

جلسه اول ذخیره و بازیابی اطلاعات

حافظه:

هر دستگاهی که قادر به ذخیره سازی و بازیابی اطلاعات باشد.

انواع حافظه:

اصلی و جانبی.

سیستم های فایل لینک:

ذخیره سازی اطلاعات در محیط برون ماشین و بررسی مکانیزم دستیابی و بازیابی آن ها.

علت به کارگیری و استفاده از حافظه های جانبی این است که حافظه های درون ماشین گران اند و ظرفیت محدودی دارند.

چگالی ذخیره سازی اطلاعات به تعداد شیار در واحد طول بستگی دارد.

از نظر تعداد شیار ها دو نوع نوار وجود دارد: ۷ شیاره و ۹ شیاره

تعریف ۱- گپ:

فضای بلا استفاده بین دو گروه کاراکتر یا بلوک یا رکورد ضبط شده می باشد.

تعریف ۲- گپ:

از نظر ذخیره سازی اصطلاحاً حافظه Waste می باشد.

پارامترهای اساسی نوار:

سرعت، چگالی، نرخ انتقال.

سرعت حس:

برای این که هد خواندن و نوشتن بتواند داده‌ای را حس کند باید پس از توقف به سرعتی مطلوب و یکنواخت موسوم به سرعت حس برسد که برای این کار فضای خالی گپ مورد نیاز است. همچنین برای رسیدن سرعت حس تا توقف کامل نیز فضای خالی گپ لازم می باشد.

کلاستر:

به چند سکتور پشت سر هم یک کلاستر می گویند.

دیسک‌ها با هد ثابت سریع تر و گران تر از دیسک‌ها با هد متحرک می باشد.

طبله:

رسانه‌ای منطقاً معادل دیسک با نوک ثابت متشکل از یک استوانه با یک یا چند هد خواندن و نوشتن است و در قدیم به عنوان حافظه اصلی استفاده می شد.

حافظه کش:

حافظه ایست مابین CPU و RAM و جزء حافظه اصلی می باشد.

زمان پیگرد:

زمان لازم جهت انتقال هد به سیلندر است و متوسط این زمان را با حرف S نشان می دهند. این زمان حدود ۲ تا ۱۰ میلی ثانیه است.

زمان درنگ دورانی:

پس از آن که هد به سیلندر مورد نظر رسید زمانی برای چرخش دیسک لازم است تا سکتور مورد نظر در زیر هد قرار بگیرد که به آن زمان درنگ دورانی می گویند. متوسط این زمان را با حرف R نشان می دهند که نصف زمان لازم جهت یک دور چرخیدن دیسک می باشد.

موجودیت:

نوع موجودیت به فرد، شیء، پدیده یا مفهومی که می خواهیم در رابطه با آن اطلاعات داشته باشیم گفته می شود.

محیط عملیاتی:

به محیطی که در رابطه با آن می خواهیم یک سری داده ها را در آن ذخیره، بازیابی یا پردازش کنیم گویند. مثلاً محیط عملیاتی دانشگاه از موجودیت های دانشجو، استاد، درس، کارمند، کلاس و ... تشکیل یافته است. انواع موجودیت ها توسط صفحات خاص مربوط به هر یک از سایر موجودیت ها متمایز می گردد. مثلاً موجودیت استاد میتواند صفحات خاصی مدرک، نام، آدرس، سابقه تدریس و ... را داشته باشد.

فیلد:

مکان ذخیره شدن یک واحد معنادار یا یک فقره اطلاعات را فیلد گویند که کوچکترین واحد اطلاعات در فایل است.

اطلاع:

هر صفت خاصه از دو مؤلفه تشکیل یافته است. یکی اسم صفت خاصه و دیگری مقدار صفت خاصه. به مجموع این دو مؤلفه اطلاع گفته می شود. اطلاع توسط انسان یا ماشین تولید، ذخیره، بازیابی و پردازش می شود. مثلاً نام خانوادگی صفت خاصه است و مثلاً احمدی مقدار صفت خاصه است.

رکورد:

مجموعه ای از فیلدها تشکیل رکورد را می دهند و مجموعه ای از رکوردها فایل را تشکیل می دهند.

ساختارهای فیلد:

برای مشخص ساختن فیلدها در طول رکوردها راه حل‌های مختلف زیر وجود دارد:

۱- قرار دادن فیلدها در طول‌های از قبل تعیین شده.

