

Architektury SŘBD

Po absolvování modulu Architektury SŘBD budeme umět

- vyjmenovat jednotlivé typy architektury SŘBD
- vysvětlit, v čem spočívá základní rozdíl mezi architekturami
- stručně charakterizovat jednotlivé typy z hlediska zpracování dat
- vysvětlit základní výhody a nevýhody jednotlivých typů architektur

Centralizované platformy

Zásadní rozdíl mezi jednotlivými architekturami SŘBD je určen místem, kde probíhá vlastní zpracování dat. Typ architektury zpravidla přímo nepopisuje, na jakých počítačích (PC, ..., mainframe) bude databáze provozována.

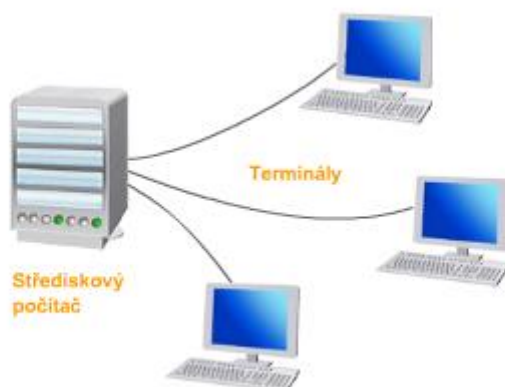
V centralizovaných systémech budeme rozlišovat dva základní prvky:

- hostitelský počítač
- terminály

Na **hostitelském počítači** se zpracovávají všechny programy - **SŘBD, aplikace, uživatelské rozhraní** a komunikační software přenášející informace na uživatelské terminály a z nich.

Uživatelé přistupují k databázi z terminálů.

Terminály zpravidla nemají žádné, nebo jen **omezené možnosti lokálního zpracování**. V nejjednodušších případech se skládají pouze z monitoru, klávesnice a hardware pro komunikaci s hostitelským počítačem. Inteligentní terminály mohou přejímat část zodpovědnosti za manipulaci s obsahem obrazovky a zpracování vstupních dat uživatele.

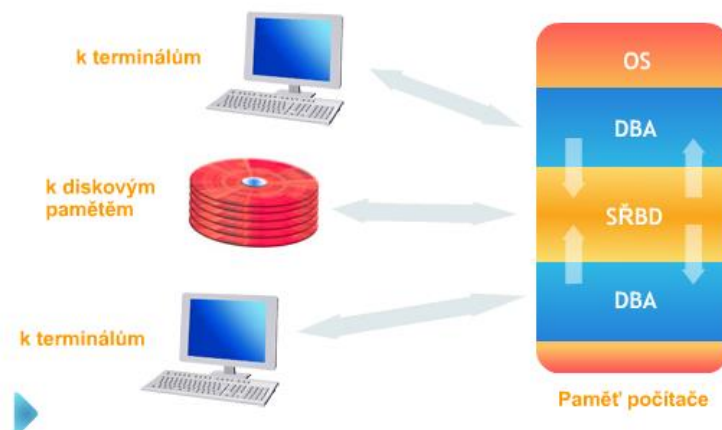


V centralizovaném systému probíhá veškeré zpracování na hostitelském počítači.

Na hostitelském počítači musí být spuštěn SŘBD.

Uživatel po zapnutí terminálu musí zadat své přihlašovací jméno a heslo a tím získá přístup k aplikacím na hostitelském počítači. Po spuštění databázové aplikace jsou obsahy jednotlivých obrazovek vysílány z hostitelského počítače na terminál.

Aplikace i SŘBD jsou spuštěny na hostitelském počítači a komunikují spolu navzájem v operační paměti hostitelského počítače:



SŘBD odpovídá za přenos dat na disk a z disku s využitím služeb operačního systému. SŘBD může být založen na libovolném databázovém modelu - nejčastěji ale na síťovém nebo relačním.

Výhody centralizovaných systémů:

- centrální zabezpečení
- uložení velkého množství dat
- možnost současné práce mnoha uživatelů (i více než 1 000)

Nevýhody centralizovaných systémů:

- vysoké náklady na pořízení a údržbu systémů
 - nutnost specializovaných technických zařízení (klimatizace, ...)
 - tým kvalifikovaných systémových administrátorů

1. Kde probíhá zpracování databázové aplikace v centralizovaném systému?

Systémy na osobních počítačích

Běží-li SŘBD na PC, pracuje PC současně jako hostitelský počítač i terminál - **veškeré zpracování probíhá na PC**. V nejjednodušších případech bývají SŘBD a databázová aplikace spojeny do jednoho programu. Databázová aplikace často zpracovává vstup od uživatele, výstup na obrazovku i přístup k datům na disk.

