فصل ۱

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 و رکورد مجموعه ای از فیلدهای به هم مرتبط است و 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.

تعاریف اولیه

داده	نمایش پدیده هاومفاهیم به صورت صوری و مناسب برای برقراری ارتباط یا پردازش.
اطلاع	داده پردازش شده می باشد.
شناخت	نمایش نمادین جنبه هایی از بخشی از جهان واقع می باشد.به عبارتی نوعی اطلاع سطح بالاتر است.
پایگاه داده ها	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.

داده همان مقدار واقعا ذخیره شده و اطلاع معنای داده است. یعنی اطلاع و داده با هم فرق دارند. اطلاع دارای خاصیت ارتباط دهندگی و انتقال دهندگی دارد، در حالیکه داده این خواص را ندارد.

- اطلاع و شناخت حاصل عملیاتی روی داده هستند ولی نوع عملیات لازم برای به دست آوردن آنها متفاوت است.

- منظور از پایایی داده ها، این است که پس از پایان اجرای برنامه کاربر، داده ها در سیستم باقی می مانند.

فرادرس

روش های ایجاد سیستم های کاربردی

یک سیستم کاربردی را می توان به دو روش ایجاد کرد:

۱- روش فایلینگ(ناپایگاهی)

در روش فایلینگ(سنتی)، نیازهای اطلاعاتی و پردازشی هر قسمت از محیط برآورده می شوند. 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. در این روش، برای ایجاد محیط ذخیره سازی اطلاعات از یک سیستم فایل (FS) و برای برنامه سازی از یک زبان سطح بالا استفاده می شود و در نهایت برای هر قسمت، یک سیستم کاربردی ایجاد می شود.

معایب روش فایلینگ

۱- عدم وجود محیط مجتمع ذخیره سازی

۲- عدم وجود سیستم کنترل متمرکز

۳- عدم وجود ضوابط ایمنی کارا

۴- عدم امکان اشتراکی شدن داده ها

۵- تکرار در ذخیره سازی اطلاعات

۶- مصرف نامناسب امکانات سخت افزاری و نرم افزاری

۷- وابسته بودن برنامه های کاربردی به محیط ذخیره سازی داده ها

۸- حجم زیاد برنامه سازی

۲- روش پایگاهی

در این روش نیازهای اطلاعاتی تمامی قسمتها مورد مطالعه قرار می گیرد تا بتوان یک سیستم یکپارچه (integrated) طراحی کرد. داده های سازمان مدلسازی معنایی (SDM) می شوند و مشخصات سیستم یکپارچه تعیین می شود. برای سیستم مدیریت متمرکز از یک یا چند DBMS استفاده می شود. طراحی پایگاه داده ها در سطوح لازم انجام می شودو کاربران هر قسمت، پایگاه داده های خود را تعریف می کنند و با آن کار می کنند.

در واقع در روش پایگاهی 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.

- کاربران در روش پایگاهی بطور همزمان از سیستم استفاده می کنند.
- در روش پایگاهی نسبت به داده های ذخیره شده، تنوع و کثرت دید وجود دارد.

دانلود رایگان مجموعه کتب کارشناسی و کنکور ارشد کامپیوتر http://faradars.org/computer-engineering-exam

- در روش پایگاهی نسبت به روش فایلینگ، حجم برنامه ها کمتر و برنامه سازی آسانتر است.

- پایگاه داده ها بر حسب تعداد رکوردهای آن، به دسته های کوچک، متوسط، بـزرگ (LDB) و خیلـی بـزرگ (VLDB) تقسیم می شوند.

فرادرس

نسل های ذخیره و بازیابی اطلاعات

پنج نسل ذخیره و بازیابی اطلاعات عبارتند از:

۱- نسل فایلهای ساده ترتیبی

در این نسل از نوار به عنوان رسانه خارجی استفاده می شد. ساختار فایلها ترتیبی بود و ساختار فیزیکی و منطقی فایل یکسان بود. تنها روش پردازش فایلها، یکجا بود و نرم افزار واسطی برای مدیریت پردازش فایلها وجود نداشت امکان اشتراکی کردن داده ها وجود نداشت و تکرار در ذخیره سازی داده ها در بالاترین حد بود. هر نوع تغییر در ساختار داده ها یا رسانه ذخیره سازی موجب بروز تغییر در برنامه می شد و نسخه های متعددی از یک فایل نگهداری می شد.

