
KAPITOLA 2

Architektura, modelování a implementace Business Intelligence procesů v SQL Serveru 2008

V této kapitole:

- ◆ Architektura Business Intelligence na platformě SQL Serveru 2008
 - ◆ Modelování procesů Business Intelligence
 - ◆ Nástroje pro práci s analytickými službami
-

Při analýze možností integrace jednotlivých podnikových systémů a procesů, ať už ekonomických, nebo výrobních, se hlavním integračním prvkem stává databáze, kde všechny integrované a konsolidované systémy ukládají svoje data. Firma Microsoft v této oblasti profituje z tradice jednoho z nejvýznamnějších dodavatelů databázových systémů. Její nový databázový produkt SQL Server 2008 je moderní komplexní serverová platforma pro ukládání a správu dat v databázích a datových skladech zahrnující sady nástrojů pro Business Intelligence včetně pokročilého reportování. IT ekosystém podnikových aplikací využívající SQL Server 2008 by měl být koncipovaný tak, že nejvyšší úroveň architektury je SharePoint Server pro spolupráci a vrstva Content management pro vyhledávání a správu obsahu.

SQL Server 2009 umožňuje prolínání a integraci podnikových procesů i na úrovni analytických služeb. Analytické služby implementované přímo na databázovém serveru tak mohou být využívány buď izolovaně v jednotlivých odvětvích, nebo komplexně. Nemusí jít jen o podporování rozhodování na strategické úrovni týkající se top managementu. Na jedné straně analytických procesů jsou podklady pro analýzy uložené v databázích nebo datových skladech. Stejně důležitý je na druhé straně i způsob prezentování výsledků analýz. Pracovníci dávají před specializovanými aplikacemi přednost používání nástrojů, na které jsou zvyklí ze své běžné práce, například programy kancelářských balíčků.

Architektura Business Intelligence na platformě SQL Serveru 2008

Implementaci Business Intelligence na platformě SQL Serveru 2008 na příslušné hierarchické úrovni rozdělujeme na tři částečně nezávislé, ale zpravidla do značné míry na sebe navazující bloky:

- ◆ **Integrační služby** – získání dat z různých nehomogenních zdrojů a jejich případná transformace, integrace a syntéza.
- ◆ **Datový sklad** – repozitář pro ukládání dat s využitím relačních databázových technologií.
- ◆ **Analytické služby** – v této fázi jsou data obohacena o výsledky analýz, různé predikce z data miningu a podobně, a tím se vlastně stávají cennými informacemi pro podporu rozhodování.
- ◆ **Reportovací služby** – slouží pro zpřístupnění dat a výsledků analýz uživatelům ve vhodné formě, obsahu a rozsahu.

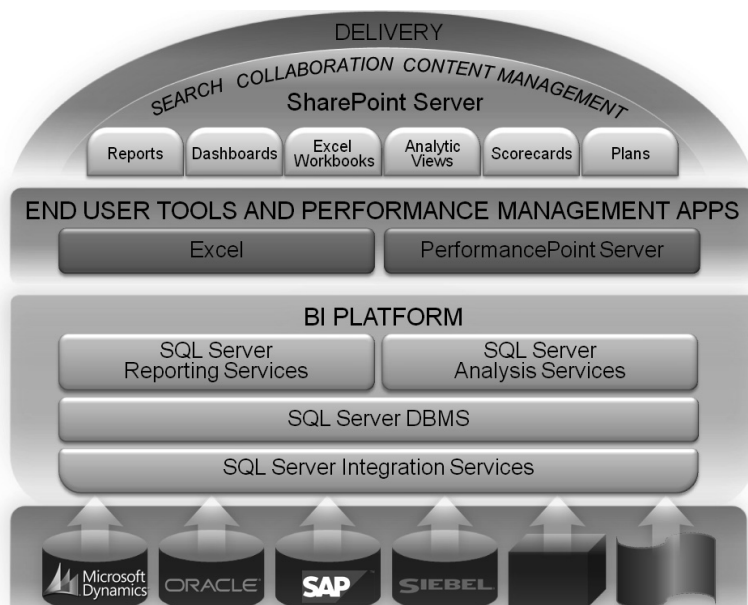
Situaci přehledně znázorňuje schéma na obrázku 2.1.

