First Out). Adresy jsou generovány automaticky vnitřními obvody v **pevně stanoveném pořadí**, a proto vnější adresa neexistuje.

Podle trvanlivosti dat po vypnutí napájení:

- Paměť **energeticky závislá** (volatile) - data se ztrácí po vypnutí napájení.
- Paměť **energeticky nezávislá** (non-volatile) - data se neztrácí po vypnutí napájení. Toto hledisko se vztahuje na **samotné** paměťové obvody. Celý paměťový systém totiž lze vybavit záložní baterií a pak i paměti energeticky závislé neztrácejí data.

Podle snadnosti a rychlosti zápisu:

- Paměť s možností čtení i zápisu **RWM** (angl. Read-Write Memory) - zápis i čtení jsou stejně snadné a rychlé. Paměť RWM je též (dokonce častěji) označována zkratkou **RAM** z (Random Access Memory), což doslova znamená „paměť s libovolným přístupem“. Tato zkratka je významově dosti posunutá, avšak všeobecně se ujala.
- Paměti **permanentní** - zápis nových dat buď není možný (jsou zapsána při výrobě), nebo je velmi pomalý vzhledem ke čtení.

## 2. Důležité parametry pamětí

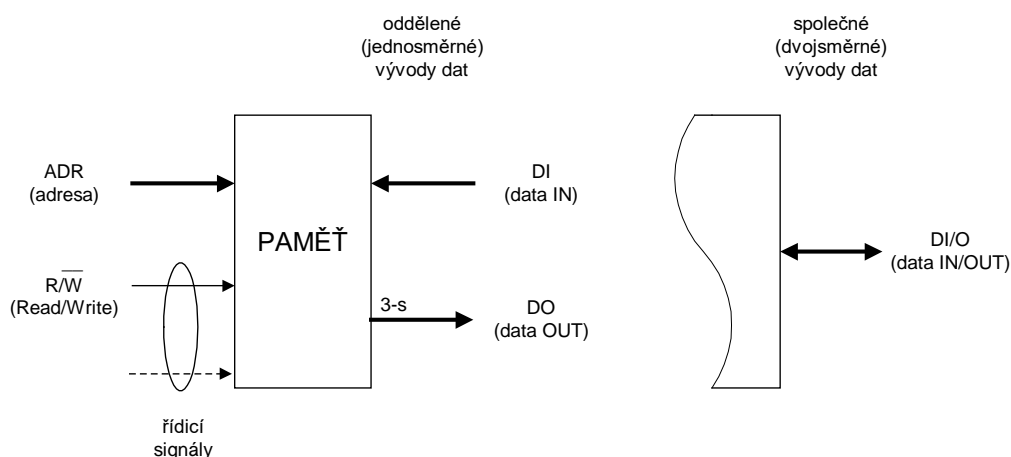
Důležitým údajem, charakterizujícím rozsah paměti, je její **kapacita**. Udává se v celkovém počtu bitů. Jednotkou je kilobit Kb ( $1K = 2^{10} = 1024$ ), megabit Mb ( $1M = 2^{20} = 1048576$ ), nebo gigabit Gb ( $1G = 2^{30}$ ). Kapacitu paměti lze udávat i v počtu slov, např. 256KB (zde „B“ znamená „byte“). Důležitá je i **organizace paměti**, tj. počet adres a délka slova - zpravidla 1 nebo 2 byte. Tak např. údaj 1M×8 znamená paměť o  $2^{20}$  adresách, na každé adrese je slovo o délce 1 byte. Celková kapacita by byla 8Mb, případně 1 MB..

Důležitým údajem, charakterizujícím rychlost paměti, je **vybavovací doba**, značena  $t_{AC}$  (access time). Uvádí se v nanosekundách ( $1ns = 10^{-9} s$ ). Je to doba od začátku čtecího cyklu do vydání platných dat. Začátkem cyklu může být změna adresy, ale podle typu paměti i

změna jiných signálů. Přesnou definici vybavovací doby je třeba nalézt v údajích výrobce. V závislosti na typu paměti existuje řada dalších dynamických parametrů.

### 3. Paměti paralelní a sériové

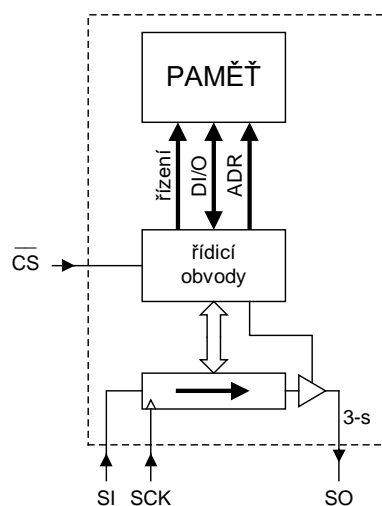
Ke spolupráci s dalšími obvody v číslicovém systému se u pamětí využívají typické soustavy signálů. Obr.1 ukazuje paralelní přístup, kdy adresa i data jsou vícebitové. Vstupní data (DI) a výstupní data (DO) mohou být rozdělená - jednosměrná, nebo častěji sdružená - dvojsměrná (DI/O). Soustava řídicích signálů určuje činnost paměti (čtení - zápis, atd.) a řídí směr přenosu dat. Podle typu paměti se soustavy řídicích signálů mohou vzájemně lišit.



Obr. 1: Paměť s paralelním přístupem

Vedle paralelního přístupu existuje i přístup sériový. Sériové paměti jsou využívány tam, kde je vhodné snížit počet spojů a pomalejší přenos není na závadu. Typickou podobu sériové paměti ukazuje obr. 2.

Jedná se o tzv. **třívodičové spojení** (výběrový signál  $\overline{CS}$  se nepočítá), zavedené pro počítačové periferie SPI. Do vnitřního posuvného registru se ze vstupu SI bit za bitem přenesou povel, kterým je dána následující operace paměti (čtení, zápis, atd.). Pak se přenesou adresa. Jedná-li se o zápis dat, přenesou se ještě zapisovaná data. Řídicí obvody provedou požadovanou operaci s vnitřní pamětí. Jedná-li se o operaci čtení, jsou ihned po zadání adresy vydávána sériová data na výstupu SO. Výstupní třístavový obvod je ovládán řídicími obvody tak, že kromě operace čtení je vždy ve vysokoimpedančním stavu.



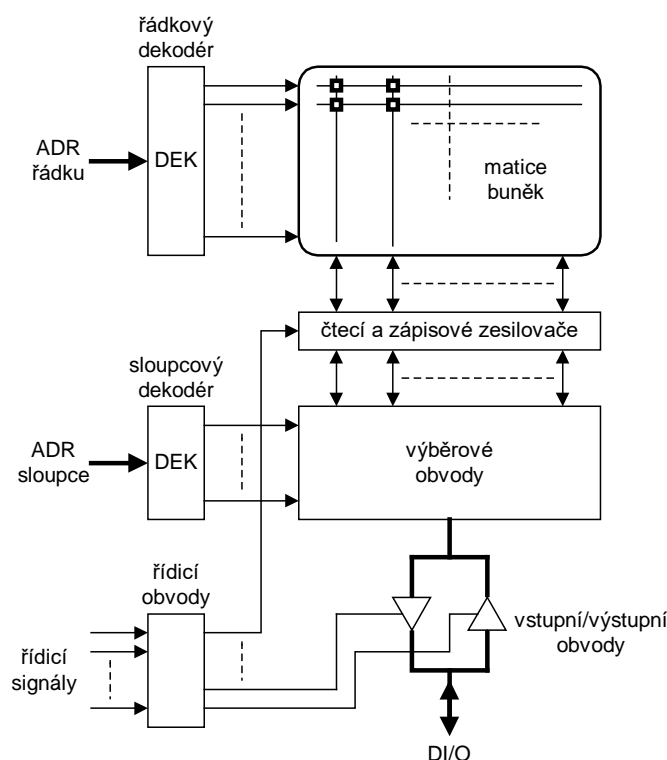
Obr. 2: Paměť se sériovým přístupem

Hodinové impulsy pro posuvný registr musí být generovány vnějšími obvody. Sériová paměť má pouzdro s malým počtem vývodů. Nevýhodou je však zpoždění způsobené přenosy po jednotlivých bitech. Vedle třívodičového spojení existuje i **spojení dvouodičové**, zavedené

pro počítačové periferie IIC. U obvodů tohoto typu se nepoužívá výběrový signál a sériová data jsou dvojsměrná.

#### 4. Vnitřní uspořádání paměti s adresovým výběrem

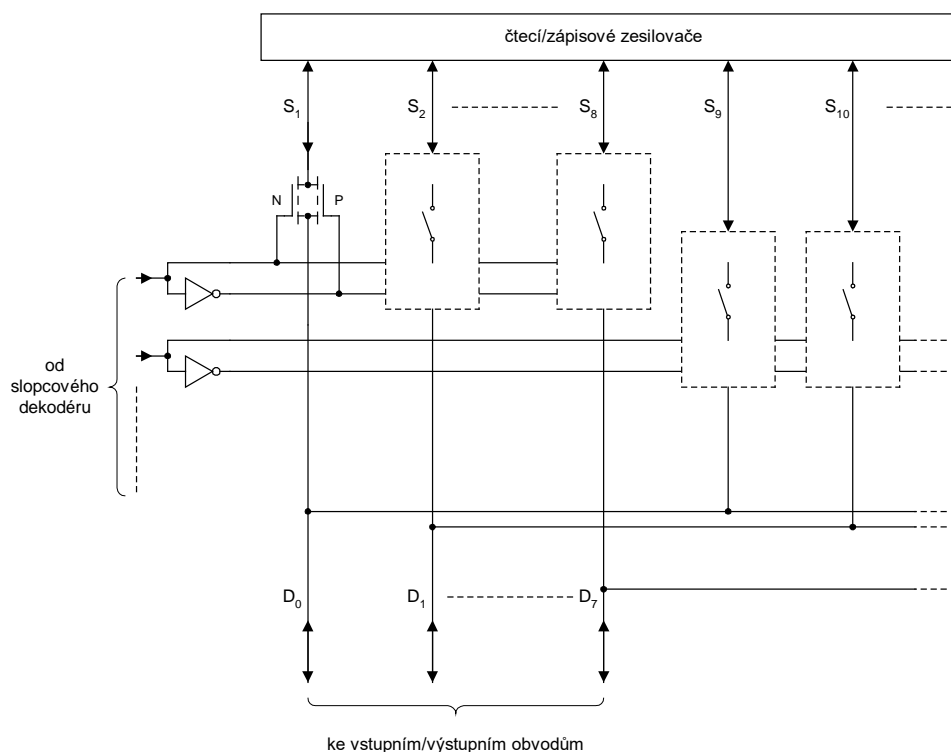
Obr. 3 ukazuje vnitřní obvody paměti bez detailů. Jádrem je **matice paměťových buněk**, umístěných v křížení řádkových a sloupcových vodičů. Jedna část bitů adresy vybírá prostřednictvím **řádkového dekodéru** vždy jeden z řádků matice. Tím jsou vybrány všechny buňky daného řádku a připojeny na sloupcové vodiče. Druhá část bitů adresy je zavedena do **sloupcového dekodéru**. Ten ovládá **výběrové obvody**, což je skupina CMOS spínačů zapojených jako multiplexer/demultiplexer. Na výběrové obvody pak navazují **vstupní a výstupní obvody**. Čtecí a zápisové **zesilovače** jsou nutné pro zesílení signálu z paměťových buněk a u dynamických pamětí i pro obnovování obsahu buněk. V anglické literatuře jsou řádkové vodiče nazývány "**word lines**" a sloupcové "**bit lines**".



Obr. 3: Vnitřní uspořádání paměti s adresovým výběrem

Řídící signály určují pracovní cyklus paměti (čtení - zápis, apod.), ovládají třístavové výstupní obvody, vyvolávají úsporný režim sníženého příkonu, atd. Soustava řídících signálů je charakteristická vždy pro daný typ paměti a bude vysvětlena v dalším textu.

Spínače ve výběrových obvodech jsou tak seskupeny, že sloupcový dekodér vybírá celé skupiny sloupců - tomu odpovídá délka slova paměti. Obr. 4 ukazuje princip výběrových obvodů pro 8-bitové slovo. Výběrové obvody slouží pro oba směry přesunu dat, tj. pro čtení i pro zápis. Jedná se vlastně o skupinový multiplexer/demultiplexer.

