

عنوان مقاله : سیستم های مدیریت پایگاه داده

چکیده

امروزه بیشتر از همیشه، موقعيت یک سازمان وابسته به قابلیت بدست آوردن داده های دقیق و زمانی درباره عملکرد هایش است تا این داده ها را با بازدهی بیشتر بکار گیرد و از آنها استفاده کند تا فعالیت هایش را بررسی و هدایت کند. الفاظی مانند بزرگراه اطلاعات---- و پردازش اطلاعات یک صنعت چند بیلیون دلاری در حال رشد است.

مقدار اطلاعات در دسترس برای ما علاوه در حال فوران است و ارزش داده ها به عنوان یک امتیاز مدیریتی به وسعت شناخته شده است.

كلمات کلیدی: DBMS، انزواع، پایگاه داده، ساختمان داده.

مقدمه

یک پایگاه داده مجموعه ای از داده ها است که معمولاً فعالیت یک یا چند سازمان مرتبط را توضیح میدهد. برای مثال: پایگاه داده دانشگاه ممکن است اطلاعاتی درمورد موارد زیر داشته باشد:

- موجودیت هایی مانند دانش آموزان، هیئت علمی ، دوره ها و کلاس ها.
- روابط بین موجودیت ها مانند ثبت نام های دانشجویان در دوره ها و استفاده کلاس ها برای هر دوره.

یک سیستم مدیریت پایگاه داده (Database Management System) یا به اختصار DBMS نرم افزاری است که به منظور یاری در حفظ کردن و بهره برداری از مجموعه های بزرگی از داده ها طراحی شده است که نیاز به آنها به اندازه کاربردشان در حال افزایش روزافزون میباشد. راهکار جایگزین برای استفاده از DBMS استفاده از راهکار های خاص منظوره (ad hoc) است. برای مثال: داده را در اسناد قرار دهیم و سپس کد وابسته به نرم افزار را بنویسم.

مزیت های DBMS

استفاده از DBMS برای مدیریت داده مزیت های زیادی دارد.

استقلال داده:

برنامه های اجرایی باید تا بیشترین حد ممکن از جزئیات قرار گیری داده و حافظه مستقل باشند. DBMS یک نمای انتزاعی خوب از داده ارائه میدهد تا که اجرایی را از چنین جزئیاتی مجزا کند.

دسترسی موثر به داده:

DBMS از تکنیک های پیچیده ای برای ذخیره و گرفتن کارآمد داده ها استفاده میکند . این قابلیت به طور مخصوص زمانی حائز اهمیت است که داده روی یک دستگاه خارجی ذخیره میشود.

یکپارچگی و امنیت داده:

وقتی داده همواره از طریق DBMS مورد دسترسی قرار میگیرد، DBMS میتواند با اعمال محدودیت هایی یکپارچگی داده را الزام کند. برای مثال قبل وارد کردن اطلاعات حقوقی برای یک کارمند DBMS میتواند چک کند که از بودجه آن بخش پا فرا نگذارد. همچنین میتواند کنترل دسترسی اعمال کند تا داده هایی که مورد دسترسی کلاس های مختلفی از کاربران است را مدیریت کند.

اداره کردن داده ها:

زمانیکه چندین کاربر داده ای را به اشتراک میگذارند ، مرکزیت سازماندهی داده ها میتواند بهبود بسزایی در روند حاصل کند. صاحب تجربگان حرفه ای که درک درستی از اصول مدیریت داده و تنوع استفاده کاربران از آن دارند برای به حداقل رساندن افزونگی آن و تنظیم درست داده به منظور بازیابی کارآمد، میتوانند مسؤول مرتب سازی نمایش داده باشند.

دسترسی و بازیابی از کوش متقارن:

یک DBMS برنامه تقارن دسترسی ها به داده را به گونه ای برنامه ریزی میکند که دسترسی برای کاربران بصورت نوبتی بنظر میرسد و از کاربران دربرابر شکست های سیستم محافظت میکند.

