ذخیره و بازیابی اطلاعات

رسانه های ذخیره سازی

تعریف حافظه

به هر وسیله ای که توانایی حفظ اطلاعات را داشته باشد بگونه ای که کاربر بتواند در هر زمان به داده های آن دستیابی داشته باشد را گویند.چنین سیستمی دارای ۲ محیط درون ماشینی و برون ماشینی می باشد.محیط درون ماشینی از کامپیوتر با اجزا و عناصر داخلی اش و محیط برون ماشینی از دستگاههای جانبی تشکیل شده است.

محیط درون ماشینی شامل دو قسمت است:

1. **MEMORY** (ROM , RAM , Data Instructions)

2. **CPU** (CU , ALU , Local Memory , Registers , Buffers , Flags , Index , Accumulator)

و محیط برون ماشینی شامل ۴ قسمت می باشد:

1. TERMINALS (Keyboard , Video display)

2. Extended Memory (Hard disk , Floppy disk , DVD , CD)

3. Output Device (Printer , Plotter , Card punches)

4. Input Devices (Laser disk drives , Card readers , Magnetic cassette)

خصوصیات حافظه:

1) **خواندن و نوشتن** : هر حافظه باید این قابلیت را داشته باشد که در آن بتواند اطلاعات را نوشت (درج کرد) و یا بتوان آن اطلاعات را خواند (بازیابی کرد) مانند هارد دیسک.
2) **قابلیت آدرس پذیری** : (به هر قسمتی از حافظه بتوانیم دستیابی داشته باشیم) هر حافظه را می توان با استفاده از مکانیزم آدرس دهی مورد دستیابی قرار داد. دستیابی ممکن است به منظور خواندن یا نوشتن در حافظه صورت گیرد. مدت زمانی را بین لحظه دستور خواندن یا نوشتن در حافظه و زمانی که داده دستیابی می شود را زمان دستیابی یا access time گفته می شود.
3) **آدرس پذیر بودن**: هر حافظه مجهز به یک مکانیزم آدرس است به تعبیر دیگر می توان به کمک این قابلیت به مکانی از حافظه دستیابی پیدا کرد.
4) **نرخ انتقال یا سرعت انتقال**: به میزان اطلاعاتی که در واحد زمان در حافظه قابل انتقال است نرخ انتقال گفته می شود.
5) **ظرفیت**: هر حافظه دارای ظرفیتی است که با بیت یا بایت یا با آدرس پارامترهای وابسته بیان می شود.
6) **مانا یا نامانا بودن اطلاعات:** برخی از حافظه ها اطلاعات را به صورت دائمی نگهداری می کنند (مانا) و برخی حافظه ها اطلاعات را به صورت موقت در خود نگهداری می کنند (نامانا).

هرچه از پایین به بالا می رویم سرعت دستیابی افزایش و هزینه یک بیت نیز افزایش پیدا می کند.
هر چه از بالا به پایین می آییم ظرفیت ها بیشتر می شود.

واحدهای حافظه عبارتند از:

$$1KB = 2^{10}B,\ 1MB = 2^{20}B\ ,1GB = 2^{30}B\ ,1TB = 2^{40}\text{B}$$

مقایسه حافظه های درون ماشینی و برون ماشینی

۱- ظرفیت حافظه‌های برون ماشینی بیشتر است.

۲- سرعت حافظه‌های برون ماشینی کمتر است.

۳- حافظه‌های درون ماشینی نامانا می باشند. (اطلاعات ذخیره شده می‌تواند از بین برود)

۴- حافظه‌های درون ماشینی گران می باشند.

علت استفاده از حافظه‌های برون ماشینی

۱- محدود بودن ظرفیت حافظه‌های درون ماشینی

۲- گران بودن رسانه‌های ذخیره سازی سریع

۳- عدم لزوم ذخیره همه اطلاعات در حافظه‌های درون ماشینی

۴- نامانا (non volatile) بودن حافظه‌های درون ماشینی

۵- نیاز به دسترسی همروند (concurrent) به داده‌ها توسط چند پروسس.

ذخیره و بازیابی اطلاعات به نحوه ی ذخیره ی داده ها در فایل ها و عملیات لازم جهت دستیابی به داده ها بر روی حافظه های جانبی می پردازد. هدف از اینکار صرفه جویی در مصرف حافظه و بالا بردن سرعت عملیات بر روی حافظه های جانبی است.

نوار مغناطیسی :

رسانایی است از جنس پلاستیک یا غشای مغناطیسی ؛ اطلاعات بر روی شیارهایی که بر سطح نوار قرار گرفته اند ذخیره می شود. این نوار توسط دستگاه نوار خوان به حرکت در می آید . هر دستگاه نوار خوان دارای یک نوک (head) خواندن و نوشتن می باشد. نوار مغناطیسی اصولا برای پردازش ترتیبی یا پی در پی اطلاعات استفاده می گردد.

انواع نوار مغناطیسی از نظر تکنولوژی:ریل به ریل، کارتریج، کاست و صوتی.

- انواع نوار از نظر تعداد شیار عبارتند از: نوارهای ۷ شیاره و نوارهای ۹ شیاره.
- در هر نوار یکی از شیارها، کنترل کننده پاریتی (Parity) می‌باشند.
- داده ها روی نوار مغناطیسی به صورت رشته های بیتی روی شیارهایی که در سطح نوار وجود دارد، ذخیره می شوند.

چگالی نوار (Density)

تعداد بیت‌های قابل ضبط در هر اینچ نوار را چگالی نوار می‌گویند. که واحد آن بیت در اینچ(bpi) است، که با توجه به نحوه نشست کاراکترها بر روی شیارها همان بایت در اینچ یا کاراکتر در اینچ است.

پارامترهای اصلی نوار

سرعت : مسافت طی شده در واحد زمان (اینچ بر ثانیه)

چگالی : میزان اطلاعات ذخیره شده در ایچ

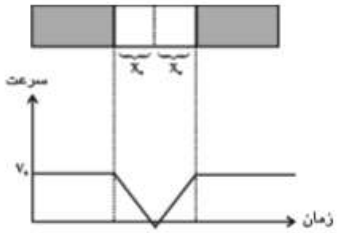
طول نوار : بر حسب ایچ یا فوت

نرخ انتقال : میزان اطلاعات منتقل شده در واحد زمان

ظرفیت اسمی : میزان گنجایش اعلام شده توسط کارخانه سازنده نوار که روی بسته قید شده است.

گپ (Gap)

فضایی بدون استفاده (هرز) (waste) بین دو رکورد (IRG) یا دو بلاک (IBG) ، که برای متوقف کردن نوار و یا حرکت دوباره آن بکار می‌رود. یعنی برای اینکه نوک خواندن / نوشتن بتواند داده‌های ذخیره شده را حس(sense) کند باید بعد از توقف به سرعتی مطلوب و یکنواخت (سرعت حس) برسد که تکه‌ای از نوار از نوک R/W خواهد گذشت. همچنین در هنگام کاهش سرعت حس تا توقف نیز تکه‌ای از نوار از زیر نوک خواهد گذشت. این دو تکه، گپ را تشکیل می‌دهند.