۲- نسل AM

در این نسل یعنی شیوه های دستیابی (Access Methods)، نرم افزاری به نام شیوه های دستیابی ایجاد شد و برنامه کاربر را از پرداختن به جنبه های فیزیکی محیط ذخیره سازی مستقل کرد. رسانه این نسل دیسک بود و امکان دسترسی ترتیبی و مستقیم به رکوردها (نه فیلدها) وجود داشت و تا حدودی ساختار فیزیکی و منطقی فایلها از یکدیگر جدا شدند.

۳- نسل DMS

در این نسل یعنی نسل سیستم مدیریت داده ها ، نرم افزار کاملتری نسبت به نرم افزار دستیابی ایجاد شد و واسط بین برنامه های کاربردی و فایلهای محیط فیزیکی شد. در این نسل:

۱- از داده های اشتراکی در برنامه استفاده شد.

۲- آدرس دهی به داده ها در سطح فیلد ممکن شد.

۳- امکان بازیابی براساس چند کلید مهیاگشت.

۴- میزان افزونگی کاهش یافت.

۵- برنامه های کاربردی در قبال رشد فایلها مصون شدند.

۴- نسل DBMS

در نسل چهارم، برنامه های کاربردی از ویژگی های محیط فیزیکی ذخیره سازی مستقل شدند. کاهش افزونگی در ذخیره سازی داده ها، افزایش سرعت دستیابی به داده ها، بالا رفتن امنیت، امکان استفاده اشتراکی از داده ها و وجود یک محیط انتزاعی (Abstractive) از دیگر ویژگی های این نسل می باشد.

۵- نسل KBS

در این نسل یعنی نسل بانک معرفت(شناخت) به کمک مفاهیم هوش مصنوعی و سیستم های خبره، سیستمی طراحی شد که قادر به استنتاج منطقی از داده های ذخیره شده می باشد. در نسل پنجم سیستم بانک معرفت (KBS) ایجاد شد که مسئولیت ذخیره سازی شناخت، تامین جامعیت و امنیت بانک و تامین نیازهای کاربران را بر عهده دارد.

دانلود رایگان مجموعه کتب کارشناسی و کنکور ارشد کامپیوتر http://faradars.org/computer-engineering-exam

بانک معرفت، بانکی حاوی واقعیت های ساده و قواعد عام است که به طور صریح بیان شده باشند.

پایگاه بصیرت(شناخت) پویا و پایگاه داده ها ایستا می باشد.

میزان وابستگی برنامه ها به داده ها در نسل DBMS ها از نسل DMS ها کمتر است و در نسل DMS ها از نسل FS ها کمتر است.

عناصر محیط پایگاه داده ها

عناصر چهارگانه محیط پایگاه داده ها عبارتند از:

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

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

۳- کاربر(موردی (نامنظم)- همیشگی (منظم))

۴- داده

تذکر: اصلی ترین سخت افزار ذخیره سازی داده ها، دیسک است. همچنین از نوار مغناطیسی هم به عنوان رسانه کمکی برای تهیه Backup استفاده می شود.

به هر استفاده کننده از سیستم پایگاه داده ها، کاربر می گویند.

کاربر منظم یا نامنظم خود بر دو نوع برنامه ساز و نابرنامه ساز تقسیم می شوند. کاربر نابرنامه ساز از طریق منو، هدایت می شود. کاربر برنامه ساز می تواند سیستمی یا کاربردی باشد.

برنامه ساز کاربردی، برنامه های بهره برداری از پایگاه داده ها را می نویسد و برنامه ساز سیستم، برنامه های ایجاد و کنترل پایگاه داده ها را می نویسد.

یک نوع کاربر نیز به نام کاربر پایانی (End–user) وجود دارد که به هر دو نوع کاربر برنامه ساز کاربردی و کاربر نابرنامه ساز گفته می شود.

فرادرس

ساختارهای داده ای

یک مدل داده ای شامل یک ساختار داده (DS) است. در واقع طراحی منطقی پایگاه داده ها به کمک مفاهیم اساسی یک مدل داده ای و در چار چوب ساختار داده ای آن مدل انجام می گیرد. ساختار داده ای امکانی است برای نمایش داده های موجودیت ها و انواع ارتباطات بین آنها.