Ve většině komplexnějších úloh z oblasti BI využíváme všechny tři základní pilíře implementace na platformě 2008, a to v naznačeném pořadí.

Při pohledu na schéma si všimněte, že BI platforma je situována mezi zdroje údajů a nástroje pro koncové uživatele, přičemž blok integračních služeb je v pozici vstupní brány do modulu BI.

Modelování procesů Business Intelligence

Primárním požadavkem koncových uživatelů je přístup k informacím z nejrůznějších firemních procesů, podobně jako jsou na to zvyklí z různých jednoúčelových aplikací a systémů, jako například Enterprise Resource Planning (ERP). Znovu zdůrazňujeme, že pro tyto uživatele je důležitý přístup k informacím, ne přístup k datům. Zařídit přístup k datům je úlohou integračních služeb, ale proces získávání informací z dat se táhne napříč celým BI procesem. U mnoha starších řešení je architektura BI procesů poměrně chaotická, budovaná buď metodou „velkého třesku“, nebo neustálým přidává-



Obrázek 2.1. Základní schéma Business Intelligence na platformě SQL Serveru 2008

ním nových vstupů dat. Data jsou i po prvotní konsolidaci do datového skladu umístěna v různých typech zdrojů, v relačních databázích, nebo kostkách. Při takovém způsobu ukládání dat docházelo velmi často k duplikacím. Například podklady pro vytvoření dimenze jsou jednak v relační vysoce normalizované produkční OLTP databázi, jednak i v datovém skladu, kde jsou již částečně vyextrahována data v tabulkách faktů a dimenzí. Duplicitní mohly být nejen data, ale i modely. Jeden model sloužil pro datový sklad, jiný model pro OLAP, jiný pro generování reportů prostřednictvím reportovacích služeb. O výhodách oddělení relačních a analytických databází jsme se zmínili v teoretickém úvodu. Když bychom ale oddělili multidimenzionální OLAP databázi od relačního reportování, nastane problém s prezentováním dat a s výsledky analýz.

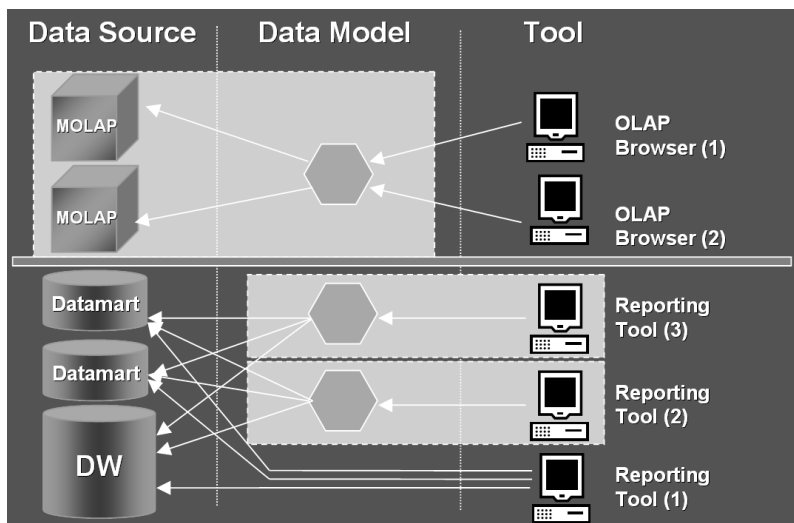
Jedním z preferovaných trendů pro budování podnikových projektů je model **Driven Development**, podle kterého je budování projektu zahájeno fází modelování. Na základě vytvořeného modelu je následně vygenerovaná struktura pro datový sklad a jsou vytvořeny úlohy pro jeho naplnění. Budování projektu pokračuje návrhem měřítek, dimenzí, kostek a podobně. U SQL Serveru 2008 je pro modelování implementovaná technologie UDM (Unified Dimension Model).