کاهش زمان توسعه برنامه:

DBMS به وضوح از بسیاری از توابع مورد معمول و مورد استفاده برنامه هایی که به داده های DBMS دسترسی دارند، پشتیبانی میکند که این مسئله در کنار فراهم کردن یک رابط سطح بالا برای داده ها، توسعه سریع برنامه ها را آسان میسازد. همچنین این برنامه نسبت به برنامه هایی که از کاملا از صفر نوشته شده اند پتانسیل بیشتری برای خوب شدن دارند چرا که در انها بسیاری از کارهای مهم توسط DBMS انجام میشود.

چه کارهایی انجام میدهد:

DBMS ها علاوه بر فراهم سازی کردن یک مدیریت فایل ساده عملکرد های دیگری نیز دارند که عبارتند از:

- ممکن سازی فعالیت متقاضی
- مدیریت امنیت
- حفظ یکپارچگی داده
- فراهم کردن پشتیبان گیری و بازیابی
- مدیریت افزونگی
- ممکن سازی استقلال داده
- فراهم سازی زبان های پرسمان غیر روابطی گرا
- انجام خودکار بهینه سازی پرسمان

نمایش داده:

یک پایگاه داده ، مجموعه ای از اسناد درهم متصل و دسته ای برنامه است که به کاربران اجازه دسترسی و تغییر اسناد را میدهد. از اهداف بزرگ سیستم پایگاه داده فراهم کردن یک نمای انتزاعی از داده برای کاربران است . این به این معنیست که سیستم یک سری از جزئیات شیوه ی حفظ و ذخیره سازی داده را پنهان میکند.

1. انتزاع داده: برای مورد استفاده بودن سیستم باید شیوه موثری برای بازیابی داده داشته باشد و این نیاز برای موثر بودن طراحان را به سوی استفاده ساختار های پیچیده از داده برای نمایش داده در پایگاه داده سوق داده است. از آنجا که بسیاری از کاربران سیستم ها آموزش دیده کامپیوتر نیستند، توسعه دهندهان پیچیدگی را با استفاده از چندین سطح انتزاع از کاربران مخفی میکنند تا تعامل کاربران با سیستم را اسان سازند.

- سطح فیزیکی: پایین ترین سطح داده که معین میکند داده ها در واقع چطور قرار گرفته اند.
- سطح منطقی: سطح بالاتر انتزاع که توصیف میکند چه داده هایی به چه شیوه ای در پایگاه داده ذخیره شده اند، و چه روابطی بین آنها وجود دارد.
- سطح نمایش: بالاترین سطح انتزاع که تنها بخشی از کل پایگاه داده را توضیف میکند. با وجود اینکه سطح منطقی از ساختار های ساده تری استفاده میکند، پیچیدگی به دلیل تنوع اطلاعات ذخیره شده در یک پایگاه داده بزرگ همچنان باقی میماند.

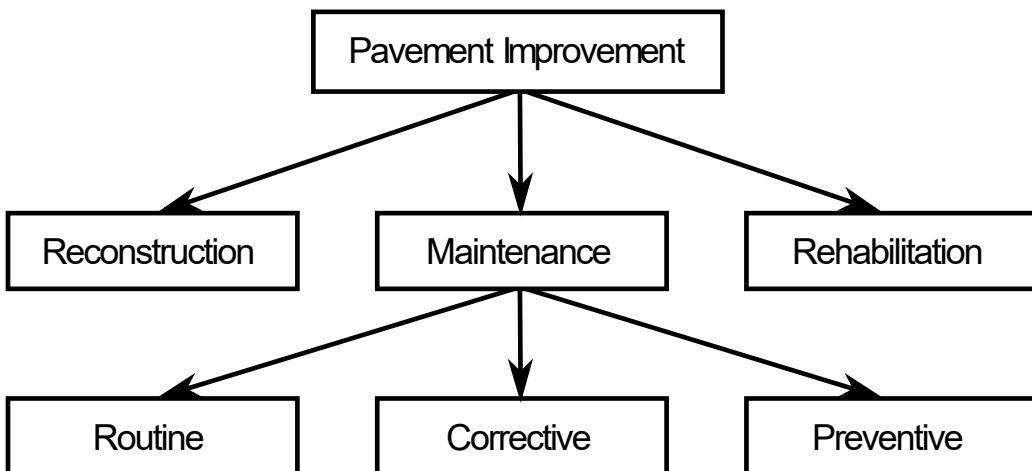
2. نمونه ها و شما ها: به مجموعه اطلاعاتی که در لحظه در پایگاه داده قرار دارد نمونه و به طراحی کلی پایگاه داده ، شمای پایگاه داده گفته میشود. شما ها به ندرت عوض میشوند اما نمونه ها بدلیل تغییرات، افزوده شدن و حذف اطلاعات از پایگاه داده دائما در تغییرند.