عناصر تشکیل دهنده هر مدل

۱- ساختار داده ای

۲- امکانات عملیات در پایگاه داده ها

۳- امکانات کنترل جامعیت پایگاه داده ها

انواع ساختارهای داده ای

مدل های داده ای بر پایه رکورد بر سه نوع می باشند:

۱- رابطه ای (RDS)

۲- سلسله مراتبی (HDS)

۳- شبکه ای (NDS)

تذکر: مدلهای داده ای بر پایه شیء عبارتند از: (ER , Semantic , Functional)

مفهوم ساختار داده ای بخشی از مفهوم مدل داده ای است.

ساختار رابطه ای

ساختار رابطه ای (RDS) ، دارای ویژگی های زیر می باشد:

۱- مبنای تئوریک قوی دارد. (تامین کننده محیط انتزاعی به طور کامل)

۲- مسطح بودن محیط (مانند یک فایل ترتیبی ساده)

۳- دارای نمایش ساده از نظر کاربر

۴- دارای فقط یک عنصر ساختاری اساسی (جدول)

۵- امکان نمایش ارتباطات 1:1, 1:N, N:M

دانلود رایگان مجموعه کتب کارشناسی و کنکور ارشد کامپیوتر http://faradars.org/computer-engineering-exam

۶- ساده بودن منطق و دستور بازیابی

۷- دارای رویه پاسخگوی قرینه برای پرسشهای قرینه

فرادرس

فرادرس

فرادرس

ساختار سلسله مراتبی

ساختار سلسله مراتبی(HDS) ، قدیمی ترین ساختار داده ای برای طراحی منطقی پایگاه داده ها می باشد که دو عنصر اساسی دارد: نوع رکورد و نوع پیوند پدر– فرزندی (PCL).

بین هر دو رکورد پشت سرهم در یک مسیر درخت، پیوند پدر فرزندی وجود دارد که ارتباط 1: N را نمایش می دهد. در این ساختار، هر رکورد فرزند، تنها یک رکورد پدر دارد، یعنی در یک نوع PCL شرکت دارد.

ویژگی های ساختار سلسله مراتبی عبارتند از:

۱- مبنای ریاضی ندارد.

۲- مناسب برای ارتباط 1 : N .

۳- سادگی نمایش ساختار رابطه ای را ندارد.

۴- نمایش ارتباط N : M در آن مشکل است.

۵- جستجو حتماً باید از ریشه انجام شود.

۶- نداشتن تقارن ساختار جدولی.

۷- داشتن تعدادی محدودیت جامعیت.

۸- مشکل بودن دستور بازیابی در آن نسبت به ساختار جدولی.

۹- نمایش ارتباط با درجه بیشتر از دو.

- پایگاه داده سلسله مراتبی، مجموعه ای منظم از نمونه های یک یا چند نوع سلسله مراتب می باشد.

- در HDS ، برای نمایش ارتباط چند به چند بین دو نوع موجودیت، می توان:
  ۱- دو نوع سلسله مراتب جدا طراحی کرد.
  ۲- دو نوع سلسله مراتب به هم مرتب طراحی کرد.
  ۳- یک نوع سلسله مراتب طراحی کرد.

- در HDS ، یک نوع رکورد می تواند چند نوع رکورد فرزند داشته باشد.

فرادرس

در HDS ، صفات نوع ارتباط R با چندی 1:N بین دو نوع موجودیت E وF با فیلدهایی در نوع رکـورد **فرزنـد** نمـایش داده می شوند.

فرادرس

ساختار شبکه ای

ساختار شبکه ای (پلکس) توسط گروه DBGT پیشنهاد شد و اولین سیستم مدیریت پایگاه داده ای شبکه ای IDMS نام دارد. این سیستم گاهی به نام کوداسیل نامیده می شود. شبکه، نوعی گراف جهت دار است. در ساختار شبکه هر گره فرزند می تواند بیش از یک گره پدر داشته باشد، بنابراین گسترش یافته ساختار سلسله مراتبی است.

ساختار شبکه ای از دو عنصر تشکیل شده است:

۱- نوع رکورد

۲- نوع مجموعه (مجموعه کوداسیلی)

مجموعه کوداسیلی از سه جزء تشکیل شده است:

۱- نام مجموعه

۲- یک نوع رکورد مالک

۳- یک نوع رکورد عضو

درج مالک، بدون عضو، ممکن نمی باشد و درج یک عضو بدون وجود مالک ممکن نمی باشد.