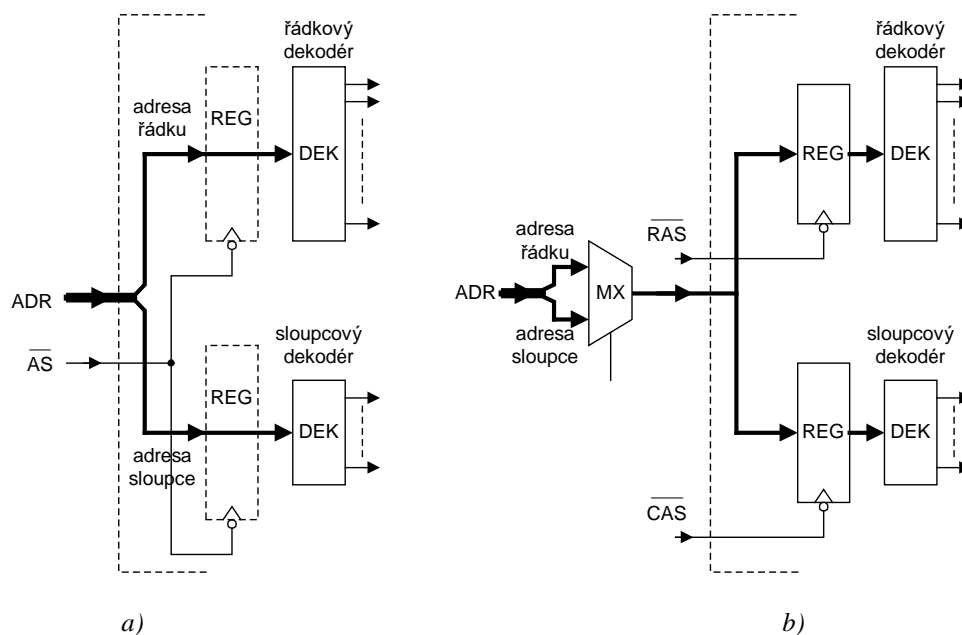


Obr. 4: Princip výběrových obvodů sloupců

Zavedení adres do řádkového a sloupcového dekodéru je v obr. 3 jen naznačeno a u konkrétních typů pamětí se liší. V podstatě se využívají tři možnosti - viz obr. 5. Na obrázku vlevo je nejjednodušší a nejčastější možnost - do paměti (její hranice je naznačena čárkovaně) je přiváděna **současně celá adresa** a až uvnitř se rozvětjuje na dvě složky, které jsou zavedeny do příslušných dekodérů. To zajišťuje nejmenší zpoždění. Typicky je tento způsob zavádění adresy využíván u **statických pamětí SRAM** a u **permanentních pamětí** (viz další text). Jistou nevýhodou tohoto řešení je velký počet adresových vývodů pouzdra u pamětí větší kapacity.

Jako druhá možnost jsou před dekodéry předřazeny registry (v obrázku 5 vlevo čárkovaně vyznačeny), do kterých se adresa zapíše impulsem  $\overline{AS}$  (Address Strobe). Někteří výrobci tak zjednodušují konstrukci paměťového systému v systémech s multiplexní sběrnici.

Třetí možnost je naznačena v pravé polovině obrázku 5, kdy adresa je do paměti zavedena **postupně ve dvou krocích**. Paměť je vybavena dvěma registry, ovládanými individuálně. Nejprve se první část adresy (adresa řádku) zapíše do řádkového registru impulsem  $\overline{RAS}$  (Row Address Strobe), pak se druhá část (adresa sloupce) zapíše do sloupcového registru impulsem  $\overline{CAS}$  (Column Address Strobe). Obě složky adresy je nutné ve vhodném čase přepínat v obvodech mimo vlastní paměť (v adresovém multiplexeru). Celkové zpoždění, vzniklé postupným zápisem do adresových registrů, lze tolerovat tehdy, když zavádění druhé části adresy (tj. adresy sloupce) probíhá v době, kdy v paměti probíhají takové vnitřní děje, které ještě adresu sloupce nepotřebují.



Obr. 5: Možnosti zavedení adresy do paměti - a) jako celek, b) postupně

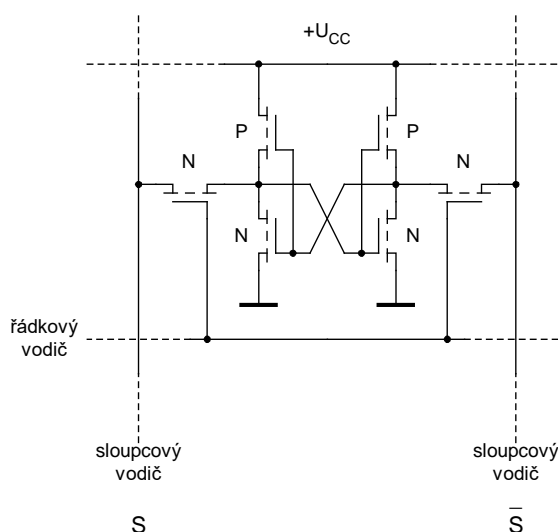
Typicky je tento způsob zavádění adresy využíván u **dynamických pamětí DRAM** o velké kapacitě (viz další text). Pro zavedení adresy ve dvou krocích je zapotřebí jen poloviční počet vývodů pouzdra, což snižuje cenu integrovaných paměťových obvodů.

## 2 Paměti RAM

### 2.1 Statické paměti RAM

Ve statických pamětech RAM (SRAM) se k zapamatování číslicových (dvojestavových) signálů využívají **klopné obvody**. Ten lze vnějším signálem během zápisu ovládat (překlápět) a naopak jeho výstupní signál lze během čtení dalšími obvody zpracovávat. Zapamatovaný bit zůstane v klopném obvodu uložen do té doby, než je eventuálně přepsán, nebo do vypnutí napájení.

Celkové uspořádání odpovídá obr. 3, 4 a 5 a), adresa je do paměti zaváděna jako celek a zpravidla bez registrů. Tím je minimalizováno zpoždění. Paměťová matice je složená z šestitransistorových buněk dle obr. 6. Klopný obvod je implementován nejjednodušším způsobem jako dva invertory CMOS (s tranzistory NMOS a PMOS). Používají se i čtyřtranzistorové buňky, kde jsou tranzistory PMOS nahrazeny rezistory o extrémně vysoké hodnotě. Klopné obvody jsou dalšími dvěma tranzistory, řízenými řádkovým vodičem, připojovány na dva sloupce. Rozdíl napětí mezi oběma sloupci je při čtení vyhodnocen rozdílovým čtecím/zápisovým zesilovačem - viz obr. 3. Při zápisu je naopak tímto zesilovačem, přes dvojici sloupcových vodičů a otevřené spínače, klopný obvod překlápen do patřičného stavu. Uspořádání se dvěma vodiči na sloupec sice komplikuje konstrukci paměťové matice, ale současně podstatně snižuje vzájemné rušení mezi sousedními sloupci.



Obr.6: Buňka statické paměti RAM

Soustava řídicích signálů SRAM je velmi jednoduchá:

- **výběrový signál**  $\overline{CS}$  (Chip Select) stavem 1 blokuje zápis do paměti, uvádí výstupy do vysokoimpedančního stavu a omezuje příkon paměti v době, kdy je nevyužitá. Příkon se omezí tím, že se sníží vnitřní napájecí napětí matice na hodnotu, při které ještě klopné obvody spolehlivě udrží zapsaná data. Při stavu 0 se naopak paměť uvede do provozu - to však vyžaduje relativně dlouhou dobu. Pro rychlé připojování a odpojování výstupů se tento postup nehodí.
- signál  $\overline{OE}$  (Output Enable) působí jen na výstupní obvody - stavem 1 je uvádí do vysokoimpedančního stavu a stavem 0 je připojuje. V obou případech ve velmi krátké době. Lze jej přibližně charakterizovat jako „**čtecí impuls**“.
- signál  $\overline{WR}$  (Write) nebo též  $\overline{WE}$  (Write Enable) rozhoduje o operaci zápisu (stavem 0) nebo čtení (stavem 1). Lze jej přibližně charakterizovat jako „**zápisový impuls**“. Při zápisu jsou výstupní obvody blokovány bez ohledu na  $\overline{OE}$ .

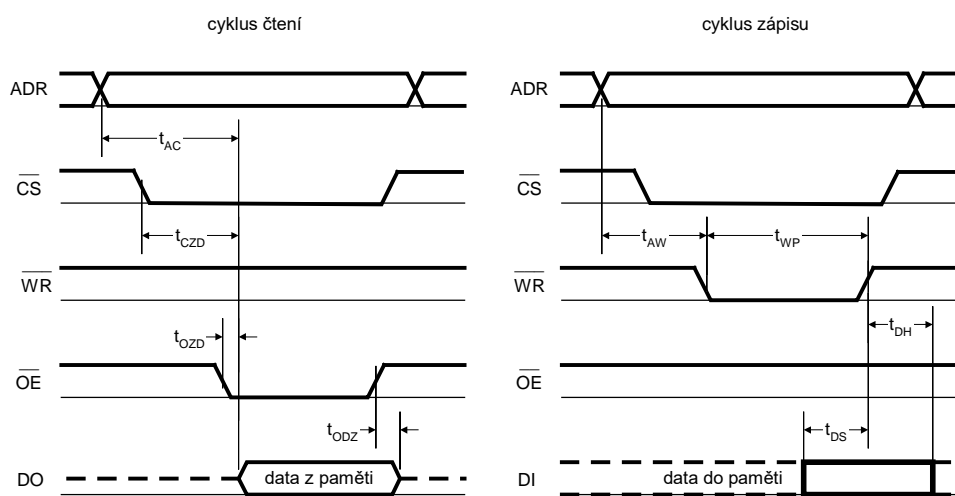
Tabulka Tab.12.1. shrnuje funkci uvedených řídicích signálů (– značí libovolný stav).

$\overline{CS}$	$\overline{OE}$	$\overline{WR}$	operace	DI/O	napájecí proud
H	–	–	žádná	vysoká impedance	snížený
L	L	H	čtení	DO	plný
L	–	L	zápis	DI	plný
L	H	H	žádná	vysoká impedance	plný

Tab. 1: Význam řídicích signálů SRAM

Časový diagram a hlavní dynamické parametry ukazuje obr. 7. V cyklu čtení je podstatná zvláště **vybavovací doba**, doba zablokování od  $\overline{CS}$  ( $t_{CZD}$ ), a podstatně kratší doba zablokování od  $\overline{OE}$  ( $t_{OZD}$ ) a odblokování od  $\overline{OE}$  ( $t_{ODZ}$ ). V cyklu zápisu je podstatná doba odstupu zápisového impulsu od změny adresy ( $t_{AW}$ ), minimální délka zápisového impulsu ( $t_{WP}$ ), a předstih a přesah zapisovaných dat kolem konce zápisového impulsu ( $t_{DS}$  a  $t_{DH}$ ). Interval, po který musí být zapisovaná data stabilní, je orámován tlustou čarou.

Signál  $\overline{CS}$  se nemusí měnit v každém cyklu - může být trvale ve stavu 0. Činnost SRAM se tím urychlí, neboť odpadá zpoždění  $t_{CZD}$ . Paměť pak ale má větší spotřebu, neboť odpadají doby nečinnosti se sníženým příkonem.



Obr. 7: Průběh cyklu čtení (vlevo) a cyklu zápisu (vpravo) paměti SRAM

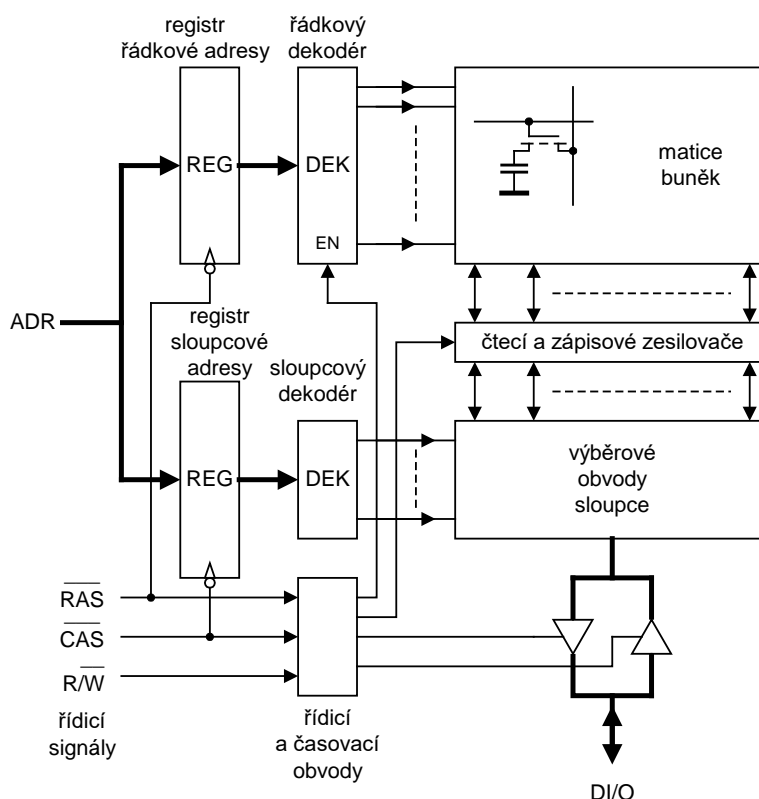
Paměti SRAM jsou velmi rychlé, vybavovací doby pod 10 ns nejsou výjimečné.