Úlohou UDM (Unified Dimension Model) je vytvoření mostu mezi daty a jejich uživateli.

Výstižný může být i cíl „Jediná verze pravdy“. I když to zní trochu jako marketingový slogan, vystihuje podstatu unifikovaného dimenzionálního modelu. Měla by existovat jen jedna uživatelsky přístupná a dostupná cesta ke všem zdrojům dat. Všechno může být unifikované počínaje cestou ke zdrojům dat, přes jednotný repozitář pro business logiku, až po postupy a kalkulace. To s sebou přináší jednotnou množinu obchodních metrik a uživatelů, a co je nejcennější – jednotnou verzi obchodních informací.

Pro pochopení výhod modelování BI může být poučné nahlédnout do historie, konkrétně až k verzi SQL Server 2000. Na obrázku 2.2 vidíte, že vrstvy databázových, analytických a reportovacích služeb

jsou poměrně striktně oddělené, přestože už i zde můžeme najít mnoho sjednocujících prvků – lze analyzovat data z relačních databází, případně generovat reporty z relačních i analytických databází. Případně lze pomocí MDX dotazu (zkratka MDX je akronymem slovního spojení Multidimensional Expressions – multidimenzionální výrazy) specifikovat určitou podmnožinu dat z multidimenzionální struktury (OLAP kostky) a vypsat ji ve formě dvojrozměrné tabulky, která obsahuje množinu buněk.



Obrázek 2.2. Oddělení vrstvy OLAP a reportovacích služeb u SQL Serveru 2000

Hlavní nevýhodou této architektury je redundance dat a modelů, která je zřejmá už na první pohled. Ta samá data jsou jednak v relačních, jednak v multidimenzionálních databázích. Při hlubší analýze datových toků zjistíte, že situace z hlediska redundance je ještě komplikovanější. Ta samá data se přenášejí z relačních databází do datových skladů, odtud často do datových trhů a dále do multidimenzionálních OLAP kostek. Další nezanedbatelná redundance a s ní množství práce vysoce kvalifikovaných specialistů se objevuje ve vrstvě modelů. Jiným problémem takto oddělené architektury jsou rozdílné návyky uživatelů při práci s daty.

Každá kostka mohla být vytvořena jen z jediného datového zdroje. Ke kostce se potom připojovali klienti.

Dalším krokem pro pochopení výhod UDM modelování je porovnání vlastností relačních a multidimenzionálních databází. Uvedeme je jen v heslovitém přehledu:

Relační databáze

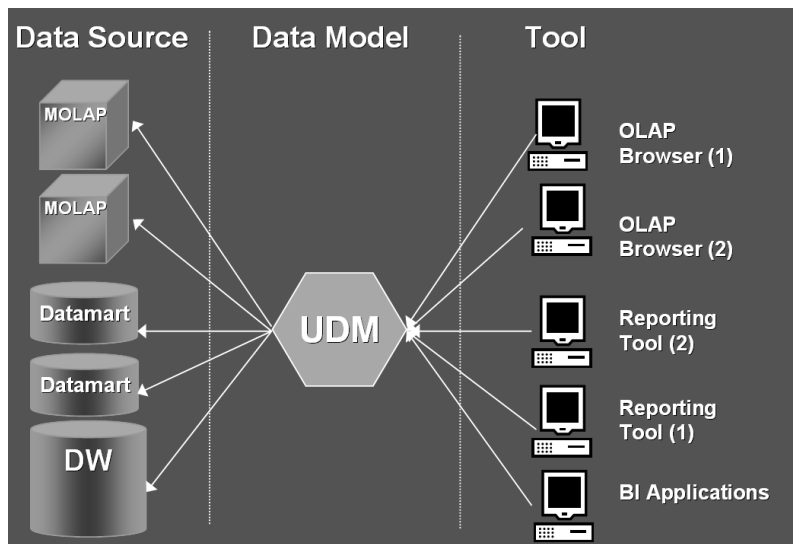
- ◆ Vícenásobné tabulky faktů
- ◆ Žádné omezení pro atributy dimenzí
- ◆ Možnost transakčních přístupů
- ◆ Možnost použití vyšších normálních forem
- ◆ Relační vztahy mezi objekty v databázi
- ◆ Možnost definování složitých vztahů typu many-to-many