و می توان نوشت : $t_0 = \frac{2x_0}{V_0} = \frac{IBG}{V_0}$

V_0: سرعت لازم برای حس کردن داده‌ها با واحد اینچ بر ثانیه

IBG : حافظه هرز بین دو بلاک با واحد اینچ یا فوت

t_0 : زمان توقف/ حرکت

به عبارتی : زمان لازم برای صفر شدن سرعت حس یا زمان لازم برای رسیدن به سرعت حس یا شروع از وسط گپ .

هر فوت برابر ۱۲ اینچ می باشد.

مثال:

اگر سرعت حس ۴۵۰ اینچ بر ثانیه و طول حرکت ۰,۹ اینچ باشد زمان را محاسبه کنید.

جواب: ۲ میلی ثانیه

تمرین:

- اگر سرعت حس نوک یک نوار برابر ۱۵۰ اینچ بر ثانیه و زمان رسیدن به این سرعت ۲ میلی ثانیه باشد ،طول گپ را محاسبه کنید.
- نواری با سرعت حس ۱۵۰ اینچ بر ثانیه ب ابتدای گپی بطول ۰,۷۵ اینچ رسیده است . چه زمانی نیاز است تا مجددا به سرعت حس برسد ؟

دیسک مغناطیسی :

نوار مغناطیس دارای محدودیت هایی است ، یکی از آن ها نوع قرار گرفتن داده ها بر روی نوار به صورت ترتیبی می باشد که امکان دسترسی مستقیم به اطلاعات را مقدور نمی سازد. دیسک مغناطیس رسانه ای است با امکان دستیابی مستقیم به داده های ذخیره شده.

دیسک مغناطیس شامل صفحه ای است دوار حول یک محور عمودی ، بر روی این صفحه دوایر با شیارهای متحدالمرکزی وجود دارد که داده ها به صورت سریالی در این شیارها قرار می گیرند. اطلاعات بر روی این شیارها توسط بازوی خواندن و نوشتن نوشته و یا از آن خوانده می شود.

تقسیمات دیسک:

استوانه یا سیلندر: شیارهای هم شعاع بر روی صفحه های هم مرکز بر روی رویه های مختلف را سیلندر گویند.

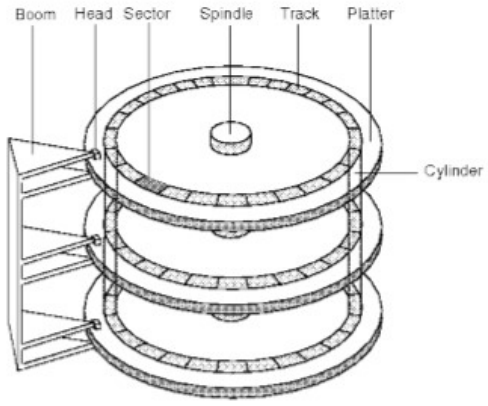
شیار یا track : به دوائر متحدالمرکز که اطلاعات بر روی آن ها در قالب صفر و یک ذخیره می شود شیار گفته می شود.

سکتور یا قطاع : تقسیمات از شیار به اندازه های مساوری را سکتور می گویند.

دو نوع سکتور وجود دارد:

1- **سکتور سخت افزاری** : توسط خود سیستم تعیین می شوند و ما دسترسی به آن ها نداریم.
2- **سکتور نرم افزاری** : ضریب صحیحی از سکتورهای سخت افزاری هستند که توسط کاربر و برنامه نویس تعیین می‌شود.

سکتور کوچکترین بخش از یک دیسک است که قابل آدرس دهی می‌باشد.



دسته بندی دیسک ها

از نظر جا به جا شدن : دیسک های ثابت و دیسک های جا به جا شدنی.

از نظر ثابت بودن یا متحرک بودن : هد خواندن و نوشتن که به دو دسته ی دیسک ها با نوک ثابت و دیسک ها با نوک متحرک تقسیم می شوند.

از نظر تعداد صفحات که روی صفحه عمود قرار می گیرند به دو دسته تقسیم می شوند : دیسک های تک صفحه ای و دیسک های چند صفحه ای ، به دیسک های چند صفحه ای اصطلاحا پک می گویند.

یک پک pack با n صفحه رویه دارای 2n رویه است . که 2n -2 رویه برای ذخیره سازی و دو رویه بالایی و پایینی برای حافظت و کنترل بیت ها استفاده می شود.

از نظر جنس صفحه : به دو دسته دیسک سخت (hard disk) و دیسک انعطاف پذیر یا نرم (floppy disk) تقسیم می شوند.

پارامترهای دیسک

زمان استوانه جویی(seek time)S : زمانی است که جهت رسیدن نوک خواندن و نوشتن و قرار گرفتن بر روی استوانه ای که اطلاعات مورد نظر ما بر روی آن قرار دارد صرف می شود و آن را با **S** نشان می دهند و واحد آن میلی ثانیه است.

زمان انتظار دورانی: زمانی است که پس از رسیدن نوک خواندن و نوشتن بر روی استوانه مورد نظر باید سپری شود تا داده‌ی مورد نظر به زیر نوک خواندن و نوشتن برسد و واحد آن میلی ثانیه است.

متوسط آن را با r نشان می دهیم که نصف زمان لازم برای یک دور دیسک 2r می باشد.$(0\leq r<2r)$

سرعت گردش دیسک : مقدار دورهای دیسک در دقیقه می باشد و با rpm(Rotation per minute) نمایش داده می شود و واحد آن دور به دقیقه است .

نرخ انتقال : تعداد بایتی است که در یک ثانیه به رسانه منتقل و یا از رسانه به سیستم منتقل می شود.

سرعت خواندن دسترسی مستقیم با نرخ انتقال دارد.

دیسکهای نوری

برای افزایش فضای ذخیره‌سازی، سرعت دستیابی و کاهش هزینه از نور به جای مغناطیس استفاده شده است. از انواع دیسکهای نوری می‌توان موارد زیر را نام برد:

1- DVD
2- CD
3- CD-ROM
4- WORM : (Write-Once Read–Many)
5- EOD : (Eraseable Optical Disk)

دیسکهای نوری – مغناطیسی

از تلفیق دو تکنولوژی مغناطیس و نور، دیسکهای نوری– مغناطیسی ایجاد شده اند، که دارای خاصیت دیسک نوری یعنی بالا بودن چگالی و خاصیت دیسک مغناطیسی یعنی قابل پاک شدن می‌باشد.