## 2.2 Dynamické paměti RAM

V dynamických pamětech RAM (DRAM) se k zapamatování číslcových (dvojstavových) signálů využívá **náboj na kondenzátoru**. Kondenzátor, nabitý či vybitý přes jednoduchý spínač (MOSFET) během zápisu, udrží náboj po jistou dobu. Napětí na kondenzátoru lze během čtení porovnat s referenčním napětím a rozlišit tak stav 0 a 1. Tím je realizována paměť na jeden bit informace. Napětí na paměťovém kondenzátoru se ale časem mění. Aby nedošlo ke ztrátě informace, je nutno náboj dostatečně často **obnovovat**.

Celkové uspořádání paměti DRAM ukazuje obr. 8. Adresa se do paměti zavádí ve dvou krocích. Nejprve se adresa řádku zapíše do řádkového registru impulsem  $\overline{RAS}$ , pak se adresa sloupce zapíše do sloupcového registru impulsem  $\overline{CAS}$ . Paměťová matice obsahuje buňky v nejjednodušší podobě. Na stejné ploše čipu proto má DRAM větší kapacitu, než by měla SRAM se složitější šestitransistorovou buňkou a dvojicí sloupců. Tomu odpovídá i **nižší cena za jeden bit** u paměti DRAM.



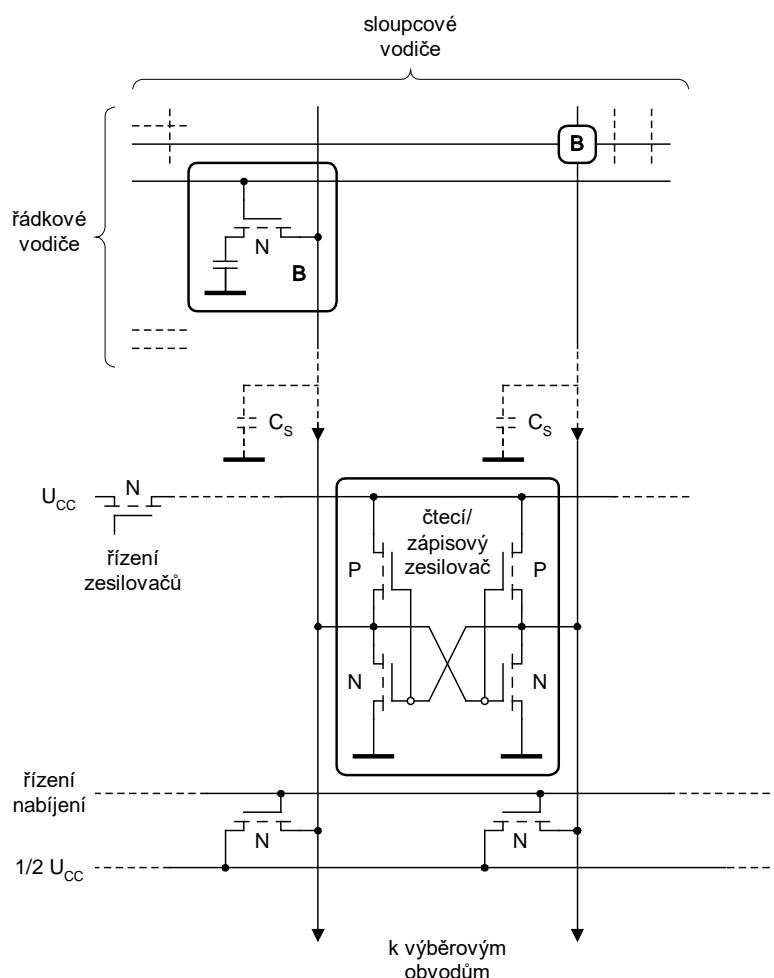


Obr.8: Vnitřní uspořádání paměti DRAM

Problém však představuje čtení stavu buňky. Paměťový kondenzátor totiž má velmi malou kapacitu, řádově desítky  $fF$  (femtofarad,  $10^{-15}$  F). Naproti tomu kapacita sloupcového vodiče proti zemi je mnohem větší. Připojením paměťového kondenzátoru na sloupcový vodič tak vzniká kapacitní dělič, který snižuje napěťový skok úměrně poměru kapacit - tedy velmi podstatně.

Výrobci DRAM vyvinuli několik vtipných řešení čtecích zesilovačů. Velmi rozšířená varianta je na obr. 9. Čtecí zesilovač je konstruován jako klopný obvod, uvedený ve vhodném okamžiku do rovnovážného stavu (tj. shodného napětí na obou výstupech). Shodné napětí se zajistí nabitím sloupcových vodičů o kapacitě  $C_S$  na referenční napětí o velikosti přibližně poloviny napájecího napětí  $U_{CC}$  přes řízené spínače - tzv. **přípravné nabíjení** (precharge). Pak se na čtecí zesilovače přivede napájecí napětí ( $U_{CC}$ ). V důsledku kladné zpětné vazby závisí překlopení KO i na nepatrné nesymetrii obou napětí. Nesymetrie je vyvolána tak, že se na jeden z obou sloupcových vodičů připojí paměťová buňka se svým paměťovým kondenzátorem, zatím co na druhý ne. Ten vlastně slouží v daném okamžiku jako zdroj referenčního napětí. Po úplném překlopení je klopný obvod již zdrojem napětí o maximální amplitudě ( $U_L = 0$ ,  $U_H = U_{CC}$ ) pro sloupcový vodič. Přes stále ještě otevřený spínač v buňce se její paměťový kondenzátor nabije (či vybije) na krajní hodnoty napětí. Informace v buňce je tak obnovena. Buňka na referenčním sloupci je zdánlivě nevyužitá - bude však využita při jiné řádkové adrese, kdy se funkce sloupců obrátí. Střídavé využívání sloupců jako paměťových a jako referenčních je zajištěno vhodným spojením buněk s řádky a sloupci

matice - buňky na lichých řádcích jsou spojeny s lichými sloupci a buňky na sudých řádcích se sudými sloupci. Čtecí zesilovače obsluhují vždy pár sloupců „sudý-lichý“.

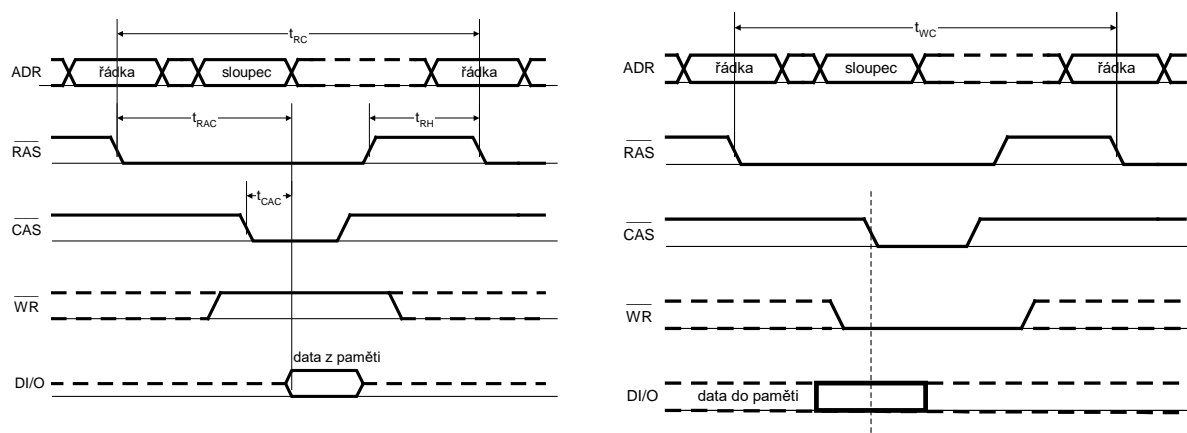


Obr. .9: Funkce čtecích zesilovačů DRAM

Uvedené uspořádání umožňuje čtení a současně i obnovení informace vždy **na všech buňkách řádku**. Výsledek čtení zůstává uložen ve čtecích zesilovačích. Na ně navazují výběrové obvody a vstupní/výstupní obvody podle obrázků 3 a 4.

Dění v paměti DRAM je podstatně složitější, než v paměti SRAM. Čtecí cyklus je zahájen hranou (doběžnou) signálu  $\overline{RAS}$  - viz obr. 8. Tím je zapsána řádková adresa do řádkového registru a současně je ukončeno nabíjení sloupců, které bylo dosud aktivní. Do čtecích zesilovačů je zavedeno napájecí napětí. Následně je řádkovým dekodérem vybrána řádka a podle stavu buněk jsou překlopeny čtecí zesilovače. To je první fáze celého čtecího cyklu. Pak je přivedena adresa sloupce a hranou (doběžnou) signálu  $\overline{CAS}$  je zapsána do sloupcového registru. Sloupcový dekodér s výběrovými obvody vybere skupinu čtecích zesilovačů (slovo z paměti). Následně jsou odblokovány výstupní třístavové obvody a zapamatovaná informace se dostává na výstupy paměti. S návratem signálu  $\overline{CAS}$  do stavu 1 se výstupní obvody blokují (vedou se do vysokoimpedančního stavu). S návratem signálu  $\overline{RAS}$  do stavu 1 se blokuje řádkový dekodér a začnou se nabíjet sloupcové vodiče, takže paměť je připravena k dalšímu cyklu. Posloupnost vhodně časovaných řídicích signálů je

generována vnitřními řídicími a časovacími obvody. Celý průběh čtecího cyklu ukazuje obr. 10 vlevo.



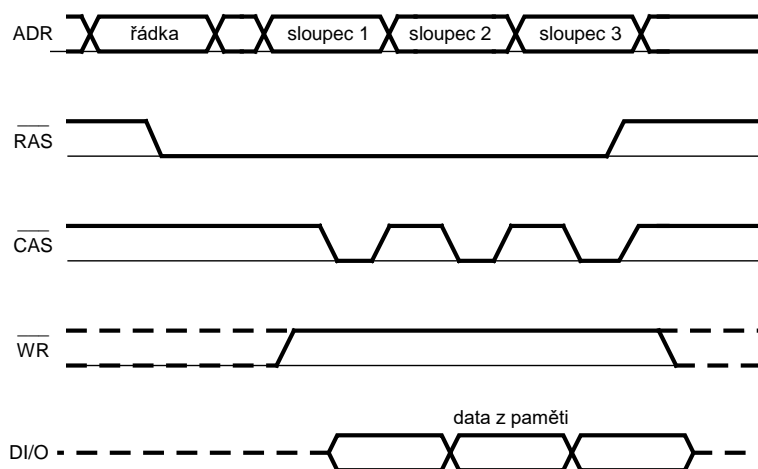
Obr. 10: Průběh čtecího (vlevo) a zápisového (vpravo) cyklu DRAM

Obrázek ukazuje důležité doby, jako je  $t_{RAC}$  (vybavovací doba od hrany  $\overline{RAS}$ ),  $t_{CAC}$  (vybavovací doba od hrany  $\overline{CAS}$ ),  $t_{RH}$  (doba nabíjení sloupců), a  $t_{RC}$  (doba čtecího cyklu). Zápisový povel  $\overline{WR}$  musí být neaktivní, tj. ve stavu 1, kolem doběžné hrany  $\overline{CAS}$ . Čtená data jsou přítomna na výstupech s malým zpožděním po dobu, kdy  $\overline{CAS} = 0$ , tedy jen po zlomek doby cyklu. Rovněž celá vybavovací doba  $t_{RAC}$  je jen částí doby cyklu. Běžně se  $t_{RAC}$  pohybuje kolem 50 ns,  $t_{RC}$  pak je kolm 100 ns. Průběh zápisového cyklu ukazuje obr. 10 vpravo. Zápisový povel  $\overline{WR}$  musí být aktivní (ve stavu 0) kolem doběžné hrany  $\overline{CAS}$ . Interval, po který musí být zapisovaná data stabilní, je orámován tlustou čarou.

Z principu paměťové buňky dynamické paměti vyplývá nutnost **obnovování informace** (refresh). Interval obnovování, tj. doba udržení neporušených dat závisí na teplotě čipu, na typu paměti a na technologii výroby. Při teplotě kolem 20 °C to mohou být i jednotky sekund - ale bez záruky. Katalogové údaje jsou podstatně opatrnější - řádově jednotky až desítky milisekund. K obnovení informace na celé řádce dojde automaticky při čtení či zápisu s adresou dotyčné řádky. Během intervalu obnovování je však nutné obnovit informaci na všech řádkách matice. Do normální činnosti paměti se proto vkládají obnovovací cykly. Jsou to zkrácené cykly, omezené jen na vložení adresy řádku, doprovázené impulsem  $\overline{RAS}$ . Tím je vyvolána činnost čtecích zesilovačů a tím i obnovení obsahu buněk daného řádku. Vkládání adresy sloupce je zbytečné, neboť žádný výstup dat z paměti není v obnovovacím cyklu předpokládán.

K obnovení celého obsahu paměti je třeba postupně vystřídat všechny adresy řádků. Tyto adresy jsou generovány **čítačem pro obnovování**, který je u moderních DRAM součástí paměti. Z vnějšího pohledu jsou pak obnovovací cykly skryté, automaticky generované. Paměti tohoto typu byly označovány jako "pseudostatické". Úplné skrytí obnovovacích cyklů však není možné, neboť během nich je paměť **nedostupná** pro běžné používání ve čtecích a zápisových cyklech.

