

SKRIPTA

RAM memorija

Student:
Ivan Mokanj

Broj indeksa:
11602

Nastavnik:
dr Veljko Malbaša

Asistent:
mr Milan Nikolić

Novi Sad, 2010.

Sadržaj

Strana

1. Uopšteno o memoriji	1
2. RAM memorija	2
2.1. Statički RAM (SRAM).....	3
2.2. Keš memorija (Cache).....	5
2.3. Multiport memorija	7
2.4. FIFO memorija.....	8
2.5. Dinamički RAM (DRAM).....	11
2.6. Memorija za grafičke kartice (GDDR).....	14
2.7. Nove SDRAM tehnologije (DDR2 i DDR3).....	15

1. Uopšteno o memoriji

Za rad PC računara neophodna je memorija pošto se u njoj tokom rada smeštaju programi koji se izvršavaju, kao i podaci koji se tim programima obrađuju. Memorisana digitalna informacija izražena je u vrednostima logičke 1 ili 0 u memorijskom elementu, odnosno u memorijskoj ćeliji. Jedna ćelija može da primi samo jedan podatak, odnosno jedinicu informacije ili jedan bit informacije. Parametar koji karakteriše veličinu memorije u pogledu maksimalne količine informacije koju ona može da primi naziva se kapacitet memorije. Osnovna jedinica za kapacitet memorije je bajt. U jedan bajt memorije, koji sadrži osam bita, može da se smesti jedan ASCII karakter. U PC računarima se koriste memorije veoma velikog kapaciteta, pa je bajt suviše mala i nepraktična jedinica. Umesto nje se češće koriste kilobajt (kB), megabajt (MB) i gigabajt (GB), pri čemu je:

$$1 \text{ B} = 8 \text{ b}$$

$$1 \text{ kB} = 1024 \text{ B}$$

$$1 \text{ MB} = 1024 \text{ kB} = 1024 \times 1024 \text{ B} = 1048576 \text{ B}$$

$$1 \text{ GB} = 1024 \text{ MB} = 1024 \times 1024 \text{ kB} = 1048576 \text{ kB} = 1073741824 \text{ B}$$

Osnovna podela memorije je na:

- ROM (Read Only Memory) – memorija koja može samo da se očitava
- RAM (Random Access Memory) – memorija sa proizvoljnim pristupom, u koju podaci mogu da se i upisuju i iz koje mogu da se i očitavaju.

2. RAM memorija

Poluprovodnička memorija, u koju se može i upisati i pročitati informacija u proizvoljnom trenutku, naziva se RAM memorija, za razliku od ROM memorija, kod koje je fizički i vremenski proces upisa različit od procesa čitanja sadržaja. Naziv RAM dolazi od engleskog naziva Random Access Memory (memorija sa slučajnim pristupom), što na neki način označava da je vreme za upis ili čitanje nezavisno od adrese na kojoj se čitanje ili upis obavlja. Poluprovodničke RAM memorije po pravilu, gube sadržaj kada se isključi napon napajanja, tako da spadaju u klasu nepostojanih memorija.

Glavni parametri RAM-a su kapacitet i brzina rada. Kao što je već rečeno, kapacitet predstavlja parametar koji karakteriše veličinu memorije u pogledu maksimalne količine informacije koju ona može da primi. Poželjno je da je RAM što većeg kapaciteta kako bi se smestilo što više podataka. Obično se nekoliko bitova skuplja i smešta na određeno mesto u memoriji. Ovo mesto naziva se memorijska lokacija.

Memorijske lokacije možemo zamisliti kao niz pretinaca od kojih svaki ima svoju adresu i može da primi jedan bit ili određenu količinu bitova, najčešće jedan bajt. Ako želimo da pročitamo bajt neke memorijske lokacije, potrebno je navesti adresu memorijske lokacije u kojoj je on smešten. Brzina rada RAM-a određena je brzinom kojom ova memorija iznalazi određenu memorijsku ćeliju ili lokaciju u memorijskom bloku. Spomenuto je već da je za čitanje nekog podatka iz memorije potrebno navesti adresu memorijske lokacije u kojoj je taj podatak nalazi. Od pojave željene adrese na adresnim magistralama pa do pojave podatka smeštenog u traženoj lokaciji na podatkovnim magistralama protekne određeno vrijeme. To vreme se zove vreme pristupa memoriji (memory access time). Vreme pristupa ograničava brzinu kojom se mogu čitati podaci iz memorije i upisivati u nju pa možemo znatno ograničiti brzinu rada celog računara. Zbog toga se u računare nastoji ugraditi RAM sa što kraćim vremenom pristupa. Tehnologija izrade poluvodičkih komponenata od kojih su izgrađeni savremeni RAM-ovi ograničava brzinu pristupa na nekoliko desetaka nanosekundi.

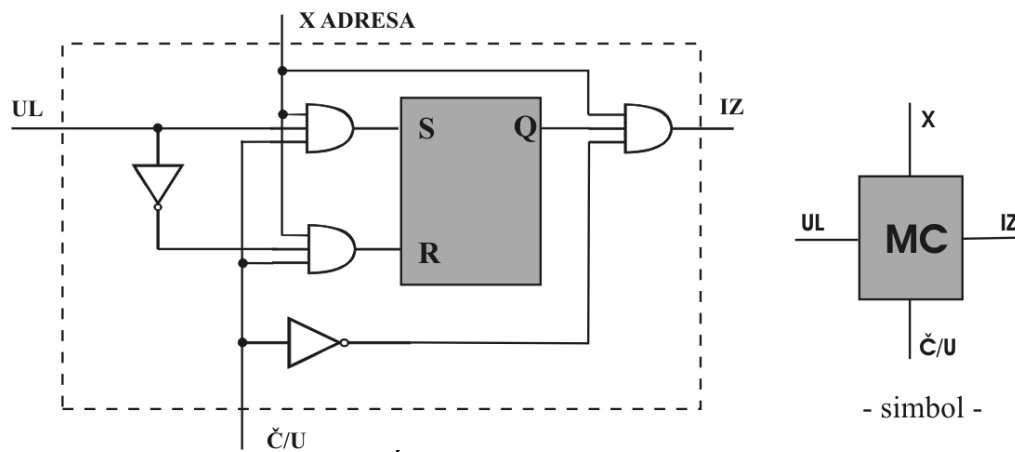
Poluprovodničke memorije mogu biti statičke (SRAM) i dinamičke (DRAM). Informacija upisana u statičku memoriju ostaje zapamćena sve dok je memorija priključena na napon

napajanja. Da bi informacija ostala zapamćenja u dinamičkoj memoriji, neophodno je periodično obavljati "osvežavanje" memorije, inače će se informacija izgubiti.