انواع پایگاه داده:

1. ساختار سلسله مراتبی:

این ساختار در مین فریم های اولیه مورد استفاده قرار گرفته شده بود. روابط بین رکوردها در آن یک مدل درخت شکل میساختند. این ساختار ساده اما غیر انعطاف پذیر است چراکه رابطه محدود به یک رابطه یک به چند میباشد. سیستم IMS که توسط IBM طراحی شده بود درکنار Mobile RDM مثال هایی از یک پایگاه داده سلسله مراتبی با چند سلسله مراتب در داده همان داده هستند. Mobile RDM یک طراحی جدید از پایگاه داده مندرج برای سیستم های کامپیوتر قابل حمل بود. ساختار سلسله مراتبی امروزه غالبا برای ذخیره اطلاعات جغرافیایی و سیستم های اسناد استفاده میشوند. در شکل زیر مثالی درختی از ساختار سلسله مراتبی داریم که اجازه نمایش اطلاعات به صورت روابط فرزند/والد را به ما میدهد: والد میتواند چندین فرزند داشته باشد اما هر فرزند میتواند تنها یک والد داشته باشد.(به همین دلیل میگویند ¹ به چند) همه خصوصیات یک رکورد بخصوص تحت نوع موجودیش تعریف شده اند.

Hierarchical Model

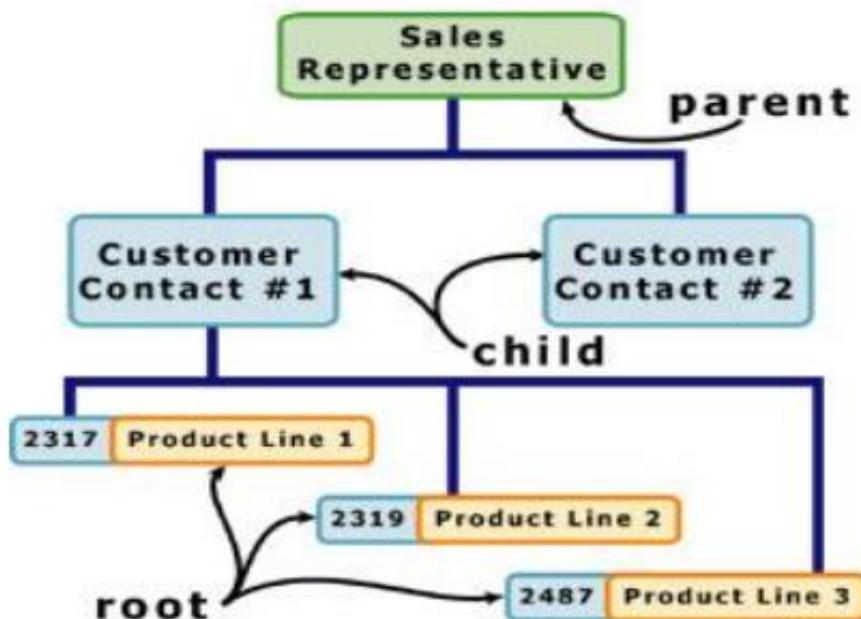


در پایگاه داده هر نوع موجودیت معادل یک جدول است. هر رکورد به تنهایی به شکل یک ردیف نمایش داده میشود و هر خصوصیت به شکل یک ستون. انواع موجودیت ها با یکدیگر با استفاده از اتصالات 1 به n با یکدیگر خویشاوندند و این مدل به عنوان اولین مدل پایگاه داده در IBM در 1960 ساخته شد.

در حال حاضر پر استفاده ترین مدل سلسله مراتبی MSIMS های توسعه داده شده توسط IBM و Windows Registry توسعه داده شده توسط Microsoft هستند.