Z obr. 10 je vidět výrazný rozdíl mezi vybavovací dobou paměti a dobou cyklu. Doba cyklu omezuje frekvenci opakovaných operací a vybavovací doba je špatně využita. Dynamické paměti umožňují za jistých podmínek podstatné zrychlení opakovaných operací. Je to **čtení po stránkách** (page mode), případně zápis po stránkách. Využívá se skutečnost, že po zadání adresy řádky a impulsu  $\overline{RAS}$  se do čtecích zesilovačů uloží obsah celé řádky. Pak stačí opakovaně přivádět postupně zvyšované adresy sloupce a doprovázet je jen impulsy  $\overline{CAS}$ , aniž by se měnila adresa řádku - viz obr. 11. V tomto režimu je trvání dat na výstupu prodlouženo vždy až do doběžné hrany  $\overline{CAS}$  a systém navazující na paměť tak má data přístupná po celou periodu impulsů CAS. Výstup se uvede do vysokoimpedančního stavu po návratu signálů  $\overline{RAS}$  a  $\overline{CAS}$  do stavu 1. Přiměřeně platí totéž pro zápis po stránkách.



Obr. 11: Čtení po stránkách

Ke zjednodušení manipulace s DRAM v režimu práce po stránkách jsou moderní paměti vybaveny **čítačem adres sloupců**, který generuje postupně zvyšované adresy. Je možné předem naprogramovat počáteční adresu a délku stránky. Samozřejmě stále existuje možnost práce po jednotlivých adresách jak řádky, tak sloupce.

Dalšího zrychlení činnosti se dosáhlo u pamětí **DDR** (double data rate), kde k vydávání dat dochází na **obě hrany**  $\overline{CAS}$ , tedy dvojnásobnou rychlostí, s periodou řádu nanosekund. Využití těchto parametrů ovšem vyžaduje velmi přesné časování řídicích signálů (RAS, CAS, WR).

## 2.3 Synchronní dynamické paměti RAM

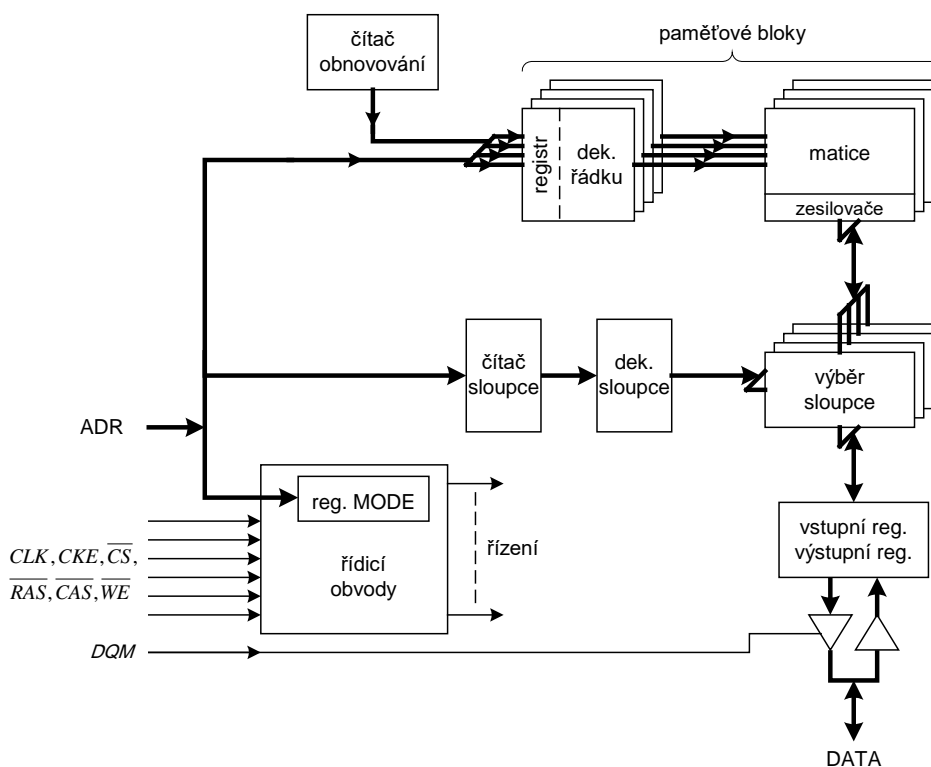
Podstatného zrychlení činnosti dynamických pamětí se dosáhlo zavedením synchronních dynamických pamětí. Synchronní dynamické paměti RAM, zkratkou **SDRAM**, využívají stejný princip paměťových buněk, čtecích zesilovačů a dekódování adres, jako běžné DRAM. Jsou však vybaveny dalšími vnitřními obvody, které podstatně urychlují čtení a zápis dat a zjednodušují časování signálů pro paměť.

Podstatnou změnou je zavedení **synchronizačních** (hodinových) **impulsů**. Všechny vstupní signály - adresy, řídicí signály, zapisovaná data - jsou hranou hodinového impulsu

vzorkovány, tj. zapsány do vnitřních registrů. Vnitřní signály se tak mění v přesně definovaných okamžicích a naopak na časování vnějších signálů jsou kladeny podstatně mírnější nároky, než u jednoduché DRAM. I výstupní data z paměti se mění v závislosti na hodinových impulsích. Celá činnost SDRAM tak může být synchronizována s dalšími obvody v systému (např. s procesorem v počítači).

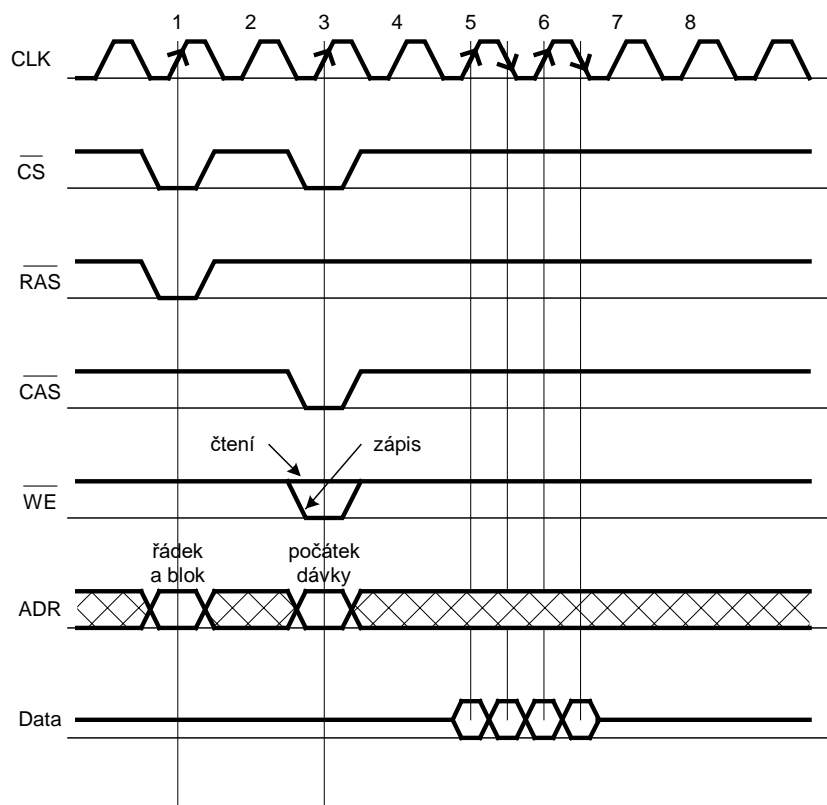
Zrychlení činnosti SDRAM je dosaženo obdobným principem, jako u výše popsané paměti DRAM. Paměť je optimalizována pro režim čtení/zápisu po stránkách, neboli **dávkového** čtení/zápisu (burst read/write), kdy vnitřní čítač adres sloupců je inkrementován hodinovými impulsy.

Obrázek 12 ukazuje vnitřní uspořádání SDRAM. Hlavní částí jsou paměťové bloky (nejčastěji čtyři) sestávající z hlavních obvodů paměti DRAM - je to řádkový registr a dekodér, paměťová matice se zesilovači, a výběr sloupce. Ostatní obvody jsou společné pro celou SDRAM. Každý blok se může nacházet v **různých režimech**: přípravné nabíjení, obnovování obsahu, čtení/zápis dat. To přispívá ke zrychlení činnosti paměti, neboť zatím co v jednom bloku probíhá operace čtení nebo zápisu, v jiném může probíhat přípravné nabíjení a v ještě jiném obnovování.



Obr. 12: Vnitřní uspořádání paměti SDRAM

Soustava řídicích signálů a hlavně jejich funkce se od DRAM liší. Některé jsou obdobné, ale některé jsou zcela nové - **CLK, CKE, CS**. Význam synchronizačních impulsů již byl vysvětlen. Základní cyklus čtení či zápisu ukazuje obr. 13.



Obr. 13: Průběh čtecího a zápisového cyklu paměti SDRAM, DDR

Zpoždění výdeje prvních dat za  $\overline{CAS}$  se nazývá **doba latence** paměti. Počet vydaných slov neboli **délka dávky** je naprogramován v řídicích obvodech paměti. Sloupcová adresa je přitom automaticky generována sloupcovým čítačem. Při zápisu dat musí být naopak data ve stejných okamžicích dodávána do paměti - o automatickém zvyšování sloupcové adresy a o délce dávky platí totéž.

Veškeré časové parametry, uváděné u DRAM v nanosekundách, jsou v případě SDRAM definovány velmi jednoduše, jen v **počtu taktů CLK**. Jsou to:

- maximální kmitočet  $CLK$
- odstup  $\overline{CAS}$  po  $\overline{RAS}$
- doba latence po  $\overline{CAS}$
- délka dávky - je v jistých mezích programovatelná
- doba přípravného nabíjení

Dynamické parametry SDRAM se specifikují jako **trojice čísel** v pořadí:

- latence
- odstup  $\overline{CAS}$  za  $\overline{RAS}$
- doba přípravného nabíjení

Dalším parametrem je maximální kmitočet  $CLK$ . Příkladem specifikace může být paměť:

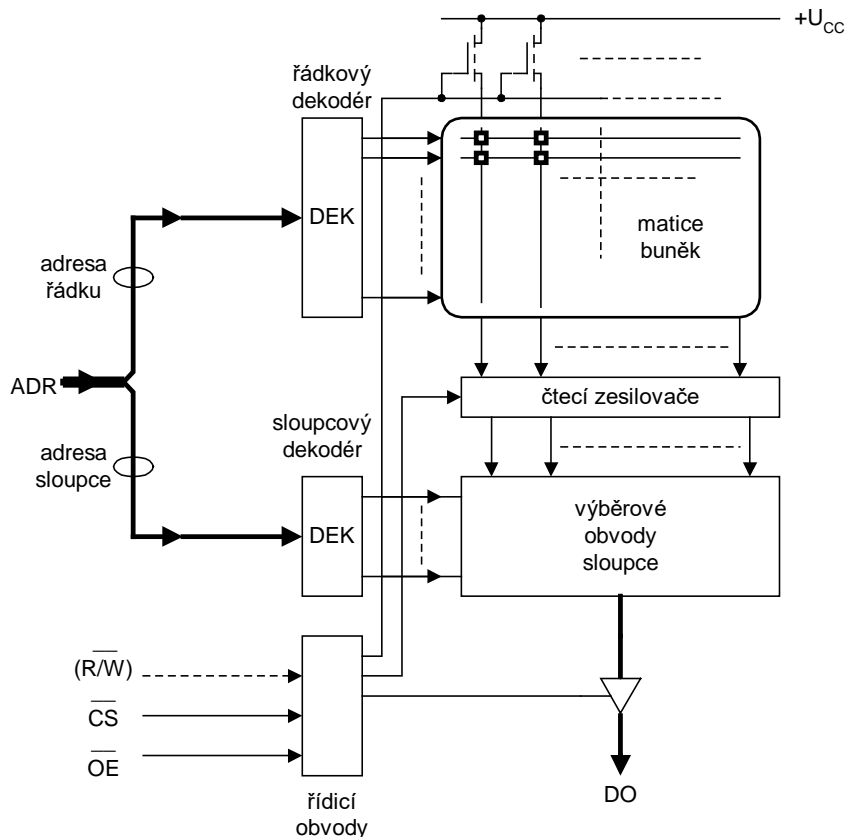
256 Mb, 32 M×8, 133 MHz, **3-2-2**.