Pored navedenih SRAM i DRAM memorija, postoje i druge vrste RAM memorija, kao na primer, FRAM memorija (ferroelectric random access memory). FRAM je feroelektrična memorija sa slučajnim pristupom. Feroelektrični efekat je osobina materijala da zadrži električnu polarizaciju i u odsustvu električnog polja, što omogućava izradu postojećih RAM memorija koje zadržavaju sadržaj i po nestanku napona.

2.1. Statički RAM (SRAM)

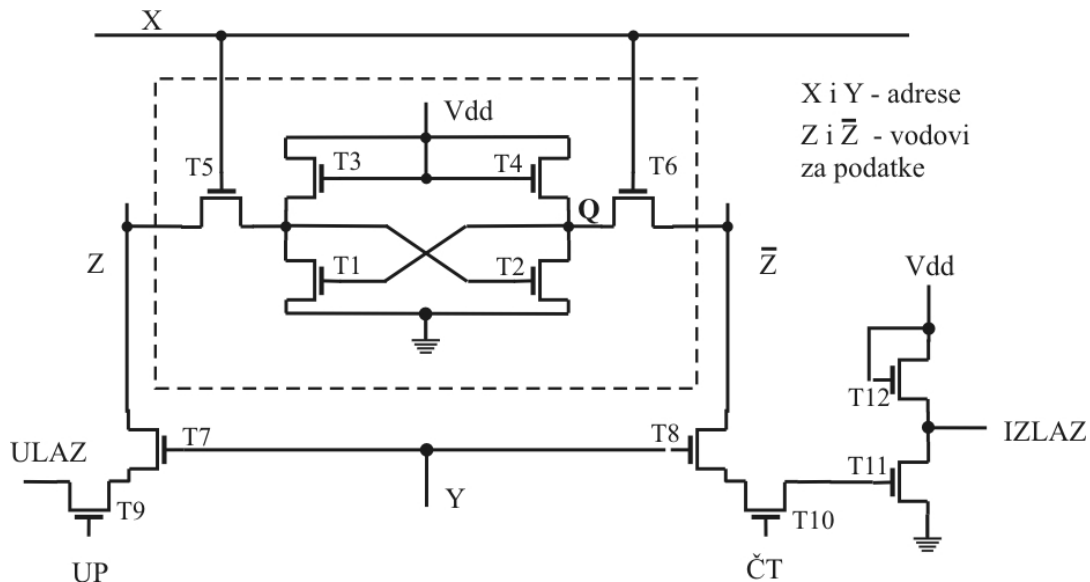
Osnovna memorijska ćelija u statičkoj RAM matrici je flipflop, čiji je sadržaj statičan, tj. veličina signala (napona) u ćeliji se ne menja u toku vremena, pa nije potrebno da se obnavlja (zadržava podatke bez spoljašnjeg "osvežavanja" dokle god je priključeno napajanje, po čemu se razlikuje od dinamičkog RAM-a koji se mora "osvežavati" puno puta u toku svake sekunde da se podaci ne bi izgubili).



Sl. 2.1.1. Ćelija bipolarne SRAM memorije

Na slici 2.1.1 je prikazana jedna ćelija bipolarne statičke RAM memorije, koja ima jednostavan ulaz radi upisa, i jednostavan izlaz radi očitavanja memorisanog podatka. Ovom memorijskom kolu može se pristupiti samo kada je pobuđen njegov priključak X za adresiranje. Pristup ulazu i izlazu je preko istog priključka Č/U, što znači, da se ovim priključkom obavlja i čitanje i upis. Upis informacije sa ulazu UL vrši se kada je $X = 1$ i $\check{C}/U = 1$. Čitanje upisanog sadržaja na izlazu vrši se kada su $X = 1$ i $\check{C}/U = 0$. Invertor u ćeliji omogućava upis podatka bez prethodnog brisanja sadržaja SR leč kola. Ovde se umesto prikazanog SR leč kola može upotrebiti i D flip flop ili bilo koji drugi. Na istoj slici je prikazan i simbol jedne ćelije SRAM memorije (MC), za linearno adresiranje sa adresnim ulazom X.

Statičke RAM memorije se realizuju i u MOSFET tehnologiji, ili u CMOS tehnologiji. Prema tome, princip rada ovih memorija je isti kao i bipolarnih, ali su tehnologije izrade, a time i karakteristike drugačije. Bipolarne memorije su brže ali znatno manjeg kapaciteta, a najbrže se izrađuju u ECL tehnici gde vreme pristupa može biti i manje od 10ns.



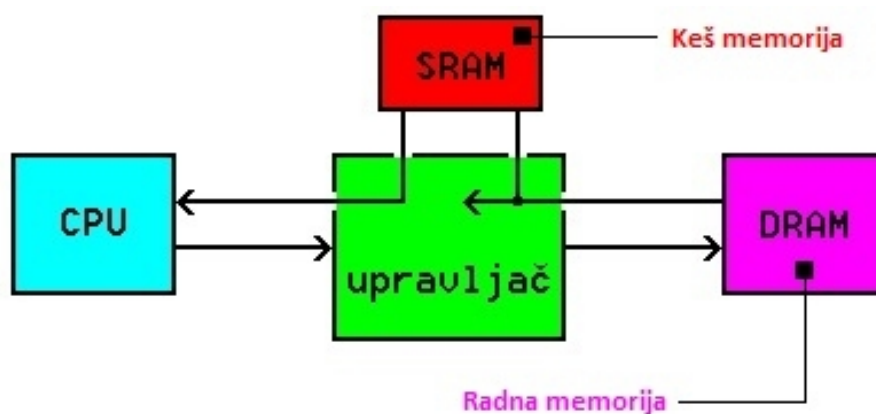
Sl. 2.1.2. Statička MOS ćelija

Na slici 2.1.2 je prikazana šema ćelije sa 6 MOSFET-ova, od kojih se prva 4 koriste kao memorijsko kolo, a MOSFET-ovi 5 i 6 služe za spregu tog kola sa vodovima za podatke Z i Z komplement. Adresu ove ćelije čine dve koordinate X i Y. Dok adresa X utiče na ćeliju MC direktno preko MOSFETA 5 i 6, koji pripadaju ćeliji, dotle koordinata Y deluje posredno preko MOSFET-a 7 i 8, koji su postavljeni u vodove podataka Z i Z po komplementu. To znači da su ovi MOSFET-ovi zajednički za sve ćelije koje su priključene na iste vodove podataka. Svaka ćelija se adresira posebno vrednostima $X_i Y_i$. Upisivanje podataka u ćeliju vrši se preko priključka ULAZ, aktiviranjem komande za upis (UP). S obzirom na to da visoki naponi na adresnim vodovima $X = 1$, i $Y = 1$, omogućavaju provođenje MOSFET-ova 5 i 6, kao i MOSFET-ova 7 i 8, ulazni signal će biti doveden na gejt MOSFET-a 2. Ako ulazni signal ima visoki naponski nivo, MOSFET 2, provodi čime se blokira MOSFET 1. Prema tome, ako je ulazni podatak koji se upisuje 1, na izlazu Q se uspostavlja komplement tj logička 0. Čitanje memorisanog podatka vrši se tako što se prvo aktiviraju adresni vodovi dovođenjem $X = 1$ i $Y = 1$, a odmah zatim se postavlja i komanda čitanje ČT na naponski nivo logičke 1. Time se izlaz Q, ćelije MC, vezuje na ulaz invertora MOSFETA 11. Pošto napon u tački Q odgovara logičkoj 0, to se na izlazu invertora, a time i na izlazu memorije, dobija logička 1, tj. ona vrednost koja je uneta u ćeliju pri upisivanju. U slučaju zahteva za malom disipacijom, umesto MOS tranzistora u memorijskoj ćeliji se mogu upotrebiti CMOS tranzistori.