Toto řešení má několik **výhod**:

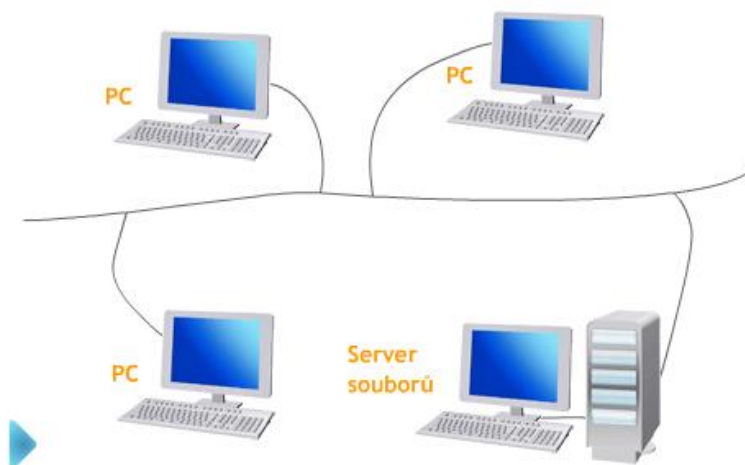
- flexibilita aplikace
- rychlost aplikace

ale i **nevýhod**:

- nízká bezpečnost a integrita dat
- malé objemy zpracovávaných informací

Většina SŘBD na PC je založena na relačním modelu, ale často v nich není implementována část relačních principů. Velmi často tyto systémy nemají integrovány mechanismy integrity dat. Jednodušší systémy často umožňují i přímý přístup k datům (bez použití příslušného SŘBD). Tak mohou být v datech prováděny různé změny narušující integritu dat, kterou jinak kontroluje SŘBD.

PC bývají velmi často propojeny do lokálních sítí (LAN). V lokální síti jsou data uložena na serveru souborů. Server zajišťuje uživatelům LAN sdílený přístup k datům na svých pevných discích. LAN umožňuje uživatelům databáze na PC sdílet společné datové soubory, ale nemění princip zpracování- veškeré zpracování dat probíhá na PC, na němž je spuštěna databázová aplikace. Souborový server pouze vyhledává požadované soubory na discích a po síťovém médiu je posílá na PC. Data zpracovává SŘBD na PC. Každá změna v databázi vyžaduje, aby PC poslalo celý soubor s daty na server, který jej uloží na disk:



Nevýhody víceuživatelského přístupu v systémech na PC:

- výkonnost systému není závislá na výkonnosti serveru, ale na výkonu koncového PC
- zpomalení práce sítě - pokud pracuje s databází více uživatelů, musí server poslat tytéž soubory na každé PC, které je používá
- nutnost řešit současný přístup více uživatelů k databázi. Pokud chce více uživatelů současně provádět změny dat, je nutné měněný soubor uzamknout, aby k němu měl přístup pouze jeden z nich (více podrobností o této problematice se dozvíme v kapitole věnované paralelnímu zpracování dat)

2. Jaké jsou výhody SŘBD na PC?

Databázové systémy klient/server

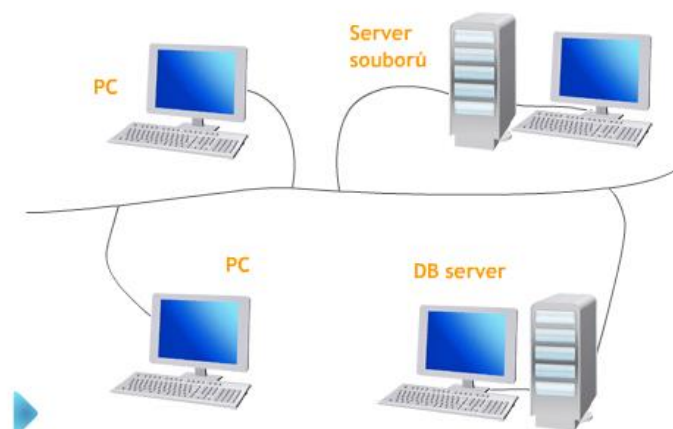
Databáze klient/server rozděluje zpracování mezi dva systémy:

- **klient** (např. PC) na němž běží databázová aplikace
- databázový **server** na němž běží SŘBD, Databázový server může běžet na témže stroji jako server souborů nebo na vlastním stroji

Klientský stroj (PC) je označován jako systém **front-end** a zajišťuje zpracování obrazovek a uživatelských vstupů a výstupů. Systém **back-end** na databázovém serveru zajišťuje zpracování dat a přístup na disk.