### 3 Permanentní paměti

Podstatnou vlastností každé permanentní paměti je to, že je energeticky nezávislá - uložená data se neztrácejí po vypnutí napájení. Existuje několik typů permanentních pamětí:

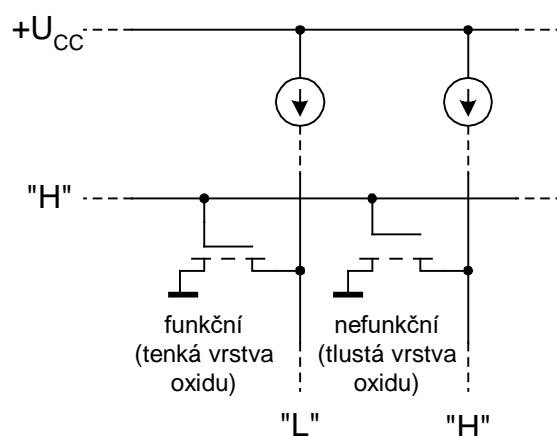
- Paměť **ROM** (Read-Only Memory) - obsah je dán při výrobě a nelze jej již změnit.
- Paměť **PROM** (Programmable ROM) - obsah může uživatel zapsat (naprogramovat), ale naprogramování je nevratné.
- Paměť **EPROM** (Erasable PROM) - paměť programovatelná a vymazatelná. Je programována elektricky, vymazána ultrafialovým zářením.
- Paměť **EEPROM** (Electrically Erasable PROM) - paměť programována elektricky na libovolné adrese bez předchozího vymazání. Spíše než o programování a mazání zde lze mluvit o zápisu dat na adrese, obdobně jako u paměti RAM. Zápis je však mnohem pomalejší než čtení.
- paměť **FLASH** - paměť programovatelná a vymazatelná. Je programována elektricky na libovolné adrese, ale elektricky mazána jako celek (nebo po velkých částech - sektorech či blocích). Před programováním musí být vymazána.

Vnitřní uspořádání permanentní paměti ukazuje obr. 14. Jsou zde vyznačeny jen cesty dat pro čtení - vnitřní signály potřebné pro naprogramování obsahu se běžně neuvádějí. Pro vložení dat se vždy využívají vývody DO, pro jejich zápis do matice buněk slouží patřičné kombinace řídicích signálů.



Obr. 14: Vnitřní uspořádání permanentní paměti

Provedení buněk paměťové matice odpovídá typu paměti. U paměti ROM má buňka nejjednodušší podobu jako tranzistor MOS, zapojený mezi sloupec a zem. V každé buňce existuje tranzistor, ale jen některé z nich jsou funkční. Funkčnost tranzistoru se ovlivní tak, že u tranzistorů je mezi řídicí elektrodou (hradlem) a polovodičem buď tenká, nebo tlustá vrstva oxidu křemíku. Po vybuzení řádkového vodiče dojde u funkčních tranzistorů k vytvoření indukovaného kanálu a tím ke zkratování sloupce na zem, zatímco u tranzistorů s tlustou vrstvou oxidu je vzdálenost hradla od polovodiče tak velká, že nedojde k vytvoření vodivého kanálu a napětí na sloupci zůstane vysoké. Princip je ilustrován obrázkem 15. Takto se dosahuje velmi vysoké hustoty integrace, ovšem zapsaná data jsou neměnná. Paměť tohoto typu se nazývá **maskou programovaná**.



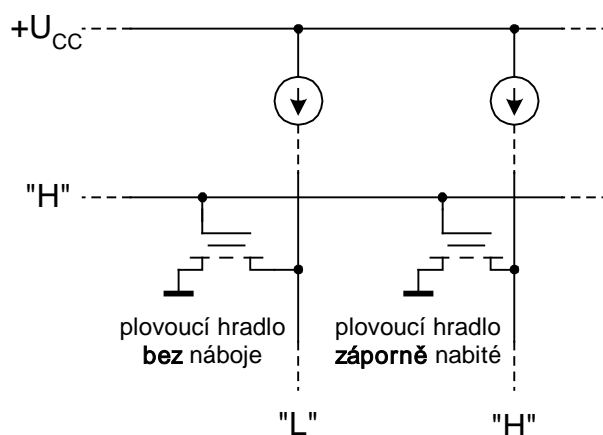
Obr. 15: Buňky paměti ROM maskou programované

Programovatelné **paměti PROM** byly v minulosti realizovány pomocí buněk s přepalovacími spojkami. Tato technologie vyžadovala velké proudy pro přepálení spojek a byla realizovatelná jen v bipolární technologii. Zanikla po nástupu technologie CMOS.

### 3.1 Elektricky programovatelné paměti

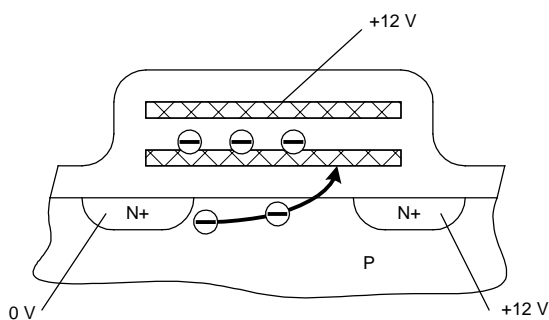
V technologii CMOS se využívají buňky, u kterých se mezi řídicím hradlem a polovodičem vytvoří další vodivá vrstva, totálně odizolovaná oxidem křemíku - tzv. **plovoucí hradlo** (floating gate). Na plovoucí hradlo lze dopravit elektrony a nabít je tak záporně. Takto nabitě hradlo vytváří elektrické pole opačného směru než je pole hradla u NMOS tranzistoru a tudíž brání vytvoření vodivého kanálu - na sloupcovém vodiči zůstane napětí  $U_H$ . Naopak při stavu bez náboje se vytvoří indukovaný kanál, sloupec je zkratován na zem a je na něm úroveň  $L$ . Jinými slovy, náboj na plovoucím hradle ovlivňuje prahové napětí MOS tranzistoru. Princip je ilustrován obrázkem 16.





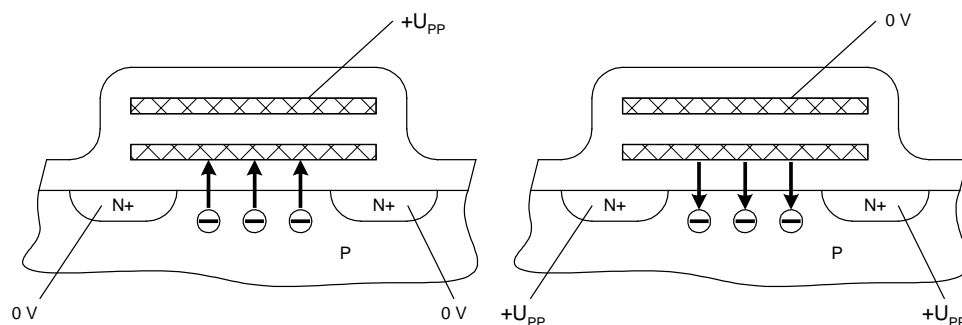
Obr. 16: Buňky paměti programované nábojem

Způsob přivedení náboje na plovoucí hradlo a jeho odvedení je v podstatě dvojitý. Z něj pak vyplývá i způsob naprogramování (případně mazání) buněk v matici. Za prvé je to **lavinová injekce** elektronů. Při ní jsou elektrony ve vodivém kanálu urychleny elektrickým polem mezi konci kanálu a je jim dodána energie dostatečná k vytvoření omezeného lavinového průrazu. Současně je jejich dráha vychýlena vlivem kladného napětí na řídicím hradle. Urychlené elektrony („hot electrons“) nabydou dostatečné energie k průletu tenkou vrstvou oxidu pod plovoucím hradlem. Tento mechanismus injekce „horkých“ elektronů je velmi účinný a umožňuje nabít hradlo záporně během krátké doby, řádově mikrosekund. Technologie umožňující tuto funkci má název FAMOS (Floating-gate Avalanche-injection MOS). Obrázek 17 ukazuje princip tranzistoru FAMOS. Tlustá šipka značí dráhu elektronů.



Obr. 17: Řez tranzistorem FAMOS

Druhý způsob využívá kvantový jev - přemístění elektronu přes vrstvu izolantu při velmi vysoké intenzitě elektrického pole, známý jako Fowler-Nordheimův mechanismus **tunelování elektronů**. Aby nebylo nutné používat vysoké napětí, musí být vrstva izolantu velmi tenká. Při vrstvě oxidu křemíku o tloušťce kolem 10 nm postačuje napětí do 20 V. Opět se využívají dvě hradla - řídicí a plovoucí (totálně odizolované) - viz obr. 18. Programovací napětí je v něm označeno jako  $+U_{PP}$ . Obrázek vlevo ukazuje uspořádání pro nabíjení plovoucího hradla záporným napětím, obrázek vpravo pak vybití náboje plovoucího hradla. Manipulací s napětím tedy lze měnit stav paměťové buňky. Celý mechanismus však je relativně pomalý - změna stavu buňky trvá řádově milisekundu.



Obr. 18: Uspořádání pro nabití a vybití plovoucího hradla

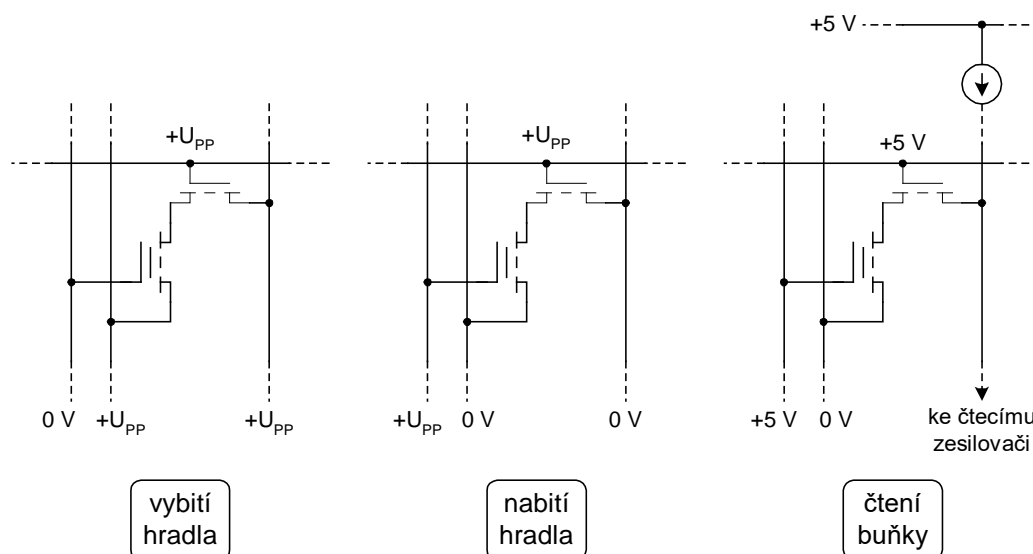
Zbývá vysvětlit, jak se při programování dostane patřičné napětí na totálně odizolované plovoucí hradlo. Je to kapacitní vazbou mezi oběma hradly a mezi plovoucím hradlem a polovodičem. Vzniká tak kapacitní dělič, kterým se impuls na řídicím hradle přenesle (s patřičným dělicím poměrem) na plovoucí hradlo.

Trvanlivost dat ve správně naprogramované buňce záleží hlavně na teplotě. Je pro běžné účely dostatečná, řádově desetiletí.

Oba výše uvedené mechanismy jsou základem pamětí typu EPROM, EEPROM a FLASH.

**Paměť EPROM** je nejstarší a využívala buňku FAMOS. Podstatnou nevýhodou je obtížné mazání obsahu, neboť buňku FAMOS lze mazat jen ultrafialovým zářením. Pouzdro obvodu musí být opatřeno okénkem propouštějícím UV záření. Paměťový obvod musí být umístěn v objímce, aby mohl být vyjmut k programování i k mazání. To vše zdražuje výrobu. Pro tyto a další důvody paměti EPROM již dávno nejsou používány. Výjimkou jsou paměti **OTP** (One-Time Programmable - jednou programovatelná). Jedná se o paměť EPROM s pouzdrém bez okénka, tedy s nemožností vymazání obsahu. To je řešení relativně levné, paměť nevyžaduje objímku (nebude se přeprogramovávat) a může být v plošném spoji zapájena.

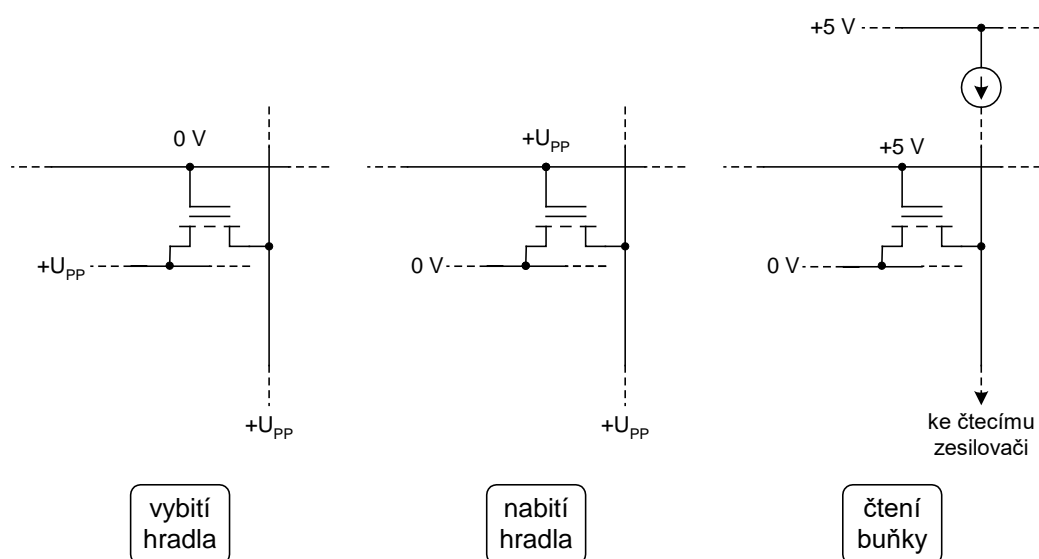
**Paměť EEPROM** využívá tunelování elektronů v obou směrech. Buňka matice je složena ze dvou tranzistorů - jeden výběrový a jeden paměťový. Paměťový tranzistor má plovoucí hradlo a řídicí hradlo. To se uplatňuje jen při zápisu dat a běžně není ve zjednodušených schématech kresleno. Princip ukazuje obr. 19 s vepsanými hodnotami napětí pro vybití hradla, nabití hradla a čtení buňky. Programovací napětí je v něm označeno jako  $+U_{PP}$ .