پایگاه داده های سلسله مراتبی(DBMS):

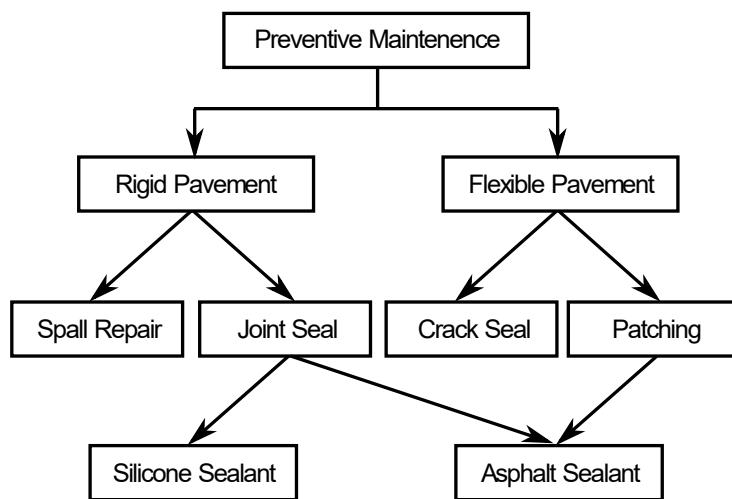
به طور معمول روی کامپیوتر های مین فریم استفاده میشند و برای مدت طولانی در صحنه حضور داشتند. یکی از قدیمی ترین متد ها برای ذخیره و مرتب سازی اطلاعات که هنوز در برخی سازمان ها برای رزرو های مسافرتی استفاده میشود. یک پایگاه داده سلسله مراتبی به صورت هرمی مرتب میشود. در بالاترین نقطه هرم کورد والد رکورد ریشه نام میگیرد و جستجو و عملیات نظیر آن در این ساختار همواره با حرمت از بالا به پایین صورت میگیرد. مزیت این نوع پایگاه های داده این است که میتوانند به سرعت به روز رسانی شوند چرا که ساختار درختی و روابط میان رکورد ها پیش ایش تعریف شده اند اما این مسئله بیشتر مانند یک شمشیر دو لبه است چرا که هر فرزند در این درخت فقط اجازه دارد یک والد داشته باشد و نتیجتاً امکان ارتباط یا اتصال بین دو فرزند وجود ندارد ، حتی اگر از دیدی منطقی امکان پذیر باشد. این ساختار به حدی سختگیر و جدی است که اضافه کردن یک فیلد یا رکورد جدید در این نوع پایگاه داده ها نیازمند تعریف کل دوباره آن میباشد.



2. ساختار شبکه ای:

ساختار شبکه ای از روابط پیچیده تری تشکیل شده است. برخلاف ساختار سلسله مراتبی میتواند با تعداد زیادی از رکورد ها ارتباط داشته باشد و با استفاده از هر یک از چندین مسیر پیش رویش به آنها دسترسی داشته باشد. به عبارت دیگر این ساختار امکان ارتباطات چند به چند را میسر میسازد.