با حذف مالک، اعضای آن نیز حذف می شوند.

ویژگی های ساختار شبکه ای

۱- تقارن دارد.

۲- مبنای ریاضی ندارد.

۳- سادگی ساختار رابطه ای را ندارد.

۴- عدم آنومالی در عملیات ذخیره سازی.

۵- فقط مخصوص نمایش ارتباطات 1:N نیست و هر نوع ارتباطی را می تواند نمایش دهد.

۶- امکان ناسازگاری داده ها در آن کمتر از سلسله مراتبی است.

۷- قواعد جامعیت ذاتی دارد.

۸- فزونکاری (به علت ایجاد و اصلاح اشاره گرها) دارد.

۹- دستور بازیابی پیچیده تری دارد.

۱۰- رویه جستجوی رکورد در آن نسبت به ساختارهای دیگر پیچیده تر است.

۱۱- اصل وحدت عملگر در یک عمل واحد مانند درج رعایت نمی شود.

دانلود رایگان مجموعه کتب کارشناسی و کنکور ارشد کامپیوتر http://faradars.org/computer-engineering-exam

- در NDS ، نوع رکورد مالک یک نوع مجموعه، می تواند عضو یا مالک در نوع مجموعه دیگر باشد.

- در NDS اگر n نوع موجودیت و احیانا دارای صفات چند مقداری، در یک نوع ارتباط شرکت داشته باشند، حداقل n مجموعه کوداسیلی برای طراحی پایگاه داده شبکه ای (NDB) لازم است.

- در نمونه مجموعه مدل شبکه ای بر خلاف مجموعه ریاضی:
  ۱- حداقل یک عنصر وجود دارد.
  ۲- یک عنصر وجود دارد که از نظر نوع با عناصر دیگر فرق دارد.
  ۳- بین بعضی از عناصر نوعی نظم وجود دارد.

- اگر n نوع موجودیت در یک نوع ارتباط شرکت داشته باشند و این نوع ارتباط a صفت داشته باشد، رکورد پیوند دهنده در NDS ، دارای n+a فیلد در سطح پیاده سازی است.

- در NDS ، صفات نوع ارتباط R با چندی 1:N بین دو نوع موجودیت E و F با فیلدهایی در نوع رکورد **عضو** نمایش داده می شوند.

- فقط در RDS ، مفهوم نظم مطرح نیست.

- ساختار سلسله مراتبی تقارن ندارد، اما ساختار شبکه ای دارای تقارن است.

- تعداد عناصر ساختاری اساسی در RDS برابر 1 و در HDS و NDS برابر 2 می باشد.

- فقط در RDS ، جستجو در پایگاه داده ها به صورت اتوماتیک(خودکار) است.

- با توجه به سه نوع ساختار رابطه ای، سلسله مراتبی و شبکه ای، سه دسته DBMS به ترتیب زمان پیدایش آنها به نام های HDBMS , NDBMS و RDBMS وجود دارد. بعد از آنها OODBMS است.

فرادرس

معماری پایگاه داده ها

معماری استاندارد پایگاه داده ها که توسط ANSI پیشنهاد شد، یک معماری سه سطحی می باشد. این سطوح عبارتند از:

