

پایگاه داده

جلسه ۱

محمد علی فرجیان



مراجع درس:

۱- بانک اطلاعات علمی- کاربردی جلد اول مولف: دکتر مصطفی حق جو انتشارات: دانشگاه علم و صنعت ایران

۲- مقدمه ای بر پایگاه داده ها- دکتر سیدمحمدتقی روحانی رانکوهی_ویرایش دوم ۱۳۸۳

3- An Introduction to Database Systems , 8th Edition , by C. J. Date , Addison-Wesley
2004 Ad

4- Database System Concepts by Korth and Silberschatz 4th Edition 2002



مفاهیم مربوط به سیستم فایل

- **داده** : حقیقتی خام که هیچ مفهوم خاصی ندارد
- **فیلد** : گروهی از کاراکترهاست که خصوصیتی از یک موجودیت را توصیف میکند، مثل سن، نام، وزن و ...
- **رکورد** : مجموعه ای از فیلدها است که مربوط به یک موجودیت خاص هستند
مثلا رکورد یک دانشجو
- **فایل** : مجموعه ای از رکوردها است که مربوط به یک موجودیت هستند
- **سیستم فایل** : مجموعه ای از فایلها و برنامه های لازم برای کار با آنهاست،
مثلا سیستم فایل یک دانشگاه



✓ نسل اول نسل فایل‌های ساده ترتیبی

در این نسل رسانه خارجی معمولاً **نوار** بوده است. این نسل را می‌توان نسل بی‌نرم‌افزار واسط نیز نامید. مشخصات کلی این نسل عبارتند از:

1. ساختار فایل‌ها ترتیبی است.
2. ساختار فیزیکی همان ساختار منطقی فایل است.
3. تنها روش پردازش فایل‌ها، پردازش یکجا یا دسته‌ای (Batch Processing) است.
4. نرم‌افزار تنها عملیات ورودی / خروجی را انجام می‌دهد. نرم‌افزار واسطی برای مدیریت پردازش فایل‌ها وجود ندارد.
5. طراحی ساختار فیزیکی فایل‌ها هم، برعهده کاربر است.
6. هرگونه تغییر در ساختار داده‌ها و یا رسانه‌های ذخیره‌سازی سبب بروز تغییر در برنامه و بازنویسی و کامپایل آن می‌شود.
8. اشتراک داده‌ها (Data Sharing) مطرح نیست.
9. تکرار در ذخیره‌سازی داده‌ها در بالاترین حد است.
10. برای انجام عملیات بهنگام‌سازی، الزاماً فایل دیگری ایجاد و تغییرات را در آن وارد کرده، نسخه قدیمی را به عنوان «فایل پدر» نگهداری می‌کنند و به این دلیل نسخه‌های متعددی از یک فایل نگهداری می‌شوند.



این نسل را باید نسل شیوه‌های دستیابی (Access Methods) نامید. مهمترین ویژگی این نسل را باید پیدایش نرم‌افزارهای موسوم به «شیوه‌های دستیابی» و همچنین ایجاد رسانه‌های با دستیابی مستقیم (یعنی دیسک) دانست. نرم‌افزار شیوه دستیابی، نرم‌افزاری است که به جنبه‌های فیزیکی محیط ذخیره‌سازی و عملیات در این محیط می‌پردازد. به نحوی که دیگر برنامه کاربر نیازی به پرداختن به این جنبه‌ها را ندارد. مشخصات این نسل عبارتند از:

1. نرم‌افزار واسط برای ایجاد فایلها با ساختارهای گوناگون بین برنامه‌های کاربردی و محیط ذخیره‌سازی وجود دارد.
2. امکان دستیابی ترتیبی و مستقیم به رکوردها (نه فیلدها) وجود دارد.
3. ساختار فیزیکی و ساختار منطقی فایلها از یکدیگر جدا هستند ولی نه تا حدی که برنامه‌های کاربردی از محیط فیزیکی ذخیره‌سازی مستقل شوند.
4. تغییر در رسانه‌های ذخیره‌سازی بر روی برنامه‌های کاربردی تاثیر چندانی ندارد.
5. ایمنی و حفاظت داده‌ها مطرح بوده ولی روشهای تامین امنیت و حفاظت ابتدایی هستند.
6. تکرار ذخیره‌سازی هنوز در حد نسبتاً بالایی وجود دارد.
7. برای پیاده‌سازی فایل با ارتباط خاصی بین انواع رکوردها (مثلاً ارتباط سلسله مراتبی) خود برنامه‌ساز باید ارتباطات را در برنامه‌اش بسازد.