Statičke memorije su manjeg kapaciteta po čipu, a koriste se u sistemima gde se zahteva veća brzina pristupa memoriji i manja potrošnja struje iz izvora za napajanje. Takođe je verovatnoća greške kod statičkih memorija manja nego kod dinamičkih, tako da se koriste u sistemima gde se zahteva visoka pouzdanost.

2.2. Keš memorija (Cache)

Kako je vreme pristupa memorijskim lokcijama u dinamičkoj memoriji znato duže od brzine kojom mikroprocesor može da obradi dobijene podatke iz memorije, zaključuje se da će mikroprocesor gubiti mnogo vremena čekajući da dobije potrebne podatke iz memorije, što bi dovelo do velikog usporenja rada računara. Da bi se to sprečilo, između glavne radne memorije, koja je realizovana kao dinamički RAM, i mikroprocesora se postavlja manja količina znatno brže statičke RAM memorije. Ova memorija se naziva keš memorijom (Cache). Prema tome keš memorija ima ulogu posrednika između glavne memorije (DRAM) i procesora i to je prikazano na slici 2.2.1.



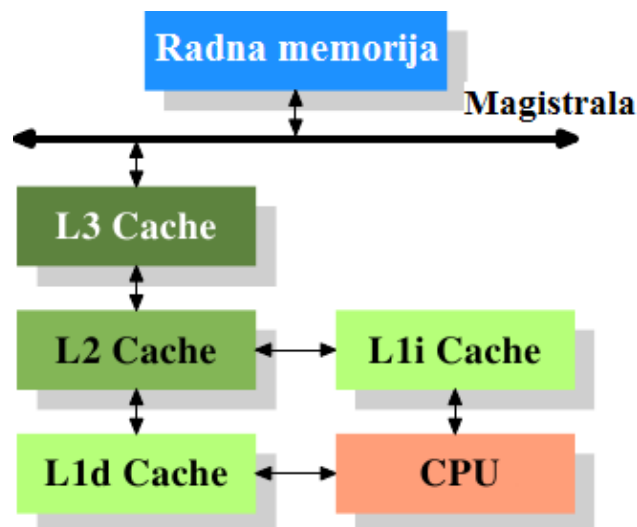
Sl. 2.2.1. Posredstvo SRAM-a između CPU-a i DRAM-a

Ovo rešenje, da se koristi keš memorija između procesora i glavne memorije, sa pratećom elektronikom osigurava da se sledeći podatak, koji je potreban procesoru, u što većem procentu slučajeva već nalazi u keš memoriji. Keš se danas izrađuje u više nivoa, koji redom imaju sve veći kapacitet (ali sve manju brzinu pristupa) kako su postavljeni dalje od procesora. Ovakva organizacija memorijske hijerarhije se ustalila jer pruža vrlo dobre performanse računara uz prihvatljivu cenu.

Primarni keš, ili keš prvog nivoa (Level 1 cache), nalazi se na procesoru i služi za privremeni smeštaj instrukcija i podataka. Ovo je najbrži vid memorijskog prostora koji postoji u računaru (brži pristup imaju samo procesorski registri). Primarni keš je uvek ugrađen na procesorski čip, i podeljen je na instrukcijski keš (L1i cache) i keš za podatke (L1d cache). Primarni keš je ograničen po veličini - u početku je standardna veličina primarnog keša bila 16kB, a potom je povećana na 32kB i kasnije na 64 kB. Primarni keš izrađuje se isključivo od SRAM čipova koji su izuzetno brzi, i podaci u njima se ne moraju osvežavati, ali su zato skuplji od DRAM-a, i zauzimaju više prostora. Međutim, za potrebe keš memorije najvažnija je izuzetna brzina, a pošto je ona mala po kapacitetu onda cena i veličina nisu veliki problem.

Sekundarni keš, ili keš drugog nivoa (Level 2 cache) koristi istu kontrolnu logiku kao i primarni keš i takođe se izrađuje pomoću SRAM čipova da bi se postigla maksimalna moguća brzina. Na današnjim procesorima on se nalazi integrisan unutar procesorskog kućišta. Ranije se sekundarni keš nalazio ili zalemljen na matičnoj ploči, ili se na nju dodavao u posebne slotove. Cilj sekundarnog keša je da isporuči podatke procesoru bez ijednog stanja čekanja.

Koliko je keš memorija važna za efikasan rada računara najbolje pokazuje kompanija Intel. Ova kompanija je na svom čipu ugradila i treći nivo keš memorije (Level 3 cache), veličine od čak 4 MB. On se, zajedno sa pratećom logikom, nalazi ugrađen unutar procesorskog kućišta da bi komunikacija sa procesorom bila što je moguće brža. Treći nivo keša je dodat da bi se eliminisalo usko grlo u komunikaciji procesor - memorija, koje predstavlja jedan od najvećih problema pri projektovanju aktuelnih računarskih sistema. Na slici 2.2.2 je prikazana hijerarhija keš memorije.