Multidimenzionální databáze

- ◆ Multidimenzionální navigace
- ◆ Možnost hierarchické prezentace
- ◆ Jména entit se vztahují k předmětu podnikání
- ◆ Možnost výkonných MDX kalkulací
- ◆ Pohled na údaje z více různých perspektiv
- ◆ Rozdělení (v orig. partition)
- ◆ Databáze obsahuje výsledky agregací
- ◆ Distribuované zdroje

Modelování s využitím UDM (Unified Dimensional Model) umožňuje využití výhod obou typů databází a do jisté míry eliminuje jejich nevýhody.

Od verze SQL Serveru 2005 jsou vrstvy Business Intelligence sjednocené do jednotného modelu Unified Dimensional Model (UDM). Schéma si můžete prohlédnout na obrázku 2.3. Tento model převzal to nejlepší z reportování a OLAP analýz. Ke sjednocování dochází na úrovni modelů, kdy je potřebný jen jeden dimenzionální model pro generování reportů i OLAP kostek. Významnou novinkou je zavedení nové skupiny služeb s názvem Integration Services. Na úrovni této vrstvy jsou integrována data z různých datových zdrojů včetně jejich požadovaných transformací.



Obrázek 2.3. Unified Dimensional Model poskytuje platformu pro sjednocení na úrovni modelů

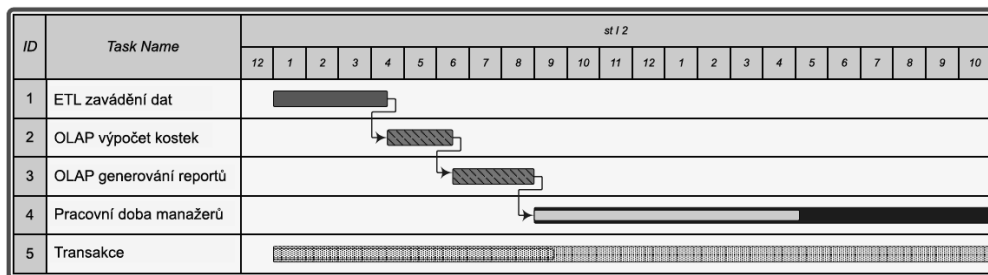
Když se čtenáři mající zkušenosti s moderními databázovými servery obsahujícími analytické a reportovací služby (například SQL Server, Oracle, IBM DB2 a podobně) zamyslí nad příklady analýz, které je možné pomocí těchto nástrojů vytvářet, uvědomí si, že vždy byly vybírány tabulky z relační databáze nebo z datového skladu, jež sloužily pro vytvoření faktů a dimenzí. Mechanismus MOLAP (multidimenzionální online analytické zpracování) naproti tomu uloží analytická data do vlastních datových struktur a sumářů.

Během tohoto procesu se spočte tolik předběžných výsledků, kolik je z technického a časového hlediska možné. Data v úložišti typu MOLAP se tedy budou ukládat jako předem vypočítané pole. Hodnoty dat i indexů se uchovávají v jednotlivých polích multidimenzionální databáze. Databáze je organizovaná tak, aby umožnila rychlé získávání příslušných údajů z více dimenzí. Proto je hlavní výhodou MOLAP maximální výkon vzhledem k dotazům uživatelů. Nevýhodou je redundance dat, protože ta jsou uložena jak v relační databázi, tak v multidimenzionální databázi. Požadavky na úložnou kapacitu mohou v případě použití více dimenzí extrémně narůstat. Pokusíme se podívat na proces Business Intelligence z časového hlediska.