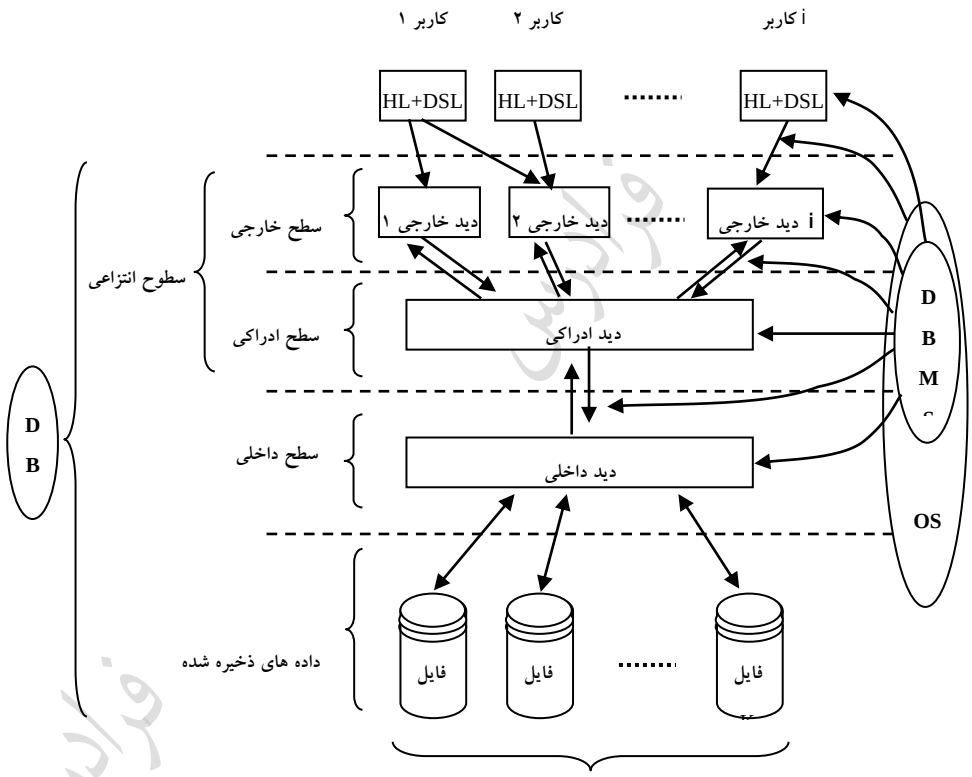
۱- سطح خارجی (External Level)

۲- سطح ادراکی یا مفهومی (Conceptual Level)

۳- سطح داخلی (Internal Level)

تذکر: به دو سطح خارجی و ادراکی، سطوح انتزاعی می گویند.

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



شرح اجزای معماری پایگاه داده ها

۱- دید ادارکی (Conceptual View)

دید، پنجره ای است که از آن کاربر می تواند محدوده پایگاه خود را ببیند و خارج از این محدوده، چیزی نمی بیند. دید ادارکی، دید طراح نسبت به داده های ذخیره شده در پایگاه است که در برگیرنده تمام نیازهای کاربران در محیط عملیاتی می باشد. دید ادراکی چون در یک محیط انتزاعی مطرح است، بر یک ساختار داده ای مشخص بنا شده است.

شمای ادراکی:

به توصیف و شرح دید ادارکی، شمای ادراکی می گویند. در واقع شمای ادراکی برنامه ای است شامل دستورات تعریف داده ها و کنترل داده ها (نه دستورات عملیات بر روی داده ها).

۲- دید خارجی (External View)

دید خارجی، دید کاربر نسبت به داده های ذخیره شده در پایگاه داده ها می باشد. این دید:

۱- یک دید جامع نمی باشد.

۲- روی دید ادراکی طراحی و تعریف می شود.

۳- در سطح انتزاعی مطرح است.

۴- مبتنی بر همان ساختار داده ای می باشد که دید ادراکی بر اساس آن طراحی می شود.

- دید خارجی تامین کننده اشتراک داده ای در یک سیستم پایگاهی است.
- سیستمی که ساختار داده ای در سطح خارجی و سطح ادراکی آن یکسان نباشد را دو ساختاره می گویند. (مثلاً ساختار داده ای در سطح خارجی و سلسله مراتبی در سطح ادراکی)
- دید ادراکی و دید خارجی هر دو جدولی هستند، اما دید خارجی یک جدول مجازی است، یعنی داده ذخیره شده خاص خود را ندارد.
- یک کاربر می تواند چند دید خارجی داشته باشد.
- کاربر خارجی پس از تعریف دید خود، همیشه بطور پیش فرض مجاز به انجام عمل بازیابی است.

شمای خارجی:

به وصف یا تعریف دید خارجی، شمای خارجی می گویند. در واقع شمای خارجی برنامه ای است حاوی دستورات تعریف داده ها و گاه کنترل داده ها.

فرادرس

۳- دید داخلی (Internal View)

دید DBMS و طراح پایگاه داده ها نسبت به داده های ذخیره شده را دید داخلی می گویند.

تعاریف ایندکس های انبوه و غیر انبوه در دید داخلی وجود دارد.