Network Model

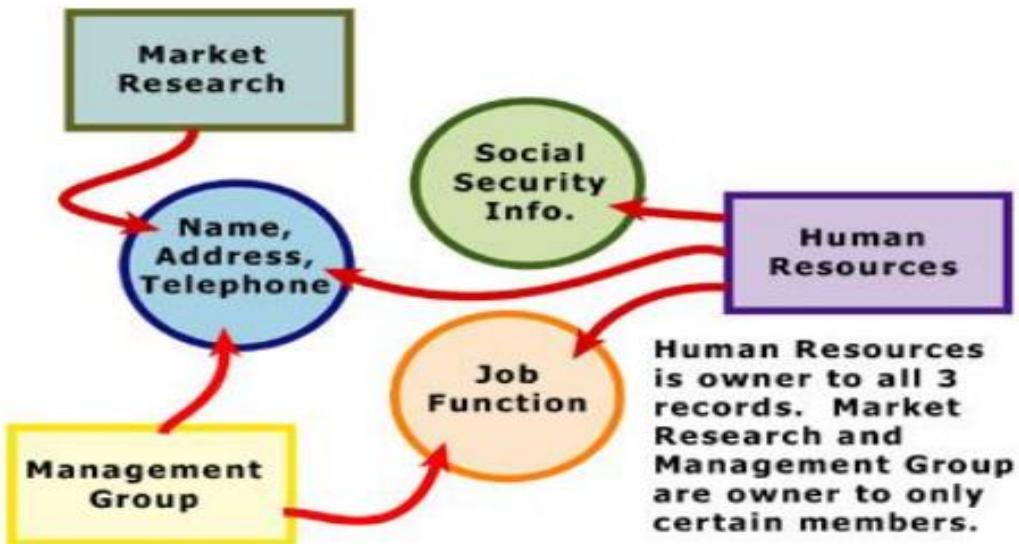


خالق اصلی مدل شبکه ای چارلز بچمن بود و در سال توسط کنسرسیوم کداسیل در ۱۹۶۹ به یک استاندارد توسعه پیدا کرد.

ساختمان داده های شبکه ای با داشتن یک ساختار سلسله مراتبی مشابه ساختمان داده های سلسله مراتبی هستند و جای شباخت به یک درخت برعکس، یک ساختمان داده شبکه ای بیشتر به یک تار عنکبوت یا شبکه هایی منسجم از رکورد ها شباخت دارد. در پایگاه های داده شبکه ای به فرزند ها (برگ ها) عضو گفته میشود و به والد ها مالک گفته میشود. مهم ترین تفاوت این است که هر فرزند یا عضو میتواند بیشتر از یک والد یا مالک داشته باشد مانند ساختمان داده های سلسله مراتبی ساختمان داده های شبکه ای اصولاً روی کامپیوتر های مین فریم استفاده میشوند.

از آنجا که ارتباطات بیشتری میتوانند بین انواع متفاوتی از داده ها ساخته شوند ساختمان داده های شبکه ای انعطاف پذیر تر تلقی میشوند اما دو محدودیت باید در حین استفاده از این ساختمان داده ها در نظر گرفته شوند. مشابه ساختمان داده های سلسله مراتبی ساختمان داده های شبکه ای باید از پیش تعریف شوند.

همچنین در تعداد رکورد هایی که میتوانند به هم متصل شوند، محدودیت وجود دارد.



3. ساختار رابطه ای:

