

## جلسه چهارم ذخیره و بازیابی اطلاعات:

### موارد استفاده ی فایل ترتیبی:

۱- در کاربردهای تجاری

۲- تغییر طول رکورد مطرح نباشد

۳- در ایجاد بعضی از ساختارها لازم است ابتدا فایل به صورت ترتیبی ایجاد گردد.

در فایل ترتیبی تمام نمونه ی رکوردها قالب از قبل طراحی شده ای دارند و رکوردها دارای قالب ثابت مکانی می باشند و در Load اولیه تمام نمونه ی رکوردها بر اساس مقادیر یکی از صفات خاصه ی موجود در فایل منظم می باشند و این نظم با همجواری فیزیکی رکوردها پیاده سازی می گردد.

### مزایای فایل ترتیبی نسبت به پایل:

ساده تر بودن قالب رکوردها و تسهیل در پردازش سریال رکوردها و وجود یک استراتژی دستیابی می باشد.

### معایب فایل ترتیبی نسبت به پایل:

کاهش انعطاف پذیری ساختار در عملیات تغییر دهنده مثل درج، حذف و به هنگام سازی — پدیده ی عدم تقارن - کاهش انعطاف پذیری ساختار از نظر طول رکورد.

عملیات تغییر دهنده (تراکنش) در فایل ترتیبی در چه فایلی صورت می گیرد؟ فایل ثبت تراکنش TLF یا سر ریزی

بعد از چه مدتی فایل TLF را در فایل اصلی ادغام می کنیم؟ مدتی که طراح تعیین می کند یا در یک پریود زمانی ثابت

متوسط اندازه ی رکورد و ظرفیت کل فایل در ساختار ترتیبی چه می باشد؟  $R=a.V$  ظرفیت کل فایل  $(n+o)a.V$

تعداد دفعات مراجعه به فایل ترتیبی در روش جست و جوی باینری چه می باشد؟  $\log_2 \frac{nR}{B}$

بازیابی رکورد بعدی در یک فایل ترتیبی هنگامی امکان پذیر نمی باشد که رکورد بعدی در فایل ثبت تراکنش یا TLF باشد.

بین رکورد در ناحیه ی اصلی و رکورد در TLF هیچ گونه ارتباط ساختاری وجود ندارد. مثلاً از طریق اشاره گر ها. بنابراین اگر رکورد بعدی در TLF باشد به دست آوردن آن ممکن نیست. مگر این که آدرس آن را داشته باشیم که در این حالت به یک عمل واکنشی در TLF تبدیل می گردد.

تکنیک ماتریس بیتی تکنیکی است برای فشرده سازی وقتی که صفت خاصه ی چند مقداری وجود داشته باشد و در شرایطی که هم طول رکوردها متغیر و هم افزونگی طبیعی تشدید می شود، تکنیک ماتریس بیتی یکی از روش های کاهش این افزونگی است و اگر  $n$  تعداد عناصر مجموعه ای باشد که مقادیر صفت خاصه ی مورد نظر از آن گرفته شده است برای ذخیره سازی تمام این صفات به  $n$  بیت حافظه نیاز است. وقتی که برای یک فایل روی یک صفت خاصه شاخص ایجاد می کنیم، در واقع یک افزونگی تکنیکی بوجود آورده ایم.

در ساختار ترتیبی نمی توان طول رکورد را تغییر داد.

## فایل عدم تقارن:

فایل هایی که بر اساس یک فیلد مشخص می شوند لزومی ندارد که بر اساس یک کلید یا فیلد دیگر مرتب شوند. مثلا اگر بر اساس شماره ی دانشجویی مرتب شود، اسم ها به هم می خورد و بر عکس.

## فایل تقارن:

فایل تقارن فایلی است که اگر مثلا بر اساس یک فیلد آن را Sort کنیم فیلدهای دیگر هم به هم نخورند.

## شاخص اصلی:

وقتی که صفت خاصه ی شاخص کلید اصلی باشد. وقتی که صفت خاصه ی شاخص کلید ثانویه باشد. کلید ثانویه کلیدی است غیر از کلید اصلی که خاصیت کلید بودن را دارد. در شاخص متراکم لزومی بر مرتب بودن فایل داده ای نمی باشد اما در شاخص غیر متراکم باید فایل داده ای مرتب باشد برای این که بتوان رکوردها را گروه بندی کرد.

## دو نوع شاخص داریم:

۱. شاخص نرم افزاری

۲. سخت افزاری

## شاخص نرم افزاری:

گروه در شاخص غیر متراکم بلاک یا باکت می باشد.

## شاخص سخت افزاری:

گروه در شاخص غیر متراکم شیار استوانه یا در حالتی که فایل روی چند دیسک ذخیره شده باشد می تواند خود دیسک باشد.

## لنگرگاه:

نقطه ای از فایل داده ای است که مدخل شاخص به آن اشاره می کند و اگر لنگرگاه رکوردی باشد شاخص را متراکم می گویند.

## اجزای تشکیل دهنده ی ترتیبی ساختار شاخص دار عبارتند از:

۱- فایل ترتیبی یا ناحیه ی اصلی

۲- ناحیه ی سر ریزی

۳- نشانه رو ها

شاخص در ساختار ترتیبی شاخص دار در چه حالتی بازسازی می شود؟ در سازمان دهی مجدد.

زمان خواندن کل فایل در ساختار ترتیبی شاخص دار به صورت پی در پی از چه رابطه ای به دست می آید؟

$$(n + o) \frac{R}{t}$$

## معایب ساختار ترتیبی شاخص دار :

۱. عدم تقارن

۲. مسئله ی درج سر ریزی ها

۳. ایستا بودن شاخص.

واکشی تک رکوردها با استفاده از شاخص سریع تر انجام می گیرد.

در شاخص متراکم واکشی تک رکوردها سریع تر از شاخص غیر متراکم انجام می گیرد.

برای انجام عمل درج یک رکورد جدید در روش Push Trough چه اعمالی انجام می گیرد؟ خواندن بلاکی که باید رکورد در آن درج شود - بازنویسی بلاک - واکنشی رکورد منطقا قبلی و تنظیم اشاره گر - بازنویسی همین رکورد.