شمای داخلی:

به تعریف دید داخلی، شمای داخلی می گویند. در واقع نوعی برنامه است که توسط خود DBMS تولید می شود. انواع رکوردها در شمای داخلی تعریف می شوند و شامل دستورهای لازم جهت ایجاد فایلها و کنترل آنها می باشد.

۴-کاربر

هر استفاده کننده از پایگاه داده ها را کاربر می نامند. کاربر می تواند موردی یا همیشگی باشد.

کاربر موردی معمولا برنامه ساز است.

۵- زبان میزبان (HL : Host Language)

زبان میزبان می تواند یکی از زبانهای برنامه سازی چون پاسکال ، C ، ایدا ، ... باشد.

DBMS ای که تعداد HL های مورد پذیرش آن زیاد است، مطلوبتر است، چون موجب تنوع کاربردها و کاربران می شود.

هر چه تعداد HL مورد پذیرش DBMS بیشتر شود، نیاز حافظه ای سیستم بیشتر می شود.

۶- 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)

DSL می تواند هم مستقل و هم ادغام شدنی باشد. (I/E . DSL)

دانلود رایگان مجموعه کتب کارشناسی و کنکور ارشد کامپیوتر http://faradars.org/computer-engineering-exam

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.

- DSL بهتر است از نظر برنامه سازی کامل باشد، یعنی تمام ساختهای اصلی یک زبان برنامه سازی متعارف را داشته باشد. در این حالت DSL را نمی توان دیگر یک زبان داده ای فرعی در نظر گرفت و یک زبان برنامه سازی با امکانات پایگاهی است یعنیDSL با HL ، پیوند قوی دارد.

- برای یک DSL استاندارد، کمال ساختاری، الزامی است، اما کمال محاسبه ای(برنامه سازی) الزامی نیست.

- هر گاه DSL از نظر محاسبه ای کامل باشد، در اینصورت اصلا نیازی به ادغام آن در HL نیست.

- DSL بهتر است کامپایلری باشد نه مفسری.

- سیستمی با DSL کامپایلری،زمانی کارایی دارد که نرخ پرسش های موردی پایین و نرخ پرسش های عادی بالا باشد.

زبان رویه ای و نارویه ای

زبان ها می توانند رویه ای و یانارویه ای باشند. زبان رویه ای حالت دستوری و زبان نارویه ای حالت توصیفی دارد. در زبان رویه ای، برنامه ساز رویه ای برای انجام عمل مورد نظرش می نویسد، اما در زبان نارویه ای برنامه ساز الگوریتم دسترسی به داده را مشخص نمی کند. مثلاً در هنگام بازیابی فقط می گوید چه داده ای را می خواهد و روش بازیابی را مشخص نمی کند.

- زبان نارویه ای باعث بهبود تولید برنامه های کاربردی و تسهیل افزایش تعداد کاربران پایانی می شود.

- از نظر کاربر برنامه ساز، وقتی DSL نارویه ای است، جستجو در پایگاه داده ها، اتوماتیک است.

دانلود رایگان مجموعه کتب کارشناسی و کنکور ارشد کامپیوتر http://faradars.org/computer-engineering-exam

۷- 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.

عوامل موثر در دسته بندی DBMS ها عبارتند از:

قیمت، نوع کاربرد، سیستم فایل ، نوع و ماهیت DSL، نوع معماری، نوع ساختار داده‌ای، محیط سخت افزاری و محیط سیستم عامل.

نرم افزار DBMS از نمای بیرونی از دو واحد تشکیل شده است:

۱- واحد پردازشگر پرسش‌ها و برنامه‌های کاربردی

۲- واحد ایجاد و مدیریت داده‌های ذخیره شده

نرم افزار DBMS از نمای درونی از سه لایه تشکیل شده است:

۱- هسته

۲- تسهیلات نرم افزاری

۳- مدیریت محیط پایگاه داده‌ها

فرادرس

در طراحی سطح ادراکی و سطح خارجی پایگاه داده ها، باید امکانات DBMS را در نظر گرفت.

DBMS می داند پایگاه داده ها از چه فایل هایی تشکیل شده، استراتژی دستیابی به رکوردهای فایل چیست و فرمت هر نوع رکورد داخلی چگونه است.