Časová postupnost sekvenčního BI

Při návrhu koncepce řešení musíme vycházet z provozu firmy a zabezpečit, aby manažeři a analytici měli data k dispozici co nejdříve. Intervaly jsou různé, například týden, den, někdy i kratší. Typickým příkladem pro zpracování s denní periodou je časová tabulka z obrázku 2.4, která zobrazuje časovou posloupnost procesů analýzy. Pozor, data v prvních třech řádcích se týkají předchozího dne, to znamená, že v našem případě se po půlnoci vždy analyzují data z předchozího dne.

Při slovním popisu celého procesu znázorněného v časové tabulce na obrázku 2.4 bude lepší postu-



Obrázek 2.4. Časová posloupnost BI procesů v rámci jednoho pracovního dne

povat zespodu nahoru.

Transakce

V posledním řádku (5) vidíme, že transakce dnešního dne běží nepřetržitě 24 hodin. Může se jednat o výrobní proces, elektronický obchod, bankovní transakce, prostě OLTP proces, který je zdrojem dat pro analýzy. Tyto analýzy dnešních dat budou prováděny až po půlnoci, aby zítra na začátku pracovní doby byly k dispozici. Pokud by byl transakční systém využíván i pro OLAP analýzy v době od 0:00 do 9:00, běžely by oba procesy souběžně, tj. dnešní transakce a analýza včerejších dat.

Pracovní čas manažerů

Začátek pracovní doby manažerů (v naší tabulce 9:00) je rozhodujícím okamžikem, kdy by měly být výsledky analýzy a příslušné reporty k dispozici. Všimněte si, že graf není scholastický, ale vychází z praktických zkušeností, kdy pracovní doba manažerů obvykle nekončí zároveň s obvyklým firemním koncem pracovní doby.

Proces přípravy dat a analýza

Aby mohli manažeři získat ve stanoveném čase kvalitní podklady pro podporu rozhodování, musí proběhnout příprava dat pro analýzu, jejich zavedení do datového skladu a následná analýza. V pří-

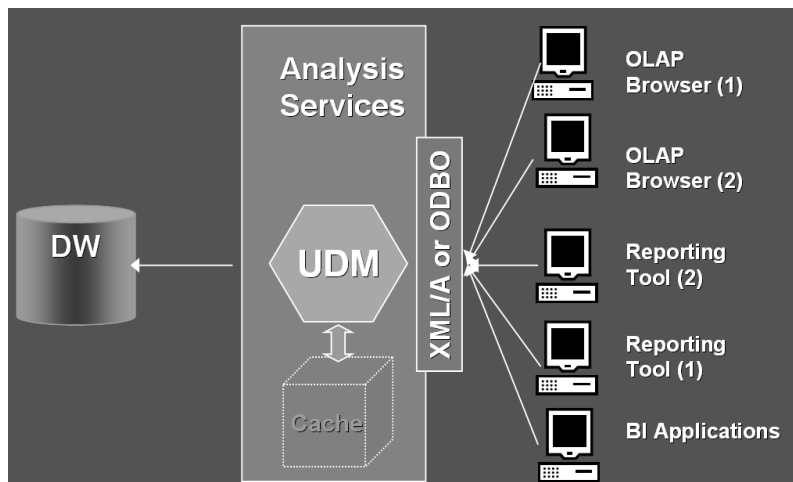
kladu znázorněném na grafu začne proces ETL (řádek 1 časové tabulky) hned po půlnoci v 0:00. Po několika hodinách je tato etapa ukončena a může začít výpočet kostek a ostatní agregační a analytické výpočty (řádek 2 časové tabulky). Na tento proces navazuje generování sestav, reportů a podobně (řádek 3 časové tabulky).

Samozřejmě musíme počítat také s určitou rezervou, například pro nepředvídatelné situace a podobně.

