

Databáze, databázové systémy

Historie, rozdělení, architektura, databázový software

Člověk potřebuje k vykonávání činnosti informace, ty vznikají zpracováním dat. Data jsou údaje zachycující fakta, musí být vyjádřena ve formě umožňující zpracování a předávání, smyslem zpracování dat je vytvořit potřebné informace. Informace jsou podkladem pro rozhodování a snižují míru nejistoty na straně příjemce. Data: Výsledek testu každého studenta; Informace: Průměrný výsledek třídy

Databáze

(Data base) místo, kde se ukládají zorganizované a strukturované údaje.

Souhrn dat, která spolu logicky souvisejí (patří k sobě), jsou nějakým způsobem svázána.

Systém řízení báze dat

Softwarové vybavení, které je schopno efektivně spravovat velké množství dat, upravovat definici struktury, ukládání, výběr a ochranu dat.

Funkce SŘBD:

Zabezpečuje a autorizuje databázi a komunikaci mezi uživatelem a systémem. Určuje strukturu složení, kontrolu integrity dat, ovládání databáze (primárně přes SQL), vazby a identifikace mezi tabulkami a indexy.

Příklady SŘBD: MySQL, Oracle, MS Access

Databáze + Systém řízení báze dat = databázový systém

Databázová aplikace

Program, který umožňuje přístup k datům v databázi prostřednictvím formulářů pro zadávání dat, dotazů a sestav.

Databázový systém

Umožňuje jednoduchou manipulaci s daty, data přijímá, uchovává a zpřístupňuje. Umožňuje zpracování datových souborů pomocí příkazů.

Historie

Předchůdcem databází byly papírové kartotéky. Umožňovaly uspořádávání dat podle různých kritérií a zařizování nových položek. Veškeré operace s nimi prováděl přímo člověk (Správa podobná dnešním databázím). Úplně první kolem 18. století.

Poté první strojové (na elektromagnetických strojích) zpracování dat bylo ke konci 19. století. Jako paměťové médium byly využity děrné štítky.

Tváří dalšího posunu databází byl o půl století později překotný vývoj počítačů, následován požadavkem na lepší zpracování dat. To se vyřešilo vydáním první verze univerzálního databázového jazyka zvaného COBOL (1960). Velice rychle se stal nejrozšířenějším jazykem pro hromadné zpracování dat.

Po roce 1965 začaly vznikat první síťové SŘBD na sálových PC (Charles Bachman). Následně byla popsána celá architektura síťového databázového systému. Mezitím byly vyvíjeny první hierarchické databáze (IMS), od roku 1970 se začínají objevovat relační databáze (E. F. Codd). Ke každému záznamu existuje přímý přístup pomocí ukazatelů, které vyjadřují vztahy mezi jednotlivými databázovými položkami. Jeden z prvních průkopníků databází byl Charles Bachman. V roce 1971 se objevily pojmy jako *schéma databáze*, *jazyk pro definici schématu*. Byla zde popsána celá architektura síťového datab. Systému.

Rok 1974 se nese ve znamení první verze jazyka SQL. Vývoj této technologie po 10 letech přinesl výkonově použitelné systémy, srovnatelné se síťovými a hierarchickými databázemi.

V 90. letech 20. století se začínaly objevovat první objektově orientované databáze, jejichž filozofie byla přebírána z oo jazyků. Tyto databáze měly vytlačit relační systémy. Tyto předpoklady se nenaplnily a vznikla kompromisní *objektově-relační technologie*.

Rozdělení databází

Síťová

Založena na síťovém modelu, kde jsou data logicky i fyzicky uspořádaná jako uzly rovinného grafu. Každý záznam může být spojený s libovolným počtem dalších záznamů.

Hierarchická

Založena na hierarchickém modelu. Logické uspořádání dat má stromovou strukturu.

Relační

Založena na relačním modelu, v němž jsou data logicky zařazena do relací (vazeb, vztahů). Je založena na tabulkách.

Objektová

Založena na objektech, jejich zapouzdření a dědičnosti. Místo tabulek jsou zde uloženy přímo objekty, vč. Svých vlastností a místo řádků se ukládají samostatné instance (konkrétní výskyty) objektů.

Systém řízení databáze

- DDL (Data Definition Language)

Prostředky pro popis dat

- DML (Data Manipulation Language)

Prostředky pro manipulaci s daty

- DCL (Data Control Language)

Prostředky pro řízení dat (přístupu k datům)

Vlastnosti databázových systémů

Struktury datových souborů jsou odděleny od aplikačních programů. Přístup je možný pouze prostřednictvím DBS. Je možný přístup více uživatelů současně. Je vyřešena ochrana dat v DB před zneužitím.

Správa databáze

Profesionální databáze mají velké množství důležitých informací, jsou navrhovány a udržovány odborníky:

Správce databáze: navrhuje strukturu, přiděluje práva uživatelům. Centrální kontrola nad daty a apk. programy.

Aplikační programátoři: pomáhají správci s údržbou databáze, vytvářejí apk. programy

Uživatelé: prostřednictvím apk. programů komunikují s databází, vkládají data, získávají z databáze informace, spravují data v databázi

Databázový software

Menší databáze: MS Access, OpenOffice Base

Velké databáze: MySQL, Oracle, PostgreSQL, Informix

Oracle

je moderní multiplatformní databázový systém s velice pokročilými možnostmi zpracování dat, vysokým výkonem a snadnou škálovatelností. Oracle funguje pod operačními systémy Windows i Linux. Je světovou špičkou a využívá se především tam, kde potřebujeme pracovat s velkým množstvím dat. Velmi drahý.