دو نوع دیگر از دیسکها

۱- دیسکهای با تغییر فاز

رویه (سطح) در دیسکهای با تغییر فاز با تابش اشعه لیزر دو حالت کریستال و یا نامشخص را به خود می گیرد. سرعت آنها دو برابر دیسکهای مغناطیسی– نوری می باشد.

۲- دیسکهای دای پولیمر

رویه در دیسکهای دای پولیمر دو لایه بوده که با لیزر گرم شده و بر آمدگی در آن ایجاد می شود و بعد از سرد کردن آن بر آمدگی ثابت می ماند.

فصل دوم

رکورد ، بلاک و روش های بلاک بندی

همه ی اطلاعات در کامپیوتر به صورت فایل ذخیره می شوند.

دو نوع ذخیره سازی برای رکوردها(هر سطر یک رکورد نام دارد.) وجود دارد:

رکورد با طول ثابت

طول رکورد = تعداد فیلدها * متوسط اندازه مقادیر فیلدها

$R=a*v$

رکورد با طول متغییر :

برای ذخیره سازی هر فیلد باید نام آن فیلد را بیاوریم مثال (نام=مریم، نام خانوادگی = محمدی ، شماره دانشجویی = 870707239)

$R=(A+V+2)*n$

2 در این فرمول برای ذخیره هر فیلد یک = و یک ، جداکننده داریم.

دلایل متغییر شدن طول رکورد

1) ممکن است که طول برخی از فیلدها متغییر باشد . مثلا فیلد مربوط به نام و نام خانوادگی یا آدرس که می تواند طول متغییر داشته باشد.
2) ممکن است که برخی از فیلدها دارای چندین صفت خاصه باشند که در نمونه های مختلف آن تعداد مقادیر متفاوتی داشته باشد به عنوان نمونه دروس انتخاب شده در یک ترم در رکورد اطلاعات دانشجو ، که تعداد دروس برای فیلد ثبت نام متفاوت می باشد.
3) ممکن است برای رکورد ، حالات مختلفی از فیلدها داشته باشیم(تعداد فیلدها متغییر باشد.) به عنوان نمونه در رکورد اطلاعاتی مربوط به کارمندان یک سازمان ، فیلدهای کارمندان رسمی و کارمندان قرار دادی با هم متفاوت اند.

سیستم فایل

سیستم فایل (file system)، نرم افزاری است که وظیفه آن ایجاد و مدیریت فایلها می باشد. این نرم افزار دارای چندین لایه است که کاربر از جزئیات داخلی آن بی اطلاع است. قبل از بررسی سیستم فایل اصطلاح های فیلد، رکورد و فایل را بررسی می نماییم.

موجودیت:

به فرد، شی و یا مفهومی که می خواهیم در رابطه با آن اطلاع داشته باشیم موجودیت گفته می شود.

محیط عملیاتی:

به محیطی که در رابطه با آن می خواهیم یکسری داده ها را ذخیره ،بازیابی و پردازش کنیم.

- محیط عملیاتی دانشگاه از موجودیت های دانشجو ، استاد ، درس و غیره تشکیل شده است.
- هر موجودیت دارای صفت خاصه attributeاست که با عث تمایز آن از سایر موجودیت ها می شود .مثلا موجودیت استاد دارای صفت نام،مدرک،سن و غیره است.
- هر صفت خاصه دارای مقدار خاصه است که اطلاع را بوجود می آورد.

عظیمی=نام خانوادگی

فیلد

مکان ذخیره‌سازی یک فقره اطلاع (یک واحد معنایی داده و نامدار) را فیلد می نامند. طول فیلد، حداقل یک کاراکتر و حداکثر وابسته به زبان برنامه نویسی و سیستم فایل می باشد.

فیلد کوچکترین واحد اطلاعات در فایل است.

رکورد

در بیان معمولی مجموعه ای از فیلدها تشکیل رکورد را می دهند .مجموعه ای از رکوردها نیز فایل را تشکیل می دهند . موجودیت استاد مجموعه فیلدهای (نام استاد-رشته-مدرک-سابقه ی تدریس) را رکورد گویند.

مفهوم رکورد را در سه سطح زیر می توان بررسی کرد:

انتزاعی	مجموعه اطلاعاتی است که در مورد هر یک از نمونه های متمایز یک یا بیش از یک نوع موجودیت از یک محیط عملیاتی در اختیار داریم.
برنامه کاربر (رکورد منطقی)	مجموعه‌ای است با ساختار مشخص که از چندین فیلد تشکیل شده است.
محیط ذخیره سازی (رکورد فیزیکی)	رکورد در سطح فیزیکی از دو قسمت داده‌ای و غیر داده‌ای تشکیل شده است.

در سطح انتزاعی رکورد مجموع اطلاعاتی است که در مورد هر یک از نمونه های متمتیز یک یا بیش از یک نوع موجودیت از یک محیط عملیاتی می خواهیم در دست داشته باشیم.

مثلا میگوییم برای موجودیت دانشجو صفات خاصه نام، فامیلی و معدل را می خواهیم و دانشجویی با نام علی و فامیلی حسینی معدل ۱۷ را دارد .رکورد در این سطح یک شی انتزاعی است.

در سطح منطقی ، رکورد از دیدگاه برنامه نویس مشخص می شود و بدین دلیل به ن رکورد از دیدگاه برنامه ساز هم گویند.

دو طرح کلی برای این ساختار داریم: یک طرح با قالب ثابت مکان و دیگری با قالب غیر ثابت مک

ثابت مکان **(Fixed Positional)**	در طرح رکورد با قالب ثابت مکان، در هر فیلد فقط مقدار صفت خاصه ذخیره می‌شود و اسم صفت خاصه ذخیره نمی‌شود.
غیر ثابت مکان **(Non Fixed Positional)**	در طرح رکورد با قالب غیر ثابت مکان هم اسم صفت خاصه و هم مقدار صفت ذخیره می‌شود و مکان فیلدها از قبل مشخص نمی‌باشد. در این طرح تعداد، طول و مکان فیلدها در نمونه‌های مختلف یک نوع رکورد، یکسان نمی باشد.

رکورد منطقی از نظر طول به دو دسته با طول ثابت و با طول متغیر، تقسیم می شود.

دلایل متغیر شدن طول یک رکورد

۱- متفاوت بودن تعداد صفات خاصه مورد نیاز برای نمونه‌های مختلف یک نوع موجودیت .

۲- متغیر بودن طول فیلدها

۳- وجود صفت خاصه مرکب چند مقداری (پدیده فقره اطلاع تکرار شونده)

مثال برای حالت اول: برای کارمند متاهل به فیلدهایی مانند نام همسر و تعداد فرزندان نیاز است که برای کارمند مجرد نیاز نمی باشد ، برای حالت دوم ، فیلد آدرس دانشجو. و برای حالت سوم ، فیلد شماره درس برای موجودیت دانشجو که در یک نمونه دارای ۲ مقدار و در نمونه ای دیگر دارای ۶ مقدار است. وجود چنین فیلدی ، فایل را نامسطح می کند.