این ساختار که عامیانه و پر استفاده ترین ساختار حال حاضر میباشد توسط سیستم های مین فریم ها، ریز پردازنده ها و کامپیوتر های میانی استفاده میشوند.. جدول های رکورد ها میتوانند بالرژش های کلیدی عامیانه متصل شوند. ای اف کاد در حالی که برای آی بی ام کار میکرد این ساختار را در سال ۱۹۷۲ طراحی کرد. این مدل برای کاربر نهایی که کوئری ها را اجرا میکند اصلا ساده نیست چرا که نیازمند ترکیبی پیچیده از تعداد زیادی جدول میباشد. در ساختمان داده های رابطه ای، رابطه بین اسناد داده ها ارتباطی است و نه سلسله مراتبی. در حالیکه پایگاه داده های سلسله مراتبی و شبکه ای الزام میکنند که کاربر یک سلسله مراتب برای دسترسی به اطلاعات مورد نیاز تعریف کند. پایگاه داده های ارتباطی داده ها را در اسناد متفاوت با استفاده از عناصر عامیانه داده یا یک فیلد کلیدی به یکدیگر متصل میکنند. داده در ساختمان داده های رابطه ای در جدول های متفاوتی ذخیره شده است که هر یک دارای یک فیلد کلیدی هستند که بطور خاصی هر ردیف را توصیف مینماید. پایگاه داده های از این نوع انعطاف پذیری بیشتری نسبت به نوع سلسله مراتبی و شبکه ای دارند. در این نوع ساختمان داده به هر جدول یا سند دارای داده یک رابطه، هر ردیف یا رکورد یک تاپل و هر سطون یک اتریبیوت یا فیلد یا خصوصیت گفته میشود

پایگاه داده های ارتباطی بر این اصل مبتنی اندکه هر جدول دارای یک فیلد کلیدی هست که هر ردیف را هویت شناسه دهی میکند و اینکه هر یک از این فیلد ها میتوانند برای اتصال یک جدول داده به جدول داده دیگر استفاده شوند.

از دلایل پر استفاده شدن این نوع پایگاه داده سهولت یادگیری و آموزش آن و قابلیت اصلاح ورودی ها بدون تغییر کل ساختار میباشد. در حالیکه از نقاط ضعف آن میتوان به کندتر بودن فرایند جستجو برای داده در آن در مقایسه با همه ساختار های قبل، اشاره کرد.

4. ساختار چند بعدی:

این ساختار مشابه با ساختار رابطه ای است. ابعاد این ساختار مکعب شکل داده هایی مرتبط با هر خانه را درخود دارند. این ساختار از دیدی به شکل یک صفحه گسترده داده را به ما نمایش میدهد، حفظ آن آسان است چرا که گزینش و ذخیره رکورد ها به عنوان خصوصیات اساسی در آن درست بر اساس شکل نمایششان انجام میگیرد. سرعت عملکرد این ساختار باعث محبوبیت آن در توانمند سازی و شروع پردازش تحلیلی آنلاین شده است.

Database 1			
1	First Name	Last Name	Social Security No.
2	John	Smith	010-22-9432
3	John	Smith	003-63-0037
4	John	Smith	020-45-9326
5	Sally	Smith	
6	Steve	Smith	
7			

Database 2		
1	Date of Birth	Social Security No.
2	6/12/82	010-22-9432
3	5/9/40	003-63-0037
4		020-45-9326
5		020-45-9326
6		020-45-9326
7		020-45-9326

Database 3	
1	Social Security No.
2	010-22-9432
3	003-63-0037
4	020-45-9326
5	020-45-9326
6	020-45-9326
7	020-45-9326

5. ساختار شی گرا:

این ساختار برخلاف ساختار های دیگر قابلیت مدیریت گرافیک، تصاویر، صدا و متن و انواعی دیگر از داده را بدون دشواری دارد. محبوبیت آن در برنامه های تحت وب چند رسانه ای میباشد و طراحی شد تا بتواند با زبان های شی گرا مانند جاوا کار کند.

مدل پراستفاده آن امروزه مدل های خاص منظوره ای هستند که در SQL تعبیه و پیاده سازی شده اند با وجود مخالفت های ایده آل خواهانی که معتقدند این مدل یک تخربی از مدل رابطه ای است چرا که برخی از اصول اساسی آنرا برای عملکرد و بازدهی بهتر نقض میکند.