Business Intelligence v reálném čase

V dnešní hektické době je požadavek na výsledky analýz v reálném čase, nebo přesněji řečeno v termínu s kontrolovatelným zpožděním. Pro tento účel se využívá takzvané proaktivní cacheování, které je implementované výhradně v Enterprise verzi. SQL Server 2008 využívá vyrovnávací MOLAP paměť Cache, kam jsou ukládány nejčastěji používané nebo očekávané výsledky agregací. Cacheování v této paměti nám automaticky umožňuje synchronizaci multidimenzionálních dat v paměti Cache vůči primárnímu datovému zdroji. Při využití této technologie už není nutné pokaždé explicitně spouštět proces vytváření multidimenzionálních kostek. Proaktivní cacheování můžeme považovat za pokus spojení výhod aktuálnosti dat v OLTP databázích s možnostmi multidimenzionálních analýz. Data tedy zůstávají v relačních databázích a spočtené agregace se ukládají do multidimenzionálních struktur. Při dotazování jsou data vybírána do multidimenzionální paměti Cache. Schematicky je proces proaktivního cacheování znázorněný na obrázku 2.5.

Princip fungování proaktivního cacheování spočívá ve sledování změn primárního datového zdroje. Když na této úrovni dojde ke změnám, řadič cache je o změně informovaný. Vymaže se původní obsah multidimenzionálních dat v paměti Cache a spočítají se nové agregace. Pro koncového uživatele je cache neviditel-



Obrázek 2.5. Unified Dimensional Model využívá proaktivní cacheování

ná. Uživatel není schopen rozlišit, zda data pocházejí z reálného prostředí, nebo z vyrovnávací paměti.

Kvalita cacheování je dána dvěma parametry:

Koherence je definovaná jako zachování konzistence dat v paměti Cache a v jejich primárním datovém zdroji.

SQL Server 2008 umožňuje trasování serveru obsahujícího primární data. Musíme vybrat, které tabulky mají být trasovány, databázový server pak automaticky sleduje a hlásí případné změny. Když změny v relační databázi způsobí aplikace, která tato data sbírá, zpracovává a ukládá, tato aplikace přesně ví, kdy způsobí změnu dat a tuto skutečnost oznámí analytickým službám pomocí volání přes SOAP. Jestliže není možné využít některou z uvedených sofistikovanějších metod, můžeme nastavit analytické služby, aby se periodicky dotazovaly datového zdroje. Takto mohou fungovat prakticky vůči libovolnému zdroji.

Latence je zpoždění mezi okamžikem změny v primárním datovém zdroji a promítnutím této změny v paměti Cache.

Jestliže je latence nulová, jde o proces v reálném čase. Proto se na první pohled může zdát snižování latence jako ta nejsprávnější strategie. Když však data budou sbírána příliš často a pokaždé se budou počítat hodnoty OLAP kostek, může se stát, že se na tuto činnost spotřebuje příliš mnoho výpočetní kapacity a celková výkonnost systému se radikálně sníží. Proto musíme vhodným nastavením parametrů zajistit kompromis mezi přiměřeným výkonem a ještě postačující dobou latence.

