

Relační databázové systémy

Vávra David · Informačné technológie

28.05.2012



Tento článok sa venuje relačným databázám. Po vysvetlení princípu relačnej databázy pokračuje článok známymi typy relácií a normalizaciami, ďalej pak priblížením jazyka SQL včetně ukážky typů a popisu nejběžnějších příkazů. Předposlední částí je představení nejznámějších databázových systémů MySQL, MS SQL, ORACLE a v závěru je jejich stručné porovnání.

Úvod

V relačnej databázi jsou data uchovávané ve formě tabulek. Jednotlivá data jsou uložena v jejich jednotlivých buňkách. Řádky tabulky představují jednotlivé záznamy, sloupce pak konkrétní vlastnosti daného záznamu. Většinou je žádoucí, aby byl každý záznam jednoznačně identifikovatelný, k čemuž slouží primární klíč. Bez primárního klíče nelze vytvářet relace s dalšími tabulkami (relace je vytvářena propojením primárního a cizího klíče dvou tabulek). Primární klíč je v dané tabulce vždy jen jeden, cizích klíčů může být oproti tomu více. Relace určují vztahy mezi jednotlivými tabulkami.

1. Typy relací

1:1 (jedna ku jedné)

Řekněme například, že modelujeme kancelářské prostory v určité budově. Budeme-li předpokládat, že v každé kanceláři sedí pouze jeden člověk, dostáváme tím vztah jedna ku jedné.



Obrázek 1. Relace 1:1

1:N (jedna ku více)

Nejběžnější vztah mezi entitami představují vztahy jedna k více. Dobře identifikovatelný vztah typu jedna k více přináší s sebou málokdy nějaké problémy.



Obrázek 2. Relace 1:N

N:M (více ku více)

V reálném světě najdeme spoustu příkladů vztahů více k více. Studenti mohou mít zapsáno několik různých přednášek a podobně jedné přednášky se může účastnit více různých studentů. Každý zákazník nakupuje v řadě obchodů a jeden obchod má mnoho zákazníků.

V relační databázi se však vztahy více k více přímo implementovat nedají. Modelují se proto s pomocí speciálního „prostředníka“, speciální samostatné relace, která má s každým z obou původních účastníků vztah jedna k více.



Obrázek 3. Relace M:N

2. Normalizace

Normalizace je technika používaná pro vytvoření sady tabulek s minimální redundancí, která podporuje datové požadavky organizace. Existuje několik normálních forem, nejužívanější z nich jsou první normální forma (1NF), druhá normální forma (2NF) a třetí normální forma (3NF). Všechny tyto formy jsou založeny na vztazích mezi sloupci tabulky. Z dlouhodobého hlediska je tvorba dobře navržených tabulek klíčová k efektivně a správně fungující databázi.

První normální forma (1 NF)

Tabulka splňuje první normální formu tehdy, pokud všechny sloupce jsou atomické, neboli dále nedělitelné. Jeden sloupec tak nesmí obsahovat více druhů dat. Je to normální forma, která je nezbytně důležitá pro vytvoření vhodných tabulek pro relační databáze.

Druhá normální forma (2 NF)

Pro to aby byla splněna podmínka druhé normální formy, je nezbytně nutné také splnění první normální formy a zároveň musí být každý sloupec v tabulce zcela závislý na primárním klíči. Z toho automaticky vyplývá, že tabulka, která má pouze jeden primární klíč automaticky splňuje 2NF, ale tabulka obsahující více primárních klíčů musí být rozdělena na dvě a více tabulek, tak aby byla splněna definice druhé normální formy.

Třetí normální forma (3 NF)

Tabulka je ve třetí normální formě tehdy, pokud splňuje první a zároveň druhou normální formu a žádný sloupec, který nepatří do primárního klíče, není tranzitivně závislý na primárním klíči. Abychom zrušili tuto závislost, musíme jako v předchozím případě, rozdělit tabulku do dvou a více tabulek tak, aby všechny neklíčové atributy byly na sobě nezávislé. [1]

3. Jazyk SQL

SQL (Structured Query Language) je standardizovaný dotazovací jazyk používaný pro práci s daty v relačních databázích. V 70. letech 20. století probíhal ve firmě IBM výzkum relačních databází. Bylo nutné vytvořit sadu příkazů pro ovládání těchto databází. Vznikl tak jazyk SEQUEL (Structured English Query Language). Cílem bylo

vytvořit jazyk, ve kterém by se příkazy tvořily syntakticky co nejbližší přirozenému jazyku (angličtině).