دلایل ایجاد سربار(فزونکاری) در DBMS عبارتند از:

۱- لزوم اعمال قواعد جامعیت

۲- لزوم انجام تبدیلات بین سطوح

۳- لزوم اعمال ضوابط ایمنی

۸- 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.

وظایف تیم DBA عبارت است از:

الف – مشارکت در :

۱- تفهیم نقش داده به مدیریت سازمان

۲- تفهیم مزایای تکنولوژی پایگاه داده ها

۳- تصمیم گیری در مورد استفاده یا عدم استفاده از تکنولوژی پایگاه داده ها

۴- انتخاب DBMS و پیکربندی سخت افزاری و نرم افزاری لازم

ب – تصمیم گیری در مورد:

۱- تعیین معماری سیستم پایگاه داده ها

۲- انتخاب و انتساب اعضاء تیمهای اجرایی

۳- زبان برنامه سازی مورد نیاز

۴- مشخصات ساختار سطح داخلی پایگاه داده ها و تعیین ساختار فایلهای مناسب

۵- نوشتن شمای داخلی (طراحی فیزیکی)

۶- چگونگی رشد دادن پایگاه داده ها

۷- چگونگی سازماندهی مجدد پایگاه داده ها

۸- چگونگی ترمیم پایگاه داده ها

ج – طراحی :

۱- سطح ادراکی پایگاه داده ها (طراحی منطقی)

۲- واسطهای کاربری

۳- برنامه های کاربردی

د- نظارت بر :

۱- تعیین دیدهای خارجی و نوشتن شماهای خارجی

۲- وارد کردن داده ها

۳- تهیه مستندات لازم

۴- عملیات انجام شونده در پایگاه داده ها

فرادرس

فرادرس

فرادرس

فرادرس

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.

استقلال داده ای بر دو نوع است:

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.

در DBMS های جدید،استقلال داده ای فیزیکی کاملا تامین است، ولی استقلال داده ای منطقی کاملا تامین نیست.

تغییر در سطح ادراکی یعنی تغییر در طراحی منطقی پایگاه داده ها و تغییر در شمای ادراکی، که این تغییر دارای دو وجه است: رشد پایگاه در سطح ادراکی و سازماندهی مجدد پایگاه در سطح ادراکی.

از مزایای مهم تکنولوژی پایگاه داده ها،استقلال داده ای است که لازمه تامین آن، معتبر ماندن شمای خارجی پس از تغییرات در شمای ادراکی و شمای داخلی است.

فرادرس

کاتالوگ سیستم

کاتالوگ سیستم حاوی داده هایی است در مورد داده های ذخیره شده در پایگاه داده های کـاربر. بـه ایـن داده هـای ذخیـره شده، متا داده می گویند.

- متا داده از دید کاربر خارجی پنهان است.
- شمای ادراکی و خارجی در کاتالوگ سیستم نگهداری می شود.

فرادرس

محتویات کاتالوگ در سیستم های مختلف یکسان نیست ولی بطور کلی شامل اطلاعات زیر می باشد:

۱- شماهای 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

- تعریف جدول های مبنا، تعریف شاخص ها و تعریف جدول های مجازی در کاتالوگ ذخیره می شود.
- دیکشنری فعال توسط DBA مورد استفاده قرار می گیرد.
- از دستورات DDL برای ایجاد کاتالوگ استفاده می شود.
- برای تغییر محتوای کاتالوگ از دستورات DML (مانند Insert ، Delete ، Update) استفاده می شود.
- استفاده از کاتالوگ باعث افزایش استقلال از داده می شود.
- استفاده از کاتالوگ، تغییری در سرعت ذخیره یا بازیابی اطلاعات نمی گذارد.
- مقادیر داده های درون جداول، در کاتالوگ سیستم ذخیره نمی شود.

فرادرس

تراکنش (TRANSACTION)

تراکنش به برنامه ای گفته می شود 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 که به ترتیب معرف Atomicity وConsistency و Isolation وDurability می باشند.

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.

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- پایائی (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.

سیستم مدیریت پایگاه داده ها در فایل ثبت تراکنش ها، ،شناسه تراکنش،زمان انجام عمل و نوع عمل را ذخیره می کند.