۱۴ بایت	۲۰ بایت	۱۲ بایت
مدرک	فامیل	نام
لیسانس	حسینی	علی

یک ایراد این روش این است که برای رساندن فیلدها به طول معین می‌بایست از فاصله خالی استفاده شود و فضای خالی باعث بزرگ شدن اندازه فایل و اتلاف حافظه دیسک می‌گردد.

۲- قرار دادن طول فیلد در ابتدای هر فیلد.

03Ali 06Javadi 06Doctor 07Physics

۳- استفاده از یک کاراکتر ویژه به عنوان حد فاصل در انتهای هر فیلد.

Ali, Javadi, Doctor, Physics

۴- بکار بردن نام هر فیلد در مقابل مقدار هر فیلد. به عبارت دیگر استفاده از یک عبارت کلیدی برای شناسایی هر فیلد.

Name=Ali, Family=Javadi, City=Tehran

مزیت این ساختار آن است که فیلد خودتوصیف بوده و فیلدها می‌توانند جایجا شوند؛ همچنین مقادیر بعضی از فیلدها در صورت عدم وجود ذخیره نمی‌گردد. ایراد این روش اتلاف حافظه ایست که در اثر ذخیره ی نام فیلدها با آن مواجه می‌شویم.

ساختار رکوردها:

بعضی از روش های سازمان دهی رکوردها به صورت زیر می باشد:

۱- رکوردهایی با طول ثابت: در این روش طول همه رکوردهای فایل با هم برابر می باشد و این روش متداول ترین سازمان دهی رکوردهاست. ثابت بودن طول رکورد الزاماً به منظور ثابت بودن طول فیلدهای تشکیل دهنده آن نیست.

۲- تعیین طول رکوردها بر حسب تعداد فیلدهای آن: در این روش هر رکورد از n فیلد تشکیل یافته است و n برای کل فایل ثابت است. مثلاً اگر $n = 4$ آنگاه فایل می تواند به صورت زیر باشد:

Ali, Javadi, Doctor, Physics, Mohammad, Hussein, Doctor, Computer

۳- ذخیره طول رکورد در اول هر رکورد: در این روش در ابتدای هر رکورد طول آن ذخیره می شود. این روش اغلب برای کار با رکوردهای با طول متغیر بکار می رود.

۴- استفاده از اندیس برای آدرس های هر رکورد نسبت به اول فایل.

0	26
---	----	------

۵- ذخیره یک علامت ویژه فاصل در انتهای هر رکورد.

از یک نظر می توان گفت دو ساختار کلی جهت پیاده سازی رکوردها وجود دارد:

الف) رکورد با قالب ثابت و مکانی که تعداد، مکان و طول فیلدها در نمونه های مختلف ثابت بوده و تعریف این ساختار از قبل مشخص شده است.

نام	فامیل	رشته
علی	کریمی	برق
حسین	محمودی	فیزیک

ب) رکورد با قالب غیر ثابت و غیر مکانی که در هر فیلد، اسم فیلد به همراه مقدار آن ذخیره می شود.

نام=علی، فامیل=کریمی، رشته=برق
نام=حسین، رشته=فیزیک

طول یک رکورد بنا به دلایل زیر ممکن است متغیر شود:

الف) طول بعضی فیلدها مثل آدرس ممکن است متغیر باشد.

ب) تعداد فیلدهای نمونه های یک رکورد (موجودیت) ممکن است متغیر باشد. مثلاً موجودیت استاد ممکن

است به دو دسته ی " رسمی با حقوق ثابت " و " حق التدریس " تقسیم گردد.

نوع رسمی: نام استاد، مدرک، رشته، حقوق ماهانه

نوع حق التدریس: نام استاد، مدرک، رشته، تعداد ساعات تدریس، حق الزحمه هر ساعت

ج) ممکن است در رکورد، فیلد (فقره اطلاع) تکرار شونده داشته باشیم.

نام	مدرک	دانشکده‌ای که تدریس می کند
اکبری	دکترا	برق، کامپیوتر
حسینی	دکترا	ریاضی، برق، کامپیوتر

از سه دیدگاه می توان به رکورد نگاه کرد:

الف) رکورد در سطح انتزاعی که رکورد را مستقل از جنبه های نمایشی آن و بصورت کلی نگاه می کنیم.

ب) رکورد در سطح منطقی که رکورد را از دیدگاه برنامه نویس مشخص می سازد و Sort شده است.