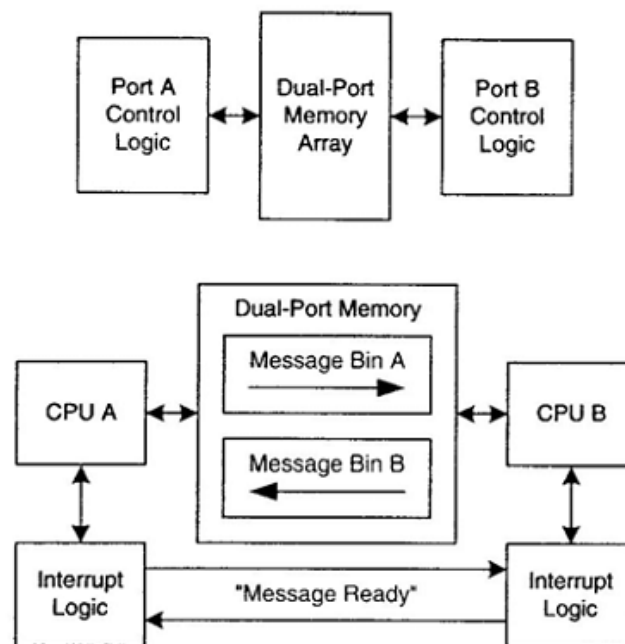


Sl. 2.2.2. Raspored keš memorije

2.3. Multiport memorija

Većina memorijskih uređaja, bilo kratkotrajne ili dugotrajne, sadrže jedan interfejs preko koga se pristupa njihovom sadržaju. U kontekstu sa osnovnim kompjuterskim sistemima sa jednim procesorom, ova jedno-portna arhitektura je sasvim zadovoljavajuća. Postoje neke arhitektura kod kojih više procesora ili logičkih blokova traže pristup istom delu memorije. Deljeni deo memorije može da bude konstruisan na više načina. Prvo, konvencionalni DRAM ili SRAM može da se kombinuje sa eksternom logikom koja uzima zahteve od više odvojenih entiteta (npr. mikroprocesora) i dozvoljava pristup jednom zahtevu u jednom trenutku. Kada je deljena memorija velika, i kada nije potreban simultani pristup većem broju zahteva, arbitracija je efikasan mehanizam. Ipak, kompleksnost kompjuterske logike za arbitraciju je prevelika za male deljive memorije i arbitracija ne omogućava simultani pristup. Sredstvo deljenja memorije bez arbitracione logike i sa mogućnošću simultanog pristupa je konstruisanje pravog višeportnog memorijskog elementa.

Višeportna memorija dozvoljava simultani pristup nekolicini spoljnih entiteta. Svaki port može da bude sposoban za čitanje/pisanje, samo za čitanje ili samo za pisanje, zavisno od implementacije i primene. Višeportne memorije se prave uglavnom male, zbog svoje kompleksnosti, i dakle, cena raste značajno kako se dodaju portovi, svaki sa svojom logikom za dekodiranje i kontrolu. Većina višeportnih memorija su dvo-portni elementi kao što je prikazano na slici 2.3.1.



Sl. 2.3.1. Primeri dvo-portnih memorija

Prava dvo-portna memorija ne stavlja ograničenja na bilo koju od transakcija na portovima u bilo kom vremenskom trenutku. Odgovornost je inženjera da osigura da jedan potražilac ne dođe u konflikt sa drugim. Konflikti se javljaju kada jedan potražilac piše memorijsku lokaciju dok drugi pokušava da čita ili piše istu lokaciju. Ako dođe do simultanog čitanja/pisanja, koje podatke vidi onaj što čita? Da li su to podaci pre ili posle pisanja? Slično, ako se dva procesa upisa jave u isto vreme, koji pobeđuje? Dok se ovakvi problemi mogu rešiti za našu specifičnu primenu posebnom logikom, bezbednije je da se ne brine o takvim graničnim slučajevima. Umesto toga, dizajn sistema bi trebalo da izbegava takve konflikte osim ako ne postoje važni razlozi da se to učini. Jedna uobičajena primena dvo-portnih memorija je deljenje informacija između dva mikroprocesora kao što je prikazano na slici. Dvo-portna memorija leži između mikroprocesora i može da se izdela u odvojene korpe za poruke, ili memorijska polja za svaku stranu. Korpa A sadrži poruke koje je napisao CPU A a čita ih CPU B. Korpa B sadrži poruke koje je napisao CPU B a čita ih CPU A. Obaveštavanje o poruci koja čeka na čitanje se izvodi preko prekida CPU-a, pa samim tim, oslobađa CPU stalnog nadgledanja memorije čekajući da stigne poruka. Čitav proces može da radi ovako:

1. CPU A piše poruku za CPU B u Korpu A.
2. CPU A postavlja prekid na CPU B indicirajući da postoji poruka koja čeka u Korpi A.
3. CPU B čita poruku iz Korpe A.
4. CPU B potvrđuje prekid koji je postavio CPU A.
5. CPU A skida prekid sa CPU-a B.

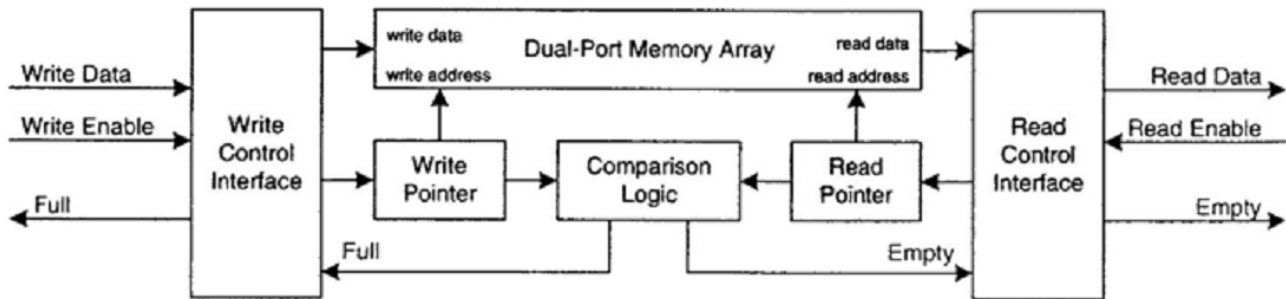
Ovakva implementacija sprečava konflikte dvo-portne memorije zato što jedan CPU neće čitati poruke pre nego što je potpuno napiše drugi CPU i nijedan CPU ne piše u obe korpe.

2.4. FIFO memorija

Memorijski uređaji o kojima se do sada diskutovalo su u osnovi linearni nizovi bitova okruženi minimalnom količinom logike interfejsa da se pomeraju bitovi između porta i niza. First-in-first-out (FIFO) memorije su uređaji posebne namene koji implementiraju osnovnu strukturu reda koja ima široku primenu u računarstvu i komunikacionoj arhitekturi. Za razliku od memorijskih uređaja, tipičan FIFO ima dva jednodirekciona porta bez ulaza za adrese: jedan za pisanje i drugi za čitanje. Kao što samo ime kaže, prvi podatak upisan se prvi čita, a poslednji podatak koji se upiše se čita poslednji. FIFO nije memorija sa slučajnim pristupom već sekvencijalna memorija. Zato, za razliku od konvencionalnih memorija, jedanput kada se element podataka pročita, ne može da se pročita ponovo, zato što će sledeće čitanje vratiti

sledeći element podataka upisan u FIFO. Po svojoj prirodi, FIFO-i su podležni uslovima overflow (prepunjenju) i underflow (ispražnjenju). Njihova konačna veličina, često nazivana dubina, znači da mogu da se napune ako ne dolazi do čitanja i pražnjenja već napisanih podataka. Overflow se dešava kada se pokušava pisanje novih podataka u pun FIFO. Slično, prazan FIFO ne može da da podatke pri čitanju, što dovodi do underflow-a.

FIFO se stvara okruživanjem dvo-portnog memorijskog niza – uglavnom SRAM-a, ali može da se radi i sa DRAM-om za određene primene – sa pokazivačem za pisanje, pokazivačem za čitanje i kontrolnom logikom kao što je prikazano na slici 2.4.1.