Princip fungování systému:

- uživatel na počítači front-end vytvoří požadavek (dotaz) na data z databázového serveru, aplikace front-end vyšle požadavek po síti na server
- databázový server provede vlastní vyhledání dat a pošle zpět pouze data, která odpovídají uživatelskému dotazu



Po síti tedy neputují celé datové soubory, ale pouze dotazy (nejčastěji SQL příkazy - o velikosti řádově desítek bajtů) a odpovědi na tyto dotazy (nejčastěji jeden nebo několik datových záznamů z databáze). Klientské systémy obvykle běží na PC, jako databázový server může být použito PC i počítače vyšší kategorie.

V současné době mají obrovský přínos webové služby (web service), které se dají využít téměř při jakékoli úrovni distribuce aplikace mezi klienta a server. Webová služba zpravidla obaluje datový zdroj a komunikuje s uživatelem. Tento koncept je výhodný z hlediska zabezpečení.

Mezi největší **výhody** architektury klient/server patří:

- rozdělení práce mezi klienta (aplikaci) a databázový server
- redukce množství dat posílaných po síti
- udržení integrity dat

Nevýhodou systémů klient/server je skutečnost, že ve své základní variantě:

- nepodporuje rozdělení dat mezi několik lokalit, tzn. data musí být uložena na jediném databázovém serveru

3. Jaký je rozdíl mezi objemem údajů posílaných po síti mezi architekturou databází na PC (v síti) LAN a architekturou klient/server?

Systémy distribuovaného zpracování

Systémy distribuovaného zpracování umožňují transparentní přístup k datům.

V takovémto systému požádá uživatel o data svůj lokální databázový server. Pokud server nemá požadovaná data k dispozici, pošle po síti požadavek počítači, na němž jsou tato data uložena. Tento počítač (server) pošle data zpět uživateli, aniž by se uživatel dozvěděl, že data pocházejí z jiného než jeho lokálního Serveru.

Mezi hlavní výhody patří dostupnost (např. pobočky pojišťovny), výpadky většinou neohroží korektní přístupnost na další uzly (transparentnost). Využívají se převážně v bankovníctví či ve zdravotnictví. V případě bankovníctví jsou tak účty klientů spravovány lokálně, ale jsou přístupné z jiných poboček.

Distribuovaná databáze je tedy jediná logická databáze, která je fyzicky rozmístěna na více počítačích spojených komunikační sítí.

Distribuované databázové systémy tedy kombinují dvě technologie - databáze a komunikace.

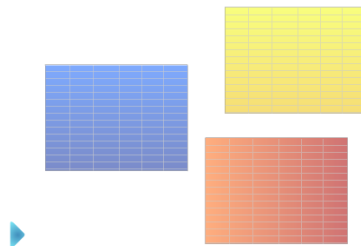
Distribuované databázové systémy je možné posuzovat ze tří základních pohledů:

- *autonomie*: autonomie se týká distribuce řízení
 - může se jednat o těsné spojení, kdy je distribuovaná databáze vytvořena nad lokálními databázemi a každé místo má úplnou znalost o všech datech v celé distribuované databázi
 - při volnější autonomii jsou sdílena pouze určitá data z lokálních databází

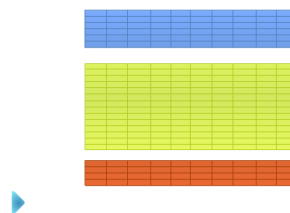
- v případě úplné izolace jsou jednotlivé lokální databáze autonomní a navzájem o sobě neví, distribuovaná databáze pak tvoří speciální vrstvu nad lokálními systémy
- *transparence distribuce*: určuje, do jaké míry je distribuce dat v jednotlivých místech viditelná uživatelům
- *heterogenost*: může se jednat o různorodý hardware v různých místech, o různé typy lokálních SŘBD nebo dokonce i různé databázové modely

Různé systémy mohou využívat různé druhy distribuce:

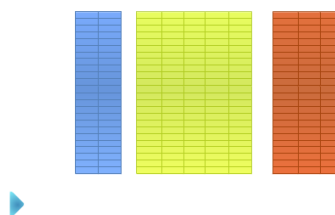
- replikace: na všech místech jsou uložena stejná data - kopie jedné a téže databáze



- objektové (horizontální) členění - na jednotlivých místech jsou uloženy informace o různých objektech, ale se stejnou strukturou



- funkční (vertikální) členění - na jednotlivých místech jsou uloženy informace s různou strukturou



4. Pokuste se definovat pojem distribuovaná databáze?

Databázové aplikační programovací jazyky

Aby měl SŘBD praktický význam, musí uživatelům umožnit přístup k datům. Možnosti, které SŘBD poskytuje pomocí vlastních prostředků, zpravidla nejsou příliš komfortní a proto uživatel využívá databázovou aplikaci.

Databázová aplikace je program umožňující komunikaci uživatele s daty (zadáání,změny, rušení a vyhledávání uložených údajů). Aplikace sestavují programátoři v univerzálních nebo specializovaných programovacích jazycích.