تراکنش های هم روند

تراکنش ها می توانند اصل سریالیتی را رعایت نکنند و به طور هم روند اجرا شوند. به طور نمونه دو تراکنش A و B را در نظر بگیرید که A دو عمل a1 و a2 و B دو عمل b1 و b2 را انجام می دهند. این اعمال در اجرای هم روند می توانند به ترتیب زیر انجام گیرند:

زمان	تراکنش A	تراکنش B
t_1	a1	-
t_2	-	b1
t_3	a2	-
t_4		b2

تذکر: مبحث هم روندی در درس سیستم های عامل (بخش همزمانی) کتاب اینجانب بررسی شده است.

تذکر: برای جلوگیری از دخالت تراکنش ها در اجرای یکدیگر، می توان از قفل استفاده کرد. مبحث قفل در درس پایگـاه داده پیشرفته در مقطع کارشناسی ارشد بررسی می شود.

فرادرس

فرادرس

فرادرس

معماری سیستم پایگاه داده‌ها

نحوه ترکیب اجزای سیستمی، شامل حداقل یک پایگاه داده‌ها، یک DBMS ، یک سیستم عامل و یک کامپیوتر با دستگاههای جانبی و تعدادی کاربر را، معماری سیستم پایگاه داده‌ها می‌گویند.

انواع معماری‌ها عبارتند از:

۱-معماری متمرکز

یک پایگاه داده‌ها روی یک سیستم کامپیوتری ایجاد می‌شود و به سیستم کامپیوتری دیگری ارتباط ندارد.

۲- معماری مشتری خدمتگزار

در معماری مشتری خدمتگزار (Client-Server) قسمتی از پردازش توسط یک ماشین و قسمتی دیگر توسط ماشین دیگر انجام می‌شود. یعنی مسئولیت ها بطور منطقی تقسیم شده است. داده‌ها در سایت client ذخیره می‌شوند و برنامه‌های کاربردی در سایت server اجرا می‌شوند.

انواع معماری مشتری خدمتگزار(C/S DB) :

الف- معماری چند مشتری/یک خدمتگزار (MC/S DB)

ب- معماری یک مشتری/چند خدمتگزار(C/MS DB)

ج- معماری چند مشتری/چند خدمتگزار(MC/MS DB)

مزایای معماری مشتری- خدمتگزار نسبت به معماری متمرکز

الف- تقسیم پردازش

ب- اشتراک داده‌ها

ج- کاهش ترافیک شبکه

د- استقلال ایستگاههای کاری

۳-معماری توزیع شده

این معماری از ترکیب دو تکنولوژی پایگاه داده‌ها و شبکه معماری توزیع شده حاصل می‌شود. پایگاه داده‌های توزیع شده (DDB) یعنی مجموعه‌ای از چند پایگاه دادهٔ بهم مرتبط و توزیع شده روی یک شبکه که از نظر کاربران، پایگاه واحدی است.

در معماری DDB ، یک شمای ادارکی سراسری وجود دارد.

در معماری DDB ، وجود یک طرح واحد توزیع داده ها،الزامی است.

تذکر: صرف وجود شبکه در یک محیط، به معنای وجود پایگاه توزیع شده در آن محیط نیست.

مزایای معماری توزیع شده	معایب معماری توزیع شده
۱- اشتراک داده‌ها	۱- هزینه بالا
۲- کاهش هزینه ارتباطات	۲- مصرف حافظه بیشتر
۳- دستیابی بهتر به داده‌ها	۳- پیچیدگی پیاده سازی
۴- تسهیل گسترش سیستم	۴- پیچیدگی طراحی سیستم
۵- استفاده از پایگاه داده‌های از قبل موجود	۵ – وجود تهدیدهای بالقوه برای امنیت سیستم
۶- سازگاری با سازمانهای جدید	

۴-معماری با پردازش موازی

نوع گسترش یافته معماری توزیع شده است که برای دستیابی پذیری بالا طراحی می‌شود. سیستم‌های موازی قادر به اجرای موازی تعداد زیادی تراکنش در ثانیه می‌باشند.

مدل های ممکن برای معماری موازی

الف- معماری با حافظه اصلی مشترک

ب- معماری با دیسک مشترک

ج- معماری بی اجزا مشترک

مزایای معماری موازی با دیسک های مشترک

الف- کاهش مصرف حافظه جانبی

ب- تسهیل گسترش سیستم

ج- تسهیل تحمل خرابی

فرادرس