Sl. 2.4.1. FIFO memorija

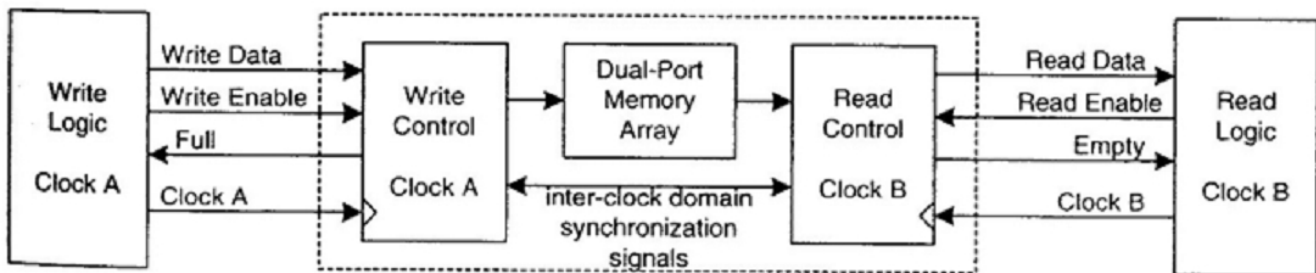
FIFO se ne adresira linearno; već se pravi da formira neprekidan krug memorije koja je adresirana pomoću dva pokazivača. Popunjenost FIFO-a se određuje ne po apsolutnim vrednostima pokazivača, već po njihovim relativnim vrednostima. Prazan FIFO počinje sa pokazivačima za čitanje i pisanje na istoj vrednosti. Kako se pišu podaci, pokazivač za pisanje se povećava. Kako se podaci čitaju, pokazivač za čitanje se takođe povećava. Ako pokazivač za čitanje stigne pokazivač za pisanje, tako da se oba poklapaju, FIFO je ponovo prazan. Ako pokazivač za čitanje ne uspeva da se pomera, pokazivač za čitanje će napraviti krug i u nekom slučaju će postati isti kao pokazivač za čitanje. U ovom trenutku, FIFO je pun i ne može da primi nikakve nove informacije, dok se čitanje ne nastavi. Flegovi za pun i prazan FIFO se generišu da bi se obezbedio status logici za pisanje i čitanje. Neki FIFO-i sadrže detaljnije statuse o popunjenosti, kao što su signali koji predstavljaju programabilni prag popunjenosti.

FIFO interfejs može biti asinhron (bez takta) ili sinhron (sa taktom). Ako je sinhron, dva porta mogu da se dizajniraju da rade sa zajedničkim taktom ili različitim taktovima. Iako se stariji asinhroni FIFO-i i dalje proizvode, sinhroni FIFO je sada češći. - 47 - Sinhroni FIFO ima prednosti poboljšanog tajminga interfejsa, zato što flopovi smešteni na ulaz i izlaz uređaja smanjuju vremenske potrebe za poznata podešavanja, čekanje i clock-to-out specifikacije. Bez takvog registrovanog interfejsa, specifikacija tajminga postaje funkcija unutrašnje logike uređaja.

Jedna uobičajena uloga koji FIFO ima je prelaženje iz domena takta. U takvoj primeni, nema potrebe da se komunicira serijom vrednosti podataka sa logičkog bloka na jednom taktu sa drugim blokom koji radi na drugom taktu. Razmena podataka između domena taktova zahteva posebnu pažnju, zato što normalno ne postoji nijedan način da se izvrši konvencionalna analiza

tajminga između dva različita takta da bi se garatovala adekvatna podešavanja i vreme zadržke na odredišnom flopu. Može se koristiti ili asinhroni FIFO ili sinhroni dual-clock FIFO da se reši ovaj problem, kao što je prikazano na slici 2.4.2.

Dvoportna memorija u srcu FIFO-a je asinhroni element kojoj može da pridelogika koja radi na bilo kome od dva takt domena. Dual-clock sinhroni FIFO je dizajniran da radi sa razlikom u taktovima između polovina uređaja. Kada se jedan ili više bitova upisuje na taktu A, informacija o pokazivaču pisanja se pažljivo prenosi na drugu stranu na domen takta B kroz FIFO-a preko inter-clock sinhronizacije logike. Ovo omogućava kontroli interfejsa za čitanje da zna da postoje podaci koji čekaju da budu pročitani. Logika na taktu B može da čita te podatke dugo nako što su oni sigurno upisani u memorijski niz i postavljeni u stabilno stanje.



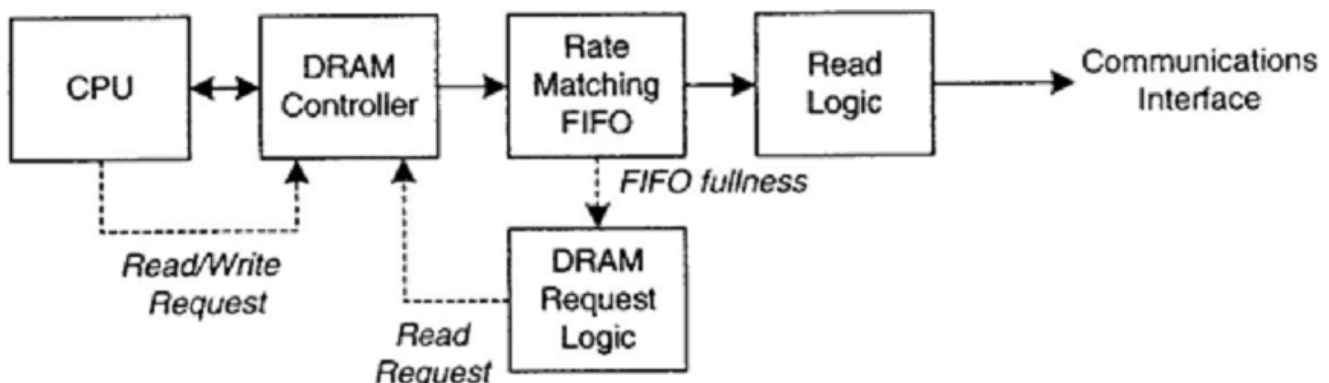
Sl. 2.4.2. Sinhroni dual-clock FIFO

Druga uobičajena primena FIFO-a je podešavanje brzina gde neki izvor podataka daje veliku količinu podataka odjednom a korisnik podataka uzima podatke redovnije. Jedan primer je situacija gde je sekvenca podataka smeštena u DRAM i potrebno je da se pročita i pošalje na komunikacioni interfejs bajt po bajt. DRAM je deljen sa CPU-om koji se nadmeće sa komunikacionim interfejsom za memorijski protok. Poznato je da je DRAM efikasniji kada se radi u modu straničenja. Zato, da se ne bi vršila čitava transakcija čitanja svaki put kada je potreban jedan bajt komunikacionom interfejsu, tok podataka se čita i smešta u FIFO. Svaki put kada je interfejs spreman za novi bajt, čita podatke iz FIFO-a. U ovom slučaju, potreban je FIFO sa samo jednim taktom, zato što ovi uređaji rade na zajedničkom takt domenu. Da bi ovaj proces išao glatko, kontrolna logika mora da nadgleda stanje FIFO-a i vrši prenos iz DRAM-a kada FIFO-u ponestane podataka. Šema je ilustrovana na slici 2.4.3.