Pro tvorbu databázových aplikací je možné využívat:

- objektově orientované jazyky
- SQL
- ostatní jazyky

Běžné hybridní jazyky (jako Delphi, C++) mohou být využity pro tvorbu databázových aplikací prostřednictvím **API** (Application Programming Interface - aplikační programové rozhraní). API se skládá z množiny funkcí rozšiřujících jazyk o prostředky přístupu k datům s využitím SŘBD.

Vyšším programovacím jazykům, které mohou být použity pro tvorbu programu (a tedy i databázové aplikace) se říká 3GL (jazyky 3.generace).

Některé programovací jazyky jsou specifické pro určitý SŘBD. Takovýmto jazykům se obvykle říká 4GL (jazyky 4.generace).

Pomocí **objektově-orientovaných jazyků** (C#, Java, ...) se program nepíše jako řada procedur, ale definují se objekty a akce nad těmito objekty.

SQL (Structured Query Language - strukturovaný dotazovací jazyk) je neprocedurální jazyk, který nepopisuje, jak je třeba operace provést, ale co je od databáze požadováno. Pomocí jazyka SQL je možné definovat data a manipulovat s nimi. SQL neobsahuje žádné prostředky pro manipulaci s obrazovkami a pro uživatelský vstup a výstup.

SQL má trojí využití:

- SQL jako dotazovací (manipulační) jazyk pro relační databáze
- SQL jako složka hostitelského jazyka pro programování databázových aplikací
- SQL jako jazyk komunikace mezi různými zdroji dat

Mezi ostatní jazyky patří například jazyky **maker**. Makrojazyky umožňují nahradit posloupnost kláves, kterou by uživatel musel stisknout, kratší posloupností - makrem. Tímto způsobem je možné zautomatizovat některé jednoduché operace. Makra se používají spíše u jednodušších SŘBD.

Dalším představitelem této skupiny je **QBE** (Query-By-Example). QBE je uživatelské rozhraní, které umožňuje uživateli snadno vytvářet dotazy, aniž by znal syntaxi příkazů SQL. Zpravidla je jej možné využít i k tvorbě jednoduchých výstupních sestav.

Při tvorbě aplikací je velmi vhodné využívat jmennou konvenci DB schémat:

- názvy tabulek v jednotném čísle, malými písmeny, bez diakritiky
- názvy atributů malými písmeny, oddělovat podtržítkem

- sloupec reprezentující jednoznačný identifikátor (umělý PK) by se měl patřičně jmenovat - název tabulky + _id (employee_id)
- cizí klíč by se měl jmenovat stejně jako primární, včetně postfixu id – pokud to dává smysl, jestliže je třeba vyjádřit skutečnost, že se jedná o zodpovědného zaměstnance projektu, určitě by nemělo zmizet minimálně _id (responsible_employee_id)
- názvy uložených procedur, pohledů... s odpovídajícím prefixem (sp_insert_employee, vw_employee_programmers)

5. K jakému účelu se používají databázové aplikační jazyky?

Doporučená literatura

- Salemi J. - Databáze klient/server, str. 17 - 29
- Pokorný J. - Databázová abeceda - str. 61 - 64; 149 - 152; 169 - 174;
- Connolly T. - Database systems, str. 383 - 526
- Date C.J. - An Introduction to Database systems, str. 42-43; 593-622

Řešení úkolů

1. Kde probíhá zpracování databázové aplikace v centralizovaném systému?

V centralizovaném systému je databázová aplikace (stejně jako SŘBD a uživatelské rozhraní) spuštěna na hostitelském počítači.

2. Jaké jsou výhody SŘBD na PC?

Uživatelské rozhraní, databázová aplikace i SŘBD běží na PC. Mezi výhody tohoto řešení patří flexibilita a rychlost aplikace. Tyto výhody ale platí pouze pro malé aplikace.

3. Jaký je rozdíl mezi objemem údajů posílaných po síti mezi architekturou databází na PC (v síti) LAN a architekturou klient/server?

V případě architektury databází na PC se po síti posílají celé databázové soubory. V architektuře klient/server se po síti posílají dotazy a odpovědi na ně (tzn. vybrané údaje z databáze). Objem zasílaných informací je u architektury klient/server mnohonásobně menší.

4. Pokuste se definovat pojem distribuovaná databáze?

Distribuovaná databáze je jediná logická databáze, která je fyzicky rozmístěna na více počítačích. Uživatel, který přistupuje k distribuované databázi nemusí vědět, na kterém z propojených počítačů jsou data uložena, vyhledání dat zajistí distribuovaný databázový systém.

5. K jakému účelu se používají databázové aplikační jazyky?

Databázové aplikační se používají pro tvorbu uživatelského rozhraní.