در این نسل نرم‌افزاری کاملتر از نرم‌افزار دستیابی به عنوان واسط بین برنامه‌های کاربردی و فایل‌های محیط فیزیکی طراحی و ایجاد شد. در این نسل دریافتند که می‌توان برنامه‌های کاربردی را در قبال رشد فایلها (File Growth) مثلاً افزودن یک فیلد به یک نوع رکورد از یک فایل مصون نگاه داشت. تا قبل از این نسل برنامه‌های کاربردی فقط در قبال تغییرات سخت‌افزاری و رشد کمی فایلها (یعنی افزایش حجم داده‌های فایل) مصون بودند. مشخصات کلی این نسل عبارتند از:

1. نرم‌افزار نسبتاً پیچیده‌ای به نام سیستم مدیریت داده‌ها واسط بین برنامه کاربردی و محیط فیزیکی ذخیره‌سازی است.
2. میزان تکرار ذخیره‌سازی کاهش یافته است.
3. صحت داده‌های ذخیره شده تا حدی تامین می‌شود.
4. نشانی‌دهی به داده‌ها در سطح فیلد یا گروهی از فیلدها امکان‌پذیر است.
5. تسهیلاتی برای پردازش فایلها پیش‌بینی شده است.
6. بازیابی به کمک چند کلید (Multikey Retrieval) امکان‌پذیر است.



IASBS
1992 - 2012

نسل چهارم .نسل DBMS

این نسل از اواخر دهه 60 آغاز شد و هم‌اکنون نیز ادامه دارد. مهمترین خصیصه این نسل مستقل شدن برنامه‌های کاربردی (Application Program) از جنبه‌ها و خصوصیات محیط فیزیکی ذخیره‌سازی است (استقلال داده‌ای)



برای ایجاد سیستم بانک اطلاعاتی دو شیوه عمومی وجود دارد :