Da bi podešavanje brzine protoka podataka radilo pravilno, prosečan protok tokom vremena na ulaznom i izlaznom portu FIFO-a mora biti jednak, zato što je kapacitet FIFO-a konačan. Ako se podaci stalno brže upisuju nego što mogu da budu pročitani, FIFO će se nekada prepuniti i podaci će početi da se gube. Slično, ako se stalno podaci čitaju brže nego što se upisuju, FIFO će se isprazniti i nepravilni bajtovi će početi da se ubacuju u izlazni tok. Dubina FIFO-a pokazuje koliko velika razlika čitanja/pisanja može da se toleriše bez gubljenja podataka. Razlika se izražava kao proizvod razlike brzina i vremena. Mala razlika može da se toleriše duži vremenski period, i veća razlika može da se toleriše kraće.

U primeru podešavanja brzina, velika razlika kratkog trajanja se balansira tokom dužeg trajanja. Kada se DRAM čita, velika količina podataka se odjednom upisuje u FIFO, stvarajući privremeno veliku nejednakost. Tokom vremena, komunikacioni interfejs čita jedan bajt u trenutku dok ne dolazi do pisanja, tako kompenzuje nejednakost tokom vremena.

Čitanje DRAM-a za ponovnoo punjenje FIFO-a mora da se vremenski pažljivo podesi, da se simultano spreči uslovi prepunjenja i ispražnjenja. Mora da se postavi prag popunjenosti FIFO-a, ispod koga se pokreće čitanje DRAM-a. Ovaj prag mora da garantuje da postoji dovoljno mesta u FIFO-u da bi se primila količina podataka iz DRAM-a, izbegavajući prepunjenje. Takođe mora da se garantuje da u najgorem mogućem vremenu odziva DRAM-a postoji dovoljna količina podataka u FIFO-u da se zadovolji komunikacioni interfejs, sprečavajući ispražnjenje. U većini sistema, vreme između postavljanja zahteva za čitanje DRAM-a i stvarnog dobijanja podataka je promenljivo. Ova promenljivost se javlja usled nadmetanja FIFO-a sa ostalim zahtevima (npr. CPU) i čekanjem da se završe skupe operacije (npr. osvežavanje).

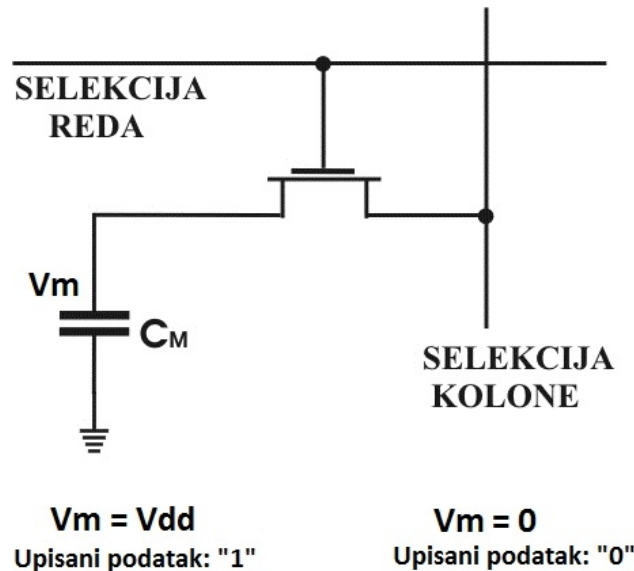


Sl. 2.4.3. Odloženo čitanje iz FIFO-a bajt po bajt

2.5. Dinamički RAM (DRAM)

Da bi se realizovala memorija sa većom gustinom pakovanja konstruisana je memorija sa samo jednim tranzistorom i jednim kondenzatorom po memorijskoj ćeliji. Ovakva memorija bazira pamćenje informacije na električnom punjenju kondenzatora. Na slici 2.5.1 je prikazana jedna ćelija DRAM memorije, kod koje se informacija pamti na taj način što se prilikom upisa, na liniju "selekcija kolon" dovede ili napon VDD ili logička 0, a kada se selektuje red, MOSFET tranzistor postaje provodan tako da se kondenzator CM napuni na napon VDD ili se isprazni na 0V. Prilikom čitanja memorijske ćelije, takođe se selektuje red tako da se, kroz provodni tranzistor, napon sa kondenzatora CM prenosi na liniju selekcije kolone i može se pročitati. Da bi se postigla velika gustina pakovanja, kondenzator u memorijskim ćelijama je veoma malih dimenzija, pa je i kapacitivnost kondenzatora veoma mala reda (10-13F). Kada bi

otpornost MOS tranzistota, za vreme dok je neprovođan, bila beskonačno velika, napon na kondenzatoru bi ostao nepromenjen sve do ponovne selekcije reda. Zbog konačne otpornosti neprovođnog MOS tranzistora, a i zbog male kapacitivnosti kondenzatora C_M , zapamćeni napon na kondenzatoru, kada je zapamćena logička 1, eksponencijalno opada i nakon nekoliko ms, zapamćena informacija bi se izgubila. Da se ovo ne bi dogodilo, svakih 2-4 ms, treba ponovo upisivati informaciju u memorijsku ćeliju. Ponovni upis se naziva "osvežavanje" sadržaja, a RAM memorija koja sadrži ćelije kojima je neophodno periodično osvežavanje se naziva dinamička RAM (DRAM) memorija.



Sl. 2.5.1. Jednotranzistorska memorijska ćelija

Pošto se jedan bit podatka pamti samo sa jednim tranzistorom i jednim kondenzatorom, za razliku od statičke RAM memorije koja za pamćenje jednog bita koristi četiri do šest tranzistora, dinamička RAM memorija je jednostavnija za proizvodnju, može gušće da se pakuje i osetno je jeftinija. Sa druge strane dinamički RAM je sporiji od statičkog, i mora se povremeno "osvežavati" pošto se kondenzatori koji čine ćelije podataka isprazne svaki put kada se podatak pročita.

DRAM memorija može da komunicira sinhrono i asinhrono. **Asinhrono** – Za svaki podatak koji želimo da dobijemo iz memorije, memorija izvršava nekoliko manjih naredbi. U asinhronom modu procesor bi zatražio neki podatak iz memorije i čekao da taj podatak dobije, ali za to vreme ne bi radio ništa drugo. **Sinhrono** – U sinhronom modu procesor zatraži podatak iz memorije, zna kada će taj podatak biti dostupan te za to vreme čekanja izvršava druge stvari i vraća se po podatak tačno kad on postane dostupan.