اشاره گر(نشانه رو)

توسط اشاره گرها می‌توان بین رکوردها ارتباط ساختاری برقرار ساخت و نظم منطقی موجود در برنامه را در محیط فیزیکی برقرار کرد. هر اشاره گر یک مبدا و یک مقصد دارد. از این نظر انواع اشاره گرها عبارتند از:

۱- رکورد به رکورد

۲- بلاک به بلاک

۳- رکورد به بلاک

۴- بلاک به رکورد

۵- فایل به فایل

۶- گروهی از بلاکها به گروهی دیگر

انواع اشاره گر از نظر نوع نشانی عبارتند از: نشانی در سطح فیزیکی، نشانی نسبی و نشانی ضمنی.

فایل

فایل مجموعه‌ای است نامدار از نمونه های مختلف یک یا چند نوع رکورد. در حالت اول فایل را تک نوعی و در حالت دوم چند نوعی می گوییم. البته ممکن است فایل دنباله ای از کاراکترها باشد که از نظر سیستم فایل معنای خاصی ندارد.

مفهوم فایل در معنای عام دارای سه ویژگی زیر می باشد:

۱- اندازه بزرگ (طوری که یکباره در حافظه درون ماشینی نگنجد)

۲- پایائی(داده ها از بین نمی روند ، مگر به درخواست کاربر)(Resistance)

۳- اشتراکی بودن بین تعدادی کاربر مجاز(Shared)

کلید

صفت خاصه یا ترکیبی از صفات خاصه که در تمام نمونه های مختلف یک موجودیت ، مقدار یکتایی داشته باشد را کلید گویند .طول کلید حتی الامکان باید کوتاه باشد .

مثلا در فایل اطلاعات دانشجویان فیلد شماره دانشجویی کلید است زیرا هر شماره دانشجویی فقط یک دانشجو را معرفی می کند .(شماره دانشجویی تکراری نیست)

بلاک

بلاک قالبی ساختار یافته متشکل از چند رکورد است. بلاک کمترین مقدار داده قابل مبادلـه بـین بیـرون و درون ماشین توسط سیستم فایل در یک عملI/O می باشد.

بلاک بندی :

قرار دادن چندین رکورد را در قالب بزرگتر جهت عملیات خواندن و نوشتن از رسانه ذخیره سازی ، بلاک بندی می گویند. در واقع بلاک مجموعه ای است از تعدادی رکوردها با طول ثابت.

بلاک کمترین داده ای است که در عملیات خواندن و نوشتن به یا از دیسک منقل می شود. بین هر دو بلاک فضای هرزی وجود دارد که به آن IBG (Inter Block Gap) گفته می شود . ظرفیت بلاک را با B نشان می دهیم.

ضریب بلاک بندی (blocking factor) :

به تعداد رکوردهای موجود در بلاک بندی گفته می شود و آن را با Bf نمایش می دهیم.

$$bf = [\frac{B}{R}]$$

B: اندازه بلاک

R: اندازه رکورد

مزایای بلاک بندی

۱- کاهش گپ‌های بین رکوردها

۲- کاهش زمان پردازش فایل

۳- کاهش دفعات ورودی/خروجی

۴- کاهش زمان اجرای برنامه فایل پرداز

معایب بلاک بندی

۱- مصرف بیشتر حافظه اصلی به علت نیاز به بافر بزرگتر

۲- کار نرم افزاری بیشتر برای بلاک بندی و بلاک گشایی

۳- افزایش احتمال اشتباه در تبادل اطلاعات به علت افزایش مقدار داده منتقل شونده.

روش های مختلف بلاک بندی

۱. بلاک بندی رکوردهای با طول ثابت که معمولا یکپاره یا یکپارچه هستند.

تعداد رکوردها در یک بلاک را فاکتور بلاک بندیBf گویند.

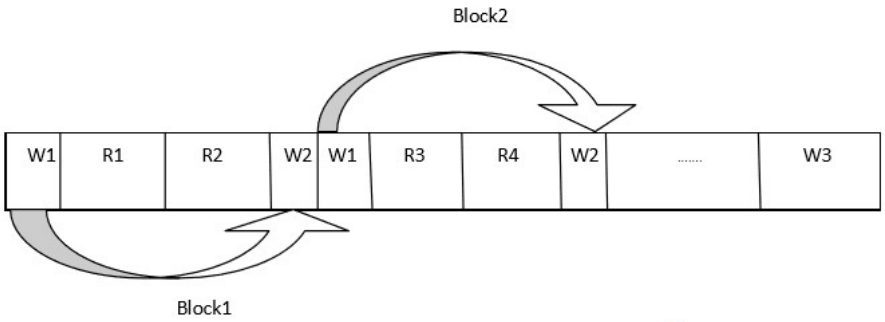
میزان حافظه ی هرزW

میزان حافظه ی هرز به ازای هررکورد WR

تعدادبلاکها در هر شیار را فاکتور تراکینگ و با Tf نشان می دهند.

B طول بلاک(بدون حافظه ی هرز)

R طول رکورد یا متوس طول رکورد (بدون حافظه ی هرز)



ضریب بلاک بندی $\Leftarrow$ $bf = [\frac{B}{R}]$

W1 : فضای هرز مربوط به Gap

W2: فضای هرز ناشی از نگنجیدن آخرین رکورد (به علت ثابت بودن طول رکورد)

W3 : فضای هرز ناشی از نگنجیدن آخرین بلاک در شیار

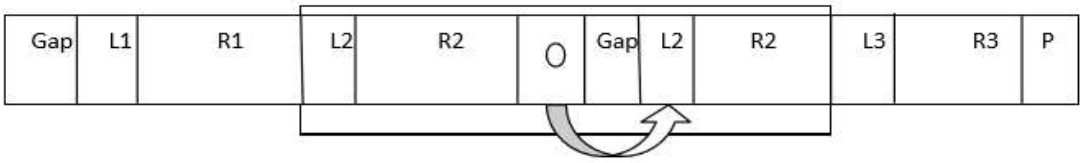
مثال:

در سیستم فایلی بلاکها بطوری انتخاب شده اند که تمام فضای هر شیار را تکمیل می نمایند.اگر طول هر بلاک و رکورد در این فایل بترتیب ۴۲۵ و ۹۰ بایت باشد مقدار Bf را محاسبه کنید.

$$\mathrm{Bf} = \lfloor \frac{B}{R} \rfloor = \frac{425}{90} = 4$$

۲. بلاک بندی رکوردهای با طول متغیر که معمولا دوپاره هستند.

در این مورد R3که به تمامی در بلاک اول نمی گنجد بصورت دوپاره در دو بلاک ذخیره میشود.



