

فرادرس

فراتر از یک کلاس درس
www.faradars.org

پایگاه داده ها

مدرس:

فرشید شیرافکن

دانشجوی دکتری دانشگاه تهران

(کارشناسی و کارشناسی ارشد : کامپیوتر نرم افزار) (دکتری: بیوانفورماتیک)

فهرست

- فصل اول : تعاریف اولیه
- فصل دوم : مدل رابطه ای
- فصل سوم : نمودار ER
- فصل چهارم : جبر رابطه ای
- فصل پنجم : SQL
- فصل ششم : وابستگی ها
- فصل هفتم : نرمال تر سازی

منابع

۱- سی جی دیت

۲- دکتر رانکوهی

۳- دکتر حق جو

۴- سیلبرشاتز

۵- شیرافکن

فصل اول: تعاریف اولیه

FaraDars.org

مفاهیم اولیه

فایل های اطلاعاتی که به نوعی به هم مرتبط هستند، تشکیل یک **بانک اطلاعاتی** را می دهند.

فایل شامل مجموعه ای از رکوردها می باشد.

رکورد مجموعه ای از فیلدهای به هم مرتبط است.

فیلد کوچکترین جزء یک بانک اطلاعاتی می باشد.

مثلا در بانک اطلاعاتی **دانشگاه** چندین فایل وجود دارد مانند فایل دانشجویان که شامل چندین رکورد است، هر رکورد شامل اطلاعات یک دانشجو می باشد که از چند فیلد مانند شماره دانشجویی، نام، آدرس، معدل و.... تشکیل شده است.

چند تعریف

داده: نمایش پدیده ها و مفاهیم به صورت صوری و مناسب برای برقراری ارتباط یا پردازش.

اطلاع: داده پردازش شده .

داده همان مقدار واقعا ذخیره شده و اطلاع معنای داده است.

پایگاه داده: مجموعه ای از داده های ذخیره شده و پایا به صورت مجتمع و بهم مرتبط، با کمترین افزونگی، تحت

مدیریت یک سیستم کنترل متمرکز، مورد استفاده یک یا چند کاربر به صورت همزمان و اشتراکی .

روشهای ایجاد سیستم های کاربردی

۱- روش پایگاهی

در این روش نیازهای اطلاعاتی تمامی قسمتها مورد مطالعه قرار می گیرد تا بتوان یک سیستم یکپارچه (integrated) طراحی کرد. در این روش یک محیط ذخیره سازی واحد، مجتمع واشتراکی، تحت کنترل متمرکز وجود دارد که کاربران براساس نیاز خاص خود، پایگاه خود را تعریف کرده و هر کاربر تصور می کند که پایگاه خود را دارد.

۲- روش فایلینگ (ناپایگاهی)

مراحل اولیه طراحی و تولید برای هر قسمت به طور کلاسیک انجام شده و بعد از طراحی، مشخصات هر سیستم همراه با وظایف آنها مشخص می شود. برای ایجاد محیط ذخیره سازی اطلاعات از یک سیستم فایل و برای برنامه سازی از یک زبان سطح بالا استفاده می شود و در نهایت برای هر قسمت، یک سیستم کاربردی ایجاد می شود.

عناصر محیط پایگاه داده‌ها

۱- نرم افزار (DBMS - نرم افزار شبکه - برنامه های کاربردی - رویه های ذخیره شده)

۲- سخت افزار (ذخیره سازی - ارتباطی - پردازشگر)

۳- کاربر (موردی - همیشگی)

۴- داده

ساختارهای داده‌ای

یک **مدل داده ای** شامل یک **ساختار داده** است. (همچنین شامل امکانات **عملیات** بر روی پایگاه داده ها و **کنترل جامعیت** پایگاه داده ها است). ساختار داده ای امکانی است برای نمایش داده های موجودیت ها و انواع ارتباطات بین آنها.

انواع ساختارهای داده ای

- ۱- رابطه ای (RDS)
- ۲- سلسله مراتبی (HDS)
- ۳- شبکه ای (NDS)

ویژگی های ساختار رابطه ای

۱- مبنای تئوریک قوی دارد.

۲- دارای نمایش ساده از نظر کاربر

۳- دارای فقط یک عنصر ساختاری اساسی (جدول)