بسیاری از سیستم های مدیریت پایگاه داده امروزه از واسطه برنامه نویسی اتصال آزاد پایگاه داده (Open Database Connectivity API) استفاده میکنند که از یک راه استاندارد برای دسترسی برنامه نویسان به DBMS پشتیبانی میکند.

قبل از استفاده از مدیریت پایگاه داده بسیاری از سازمان ها برای ذخیره، مرتب سازی و پردازش متکی بر سیستم های پردازش فایل بودند که در آن کاربران نهایی انتقاداتی بر ذخیره داده ها در فایل های متنوع با چینش های متنوع داشتند چرا که هر فایل برای استفاده از یک برنامه بخصوص، ویژه شده بود.
افزونگی داده ها

یک مشکل سیستم پردازش سند است. چرا که اسناد داده های مستقل داده های تکراری تولید میکنند در نتیجه در حین هر بروزرسانی هر سند مستقل میباشی آپدیت شود.

مسئله دیگر عدم یکپارچگی داده ها است. داده برای چینش و ذخیره به داده دیگر وابسته است. و در آخر هیچ ثبات یا استانداردسازی داده ای در سیستم پردازش نبود که کار را در حفظ و نگه داری دشوار میکرد.

ساختمان های داده:

فیلد ها، رکورد ها ، اسناد و اشیا هستند که برای رسیدگی به مقادیر بسیار زیادی از داده ها که بروی یک دستگاه حافظه بلند مدت، بهینه شده اند. (که به خودی خود درباره کندی دسترسی نسبت به حافظه فرار به ما میگوید).

زبان پرسمان پایگاه داده:

یک زبان پرسمان پایگاه داده و گزارش شی به کاربران اجازه میدهد تا به صورت تعاملی از پایگاه داده به کسب اطلاعات، تحلیل و بررسی و به روز رسانی آن مطابق با میزان آزادی کاربر پردازند(امنیت را نیز کنترل میکند). کاربران به لطف سیستم امنیت داده میتوانند با استفاده از پسوردها به تمام یا بخشی از پایگاه داده که subschema نام دارد دسترسی پیدا کنند.

مکانیزم تراکنش پایگاه داده:

به طور ایده آل خواص ACID را تضمین میکند تا از یکپارچگی داده با وجود دسترسی همزمان کاربران اطمینان حاصل کند و به نوعی حفظ یکپارچگی داده در پایگاه داده را نیز انجام میدهد که اینکار را میتواند با مجوز ندادن همزمان به بیش از یک کاربر برای بروزرسانی یک رکورد مدیریت کند DBMS. همچنین میتواند با استفاده از اندیس های خاص از تکراری شدن رکورد ها جلو گیری کند.

پیاده سازی: سیستم های مدیریت پایگاه داده:

یک سیستم مدیریت پایگاه داده در کنار اجازه دادن برای ساخت، رسیدگی، استفاده، حفظ و گرفتن اطلاعات باید نوع پایگاه داده قابل مورد پشتیبانی، عملکرد و توانایی هایش را تعریف و در حین ارائه پردازه های داخلی برای برنامه های بیرونی، فرایند های داخلی را دور از چشم کاربران نهایی نگه دارد چراکه نیاز آنها معمولاً باید با واسطه های اختصاصی که برایشان طراحی شدند حل شوند.

با توجه به حوزه های که وسیعی که یک DBMS موظف به مدیریتشان است (تخصیص حافظه، دسترسی، امنیت و پشتیبان گیری) DBMS ها میتوانند تحت مدل های پایگاه های داده، کامپیوتر ها و زبان های پرسمان مورد پشتیبانیشان دستبه بندی شوند. از آنجا که نرم افزار های پایگاه داده معمولاً از استاندارد ODBC پشتیبانی میکنند امکان یکپارچه سازی محدود یک پایگاه داده با دیگری را میدهد.

توسعه یک DBMS به طور کلی فرایندی بسیار زمان گیر است چرا که باید دائماً برای بکار گیری تکنولوژی های روز حافظه و پیشرفته کامپیوتر ها آمده باشند. ساخت برخی از DBMS ها از سال ها 1970-1980 هنوز در وضعیت در حال توسعه قرار است.