Obr. 19: Pracovní režimy buňky EEPROM

Buňka EEPROM je poměrně složitá, umožňuje však nabití či vybití plovoucího hradla na jakékoli adrese. Spíše než „programování“ či „vymazání“ má smysl zde hovořit o **zápisu** hodnoty „0“ či „1“. Tím připomíná paměť RAM, zápis je však mnohem pomalejší. Počet zápisů na jedné adrese je omezený. Různí výrobci uvádějí počty desítek až stovek tisíc.

**Paměť FLASH** je založena na buňce, využívající kombinace principu lavinové injekce elektronů pro záporné nabití hradla (naprogramování buňky), a principu tunelování elektronů pro vybití hradla (vymazání buňky). Buňky jsou jednotranzistorové, jednodušší než u EEPROM, mají také méně vývodů. Napětí přivedené do buňky v různých režimech ukazuje obr. 20.



Obr. 20: Pracovní režimy buňky FLASH

U paměti FLASH je nutné buňku vymazat dříve, než bude naprogramována. Vybití plovoucího hradla je docíleno přivedením vysokého kladného napětí na oba konce kanálu a nuly na hradlo - záporný náboj je tunelovým proudem odveden z hradla do kanálu. K nabití hradla je vyvolán lavinový průraz přivedením vysokého napětí na jeden konec kanálu a odchýlením rychlých elektronů vysokým kladným napětím na hradle. Pokud by na řídicím hradle byla nula, k nabití plovoucího hradla nedojde. Tak je vyvolán zápis nuly nebo jedničky.

Vzhledem k jednoduché konstrukci buňky však paměť FLASH neumožňuje selektivní vymazání na jednotlivých adresách, nýbrž jen **mazání bloků** (sektorů). Celý rozsah paměti je rozdělen na (téměř) samostatné jednotky - bloky či sektory - za účelem zrychlení některých operací (zvláště mazání a programování). Blok je vybírán nejvyššími bity adresy. Programování je rychlé, řádově zlomky mikrosekundy (odtud název „flash“), srovnatelné s programováním paměti EPROM. Mazání je ale pomalé, řádově milisekundy. Maže se však najednou celá velká oblast (blok). Postup při změně obsahu je takový, že nejprve se vymaže patřičný blok a pak se celý znova naprogramuje novým obsahem. Paměť FLASH se svojí jednoduchou buňkou umožňuje velmi vysokou integraci.

Počet zápisů na jedné adrese je omezený a garantované počty jsou přibližně stejné jako u EEPROM (desítky až stovky tisíc). Je nutné si uvědomit, že se jedná o počty zápisů **na jedné adrese**, nikoliv do celé paměti - počet zápisů pro celou paměť je mnohem větší.

Programování a mazání EEPROM a FLASH pamětí může probíhat v programátoru. Tak se běžně postupuje při vývoji a ověřování systému, kdy paměť je třeba přeprogramovat mnohokrát. Po dokonalém ověření funkce se pak naprogramuje nová paměť a do systému se může zapájet. Paměti EEPROM a FLASH však umožňují i programování a mazání přímo v systému, takže paměť může být zapájena. Pro manipulaci s pamětí musí některé signály být přepínány dodatečnými obvody (multiplexery), nebo musí být pro programování přímo rezervovány. Paměti EEPROM a FLASH jsou často součástí obvodů velmi velké integrace, jako jsou jednočipové mikropočítače. Pak jsou samozřejmě programovány v systému.

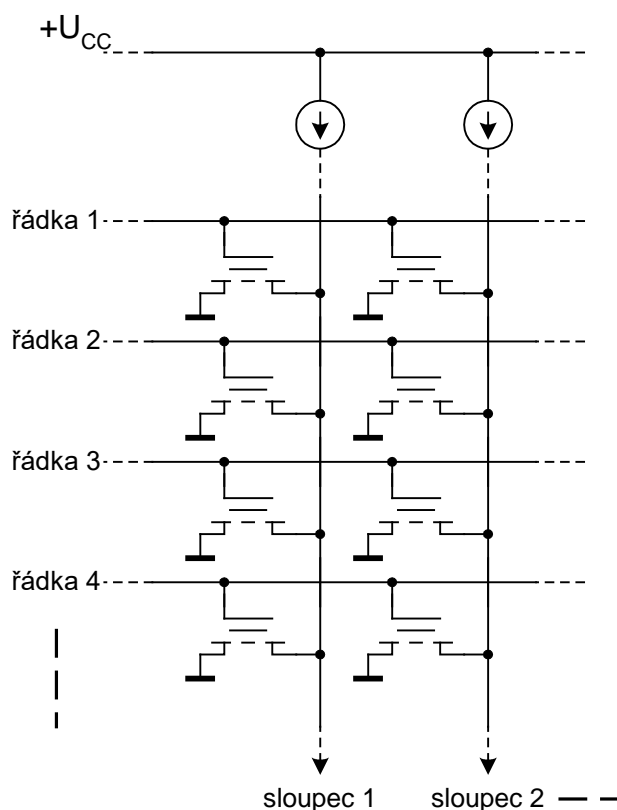
Programování moderních typů EEPROM a FLASH je velmi usnadněno tím, že přímo na čipu je vedle vlastní paměti integrován i sekvenční obvod - **řadič programování** - který řídí potřebné vnitřní operace. Rovněž zvýšené programovací napětí (např. 20V) je generováno na čipu integrovanou **nábojovou pumpou**, takže zvenku je paměť napájena jen běžným provozním napětím  $U_{CC}$  (5V, případně méně). Zápis, mazání a další operace se liší typ od typu podle výrobce.

Průběh zápisu je pomalý. Významného zrychlení bylo dosaženo u pamětí, vybavených vnitřním posuvným registrem na obsah bloku. Data se neprogramují přímo do jednotlivých buněk, ale postupně (a rychle) se zavedou do registru. Po jeho zaplnění se jeho jednotlivé klopné obvody propojí se sloupcovými vodiči a data se přesunou do celé řádky buněk najednou.

Vybavovací doba pamětí EEPROM a FLASH je poměrně dlouhá, ve srovnání s rychlou SRAM řádově desetinásobná. Tuto nevýhodu se podařilo obejít u **synchronních EEPROM a FLASH** podobným trikem jako u SDRAM - paměť je vybavena vnitřním čítačem sloupcové adresy a data jsou vydávána v rytmu hodinových impulsů.

### 3.2 Paměť NOR FLASH

Vývoj v permanentních pamětech se soustředil na EEPROM a hlavně na FLASH. Objevilo se nové uspořádání matice buněk, tzv. **NAND FLASH**. Paměť s klasickým uspořádáním buněk je nazývána **NOR FLASH** pro jistou podobnost s logickým členem, u kterého vzniká paralelním připojením tranzistorů na sloupcový vodič logická funkce NOR. S logickou funkcí však zapojení paměťových buněk nemá nic společného. Toto uspořádání vyžaduje připojení každé buňky jednak na sloupcový vodič, jednak na zem - viz obr. 21. Tyto spoje zabírají značnou část plochy buňky. Obrázek ukazuje cesty signálů pro čtení. Řádkové vodiče jsou buzeny z řádkového dekodéru a sloupce jsou zavedeny do výběrových obvodů.

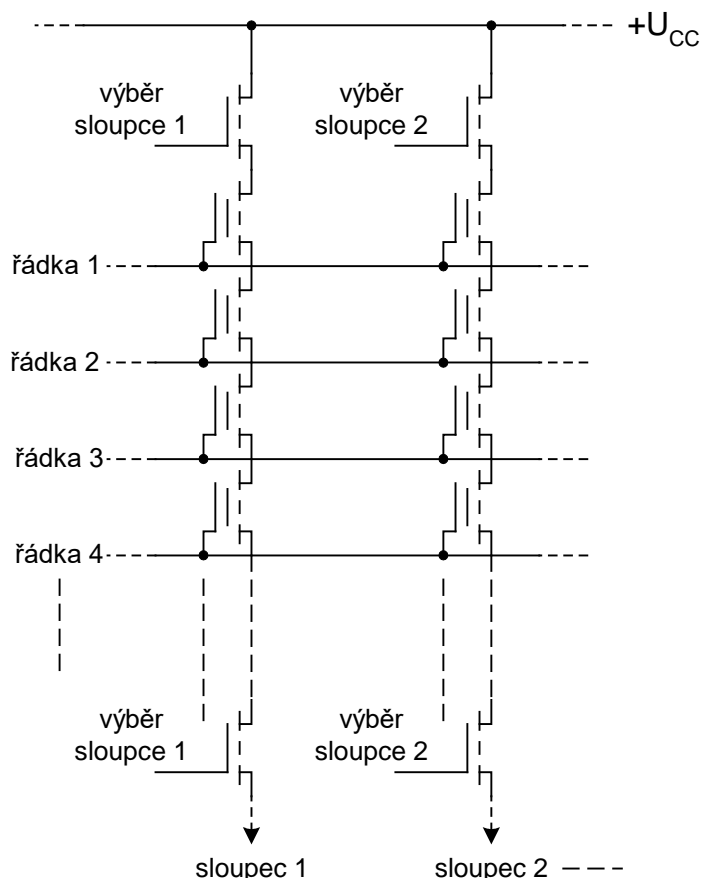


Obr. 21: Zapojení buněk u NOR FLASH

Paměť NOR FLASH je spolehlivá a osvědčila se jako **programová paměť** u jednočipových počítačů.

### 3.3 Paměť NAND FLASH

Podstatně jednoduššího propojení buněk se dosahuje při **sériovém spojení** buněk u paměti NAND FLASH. Zapojení buněk je zcela odlišné od NOR FLASH a funkce buňky také. I technologie výroby je jiná. Společný je princip přesunování náboje **tunelováním elektronů**, a to jak při **programování**, tak i při **mazání** - na rozdíl od NOR FLASH, která se programuje lavinovým průrazem. Uspořádání buněk ukazuje obr. 22.



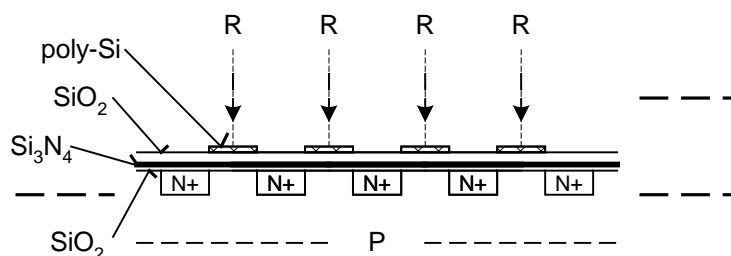
Obr. 22: Zapojení buněk u NAND FLASH

Zapojení buněk do série připomíná spojení tranzistorů v obvodech logického součinu. Kanály tranzistorů jsou navzájem propojeny a jediným vývodem buňky je řídicí hradlo. Na čipu je tak dosahováno mnohem **větší hustoty** buněk než u paměti NOR. Se sloupcem buněk je ale nutné zacházet jinak než u NOR FLASH. Aby bylo možné zjistit stav vybrané buňky, tj. náboj plovoucího hradla, musí být všechny ostatní buňky **plně vodivé**. U nevybraných buněk je proto napětí na výběru řádky tak **vysoké**, že kanál je vodivý i při záporném náboji na plovoucím hradle. Naopak u vybrané buňky se na řídicí hradlo přivádí jen tak velké napětí, že vodivost kanálu je plně ovlivněna nábojem na plovoucím hradle. Je to tedy jinak než u NOR FLASH, kde na nevybrané řádky se přivádí nula.

V průběhu vývoje byly technologické procesy tak zdokonaleny, že umožnily velmi přesně odměřit náboj dodaný na plovoucí hradlo. Při čtení z buňky pak čtecí zesilovač může rozlišit **několik úrovní napětí** na výstupu sloupce. Tímto způsobem lze do buňky zapsat ne jeden, ale 2 bity informace (při 4 úrovních), nebo 3 bity (při 8 úrovních), možná ještě více. Celková kapacita paměti se tím úměrně zvýší.