۴- امکان نمایش ارتباطات $1:1$, $1:N$, $N:M$

۵- ساده بودن منطق و دستور بازیابی

معماری پایگاه داده ها

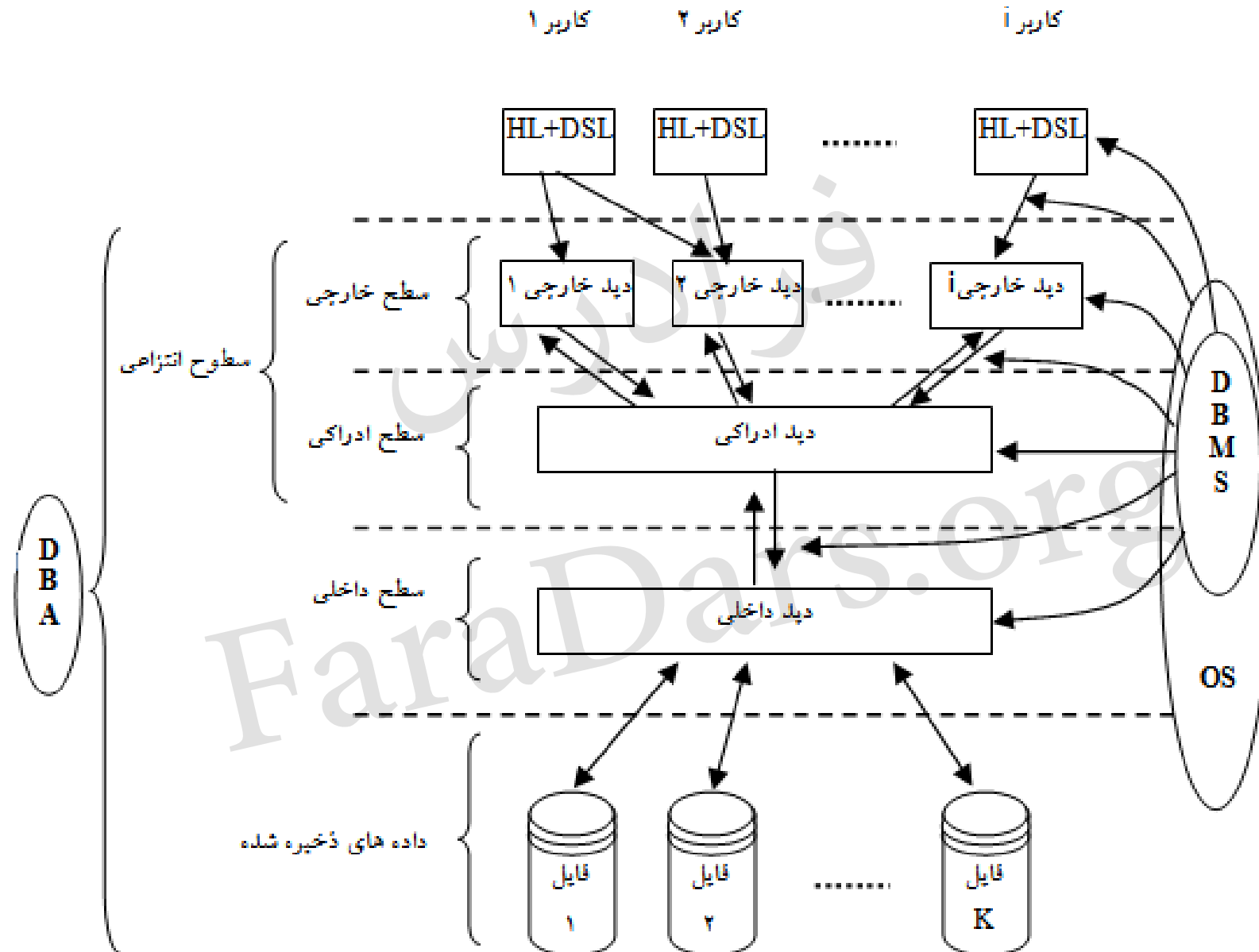
معماری استاندارد پایگاه داده ها که توسط ANSI پیشنهاد شد، یک معماری سه سطحی می باشد.

۱- سطح خارجی (External Level)

۲- سطح ادراکی (Conceptual Level)

۳- سطح داخلی (Internal Level)

شکل زیر معماری پایگاه داده ها را با همه اجزاء آن نمایش می دهد:



تعریف دیدها

دید، پنجره ای است که از آن کاربر می تواند محدوده پایگاه خود را ببیند و خارج از این محدوده، چیزی نمی بیند.

- ۱- دید خارجی : دید کاربر نسبت به داده های ذخیره شده در پایگاه داده ها
- ۲- دید ادارکی : دید طراح نسبت به داده های ذخیره شده در پایگاه داده ها
- ۳- دید داخلی : دید DBMS نسبت به داده های ذخیره شده در پایگاه داده ها

زبان داده ای فرعی (DSL : Data Sub Language)

دستورهای این زبان به سه قسمت تقسیم می شود:

۱- دستورات تعریف داده ها (DDL)

۲- دستورات کنترل داده ها (DCL)

۳- دستورات عملیات روی داده ها (DML)

انواع DSL

DSL می تواند هم مستقل و هم ادغام شدنی باشد.

۱- IDSL (مستقل از زبان میزبان)

۲- EDSL (ادغام شده با زبان میزبان)

ادغام می تواند **صریح** یا **ضمنی** باشد. در حالت ادغام صریح، عین دستورات DSL در برنامه به زبان میزبان نوشته می شود و در حالت ادغام ضمنی، دستورهای DSL به صورت توابع فرا خوانده می شوند. در حالت ادغام صریح، محیط برنامه سازی دو زبانی است و به دو کامپایلر نیاز دارد.