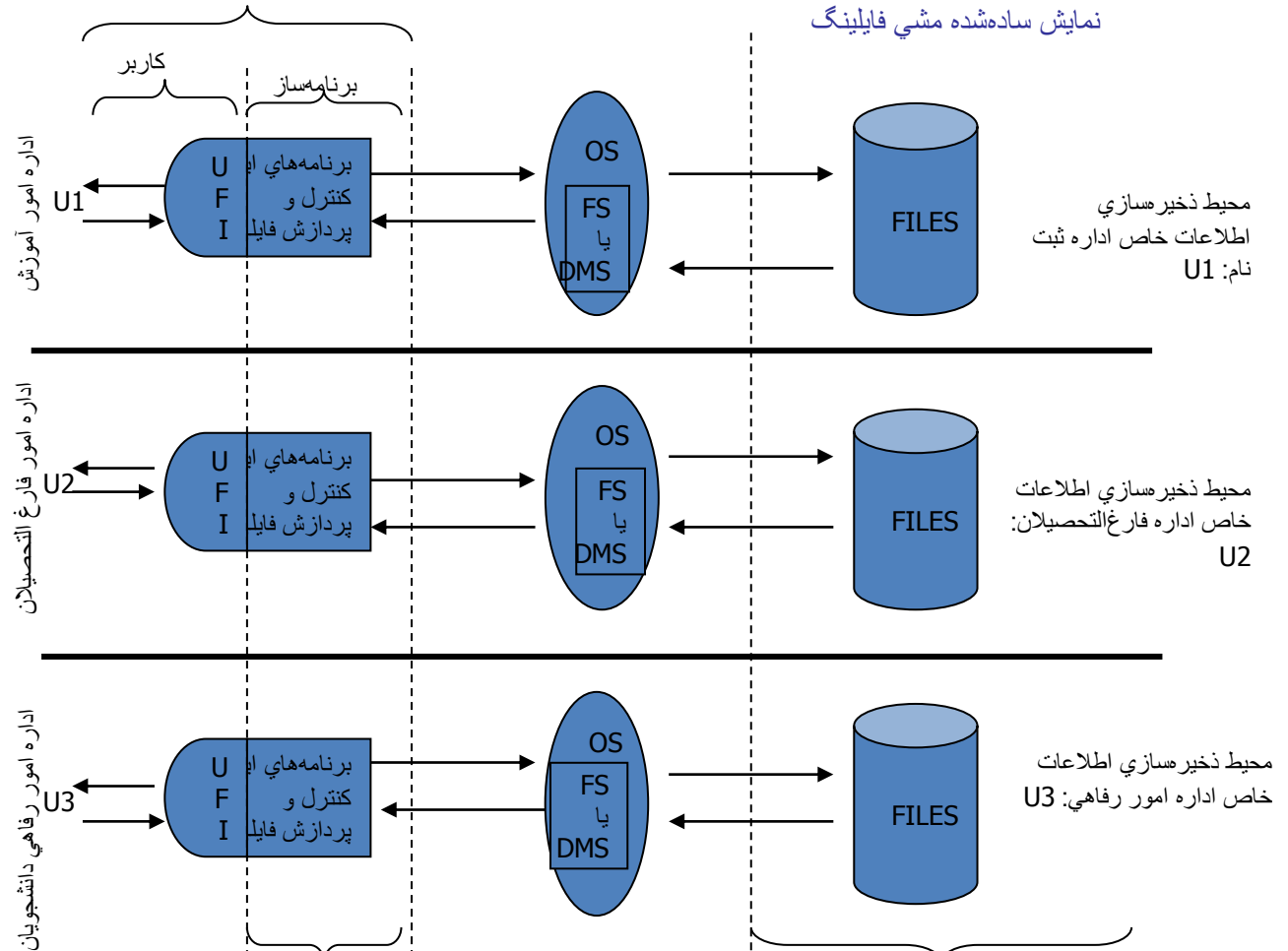
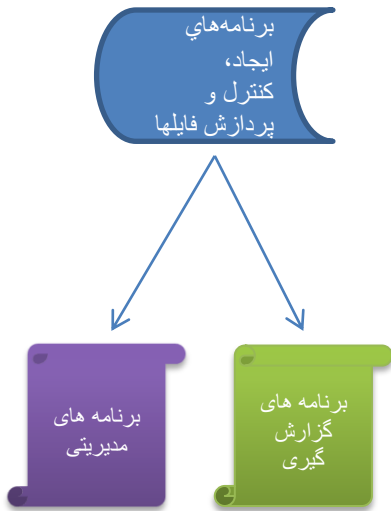
- الف : شیوه سنتی یا مشی فایلینگ
- ب : شیوه (مشی) پایگاهی



شیوه فایلینگ چیست ؟

کاربران

نمایش ساده شده مثنی فایلینگ



محیط فایلینگ

محیط فیزیکی ذخیره و بازیابی



معایب سیستم فایل

1. نیاز به برنامه نویسی زیاد و پیچیده : باید برنامه نویس برای کوچکترین برنامه نیز از جزئیات نحوه ذخیره سازی فایل ها کاملاً آگاه باشد (کدهای پیچیده و طولانی بنویسد)
2. وابستگی داده ای (data dependence) : لزوم تغییر کد برنامه های مربوط به یک فایل پس از اعمال یک تغییر در مشخصات یکی از فایل های
3. وابستگی ساختاری (Structural dependence) : لزوم تغییر کد برنامه های مربوط به یک فایل پس از اعمال هر تغییر در ساختار فایل "اضافه کردن یا حذف فایل" اطلاق میشود



معایب سیستم فایل

4. **افزونگی داده** : هر یک از بخشهای سازمان از فایل مجزایی استفاده میکردند ، یک قلم داده در چندین جا ذخیره میشود
- **هدر رفتن فضا و نیروی کار** : ذخیره یک قلم داده در چند جای مختلف علاوه بر هدر رفتن فضا حافظه بلکه باعث هدر رفتن نیروی کار نیز میشود
 - **ناسازگاری داده ها** : وجود دو یا چند مقدار متفاوت برای یک قلم داده ای (فیلد در یک رکورد) اطلاق میشود
 - **ناهنجاری اصلاح**
 - **ناهنجاری درج**
 - **ناهنجاری حذف**
5. **عدم وجود امکانات لازم برای تامین امنیت دادهها** : سطح دسترسی یکسان بوده و خطر ناامنی وجود داشته



IASBS
1992 - 2012

پایان جلسه اول