ضریب بلاک بندی $bf = [\frac{B-p}{R+p}]$

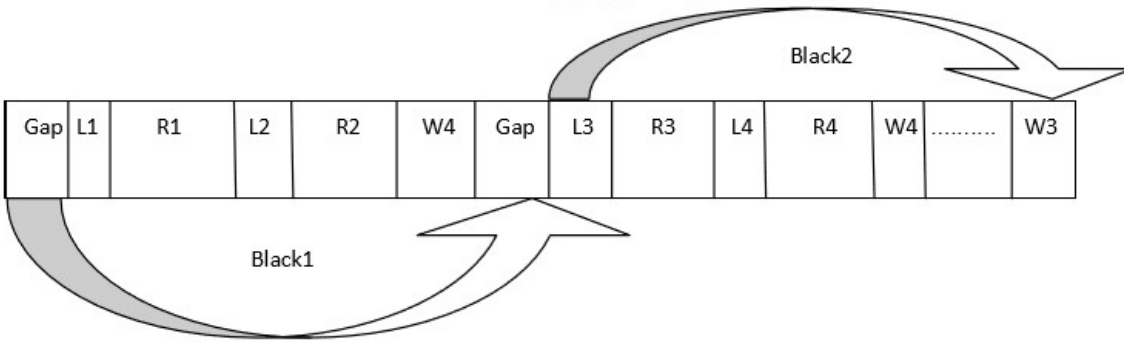
$$bf = [\frac{\text{اندازه بلاک} - point}{\text{طول رکورد} + \text{اندازه رکورد}}]$$

مثال:

در محاسبه ی فاکتور بلاک بندی روی یک رسانه ی ذخیره سازی طول بلاک ۲۰۰۰بایت و طول رکورد ۱۰۰ بایت است درصورتیکه رکوردها بلاک بندیشان بصورت متغیر و دوپاره باشد و طول اشاره گر ۱۰ بایت باشد مقدار Bf را محاسبه کنید.

$$\text{Bf} = \frac{B-P}{R+P} = \frac{2000-10}{100+10} = 18$$

۳. بلاک بندی رکوردهای با طول متغیر که معمولا یکپاره هستند.



در این روش چون جای R3 در بلاک 1 وجود ندارد ، R3 به بلاک دوم رفته در نتیجه فضای هرزی به نام W4 به وجود می آید.

به طور متوسط W4 می تواند صفر یا به اندازه ی یک رکورد باشد.

ضریب بلاک بندی $bf = [\frac{B-w4}{R+p}]$

مقایسه سه روش :

در روش اول پیاده سازی و مدیریت راحت تر است ولی در صورتی که طول رکوردها عوض نشود به همین دلیل روش اول فاقد انعطاف پذیری می باشد. اما در روش دوم و سوم نسبت به انعطاف پذیری بیشتری وجود دارد اما نرم افزار پیچیده تری نیاز می باشد.

روش سوم از نظر حافظه مقرون به صرفه تر می باشد ولی به نرم افزار پیچیده تری نیاز دارد ، در تکنیک های اول و دوم مشکلی که مطرح است این است که حداکثر طول رکورد به طول بلاک محدود می شود یعنی در مرحله ایجاد فایل باید حداکثر طول رکورد را برابر طول بلاک در نظر بگیریم در حالی که تکنیک دوم این محدودیت را ندارد و از این نظر انعطاف پذیر تر است. از نظر زمانی به علت این که روش سوم حافظه هرز بیشتری دارد طول فایل بیشتر از روش دوم می باشد. برای خواندن تمام فایل و یا پردازش آن زمان بیشتری صرف می کنیم. به عبارتی از روش سوم بیشتر از روش دوم برای خواندن فایل استفاده می کنیم.

تمرین:

اگر اندازه یا طول هر بلاک ۱۶۰۰ بایت و طول هر رکورد ۶۴۰ بایت باشد مقدار فاکتور بلاک بندی یک پارچه و دوپارچه را محاسبه کنید.

فصل ۳ :

فایل در محیط فیزیکی - مدیریت بلاکهای آزاد

پشتیبان گیری - چگالی لود اولیه - لوکالیتی

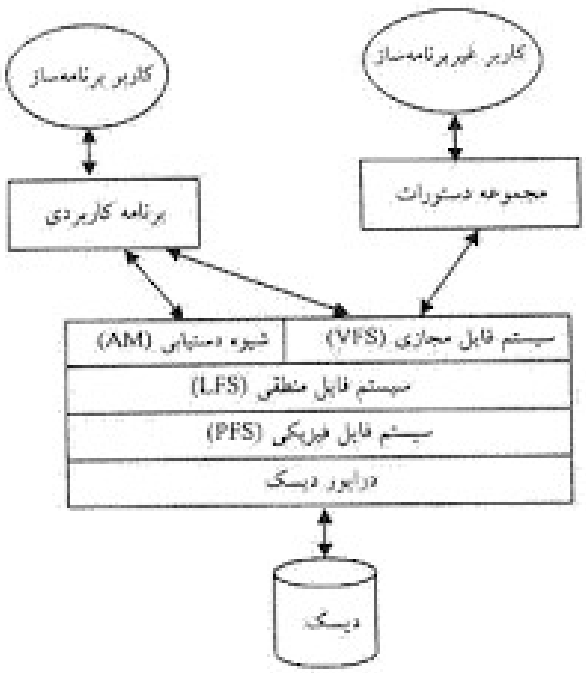
سیستم فایل

بخشی از سیستم عامل که با فایل ها سروکار دارد سیستم فایل نامیده می شود.سیستم فایل درخواست لایه های بالاتر را تبدیل به صدازدن توابعی در سیستم عامل می کند تا آن توابع عملیات فیزیکی I/O را انجام دهند.

۳ویژگی اصلی فایل:

- به اندازه ای بزرگ است که بطور کامل در حافظه ی اصلی جای نمیگیرد.
- پایایی دارند یعنی داده های آن از بین نمی روند مگر آنکه کاربر درخواست حذف آنها را بکند.
- می تواند بین چند کاربر مجاز بصورت اشتراکی دستیابی شود.

معماری کل سیستم فایل را می توان مشابه شکل زیر در نظر گرفت:



فایل در محیط فیزیکی

فایل در دیسک به دو نحوه پیوسته و ناپیوسته می‌تواند ذخیره شود. در نشست پیوسته، فایل در بلاکهای فیزیکی همجوار روی دیسک ذخیره می‌شود و در نشست ناپیوسته، تعدادی بلاک ناهمجوار به فایل تخصیص می یابد.

مزایای طرح پیوسته

۱- سادگی پیاده سازی

با داشتن آدرس بلاک اول روی دیسک، به بلاکهای دیگر می‌توان دسترسی پیدا کرد.

۲- بالا بودن کارائی

کل فایل طی یک عمل واحد از روی دیسک خوانده می شود.