U kompjuterskoj industriji prvo su se pojavile asinhronne DRAM memorije i u daljem tekstu su navedene neke od njih.

FPM (Fast Page Mode DRAM) – FPM memorija je bila najkorišćenija vrsta DRAM

memorije u personalnim računarima. Prednost ove memorije je da se red bira jednom, a kolona koliko je puta potrebno. Vrlo velik napredak pri manipulaciji nizovima podataka.

EDO (Extended Data Out) – EDO memorija je ustvari poboljšana verzija FPM memorije, a napredak je u tome da čim je selektovan prvi bit podatka memorija počinje da traži adresu sledećeg bita. Ovime se postiže preklapanje operacija tj. izvršenje više operacija u isto vreme. Korišćena je kao memorija za grafiku zbog niske cijene.

Burst EDO – Poboljšana EDO memorija koja pri selektovanju kolone automatski selektuje sledeće tri. Neuspešna memorija zbog orijentacije proizvođača na SDRAM memoriju koja se pojavila u međuvremenu.

MDRAM (Multibank DRAM) – memorija koja svoj kapacitet deli na manje blokove i pruža mogućnost manipulacije sa dva različita bloka u isto vreme. Korišćena je za grafičke kartice.

VRAM (Video RAM) – video verzija FPM memorije koja je unapređena jer je omogućen dvostruki istovremeni pristup memoriji. S tim da je jedan pristup read-only. Na primer, jednim pristupom procesor crta sliku i upisuje u memoriju, a drugi pristup čita to isto i ispisuje na ekran.

WRAM (Window RAM) – grafička memorija sa dvostrukim pristupom napravljena sa ciljem da zameni VRAM-a. Bila je 50% brža i 20% jeftinija, ali opet je SDRAM imao veću podršku. Korišćena je u Matrox Millennium grafičkim karticama.

Za razliku od asinhronih DRAM memorija koje su se ranije koristile, a koje su imale asinhroni interfejs prema procesoru, noviji modeli računara koriste sinhronu DRAM memorije. Sinhroni dinamički RAM (SDRAM) je bio najznačajniji i najzastupljeniji tip memorije na tržištu. SDR SDRAM (SDR - Single Data Rate) se pokazao kao veoma vredna tehnologija - dozvoljava efikasniju komunikaciju sa procesorom i povećava brzinu rada na 133 MHz prilagodivši se rastu brzine systemske magistrale. Međutim SDR SDRAM ne može da radi stabilno na brzinama većim od 140 MHz . Međutim razvijena je nova memorijska arhitektura koja korišćenjem obične SDRAM tehnologije može da radi na taktovima do 266 MHz - DDR SDRAM memorija (DDR - Double Data Rate). Da bi se prevazišla ograničena brzina kojom memorija može da komunicira sa ostalim komponentama, mora se napraviti memorija koja šalje mnoštvo signala u istom ciklusu takta. Rambus je firma koja je to ostvarila kroz posebnu memorijsku arhitekturu koja se naziva Rambus DRAM, ili skraćeno RDRAM. Ova memorija može bez problema da radi na čak 800 MHz. Ipak, ovaj izuzetno visok takt još uvek ne znači da je RDRAM zapravo nekoliko puta brži od SDRAM-a. Dok SDRAM podatke šalje kroz punu 64-bitnu magistralu, magistrala podataka kod Rambus memorije je široka samo 16 bita, što znači da RDRAM mora da obavi četiri memorijska ciklusa da bi postigao transfer koji SDRAM obavi u jednom ciklusu. DDR SDRAM se ponekad naziva DDR1 memorijom. Iako je prilično zastarela može se naći u prodavnicama, ali brzo je sa tržišta potiskuju nove tehnologije poput DDR2 i DDR3 koje daleko jeftinije i nude bolje performanse.

2.6. Memorija za grafičke kartice (GDDR)

Do pre nekog vremena, memorije koje su se koristile na grafičkim karticama bile su iste kao i SDRAM memorije u računaru. Ali s vremenom, grafičke kartice su počele zahtevati više performansi, pa se u te svrhe razvila GDDR (GDDR2, GDDR3, itd.) memorija, koja koristi nešto izmenjenu SDRAM memoriju. Golim okom nije moguće utvrditi da li se radi o DDR ili GDDR memoriji, osim ako nije naznačeno na čipu, ali pretragom oznake moguće je naći informacije na internetu. Neki primeri GDDR memorija su dati na slikama 2.6.1 i 2.6.2, a u daljem tekstu navešćemo vrste i karakteristike GDDR memorija.



Sl. 2.6.1. Qimonda GDDR čip



Sl. 2.6.2. Samsung GDDR čip

GDDR1

Ovo je prva generacija GDDR-a i napon VDD/VDDQ iznosi 2.5/2.5 V. Memorijski bus clock je od 183 do 500 Mhz, a latencije (kašnjenja) su 3, 4 i 5 clock ciklusa, što je rezultovalo maksimalnom stopom protoka informacija od 16 GB/s sa 128-bitnom magistralom.

GDDR2

GDDR2 predstavlja blago unapređeni GDDR1. Karakteristike GDDR2 memorije su, napon VDD/VDDQ od 2.5/1.8 V, memorijski bus clock od 400 do 500 MHz i latencije od 5, 6 i 7 clock ciklusa. Najveća brzina protoka informacija je 32 GB/s sa 256-bitnom magistralom. GDDR2 memorija nije bila kvalitetno rešenje jer sa povećanjem napona, naglo je dolazilo do porasta potrošnje, a samim time i porastom temperature pa su se ovi čipovi lako pregrevali. Brzo je zamenjena GDDR3 memorijom.

GDDR3

GDDR3 se isporučuje sa naponom VDD/VDDQ od 1.8/1.8 V. Sa clock-om stope od 500 do 800 MHz i latencije 5-9 clock ciklusa. Najveća brzina protoka informacija je 51,2 GB/s i moguća je uz 256-bitni bus. GDDR3 radi kao DDR2 Speicher-prefekt quad, tako da će preneti četiri informacije svaka dva ciklusa obrade. GDDR3 memorija je bila znatno više u upotrebi od svojih prethodnika jer je rešila problem pregrevavanja i niskog clock-a.

GDDR4

GDDR4 memorija je počela da se razvija 2005. godine od strane Samsunga. Učestanost clock-a je ograničena na 1.6 Ghz, ali je kašnjenje veće u odnosu na GDDR3 module. Teoretski maksimalni bandwidth je oko 100 GB/s sa 256-bitnom magistralom. GDDR4 radi kao DDR3-Speicher prefekt eightfold i troši oko 45% manje energije od GDDR3.