سیستم مدیریت پایگاه داده ها (DBMS)

این سیستم یکی از نرم افزارهای واسط بین محیط فیزیکی ذخیره و بازیابی و محیط منطقی برنامه سازی می باشد.

DBMS به برنامه ساز امکان می دهد تا پایگاه داده های خود را تعریف کرده و در آن عملیات خود را انجام دهد.

مدیر پایگاه داده ها (DBA)

مدیر پایگاه داده ها فردی است متخصص در پایگاه داده ها، با مسئولیت علمی و فنی که همراه با یک تیم تخصصی کار می کند.

چند وظیفه از وظایف تیم DBA :

- ۱- مشارکت در انتخاب DBMS و پیکربندی سخت افزاری و نرم افزاری لازم
- ۲- تصمیم گیری در مورد تعیین معماری سیستم پایگاه داده ها
- ۳- تصمیم گیری در مورد چگونگی سازماندهی مجدد پایگاه داده ها
- ۴- طراحی سطح ادراکی پایگاه داده ها (طراحی منطقی)
- ۵- نظارت بر تعیین دیدهای خارجی و نوشتن شماهای خارجی

استقلال داده ای

وابسته نبودن برنامه های کاربردی به داده های ذخیره شده را استقلال داده ای می نامند که مهمترین اهداف تکنولوژی پایگاه داده ها می باشد. به عبارتی استقلال داده ای عبارت است از تاثیر ناپذیری برنامه های کاربردی در سطح خارجی در قبال رشد پایگاه داده ها و تغییر در ساختار داده های عملیاتی است.

انواع استقلال داده ای:

۱- استقلال داده ای فیزیکی

مصونیت دیدهای کاربران و برنامه های کاربردی در قبال تغییرات در سطح داخلی - فیزیکی.

۲- استقلال داده ای منطقی

مصونیت دیدهای کاربران و برنامه های کاربردی در قبال تغییرات در سطح ادراکی.

کاتالوگ سیستم

کاتالوگ سیستم حاوی داده هایی است در مورد داده های ذخیره شده در پایگاه داده های کاربر. (متا داده)

استفاده از کاتالوگ باعث افزایش استقلال از داده می شود.

استفاده از کاتالوگ، تغییری در سرعت ذخیره یا بازیابی اطلاعات نمی گذارد.

تعریف جدول های مبنا، تعریف شاخص ها و تعریف جدول های مجازی در کاتالوگ ذخیره می شود.

مقادیر داده های درون جداول، در کاتالوگ سیستم ذخیره نمی شود.

محتویات کاتالوگ

- ۱- شمای خارجی، ادراکی، داخلی
- ۲- ضوابط کنترل ایمنی داده ها
- ۳- مشخصات پیکربندی سخت افزاری سیستم
- ۴- شرح سازمان فیزیکی داده های ذخیره شده
- ۵- مشخصات کاربران و حقوق دستیابی آنها به داده ها
- ۶- مشخصات برنامه های کاربردی
- ۷- مشخصات پایانه های متصل به سیستم
- ۸- قواعد جامعیت
- ۹- ارتباط بین برنامه های کاربردی و داده های ذخیره شده
- ۱۰- توابع تعریف شده توسط کاربران

تراکنش (TRANSACTION)

تراکنش : برنامه ای که یک کاربر در محیط بانک اطلاعاتی اجرا می کند.
پایان یک تراکنش یا موفق (commit) است و یا ناموفق (abort).

DBMS بر روی هر تراکنش کنترل هایی را انجام می دهد تا جامعیت بانک اطلاعاتی تضمین شود.
این کنترلها به ACID معروف می باشند :
Atomicity و Consistency و Isolation و Durability

ACID

۱- یکپارچگی (Atomicity)

به این معنی است که یا تمام دستورات یک تراکنش انجام می شود یا هیچکدام از دستورات اجرا نمی شوند. این خاصیت به همه یا هیچ موسوم است. (مثلا تراکنش انتقال مبلغی از یک حساب به حساب دیگر)

۲- همخوانی (Consistency)

یعنی هر تراکنش اگر به تنهایی اجرا شود بانک را از حالتی صحیح به حالتی صحیح دیگر منتقل می کند.

ACID

۳- انزوا (Isolation)

یعنی اثر تراکنش های همروند روی یکدیگر چنان باشد که ظاهرا هر کدام به طور مجزا و در انزوا انجام می شوند.

۴- پابائی (Durability)

به این معنی است که اثر تراکنش هایی که به مرحله انجام (commit) می رسند ماندنی است و به طور تصادفی از بین نمی رود. مثلا در تراکنش انتقال پول از حسابی به حساب دیگر ، بعد از واریز مبلغ تحت هیچ شرایطی (همچون آتش سوزی) اثر عمل انجام شده از بین نمی رود.

فرادرس پایان فصل اول

FaraDars.org

این اسلاید ها بر مبنای نکات مطرح شده در فرادرس
«پایگاه داده ها»
تهیه شده است.

برای کسب اطلاعات بیشتر در مورد این آموزش به لینک زیر مراجعه نمایید

faradars.org/fvsft105