K vývoji jazyka se přidaly další firmy. V r. 1979 uvedla na trh firma Relational Software, Inc. (dnešní Oracle Corporation) svoji relační databázovou platformu Oracle Database. IBM uvedla v roce 1981 nový systém SQL/DS a v roce 1983 systém DB2. Dalšími systémy byly např. Progres, Informix a SyBase. Ve všech těchto systémech se používala varianta jazyka SEQUEL, který byl přejmenován na SQL. Relační databáze byly stále významnější, a bylo nutné jejich jazyk standardizovat. Americký institut ANSI původně chtěl vydat jako standard zcela nový jazyk RDL. SQL se však prosadil jako de facto standard a ANSI založil nový standard na tomto jazyku. Tento standard bývá označován jako SQL-86 podle roku, kdy byl přijat.

V dalších letech se ukázalo, že SQL-86 obsahuje některé nedostatky a naopak v něm nejsou obsaženy některé důležité prvky týkající se hlavně integrity databáze. V roce 1992 byl proto přijat nový standard SQL-92 (někdy se uvádí jen SQL2). Zatím nejnovějším standardem je SQL3 (SQL-99), který reaguje na potřeby nejmodernějších databází s objektovými prvky. Standardy podporuje prakticky každá relační databáze, ale obvykle nejsou implementovány vždy všechny požadavky normy. A naopak, každá z nich obsahuje prvky a konstrukce, které nejsou ve standardech obsaženy. Přenositelnost SQL dotazů mezi jednotlivými databázemi je proto omezená. [2]

Příkazy pro manipulaci s daty

Jsou to příkazy pro získání dat z databáze a pro jejich úpravy. Označují se zkráceně DML - Data Manipulation Language („jazyk pro manipulaci s daty“).

- SELECT - vybírá data z databáze, umožňuje výběr podmnožiny a řazení dat.
- INSERT - vkládá do databáze nová data.
- UPDATE - mění data v databázi (editace).
- MERGE - kombinace INSERT a UPDATE - data buď vloží (pokud neexistuje odpovídající klíč), pokud existuje, pak je upraví ve stylu UPDATE.
- DELETE - odstraňuje data (záznamy) z databáze.
- EXPLAIN - speciální příkaz, který zobrazuje postup zpracování SQL příkazu.
- SHOW - méně častý příkaz, umožňující zobrazit databáze, tabulky nebo jejich definice

Příkazy pro definici dat

Těmito příkazy se vytvářejí struktury databáze - tabulky, indexy, pohledy a další objekty. Vytvořené struktury lze také upravovat, doplňovat a mazat. Tato skupina příkazů se nazývá zkráceně DDL - Data Definition Language („jazyk pro definici dat“).

- CREATE - vytváření nových objektů.
- ALTER - změny existujících objektů.
- DROP - odstraňování objektů.

Příkazy pro řízení dat

Do této skupiny patří příkazy pro nastavování přístupových práv a řízení transakcí. Označují se jako DCL - Data Control Language („jazyk pro ovládání dat“), někdy také

TCC - Transaction Control Commands („jazyk pro ovládání transakcí“).

- GRANT - příkaz pro přidělení oprávnění uživateli k určitým objektům.
- REVOKE - příkaz pro odnětí práv uživateli.
- START TRANSACTION - zahájení transakce.
- COMMIT - potvrzení transakce.
- ROLLBACK - zrušení transakce, návrat do původního stavu. [3]

3.1 MySQL

MySQL je databázový systém, vytvořený švédskou firmou MySQL AB. Jeho hlavními autory jsou Michael „Monty“ Widenius a David Axmark. Je považován za úspěšného průkopníka dvojího licencování - je k dispozici jak pod bezplatnou licenci GPL, tak pod komerční placenou licenci. MySQL je multiplatformní databáze. Komunikace s ní probíhá - jak už název napovídá - pomocí jazyka SQL. Podobně jako u ostatních SQL databází se jedná o dialekt tohoto jazyka s některými rozšířeními.

Pro svou snadnou implementovatelnost (lze jej instalovat na Linux, MS Windows, ale i další operační systémy), výkon a především díky tomu, že se jedná o volně šiřitelný software, má vysoký podíl na v současné době používaných databázích. Velmi oblíbená a často nasazovaná je kombinace MySQL, PHP a Apache jako základní software webového serveru.

MySQL bylo od počátku optimalizováno především na rychlost, a to i za cenu některých zjednodušení: má jen jednoduché způsoby zálohování, a až donedávna nepodporovalo pohledy, trigger, a uložené procedury. Tyto vlastnosti jsou doplňovány teprve v posledních letech, kdy začaly nejčastějším uživatelům produktu - programátorům webových stránek - již poněkud scházet. [4],[5] MySQL je výkonný a robustní relační databázový systém, jehož základem jsou entity a relace. Entitu lze chápat jako věc, o níž je uchovávána informace. Relace pak určuje vztahy mezi jednotlivými entitami.