GDDR5

GDDR5 (Graphics Double Data Rate, verzija 5) je vrsta GDDR memorije za grafičke kartice. To je naslednik GDDR4, a trenutno je dostupna široj javnosti u nekim ATI grafičkim karticama (HD4800 serija).

2.7. Nove SDRAM tehnologije (DDR2 i DDR3)

DDR2 SDRAM

DDR2 SDRAM ima nekoliko unapređenja u odnosu na DDR SDRAM. Taktne brzine DDR2 SDRAM su veće, pa se stoga povećavaju i brzine podataka u memoriji (vidi Tabelu 2). Integritet signala postaje mnogo važniji za pouzdan rad memorija sa povećanjem taktnih brzina. Sa povećanjem taktnih brzina, sve je važniji dizajn i ležaut prenosnih linija i njihovih krajeva.

Eliminisanje signala adresa, taktova i komandi je jednostavno jer su ovi signali jednosmerni i završavaju na ploči kola. Signali podataka su dvosmerni. Memorijski kontroler vodi ove signale tokom operacije upisivanja, dok ih DDR2 SDRAM vodi tokom operacije čitanja. Složenost se povećava povezivanjem višestrukih DDR2 SDRAM na iste signale podataka. Ovi višestruki DDR2 SDRAM mogu biti na istom DIMM ili na različitim DIMM u memorijskom sistemu. Kao rezultat imamo da se drajveri i prijemnici za podatke stalno menjaju usled operacija čitanja / pisanja, i zavisno od toga kojem DDR2 SDRAM se pristupa.

DDR2 SDRAM poboljšava integritet signala podataka omogućavanjem ODT opcije (eliminisanja na kraju), tako da ODT signal omogućava ovu opciju i daje sposobnost programu da odredi vrednosti za eliminisanje signala (75 oma, 150 oma,...) sa DDR2 SDRAM registrima. ODT vrednost i rad se kontrolišu preko memorijskog kontrolera i one su funkcije DDR2 SDRAM lokacije DIMM i tipa memorijske operacije (čitanje ili upisivanje). Rad ODT pruža bolji integritet signala kreiranjem velikog dijagrama oka za prozor važećih podataka, sa povećanim marginama napona, brzinama odziva, smanjenim prekoračenjem i smanjenom ISI (međusimbolskom interferencijom).

DDR2 SDRAM smanjuje snagu memorijskog sistema jer radi na 1.8 V, što je 72% od napona za DDR SDRAM (2.5 V). U nekim primenama se smanjuje broj kolona u jednom redu, što dodatno daje smanjenje u utrošku snage pri aktivaciji reda za čitanje ili upisivanje.

Još jedna prednost manjih radnih napona je smanjenje razlike u naponima logičkih nivoa. Za iste brzine odziva ovo daje povećanje brzine prelazaka iz jednog stanja u drugo i samim tim se omogućava podrška većim taktnim brzinama. Može se isprogramirati i korišćenje diferencijalnog signala. Njihovim korišćenjem se smanjuje šum, preslušavanje, dinamička potrošnja snage i EMI (elektromagnetna interferencija), dok se povećava margina šuma.

Nova osobina koja je uvedena u DDR2 SDRAM je dodatno kašnjenje, koje omogućava fleksibilnost memorijskom kontroleru prilikom slanja komandi čitanja i upisivanja odmah nakon komande aktiviranja. Ovo optimizira propusnu moć memorije i konfiguriše se programiranjem pomoću registara DDR2 SDRAM.

DDR2 SDRAM poboljšava propusni opseg za module od 1 Gb i 2 Gb korišćenjem osam grupa. Ovih osam grupa povećavaju fleksibilnost prilikom pristupa velikim DDR2 SDRAM memorijama preplitanjem različitih operacija nad memorij-skim grupama. Takođe, za velike memorije, DDR2 SDRAM podržava sve do osam kratkih sekvencijalnih signala.

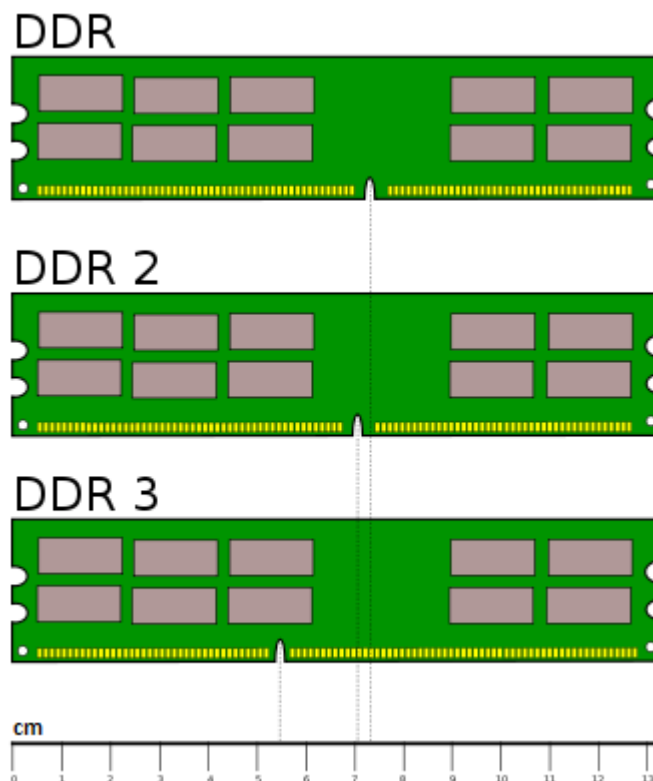
DDR3 SDRAM

DDR3 SDRAM se po razvoju performansi nalazi nakon DDR2 SDRAM-a. DDR3 SDRAM podržava narednu generaciju većih brzina podataka i većih taktnih brzina što je i prikazano u tabeli 2.7.1.

DDR3 SDRAM	Data Rate	Memory Clock
DDR3-800	800 Mb/s/pin	400 MHz
DDR3-1066	1066Mb/s/pin	533 MHz
DDR3-1333	1333Mb/s/pin	667 MHz
DDR3-1600	1600 Mb/s/pin	800 MHz

Tab. 2.7.1. Brzine podataka i taktne brzine DDR3 SDRAM

Ostale promene obuhvataju smanjenje radnog napona za DDR3 SDRAM memorije na 1.5 V, što je 83% napona korišćenog za DDR2 SDRAM od 1.8 V. DDR3. Takođe maksimalni kapacitet modula je povećan na 16 GB. DDR3 SDRAM je relativno nova tehnologija i u današnjim računarima se uglavnom još uvek koristi DDR2 SDRAM memorija. Raščuna se da će prodaja DDR3 modula preći prodaju DDR2 modula 2010. godine i da će tad DDR3 SDRAM postati defakto standard. Na slici 2.7.1 su prokazana fizička poređenja DDR1, DDR2 i DDR3 modula.



Sl. 2.7.1 Poređenje DDR1, DDR2 i DDR3 modula