معایب طرح پیوسته

۱- حداکثر اندازه فایل در مرحله ایجاد فایل باید معلوم باشد.

۲- بروز پدیده بندبند شدن فضای دیسک (با تکنیک یکپارچه سازی یا فشرده سازی بر طرف می شود.)

روشهای پیاده سازی طرح نشست ناپیوسته

۱- ایجاد لیست پیوندی

۲- ایجاد لیست پیوندی به همراه جدول راهنما

۳- تکنیک گره I (Index-node)

- دستیابی تصادفی در لیست پیوندی به همراه جدول راهنما سریع تر از روش ایجاد لیست پیوندی می باشد.
- عیب روش ایجاد لیست پیوندی به همراه جدول راهنما در این است که کل جدول باید در حافظه اصلی قرار بگیرد.

ویژگیهای روش لیست پیوندی

۱- سادگی یافتن بلاکهای فایل

۲- کند بودن دستیابی تصادفی به رکوردها

۳- تکه تکه نشدن حافظه

۴- آسان بودن خواندن پی در پی.

مدیریت بلاک های آزاد

برای مدیریت بلاکهای آزاد ۲ طرح وجود دارد:

۱) ایجاد لیستی از چند بلاک دیسک

۲) استفاده از بیت نقش

ایجاد لیستی از چند بلاک دیسک

در این روش، در هر بلاک شماره بلاکهای آزاد قرار می گیرد.

*اگر اندازه بلاکها یک کیلو بایت و هر شماره بلاک در 32 بیت نمایش داده شود، در هر بلاک شماره چند بلاک آزاد را می توان ذخیره کرد؟

حل: تعداد شماره هایی که می توان در یک بلاک ذخیره کرد برابر است با:

$$\frac{1KB}{32bit} = \frac{1024\ \ byte}{4\ \ byte} = 256$$

البته چون یک مدخل بلاک برای ایجاد نشانه رو به بلاک بعدی است، بنابراین در هر بلاک می توان شماره 255 بلاک آزاد را تخصیص داد.

*یک دیسک 200MB که دارای 200K بلاک 1KB است، حداکثر به چند بلاک برای ایجاد لیست نیاز دارد؟

(شماره بلاک در ۴ بایت نمایش داده می شود)

(در انتهای هر بلاک یک مدخل ۴ بایتی برای اشاره به بلاک بعدی در نظر گرفته شده است)

حل:چون هر آدرس ۴ بایتی است، مقدار بایت لازم برای تهیه لیست شماره بلاکهای آزاد برابر است با:

$$200K \times 4 = 200 \times 1024 \times 4 = 819200\ \text{Byte}$$

و چون هر بلاک 1KB است، تعداد بلاک لازم برابر است با:

$$\frac{819200}{1024} = 800$$

و چون به ازای هر بلاک یک مدخل چهار بـایتی بـرای اشـاره بـه بـلاک بعـدی داریـم، مقـدار 3200 بایـت (4×800) به فضای مورد نیاز افزوده خواهد شد. بنابراین تعداد بلاکهای اضافی مورد نیاز برای نگهداری این مدخلها برابر است با:

$$\left\lceil \frac{3200}{1024} \right\rceil = 4$$

بنابراین در کل به 4+800 یعنی 804 بلاک نیاز است.

استفاده از بیت نقش(Bitmap)

در این طرح برای دیسکی با n بلاک به n بیت نیاز است. بلاکهای آزاد را با 1 و بلاکهای تخصیص یافته را بـا 0 نمایش می دهیم.

*در طرح Bitmap، برای دیسکی به ظرفیت 200 مگا بایت با بلاکهای یک کیلو بایتی، به چند بیـت حافظـه نیاز است؟

حل:

تعداد بلاکها و در نتیجه تعداد بیت ها برابر است با

$$\frac{200 \ \ MB}{1 \ \ KB} = \frac{200 \times 2^{20}}{1 \times 2^{10}} = 200 \times 2^{10} = 204800$$

تکنیکهای تهیه پشتیبان

تکنیکهای تولید نسخه پشتیبان از دیسک سخت با ظرفیت زیاد عبارتند از:

۱- آینه سازی (Mirroring)

۲- استفاده از نیمه دو دیسک

۳- تولید دامپهای تدریجی (Incremental dump)

آینه سازی

در این روش از دو یا بیش از دو دیسک استفاده می‌شود. در حالت استفاده از دو دیسـک، عمـل نوشـتن در هر دو دیسک انجام می‌گیرد. نوشتن در دیسک آینه کمی با تأخیر صورت می‌گیرد. ولی عمل خواندن فقـط از یک دیسک صورت می‌گیرد. وقتی که یکی از دیسکها خراب شود، از دیسک دیگر می‌توان استفاده کرد.

استفاده از نیمه دو دیسک

در این روش، هر دیسک به دو نیمه "داده ها و پشتیبان " تقسیم می‌شود. در پایان هر روز کاری نیمه داده ای هر دیسک در قسمت پشتیبان دیسک دیگر کپی می‌شود.

تولید دامپهای تدریجی

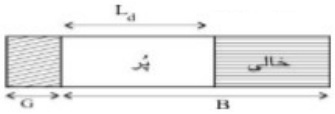
در این روش، ضمن ایجاد نسخه پشتیبان بطور هفتگی یا ماهانه، دامپ فایلهایی که از زمان تولید آخرین نسخه پشتیبان تغییر کرده اند، نیز به طور روزانه تولید می شود. این روش به حافظه زیادی نیاز دارد و معمولا از نوار استفاده می شود.

لوکالیتی (locality): میزان همسایگی (نزدیکی) فیزیکی رکورد منطقی بعدی به رکورد فعلی. به عبارتی شاخصی است که نشان می دهد رکوردهای منطقا هم جوار تا چه اندازه فیزیکی أ هم جوار اند.

هر چه لوکالیتی رکوردها قوی تر باشد زمان پردازش سریال آن کمتر خواهد بود. چرا که دیگر در هنگام خواندن ، زمان کمتری برای انتقال هد دیسک به جلو و عقب برای خواندن رکورد بعدی صرف خواهد شد.

چگالی لود اولیه

می‌توان تمامی فضای بلاک را در لود اولیه پر نکرد و قسمتی از آن را رزرو کرد تا در



عملیات ذخیره‌سازی بعدی از این فضا استفاده کرد. چگالی لود(Loading density) اولیه توسط درصدی از اندازه بلاک مشخص می‌شود که با داشتن نسبت $\frac{L_d}{B} < 1$ بدست می‌آید.

چگالی لود اولیه $\Longleftarrow$ $\frac{Ld}{B} < 1$ قسمت پر بلاک (LD) تقسیم بر کل بلاک .