ج) رکورد ذخیره شده یا رکورد در سطح فیزیکی که رکورد را به صورتی که در محیط ذخیره سازی مثل

دیسک قرار می گیرد معنی می سازد و ممکن است به آن اطلاع بیشتری اضافه شود و یا ساختار آن قدری

تغییر کند و معمولاً رکورد ذخیره شده دارای دو بخش مجزای داده‌ای و کنترلی می باشد و به بخش کنترلی،

بخش پیشوندی، بخش غیرداده‌ای یا Meta Section نیز گفته می‌شود. بخش کنترلی اغلب توسط سیستم فایل استفاده شده و از دید برنامه مخفی است.

اغلب در بخش کنترلی اطلاعات زیر ذخیره می‌شود: ۱- طول رکورد ۲- نوع رکورد ۳- اشاره گرها ۴- پرچم های عملیاتی و حفاظتی ۵- اطلاعاتی خاص، ویژه بعضی ساختارها

۱- طول رکورد:

هنگامی که طول رکوردها متغیر باشد در بخش کنترلی طول آن رکورد ذخیره میشود و رکوردهای با طول ثابت به این اطلاع نیازی ندارند.

۲- نوع رکورد:

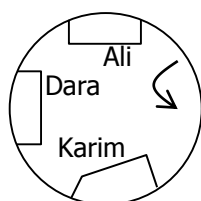
ممکن است در یک فایل اطلاعات دو یا چند رکورد ذخیره شود (فایل چند نوعی) ممکن است در یک فایل هم اطلاعات اساتید و هم اطلاعات دانشجویان ذخیره گردد در اینجا نوع هر رکورد باید در ابتدای آن مشخص گردد و فایلی را که فقط یک نوع رکورد دارد، فایل تک نوعی می‌گویند.

۳- اشاره گرها:

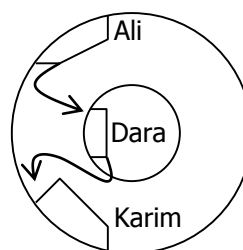
مثلاً پردازشگر فایل ممکن است رکوردهای اساتید را به ترتیب حروف الفبا مشاهده و پردازش کند؛ ولی این رکوردها که منطقاً مجاور یکدیگرند هنگام ذخیره شدن بر روی دیسک الزاماً به همان ترتیب نخواهند بود. در این حال با استفاده از اشاره گرها ارتباط منطقی بین رکوردها پیاده سازی می‌گردد.

Ali
Dara
Karim

همجواری
منطقی
رکوردها



همجواری
فیزیکی
رکوردها



ناهمجواری
فیزیکی
رکوردها

۴- پرچم (Flag):

این پرچم ها برای نشان دادن عملیاتی که قرار است روی رکورد انجام بگیرد و یا نشان دادن عملیاتی که روی آن رکورد انجام شده بکار می روند. مثلاً در بسیاری از سیستم ها حذف به دو صورت منطقی و فیزیکی صورت می گیرد؛ بدین معنا که هنگام صدور فرمان حذف (جهت بالا بردن سرعت عملیات) تنها در ابتدای آن رکورد پرچمی "I" می شود (بدون حذف واقعی) و در این حالت مثلاً هنگام نمایش رکوردها آنهایی که علامت حذف خورده اند نشان داده نمی شوند و سپس در فرصتی مناسب این اطلاعات بطور فیزیکی حذف می شوند. همچنین در محیط های اشتراکی نیاز به پرچم های کنترلی است که نحوه دستیابی افراد را به رکوردها معین می سازد. مثلاً اگر پرچم Read Only برای کاربری فعال شود آنگاه آن کاربر نمی تواند رکورد را تغییر دهد.

۵- اطلاعاتی خاص، ویژه بعضی ساختارها:

در ساختارهای مختلف فایل جهت پیاده سازی آن ها گاهی اوقات لازم است اطلاعات خاصی همراه رکوردها ذخیره گردد.

کلید (Key):

صفت خاصه یا ترکیبی از چند صفت خاصه را که در تمام نمونه های مختلف یک موجودیت، مقدار یکتایی را داشته باشد، کلید می گویند. مثلاً در فایل اطلاعات دانشجویان فیلد شماره دانشجویی کلید است، زیرا هر شماره دانشجویی فقط یک دانشجو را معرفی می کند (شماره دانشجویی تکراری نیست).

فایل:

مجموعه ای از نمونه های مختلف یک یا چند رکورد با ساختار مشخص است. فایل نیز دارای دو ساختار منطقی و فیزیکی است. فایل ها بر روی حافظه جانبی ذخیره شده و محتویات آن ها ماندگار است. اغلب فایل ها به قدری بزرگند که نمی توان آن ها را جهت پردازش بطور کامل به حافظه اصلی آورد. همچنین در حالت کلی فایل ها بصورت اشتراکی توسط چند کاربر استفاده می شوند.