



Databázové systémy zblízka

Štefan Havlíček

V naší moderní společnosti snad není člověka, který by nikdy neslyšel či nevyslovil slovo databáze, avšak odvažují se říci, že v případě dotazu, co to ta „databáze“ je, bychom dostali různé odpovědi. Co se skrývá za slovem databáze? Jsou to data? Je to software, nebo místo, kam se ukládají data, či snad způsob, jakým jsou data ukládána a posléze prohledávána a zase znovu získávána? Nebo je to něco úplně jiného?

Ať si pod pojmem databáze představíme cokoli z výše uvedeného, jisté je, že dnes jsou téměř všechny organizace, finanční instituce, veškerý průmysl, zdravotnictví, státní správa atd. závislé na databázových systémech, na datech v nich uložených a na informacích z nich získávaných. Primárním úkolem IT oddělení dnešní organizace je odpovědnost za plynulý a bezpečný chod informačního systému.

Hybnou silou vývoje správy dat a jejich zpracování byly a stále jsou rostoucí požadavky na překonání omezení, která má stávající systém, a neustálý hlad po informacích. Zprvu to byly souborově orientované systémy, které sloužily jako úložiště firemních dat. Počítačové soubory plnily obdobnou funkci jako

papírové soubory uložené ve složkách. Avšak tento koncept brzy narazil na svá omezení. Zprvu to byly sekvenční soubory, kde jednotlivé záznamy byly zapisovány do souborů v daném pořadí tak, jak vznikaly. Později byly tyto sekvenční soubory nahrazeny index-sekvenčními soubory, kdy k danému datovému souboru existoval ještě jeden soubor indexovací, který umožňoval rychlejší vyhledávání a manipulaci s daty. Nedostatky tohoto přístupu (nekontrolovatelná redundance dat, častá nekonzistentnost dat, obtíže při integraci aplikací, ...) daly vzniknout novým koncepcím databázových technologií.

Podívejme se na následující definici: databáze je uspořádaná kolekce navzájem souvisejících datových prvků za účelem vyhovět

požadavkům na informace a navržena tak, aby byla sdílitelná více uživateli. Uspořádaná kolekce navzájem souvisejících datových prvků, znamená, že se nejedná o náhodnou množinu datových struktur, nýbrž o datové prvky nějakým logickým způsobem uspořádané. Mezi těmito prvky existují nějaké vztahy. Požadavky na informace vymezují rámec, za jakým účelem jsou data do databáze vkládána. V databázi nemocničního informačního systému budou data o pacientech nemocnice. Účelem těchto dat je poskytnout lékařům a nemocničnímu personálu informace, které potřebují pro léčbu pacienta. Tyto informace jsou sdíleny všemi uživateli, kteří k nim mají mít přístup.

Tato změna koncepce od souborově orientovaných systémů k databázovým s sebou nese i změny v přístupu k vytváření aplikací. Jedná se o přechod od procesy řízenému k datacentrickému přístupu vytváření aplikací. Hlavní přínosy tohoto přístupu jsou existence datového úložiště, existence datového slovníku a databázového stroje, který umožňuje efektivně nakládat s daty v úložišti. Dále pak řízení přístupu

k datům, podpora základních operací s těmito daty a podpora transakcí.

Databázové systémy plní v organizacích různé úlohy. Někde slouží jako transakční systémy (OLTP), jinde jako podpora strategického rozhodování (OLAP). V poslední době nabízejí některé systémy též služby takzvané business intelligence, což umožňuje provádět OLAP dotazy nad OLTP daty v reálném čase.

S databázovými systémy úzce souvisí datové modely použité při návrhu databáze. Jelikož databáze by měla co nejvěrněji reprezentovat určitý výsek reality, byly navrženy různé způsoby, jak tuto realitu modelovat. V současnosti je nejčastěji používaným datovým modelem model relační. Relační model vychází ze solidního základu matematických principů odvozených z teorie množin a predikátové logiky. Definuje způsob, jakým je možné data reprezentovat (strukturu dat), způsoby jejich ochrany (integritu dat) a dále operace, které můžeme nad daty provádět (manipulace s daty).

Obecně se dá říci, že relační databázové systémy vykazují následující charakteristické vlastnosti:

- veškerá data se pomyslně dají reprezentovat v pravidelně uspořádaných strukturách s řádky a sloupci, kterým se říká relace,

- všechny hodnoty jsou skalární,
- operace v databázi se provádějí vždy nad celou relací a jejich výsledkem je opět relace.

Přestože v dnešní době relační databázové systémy převažují, je nyní již jasné, že v mnoha situacích nestačí postihnout komplexnost vztahů, které mají modelovat. Svět kolem nás není plochý a komplikované vztahy mezi modelovanými entitami nelze jednoduše postihnout tabulkami. U nových databázových aplikací se čím dál tím častěji začíná prosazovat objektová technologie, protože vývoj složitých aplikací je mnohem rychlejší a pozdější úpravy jsou mnohem jednodušší. Objektová technologie poskytuje mnoho výhod, například podporuje bohatší datové struktury, pojetí objektu jako „černé skříňky“ se zapouzdřenými prostředky, umožňuje programátorům zdokonalovat implementaci chování objektu, aniž by přitom ovlivňovali ostatní části aplikace atd. Otázka perzistence objektů je v databázových systémech pojímána různě. Některé systémy k perzistenci přistupují tak, že ukládané objekty mapují do tabulek relačního systému, jiné ukládají objekty přímo do vícerozměrných datových struktur.

Jelikož jsou informace považovány za významná aktiva organizace, musí být také náležitým způsobem spravovány a chráněny. Je

tudíž nutné se vypořádat s problémy, jak a kde mají být data, jež slouží jako zdroj informací, uskladněna. Mají být uložena na jednom, či na více místech. Oba přístupy mají své klady a zápory, a volba centralizovaného či distribuovaného úložiště závisí na konkrétních požadavcích a možnostech vlastníka dat.

Takto bychom mohli pokračovat dál a dál, zabývat se například vlivem internetu na databázové technologie, avšak z výše uvedených skutečností je již patrné, že pojem databáze, tak jak ho používají lidé, má mnoho tváří, mnoho odstínů a v každé situaci musíme být připraveni čelit scénářům, kdy na dotaz ohledně podivné hodnoty na faktuře od mobilního operátora, vám osoba na druhém konci telefonu sdělí, že „tato chyba vznikla tím, že databáze vytiskla špatná čísla“.



Autor působí jako sales engineer společnosti InterSystems.