مزایای در نظر گرفتن حافظه رزو در بلاک

1- این امر موجب افزایش لوکالیتی فایل می شود چرا که از پراکندگی نشست رکوردها روی دیسک تا حدودی جلوگیری می کند، این امر میزان پراکندگی در زمان دستیابی تصادفی و نیز زمان پردازش فایل را بهبود می‌بخشد.

2- موجب تسهیل در انجام برخی از عملیات بر روی فایل خواهد شد. مثلا در هنگام درج رکوردی که طول آن در اثر عملیات بهنگام سازی افزایش یافته دیگر نیاز به حذف منطقی از بلاک فعلی و درج در بلاک دیگر نخواهد بود و می توان با یک شیفت درون بلاکی رکورد تغییر یافته را در جای قبلی خود باز نویسی کرد.

معایب :

1- این حافظه در واقع نوعی حافظه هرز محسوب می شود که موجب افزایش اندازه حجم فایل خواهد شد و در نتیجه زمان خواندن کل فایل افزایش می یابد.

2- در صورتی که توزیع رکوردها یکنواخت نباشد ، موجب باقی ماندن حافظه هرز در انتهای بعضی از بلاک ها خواهد شد.

هرچه چگالی لود اولیه کمتر باشد تعداد بلاک ها بیشتر و حجم فایل بیشتر می شود.

مثال:

چگالی لود اولیه را می توان از رابطه $L_d = \frac{n}{b \times B_f} \times 100$ بدست آورد.

*فایلی با ۱۲ رکورد و فاکتور بلاک بندی ۶ و چگالی لود اولیه 50% مفروض است. تعداد بلاکهای اشغال شده توسط فایل کدام است؟

$$b = \left\lceil \frac{n}{l_d \times B_f} \right\rceil = \left\lceil \frac{12}{(\frac{50}{100} \times 6)} \right\rceil = 4$$

به عبارتی اگر از کل فضای هر بلاک استفاده شود، آنگاه به ۲ بلاک نیاز است، ولی چون از نصف فضای هر بلاک استفاده می شود، به ۴ بلاک نیاز است.

*فایلی دارای ۱۰ رکورد است. در صورتیکه این فایل ۴ بلاک را اشغال نماید، چگالی لود اولیه آن کدام است؟ (فاکتور بلاک بندی =۵)

حل:

$$l_d = \frac{n}{b \times B_f} \times 100 = \frac{10}{4 \times 5} \times 100 = \%50$$

فصل ۴ :

سطوح نشانی دهی - بافرینگ

سیستم فایل از دو بخش عمده تشکیل شده است:

۱- سیستم فایل منطقی

این لایه در خواست های برنامه کاربر مانند بازکردن، بستن، خواندن و نوشتن فایل را انجام می دهد.

۲- سیستم فایل فیزیکی

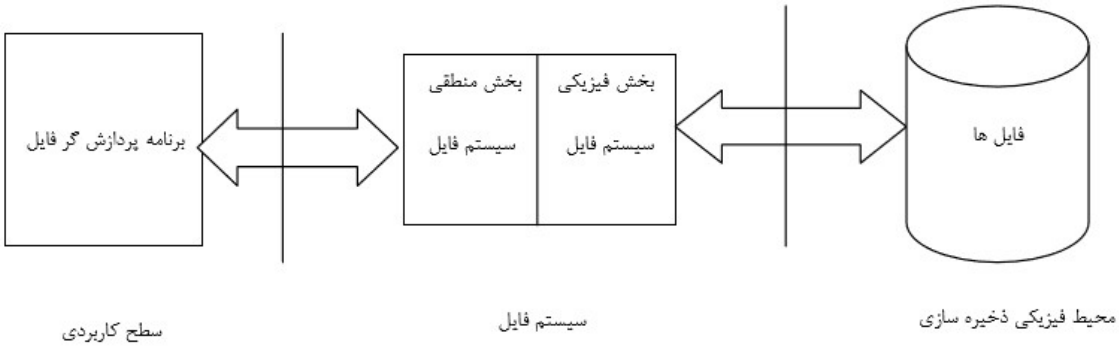
وظیفه این لایه دسترسی فیزیکی به فایلها در محیط فیزیکی می باشد. برای این منظور درخواستهای دریافتی از بخش منطقی را به دستوراتی جهت صدور به کنترولر رسانه تبدیل می کند.

سطوح آدرس دهی

الف) در سطح برنامه پردازش گر فایل

ب) در سطح سیستم فایل منطقی

ج) در سطح سیستم فیزیکی



بخش فیزیکی سیستم فایل با فایل در ارتباط است و بخش منطقی سیتم فایل با کاربر در ارتباط است.

عملیات اساسی که در محیط فیزیکی انجام می شوند عبارتد از : خواندن از رسانه، نوشتن در رسانه و پیگرد (Seek).

۱) آدرس دهی در سطح برنامه کاربردی(کاربر)

1- نسبی 2-محتوایی 3- نمادی

در آدرس دهی نسبی کاربر فایل ، فایل را به صورت یک ساختار خطی مشاهده می کند که در آن هر رکورد دارای شماره است با شروع از صفر برای اولین رکورد فایل ، در این روش کاربر آدرس رکورد RRA (Relative Record Addres) (آدرس نسبی رکورد) را به عنوان آرگومان جستجو می دهد. در آدرس دهی محتوایی کاربر یک مقدار صفت خاصه را به عنوان پارامتر جستجو می کند. در آدرس دهی نمادی کاربر رکورد مورد نظرش را به کمک یک اسم سمبلیک مشخص می‌کند مثال برای آدرس دهی محتوایی : رکوردی را پیدا کن که شماره شناسنامه 4357 داشته باشد.

۲) آدرس دهی در سطح منطقی سیستم فایل

بخش منطقی سیستم فایل کل فضای ذخیره سازی را به شکل آرایه ای از بلاک ها می بیند آدرس یا شماره بلاک را با RBA(Relative Block Address) نشان می دهند که آدرس نسبی بلاک اول برابر صفر در نظر گرفته می شود.

آدرس دهی نسبی بلاک حاوی رکورد آی ام نسبت به بلاک اول فایل ⟸ شماره بلاک آی ام

$$\left[\frac{(i-1)*R}{B}\right]$$

این آدرس نسبت به ابتدای فایل است. مثلا ادرس شماره 5 یعنی از ابتدای فایل رکورد پنجم

RBA مطلق : اگر RBA_R نسبی را بعلاوه RBA_B ابتدای فایل کنیم آدرس مطلق بدست می آید.