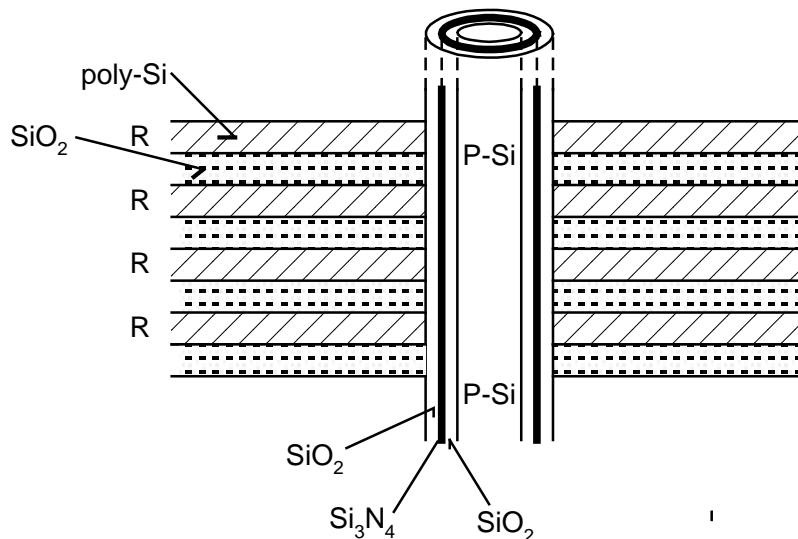
Vývoj však šel ještě dál. V NOR FLASH a starších NAND FLASH jsou obě hradla z vodivého materiálu (polykrystalický křemík) a je nutné je oddělit a izolovat nejen navzájem, ale i mezi sousedními buňkami, které jsou u NAND FLASH blízko sebe. K izolaci se používá oxid křemíku ( $\text{SiO}_2$ ). U novějších pamětí se místo vodivé vrstvy používá vrstva **nitridu křemíku** ( $\text{N}_3\text{O}_4$ ), což je výborný izolant s přibližně dvojnásobnou  $\epsilon_r$  proti oxidu křemíku. Náboj se zachycuje na rozhraní nitridu a oxidu křemíku, a vyšší  $\epsilon_r$  snižuje napětí potřebné pro

čtení buňky. Další výhodou je to, že vzhledem k tomu, že nitrid je izolant, nemusí jeho vrstva být přerušovaná mezi sousedními buňkami. Totéž samozřejmě platí i o vrstvě oxidu. Konstrukce sloupce buněk se tak výrazně zjednoduší. Tato technologie se nazývá **Charge-Trap** NAND - viz obr. 23. Řádkové výběrové vodiče R jsou naznačeny jen symbolicky; ve skutečnosti jsou z polykrystalického křemíku a jsou vedeny po povrchu.



*Obr. 23: Princip Charge- Trap NAND*

V posledních letech došlo k dalšímu skoku v technologii, zvláště ke zvládnutí přesného leptání do hloubky. Lze tak vytvářet hluboké příkopy k oddělení vrstev nebo hluboké otvory. Tím byly dány prostředky pro realizaci **vertikální NAND** paměti, V-NAND FLASH. též 3-D FLASH. Průkopníky byly firmy Toshiba a Samsung. Předchozí technologie vytvářela buňky horizontálně, na povrchu čipu. Vertikální technologie staví buňky **ve vrstvách na sebe**. Docílí se tak mnohonásobného zvýšení kapacity paměti - viz obr. 24.



*Obr. 24: Princip V-NAND*

Při výrobě se na povrchu čipu vytvoří řada vrstev polykrystalického křemíku, prostřídáných s izolačními vrstvami  $\text{SiO}_2$ . Poly-Si slouží jako vodič a současně vytvoří řídicí hradlo. Pak je všemi vrstvami proleptán **válcový otvor**. Na jeho stěny se postupně navrství oxid, nitrid, a opět oxid křemíku. Konečně se dutina vyplní křemíkem typu P. Tím se vytvoří struktura podobná té z obr. 23, jenže vertikálně. U velkokapacitních pamětí je takovýchto otvorů vytvořeno několik miliard (!) na čipu při více než 100 vrstvách. Vyleptáním hlubokých brázd

mezi skupinami otvorů lze celé paměťové pole **rozdělit na bloky**, které budou vzájemně částečně nezávislé. S využitím popsané technologie je dosažena kapacita paměti 1 Tb (terabit).

**Jednoduchý výpočet:**  $2,5 \cdot 10^9$  otvorů  $\times 10^2$  vrstev  $\times 4$  b na buňku  $= 10^{12}$  b

Nevýhodou paměti NAND FLASH je **pomalejší činnost** než u paměti NOR FLASH a větší **počet chyb**. Samotné vymazání bloku trvá stovky *ns*, zápis o něco méně. Zrychlení je možné podobným trikem, jako u DRAM - **prací po stránkách**. Přečtená stránka se uloží do vnitřního posuvného registru a jeho obsah se pak postupně vydává. Zrychlení je částečně úměrné velikosti stránky. Při zápisu je postup obdobný.

K chybám dochází poměrně často vlivem přeslechů a rozlévání náboje mezi buňkami, překročením tolerance napěťových úrovní u vícebitových buněk a dalších vlivů. Bez **korekce chyb** by tyto paměti byly nepoužitelné. Pracuje se proto se samoopravnými kódy. Paměť má vždy **záložní bloky** pro případ trvalé poruchy nebo přílišné chybovosti využívaných bloků. Všechny kontrolní a organizační funkce vykonává **řadič paměti**, který obvykle obsluhuje několik paměťových čipů.

Dalším problémem je **omezený počet zápisů** do jednotlivých buněk. Počítá se na desítky až stovky tisíc, což je pro některé funkce nepřijatelné. Stejným omezením trpí i paměti NOR FLASH, ale vzhledem k jejich převažujícímu použití jako programové paměti mikropočítačů, které se nikdy tolikrát nepřeprogramují, je to omezení jen teoretické. Paměti NAND FLASH se ale používají jako velkokapacitní paměti nahrazující disky, kde je velký počet zápisů samozřejmý.

Důležité je **rovnoměrné rozložení zátěže** na jednotlivé buňky (wear leveling). Jelikož s buňkami nelze manipulovat samostatně, týká se údaj o počtu mazání celého bloku. **Řadič paměti** optimalizuje využití jednotlivých bloků a sleduje jejich zátěž. Při potřebě přepsání dat je daný blok přepsán do rezervního bloku včetně požadovaných změn a původní blok je pak vymazán a zařazen do rezervy. Tím se výrazně zkrátí **dobu pro zápis dat** – nemusí se čekat napřed na vymazání bloku a pak na zápis dat. Při přidělování bloků řadič bere do úvahy jejich zatížení a vybírá ty méně zatížené. Řadič též sleduje počet opravovaných chyb a při jejich přílišném počtu blok zcela vyřadí. To umožňuje provozovat paměť až do úplného vyčerpání rezervy. Zvláště náročné je rozkládání zátěže u náhrady **diskových pamětí**, kde je velmi zatížena oblast s adresáři – ty se stále přepisují. Adresáře se proto u NAND FLASH postupně přesunují na různé bloky a tak se zátěž rozprostře.

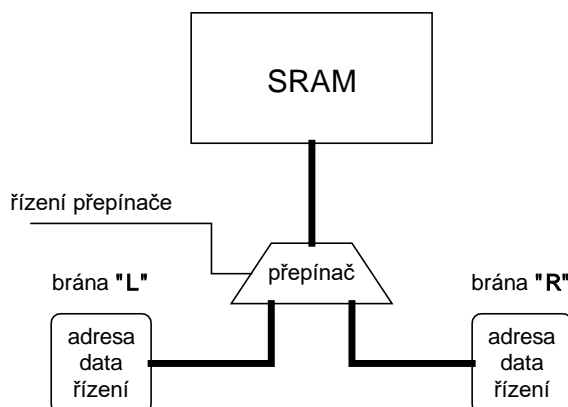


## 4 Dvoubránová paměť

Dvoubránová paměť DPM (angl. Dual-Port Memory) je statická RAM, která umožňuje přístup k datům ze dvou bran. Každá brána zahrnuje adresu, data a řídicí signály. Dvoubránová paměť je využívána k předávání dat mezi dvěma částmi systému - „uživateli společné paměti“ - které nepracují v přísném synchronismu. Velmi často je např. vložena mezi komunikační procesor a řídicí počítač. V paměti jsou na stanovených místech uloženy zprávy k odeslání a zprávy přijaté, určené ke zpracování. Komunikační procesor pracuje svým rytmem, odesílá a přijímá zprávy. Řídicí počítač pracuje též svým rytmem, zpracuje přijaté zprávy a ukládá do paměti zprávy určené k vysílání. Při přímém propojení, bez DPM jako spojovacího článku, by docházelo k časovým ztrátám, k čekání jedné části systému na druhou.

Někdy, i když méně často, je třeba zavést spojovací článek pro více než dva uživatele. Existují proto i paměti se třemi nebo čtyřmi branami. Jejich konstrukce je jen rozšířením principů dvoubránových pamětí, uvedených v dalším textu. Paměti s více než dvěma branami proto nebudou zvláště popisovány.

Existuje několik způsobů, jak ze dvou stran dosáhnout na buňky ve společné paměťové matici. Nejjednodušší je vzít za základ standardní paměť SRAM a doplnit ji **přepínačem**, který přepíná signály ze dvou bran na vstupní a výstupní vývody SRAM - viz obr. 25. Je zřejmé, že není možný současný přístup z obou bran - brány se musí vhodným způsobem střídát.



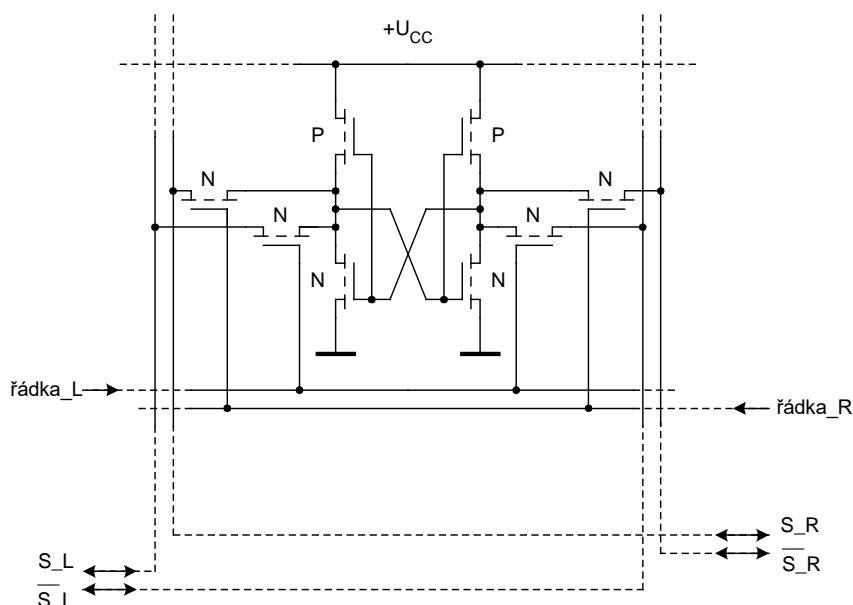
Obr. 25: Vytvoření dvoubránové paměti z SRAM

Někdy je možné zařídit střídání přístupu v **pravidelném rytmu**. Každý uživatel má svoji bránu k dispozici po přesně určenou dobu časového okna. To je možné bez časových ztrát, pokud každý uživatel pracuje v rytmu „práce s pamětí“ - „zpracování dat“. Funkce obou uživatelů a přepínače jsou vzájemně vázány a centrálně řízeny. Pak se mohou tyto dvě fáze činnosti vzájemně překrývat takto:

```

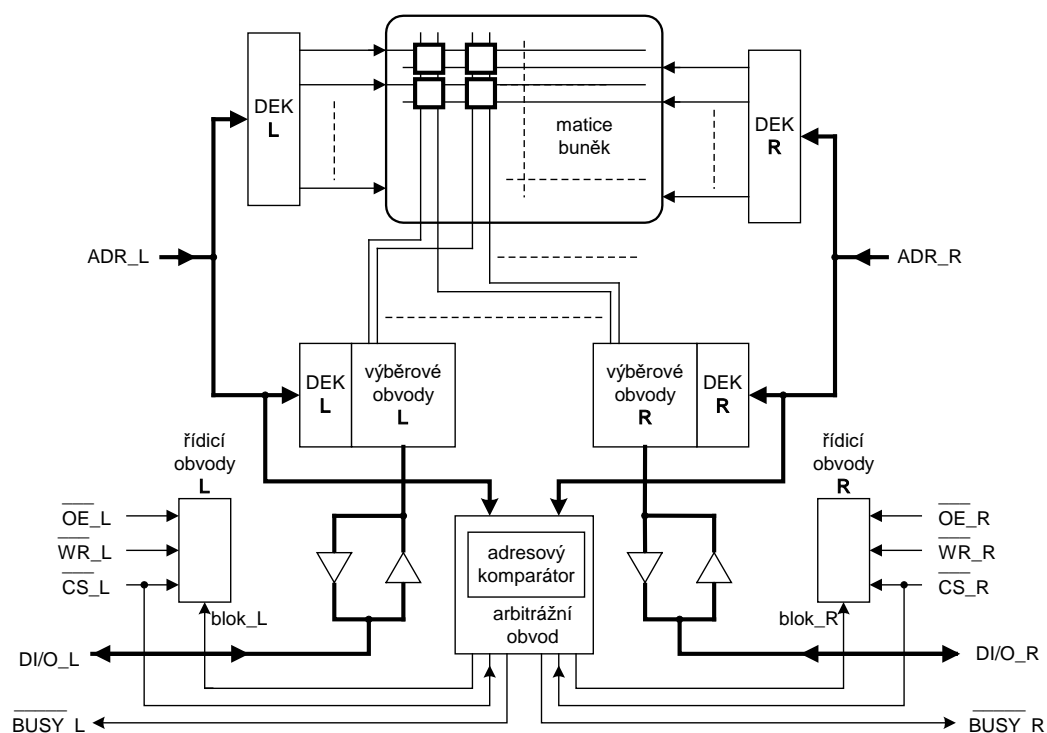
uživatel L (levý): ..... práce s pamětí → zpracování dat → práce s pamětí →
uživatel R (pravý): ..... zpracování dat → práce s pamětí → zpracování dat →
přepínač:           ..... L R L
  
```

Většinou takováto pravidelnost není možná a požadavky na přístup k paměti jsou náhodně rozložené v čase, občas i současné z obou stran. Pro náročnější aplikace existuje zvláštní typ paměťové buňky, **přístupné ze dvou stran** - viz obr. 26.