معماری : DBMS مولفه های مهم:

معماری DBMS مولفه ها (به همراه تعریف های عملگر و توابع آنها) و واسطه های آنرا مشخص میکند. این معماری متمایز از معماری پایگاه داده است. از مولفه های مهم :

- واسطه های خروجی هستند که برای ارتباط برقرار کردن با و از DBMS بع منظور اعمال لازمه میباشند که خود این اعمال میتوانند اعمال مستقیم بر روی پایگاه داده یا اعمال مربوط به DBMS باشند.

- اعمال مستقیم بر روی پایگاه داده: تعریف نوع داده، تعیین کردن سطح امنیت، بروزرسانی داده، پرسمان کردن پایگاه داده.

- اعمال مرتبط با فرایند DBMS و مدیریت:

پشتیبان گیری و بازسازی، نظارت بر امنیت، تخصیص فضای حافظه داده و نظارت بر پیکر بندی طرح بندی، بازیافت پایگاه داده، نظارت بر عملکرد و نظارت بر تنظیم.

یک واسط خارجی میتواند یک رابط کاربری (معمولاً برای یک ادمین پایگاه داده) یا یک واسط برنامه نویسی باشد.

-موتور های زبان پایگاه داده یا پردازشگر ها: اکثر اعمال روی پایگاه های داده با زبان های پایگاه داده ای انجام میشوند که علاوه بر تعریف داده، به کار گیری و پرسمان برای بخش هایی امنیتی و بیشتر مورد استفاده قرار میگیرند. یک موتور زبان با یک کامپیلر یا مفسر زبان را پردازش میکند تا اعمال مورد نیاز پایگاه داده را به شیوه ای استخراج کند که بدرد DBSM بخورد.

-بهینه ساز پرسمان: بهینه سازی هایی روی پرسمان انجام میدهد تا بهترین نقشه پرسمان (یک درخت از عملکرد ها) را برای اجرای پرسمان بیابد.

-موتور پایگاه داده: عملیات گرفته شده از پایگاه داده را روی اشیا آن انجام میدهد (معمولاً روی شکل سطح بالایشان).
- موتور حافظه: ترجمه عملیات به کد های سطح پایین بیتی که در برخی جاها به عنوان بخشی از موتور پایگاه داده تعریف میشود.

-موتور تراکنش: برای پایداری و درستی اکثر مسائل مربوط به تراکنش ها به صورت کپسوله شده انجام میشوند اما میتوانند به صورت خارجی توسط DBMS گزیده شوند تا گروه خاصی از عملیات را با توجه به قوانین تراکنش کپسوله انجام دهند.

-مولفه مدیریت و عملکرد DBMS: اجزای مهم ،کاربردی و نظارتی دیگری که با مدیریت و عملکرد DBMS مرتبط هستند مانند تنظیمات جزئی و پشتیبانی گیری و امنیت را گرد هم میآورد.

نتیجه گیری:

بدلیل همه نیاز هایی که سیستم DBMS باید رفع کند و سخت افزار ها و نرم افزار هایی که باید واسط خوبی برای آنها باشد نوشتار این سیستم کار آسانی نیست و نیاز مند دانش تخصصی درموردنوع سخت افزار و نرم افزار های ملزم به پشتیبانی میباشد. در نتیجه وارد شدن به جزئیات کار آنها امری سخت است چراکه DBMS عملاً مجموعه ای از سیستم ها درونی بسیار پیچیده میباشد، که برای پیاده سازی هر یک علاوه بر برنامه نویسان با تجربه و تخصص بالا ، زمان زیادی لازم است.

مراجع

[1]- Sweta Singh, Assistant Professor, Faculty of Management Studies, BHU, Varanasi, India. DATABASE MANAGEMENT SYSTEM