MySQL

je relační databáze typu DBMS (database management system) a vychází z deklarativního programovacího jazyka SQL. Je šířen jako Open Source. Díky své licenci a rychlosti je v poslední době téměř nejoblíbenějším systémem. MySQL bylo od počátku optimalizováno především na rychlost, a to i za cenu některých zjednodušení. Není tak robustní jako Oracle.

PostgreSQL

primárně vyvíjen pro Linux resp. pro unixové systémy obecně, nicméně existují i balíčky pro platformu Windows. Výhodou je vysoká rozšiřitelnost (vlastní datové typy, uložené procedury, domény...), bezpečnost, podpora pokročilých technologií (např. XML) a velmi dobrá podpora SQL standardů. Je to otevřený software a je možné ho používat zcela ZDARMA i pro komerční využití.

Firebird

Pro databáze od několika mála KB až po gigabajty s dobrým výkonem a skoro bez údržby, open source

MS SQL

Jedná se o produkt společnosti Microsoft, který lze provozovat pouze pod systémem Windows a který je vydáván v několika verzích, podle množství funkcí. Licence není zdarma, ale je zpoplatněna, podle druhu verze. Podporuje standardy jazyka SQL a navíc přináší další rozšíření

Architektura databázových systémů

Typ architektury databázového systému by měl odpovídat:

- Účelu a typu databáze
- Technologickým a finančním možnostem
- Počtu uživatelů sdílejících informace v databázi
- Typu zpracovávaných informací

Základní dělení

- Centralizované databázové systémy
- Systémy klient-server
- Distribuované systémy

Centralizované databázové systémy

V této architektuře je jeden centrální počítač a terminály, které volají aplikaci i data na centrálním počítači. Centrální počítač obsahuje řídicí program + databáze + aplikace.

Jednovrstevná centralizovaná architektura

V případě, že databázový systém poběží bez použití sdílení informací mezi více uživateli, pak lze použít jednovrstvou architekturu s lokální databází.

Výhody

- Rychlost
- Není potřeba DB server
- Není potřeba počítačová síť

Nevýhody

- Omezené množství dat
- Omezený počet uživatelů

Systémy klient-server

U této architektury je zpracování dat rozděleno mezi dva systémy.

Na PC běží tzv. klient, v němž je vytvořena aplikace a na datovém serveru běží vlastní ovládání databáze tzv. server.

Server může být umístěn i na PC, kde je klient, častěji však běží na samostatném PC. Protokol mezi klientem a serverem je nejčastěji SQL, proto mohou mezi sebou komunikovat i různé SŘBD. Zpracování dotazů a přístupů do databáze si řídí server. Po síti se přenášení jen požadovaná data.

Dvouvrstvá architektura Klient-server

Průběh komunikace mezi serverem a počítačem:

- Aplikace na PC formuluje dotaz, nebo požadavek na data pomocí SQL dotaz a odešle jej na server
- Server zpracuje dotaz
- Výsledek dotazu posléze odešle do PC
- Aplikace převede výsledek do výstupní podoby

K uživateli jsou přesunuty pouze uživatelské služby a získává pouze požadované informace

Aplikační a datové služby probíhají na straně serveru

Výhody:

- Snížení množství dat pohybujících se v síti
- Minimální zatížení sítě
- Vysoká pružnost aplikací
- Rozdělení zpracování záznamů

Typy serverů

- Transakční servery (SQL server) – pro relační databáze
- Datové servery – zpravidla jen pro databáze objektové

Distribuované systémy

- Tyto systémy jsou mnohem náročnější než klasické SŘBD. Uživatel požádá o data lokální počítač. Pokud tam nejsou, lokální PC zjistí po síti, kde data jsou, vyžádá je a předá uživateli, který ani nemusí vědět, odkud data pocházejí. V této architektuře jsou územně rozloženy lokální báze dat.
- Může zde docházet k tomu, že data potřebná pro aplikaci jsou na nepřístupném lokálním PC.
- Lokální báze dat je část celé (distribuované) báze, která je umístěna v jednom uzlu (počítači) počítačové sítě; je řízena lokálním SŘBD, který pracuje ve spolupráci s ostatními složkami distribuovaného databázového systému
- Distribuovaná báze dat je sjednocení všech lokálních bází dat distribuovaného databázového systému

Architektura musí obsahovat:

Globální SŘBD

- Eviduje umístění všech dat, zajišťuje převod požadavků, referenční integritu a řízení sdíleného přístupu k datům
- Kopie globálního SŘBD je umístěna na každé stanici pod názvem Distribuovaný SŘBD

Lokální SŘBD

- Vytváří exportní schéma, které definuje data sdílená s jinými uživateli

Základní vlastnosti distribuované architektury:

Transparentnost – z pohledu klienta se zdá, že všechna dat jsou zpracovávána na jednom serveru v lokální databázi

Autonomnost – s každou lokální bází dat zapojenou do distribuované databáze je možno pracovat nezávisle na ostatních databázích

Nezávislost na typu sítě – Architektura podporuje různé typy sítí

Výhody:

- Zvýšená spolehlivost a míra dostupnosti dat
- Místní řízení báze dat a snazší růst systému
- Snazší implementace dalších lokálních databází
- Menší nároky a náklady na komunikaci
- Rychlejší odezvy

Nevýhody:

- Méně snadná kontrola referenční integrity dat
- Nebezpečí pomalé odezvy, pokud jsou data nevhodně distribuována