RBA_R ⟸ $RBA_B + \left[\frac{(i-1)*R}{B}\right]$

RBA_B: آدرس نسبی بلاک اول فایل (شماره اولین بلاک ابتدای فایل)

RBA_R : آدرس نسبی حاوی رکورد مورد نظر

مثال:

* RBA آدرس نسبی رکورد دهم از کدام بایت شروع می شود و چقدر می باشد؟

$(R=250, B=1000, n=12, RBA_{BOF}=6)$

حل:

$$RRA=(i-1)\times R=(10-1)\times 250=2250$$

$$RBA_{(10)}=6+\left\lfloor\frac{(10-1)\times 250}{1000}\right\rfloor=6+2=8$$

بنابراین RBA آدرس نسبی رکورد دهم از بایت ۲۲۵۰ ام شروع می شود و برابر ۸ می باشد.

* فایلی با رکوردهای ۲۰ بایتی و بلاکهای ۶۰ بایتی را در دیسکی با ۴ شیار در هر سیلندر و ۱۰ بلاک در هر شیار ذخیره کرده ایم. مقدار RBA را برای هشتمین رکورد بدست آورید؟

حل:

$$RBA = 0 + \left\lfloor \frac{(8-1)\times 20}{60} \right\rfloor = \lfloor 2.3 \rfloor = 2$$

۳) آدرس دهی در سطح فیزیکی سیستم فایل

در این سطح، آدرس مکان داده مورد نظر در رسانه مشخص می‌شود. اگر رسانه دیسک باشد، اجزاء آدرس عبارتند از:

۱- شماره درایور

۲- شماره استوانه

۳- شماره رویه (شیار از استوانه)

۴- شماره سکتور از شیار(یا شماره بلاک)

سیستم فایل فیزیکی، RBA را به آدرس فیزیکی تبدیل می‌کند. برای این تبدیل چنین عمل می کند:

نحوه محاسبه اجزاء آدرس

شماره استوانه	$Cyl\# = RBA \ div \ (t_i \times b_i)$
شماره شیار	$Trk\# = [RBA \mod (t_i \times b_i)] \ div \ b_i$
شماره بلاک	$Blk\# = RBA \mod b_i$

تذکر: در فرمولهای جدول بالا، در حالت کلی باید به جای RBA از $RBA - RBA_{Beginofdevice}$ استفاده شود. (در روابط بالا فرض کردیم که : $RBA_{BeginOfDevice} = 0$)

تذکر : شماره اولین شیار هر سیلندر صفر می باشد.

تذکر مهم:

بجای استفاده از ۳ فرمول فوق می توان از فرمول ساده تر زیر استفاده کرد:

RBA ⌉ b
..... n ⌉ t
_____ ..... Cyl#
Blk# _____
Trk#

مثال:

* فایلی با رکوردهای ۲۰ بایتی و بلاکهای ۴۰ بایتی را در دیسکی با ۴ شیار در هر سیلندر و ۱۰ بلاک در هر شیار ذخیره کرده ایم. مقادیر CYL#,TRK#,BLK# را برای رکورد ۴۴ ام بدست آورید؟

حل :

ابتدا RBA را بدست می آوریم:

$$RBA_{(44)} = 0 + \left\lfloor \frac{(44-1)\times 20}{40} \right\rfloor = \left\lfloor \frac{43}{2} \right\rfloor = 21$$

$$cyl\# = 21 \;\; div \;\; (10\times 4) = 0$$

$$trk\# = [21 \;\; \mathrm{mod} \;\; (10\times 4)] \;\; div \;\; 10 = 2$$

$$blk\# = 21 \;\; \mathrm{mod} \;\; 10 = 1$$

بافر (Buffer)

ناحیه ای است از حافظه اصلی که جهت هماهنگی در عملیات ورودی/خروجی و عملیات سی پی یو بکار می رود. بافر ناحیه ای است که حداقل یک بلاک را بتواند در خود جای دهد که در هر بار عملیات خواندن و نوشتن از دیسک حداقل یک بلاک در بافر گذاشته و یا خوانده می شود. در صورتی که فایل را بلاک بندی کرده باشیم معمولا برای انجام عملیات ورودی و خروجی بر روی فایل ها از دو بافر ورودی و خروجی استفاده می شود.

قسمتی از Ram است که جداگانه قرار می گیرد برای این که بتواند بین دستگاههای ورودی خروجی و سی پی یو یک هماهنگی سرعتی داشته باشد(یه قسمت از Ram را در نظر می گیریم که اطلاعات قبل از این که به سی پی یو بروند در آنجا کپی شوند.)

دلیل استفاده از بافر : چون سرعت دستگاههای ورودی/خروجی خیلی کمتر از سرعت سی پی یو است این عدم هماهنگی توسط بافر بتواند کنترل شود.

انواع بافر

1- **بافر نرم افزاری** : ناحیه ای است از حافظه اصلی که توسط سیستم عامل در اختیار برنامه ها قرار می گیرد.
2- **بافر سخت افزاری** : بافر موجود در دستگاههایی مانند کارت خوان ، چاپگر و .... استفاده می شود.این بافرها معمولا حجم زیادی از اطلاعات را می توانند ذخیره کنند این نوع بافر با سرعت دستگاههای ذخیره سازی پر شده و پس از آن که بافر تکمیل شد محتوای آن را با سرعت انتقال کانال به کامپیوتر و از آن جا به بافر نرم افزاری منتقل می شوند بافرهای سخت افزاری خود به چند دسته تقسیم می شوند.

الف) بافرینگ استاندارد (single buffering) : ما یک بافر داریم.

هنگامی که بافر در حال پر شدن است، CPU در حالت عاطل (idle) قرار دارد.

بدلیل بیکار بودن cpu کارایی برنامه کاهش می یابد.البته در سیستم های چند برنامه ای از این زمان انتظار برای اجرای برنامه های دیگر استفاده می شود.

ب) بافرینگ مضاعف (double buffering): یعنی ما دو بافر داریم که یکی کمکی است. برای این که کارایی خودش را داشته باشد چیزی به نام شرط مضاعف داریم در غیر این صورت هیچ کارایی برای ما نخواهد داشت.

بعبارت دیگر با دو بافر می توان در حین انتقال یک بلاک به بافر ، بافر دیگری را که پر است پردازش کرد.

ج) بافرینگ چندگانه (multiple buffering): یعنی ما چندتا بافر داریم.

با استفاده از بافر دهی چندگانه هنگام پردازش پی در پی و انبوه فایل می توان بلوکهای فایل را جلوتر خواند و د بافر ذخیره کرد .بدین ترتیب همواره رکورد بعدی در بافر موجود خواهد بود.به این روش ، روش صف بندی و یا بافر چرخشی نیز گویند.

نکته:

اگر سرعت پردازش خیلی بیشتر از سرعت عملیات بافرینگ باشد ، حتی با بافر دهی چندگانه نیز ممکن است وضعیتی پیش بیاید که پردازنده مجبور به انتظار کشیدن باشد.البته در سیستم های چندبرنامه ای در این حالت cpu به فرآیندهای دیگر سوئیچ می کند .بنابراین در کل وجود بافرینگ کارایی سیستم را افزایش می دهد.