Přehled podporovaných vlastností:

- cizí klíče (od verze 3.23 podporovány v tabulkách typu InnoDB)
- transakce (od verze 3.23 podporovány v tabulkách typu InnoDB)
- podpora různých znakových sad a časových pásem v datech (od verze 4.1)
- poddotazy (od verze 4.1)
- uložené procedury (od verze 5.0)
- trigger (od verze 5.0)
- pohledy (od verze 5.0)
- práce s metadaty (od verze 5.0)

MySQL nabízí několik typů databázových tabulek (storage engine), které se liší svými možnostmi, použitím a způsobem ukládání dat do souborů:

- MyISAM - nejpoužívanější, bez podpory transakcí
- InnoDB - podpora transakcí
- BerkeleyDB (BDB)
- MEMORY - práci s daty v paměti

- NDB Cluster - úložiště pro clusterované databáze (od verze 5.0)
- ARCHIVE - komprimované tabulky, bez podpory indexů
- CSV - ukládání dat v prostých textových souborech

4. 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.

5. Oracle

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. Přesněji Oracle Corporation je název firmy, oficiální název databázové platformy je Oracle Database.

Tento systém podporuje nejen standardní relační dotazovací jazyk SQL podle normy SQL92, ale také proprietární firemní rozšíření Oracle (např. pro hierarchické dotazy), imperativní programovací jazyk PL/SQL rozšiřující možnosti vlastního SQL (v tomto jazyce je možné tvořit uložené procedury, uživatelské funkce, programové balíky a triggery), dále podporuje objektové databáze a databáze uložené v hierarchickém modelu dat (XML databáze, jazyk XSQL). Dále též obsahuje širokou paletu nástrojů pro podporu snadného nasazení na gridových sítích. Oracle je tedy světovou špičkou a využívá se především tam, kde potřebujeme pracovat s velkým množstvím dat. [6]

6. Stručné porovnání jednotlivých databázových systémů

Tabulka 2. Stručné porovnání jednotlivých databází

	MySQL	MS SQL	Oracle
Licence:	freeware	placená	placená
Velikost databáze:	malé až střední	střední až velké	velké
Rychlost odezvy:	vysoká	nižší	nižší

7. Závěr

V relační databázi jsou data uchovávány ve formě tabulek. Relace určují vztahy mezi jednotlivými tabulkami a jsou vytvářeny propojením primárního a cizího klíče dvou tabulek. V zásadě rozeznáváme 3 typy relací. Relaci 1:1, dále 1:N, což představuje nejběžnější vztah mezi entitami a N:M. V reálném světě najdeme spoustu vztahů N:M, ale v relační databázi je nelze implementovat přímo.

Jazyk SQL tvoří příkazy, které jsou syntakticky blízké přirozenému jazyku (angličtině). Jeho příkazy lze rozdělit do několika základních skupin. Skupina příkazů pro manipulaci s daty zahrnuje příkazy pro získávání a úpravy samotných dat. Skupina příkazů pro definici dat pak obsahuje příkazy, kterými se vytváří a upravují struktury databáze. Pro nastavování přístupových práv a řízení transakcí se používají příkazy ze

skupiny příkazů pro řízení dat. Z uvedených databázových systémů nelze jednoznačně určit, který je „nejlepší“ – záleží zde na mnoha faktorech, které je potřeba zohlednit. Zdarma je však z uvedených databázových systémů pouze MySQL.

Poděkování

Tento článek byl podpořen z Interní Grantové Agentury UTB ve Zlíně, číslo projektu IGA/FAI/2012/015.

Zdroje

1. PROKOPOVÁ, Zdenka. Databázové systémy MySQL+PHP. FAI UTB Zlín, 2006, ISBN 80-7318-486-9.
2. Wikipedie: otevřená encyklopedie [online]. [cit. 2012-07-01]. Dostupný z WWW: <http://cs.wikipedia.org/wiki/SQL>
3. Wikipedie: otevřená encyklopedie [online]. [cit. 2012-07-01]. Dostupný z WWW: http://cs.wikipedia.org/wiki/P%C5%99%C3%ADkazy_jazyka_SQL
4. Wikipedie: otevřená encyklopedie [online]. [cit. 2012-07-02]. Dostupný z WWW: <http://cs.wikipedia.org/wiki/MySQL>
5. VÁVRA, David. WWW prezentace a návrh e-obchodu soukromého subjektu - živnosti [bakalářská práce]. FAI UTB Zlín, 2009.
6. Wikipedie: otevřená encyklopedie [online]. [cit. 2012-07-02]. Dostupný z WWW: <http://cs.wikipedia.org/wiki/Oracle>
7. VÁVRA, David. Tvorba e-obchodu a vývoj linuxového skriptu pro automatické generování nových e-obchodů [diplomová práce]. FAI UTB Zlín, 2011.