Obr. 26: Statická buňka přístupná ze dvou stran

Princip je obdobný jako u buňky běžné SRAM. Buňka je však napojena na dvě dvojice sloupců - dvojice sloupců „levých“ ( $S\_L, \bar{S}_L$ ) a dvojice sloupců „pravých“ ( $S\_R, \bar{S}_R$ ). Připojení na „levé“ sloupce je řízeno „levým“ řádkovým vodičem, připojení na „pravé“ sloupce „pravým“ řádkovým vodičem. S buňkou tak lze pracovat ze dvou stran. Jedna operace je zakázaná, a to **současný zápis** z obou stran. Uspořádání DPM s využitím matice buněk se dvěma přístupy ukazuje obr. 27.



Obr. 27: Přístup k matici buněk ze dvou stran

Obrázek ukazuje shodu s obrázkem běžné SRAM, všechny obvody kromě vlastní matice jsou však zdvojeny a jsou navzájem nezávislé. Vstupní a výstupní signály jsou uspořádány do dvou skupin, tj. bran. Nezávislost obvodů dovoluje současné operace čtení či zápisu na různých adresách. Na shodných adresách je však třeba vyloučit současné operace zápisu - jiné kombinace operací nejsou na závalu. Obvody podle obr. 27 jsou proto ještě doplněny o **arbitrážní obvod** - „arbiter“.

Úkolem arbitrážních obvodů je řešení střetu dvou (obecně libovolného počtu) požadavků na přístup do jednoho místa. Nejčastěji se u DPM jedná o arbitráž symetrickou, s rovnocennými prioritami požadavků. Funkci arbitráže lze popsat dvěma pravidly:

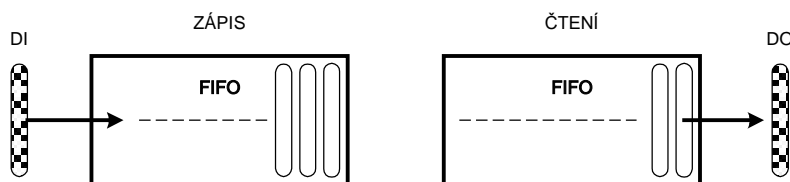
1. operace je povolena tomu uživateli, který generoval požadavek jako první
2. při současném výskytu více než jednoho požadavku je operace povolena uživateli náhodně zvolenému

Požadavek na přístup k paměti je odvozen od výběrových signálů  $\overline{CS}$ . Jediná operace, která podléhá arbitráži, je současný zápis z obou stran na stejných adresách. Arbiter proto obsahuje **adresový komparátor**. Pohotovost brány k operaci je oznámena uživateli signálem  $\overline{BUSY}$ . Uživatel brány paměti musí tento signál testovat před provedením operace a při jeho stavu „0“ počkat.

Paměti různých výrobců disponují ještě dalšími funkcemi, usnadňujícími spolupráci procesorů/počítačů, které paměť sdílejí. Trend ke zvyšování rychlosti se nevyhnul ani dvoubránovému paměti. Vzniklo tak několik typů **synchronních DPM**. Princip synchronního provozu s čítačem sloupců inkrementovaným hodinovými pulsy je obdobný jako u synchronních dynamických pamětí.

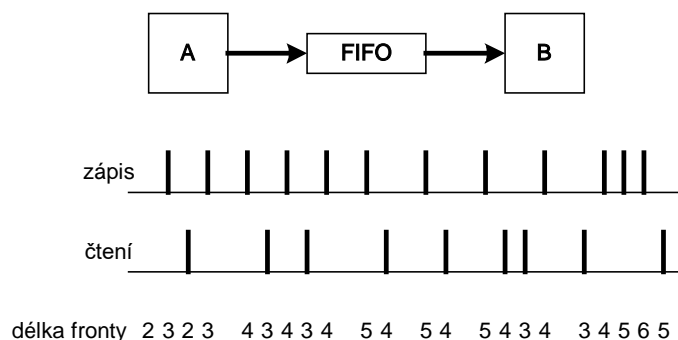
## 5 Paměť fronty

Paměť fronty FIFO (angl. First In - First Out) patří mezi paměti **bezadresové**. Adresa není zapotřebí proto, že pořadí zápisu a čtení dat je pevně dáno. Při zápisu dat se jednotlivá slova řadí do fronty a při čtení se vybírají v tom pořadí, v jakém byla zapsána. Operace zápisu a čtení jsou přitom vzájemně nezávislé, takže paměť FIFO může být zaplněna do různé míry - od zcela vyprázdňené až po zcela zaplněnou. Obrázek 28 názorně ilustruje činnost při zápisu a čtení. Posunování dat, znázorněné na obrázku, je jen zdánlivé. U moderních konstrukcí pamětí FIFO jsou data naopak na stabilních pozicích a posunují se ukazatele na ně - viz další text.



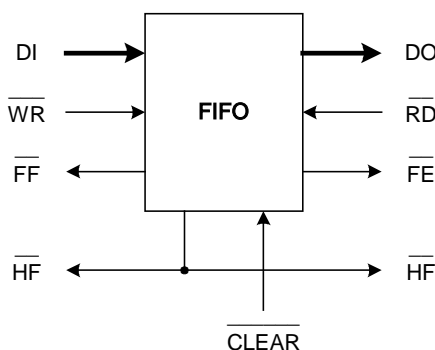
Obr. 28: Činnost paměti FIFO

Paměť FIFO se používá jako vyrovnávací člen mezi dvěma částmi systému. Předpokládejme, že jedna část (A) předává data druhé části (B). Při přímém propojení musí A i B pracovat ve shodném rytmu - jinak vždy jedna část musí čekat na druhou, aby nedocházelo ke ztrátě dat. Při vložení vyrovnávací paměti se však mohou rychlosti zpracování dat u obou částí **krátkodobě** rozcházet, aniž by docházelo k časovým ztrátám. Je-li momentálně A rychlejší než B, bude se FIFO postupně zaplňovat. Po nějaké době se situace změní a B je rychlejší, takže se FIFO postupně vyprazdňuje - viz příklad na obr. 29.



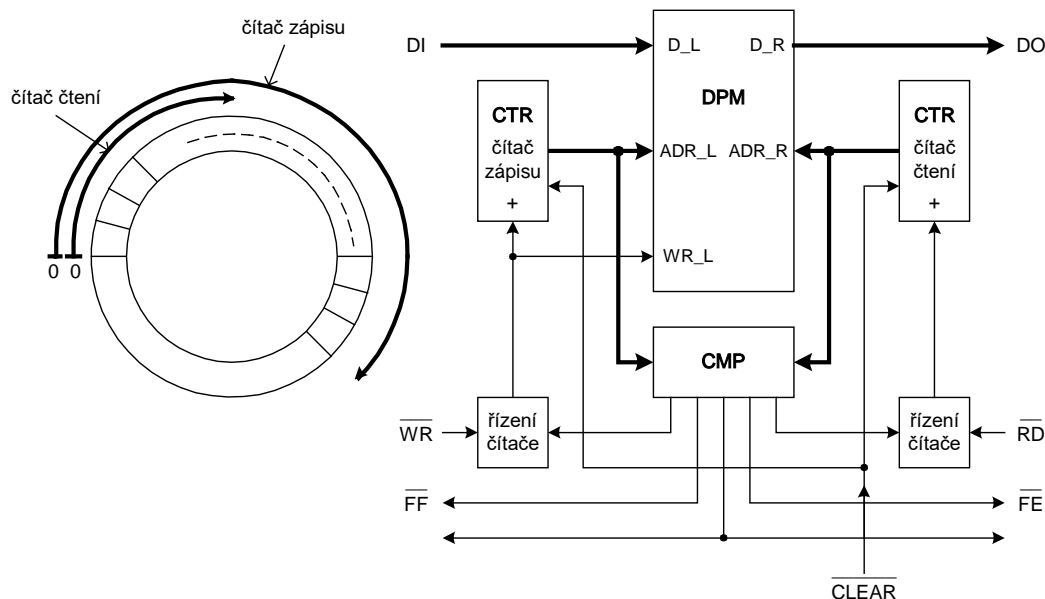
Obr. 29: FIFO jako vyrovnávací paměť

V praxi není vyloučeno vyprázdnění či přeplnění FIFO. Aby při tom nedocházelo ke ztrátám dat či čtení falešných dat, je třeba oba uživatele paměti informovat o pohotovosti paměti. FIFO proto generuje signály  $\overline{FF}$ ,  $\overline{EF}$ ,  $\overline{HF}$  - Full FIFO (plné FIFO), Empty FIFO (prázdné FIFO), Half Full (zpola plné). Signál  $\overline{FF}$  je nutný pro ochranu před přeplněním FIFO a ztrátou dat. Před zápisem je testován. Je-li ve stavu „1“, je možné zápis provést, jinak je třeba zápis pozdržet a dále signál testovat. Obdobně se zachází se signálem  $\overline{EF}$  při čtení dat. Signál  $\overline{HF}$  je užitečný jako předběžné varování před přeplněním při zápisu nebo úplným vyprázdněním při čtení. Soustava vstupních a výstupních signálů je vyznačena na obr. 30.



Obr. 30: Vstupní a výstupní signály paměti FIFO

Konstrukci paměti FIFO ukazuje obr. 31. Je založena na zjednodušené dvojbránové paměti a dvou **ukazatelích na adresy**, realizovaných čítači. Zjednodušení DPM spočívá v omezení funkce u obou bran - jedna dovoluje jen zápis, druhá jen čtení. Po každém zápisu dat se posune čítač zápisu a po každém čtení se posune čítač čtení. Vzhledem k funkci čítačů, kdy po nejvyšší adrese následuje nula, lze adresy v paměti znázornit jakoby navinuté na kružnici.



Obr. 31: Konstrukce paměti FIFO se dvěma ukazateli adres

Čítač zápisu obsahuje vyšší číslo než čítač čtení. To odpovídá částečnému zaplnění FIFO. Jakmile by čítač čtení „doběhl“ čítač zápisu, odpovídalo by to úplnému vyprázdnění FIFO. Naopak čítač zápisu nesmí „předběhnout“ čítač čtení, což by odpovídalo úplnému zaplnění FIFO. Obě adresy jsou proto srovnávány v komparátoru, který v případě zaplnění nebo vyprázdnění FIFO zablokuje příslušný čítač. Komparátor též generuje signály  $\overline{FF}$ ,  $\overline{EF}$ ,  $\overline{HF}$ . Signálem  $\overline{CLEAR}$  lze vynulovat oba čítače, což odpovídá vyprázdnění FIFO.

Na obr. 31 je vidět, že komparátor je připojen na čítače, jejichž výstupy se mohou měnit (přibližně) ve stejném okamžiku. V komparátoru, což je kombinační obvod, tak mohou vznikat falešné impulsy. Čítače proto pracují v **Grayově kódu** se změnou vždy na jediném výstupu. Je tak možné využít návrhovou metodu redundantních kombinačních obvodů.

Paměť FIFO podle výše uvedené konstrukce je **asynchronní**. Dovoluje navzájem nezávislé operace zápisu a čtení - mohou probíhat i současně. Zpoždění, které zavádí paměť FIFO vložená mezi dvě části systému, samozřejmě zpožďuje celý systém. Výrobci proto používají ty nejrychlejší technologie. Další zrychlení se dosahuje u **synchronních** pamětí FIFO. Kromě signálů z obr. 30 jsou do nich ještě zavedeny hodinové impulsy, nezávisle pro čtení a pro zápis.