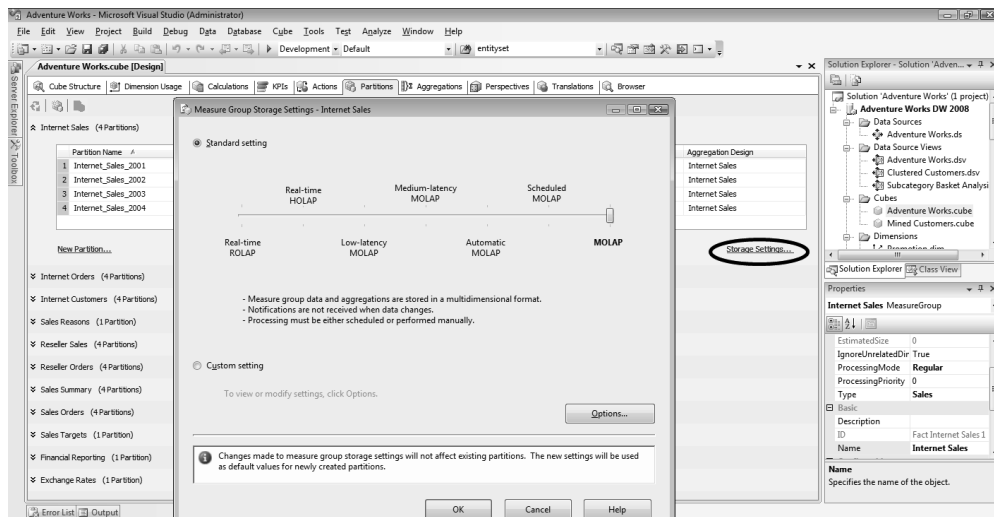
Tabulka obsahuje nejvýznamnější parametry, pomocí kterých je možné nastavovat úroveň koherence a latence:

Parametr	Popis
SilenceInterval	Doba po aktualizaci, během které nesmí dojít k další aktualizaci, aby mohlo být zahájeno zpracování OLAP kostky (-1 = ignoruje notifikace).
SilenceOverrideInterval	Když uplyne od první notifikace po posledním zpracování tento čas, spustí zpracování bez dalšího čekání (-1 = nezpracovává z tohoto důvodu).
Maximum Latency	Parametr udává, jak neaktuální může být cache bez toho, aby analytické služby začaly vybavovat požadavky přímo ze zdroje dat (ROLAP) (-1 = nikdy nepoužije ROLAP).
ForceRebuildInterval	Když uplyne od posledního zpracování doba nastavená tímto parametrem, spustí se nové zpracování bez ohledu na aktualizace (-1 = neprocesuj z tohoto důvodu).

Tyto parametry pro danou kostku můžete nastavit v záložce „partition“. Dialog umožňující jejich nastavení a změnu aktivujete pomocí odkazu „Storage Settings“ (viz obrázek 2.6).

Můžete použít jednu z úrovní standardního nastavení parametrů:

- ◆ **Real-time ROLAP** – data jsou v úložišti typu ROLAP, proaktivní cacheování je povoleno. Doba latence, protože se jedná o přístup v reálném čase, je nastavená na 0 sekund. Analytické služby přistupují přímo k relačním objektům.
- ◆ **Real-time HOLAP** – data jsou v úložišti typu HOLAP, proaktivní cacheování je povoleno. Doba latence je podobně jako v předcházejícím případě nastavená na hodnotu 0 sekund. Když se data změní, dochází k aktualizaci cache. Parametr SilenceInterval je nastavený na 0 sekund.
- ◆ **Low-latency MOLAP** – data jsou v úložišti typu MOLAP, proaktivní cacheování je povoleno. Doba latence je nastavená na hodnotu 30 minut. Parametr SilenceInterval je nastavený na 10 sekund a parametr SilenceOverride na 10 minut.



Obrázek 2.6. Nastavení parametrů koherence a latence

- ◆ **Medium-latency MOLAP** – data jsou v úložišti typu MOLAP, proaktivní cacheování je povoleno. Doba latence je nastavená na 4 hodiny. Parametr SilenceInterval je nastavený na 10 sekund a parametr SilenceOverride na 10 minut.
- ◆ **Automatic MOLAP** – data jsou v úložišti typu MOLAP, proaktivní cacheování je povoleno. Paměť Cache se aktualizuje, když se změní data. Parametr SilenceInterval je nastavený na 0 sekund a parametr SilenceOverride je zakázaný (nastavený na hodnotu -1).
- ◆ **Scheduled MOLAP** – data jsou v úložišti typu MOLAP, proaktivní cacheování je povoleno. Doba latence je nastavená na 1 den.
- ◆ **MOLAP** – data jsou v úložišti typu MOLAP

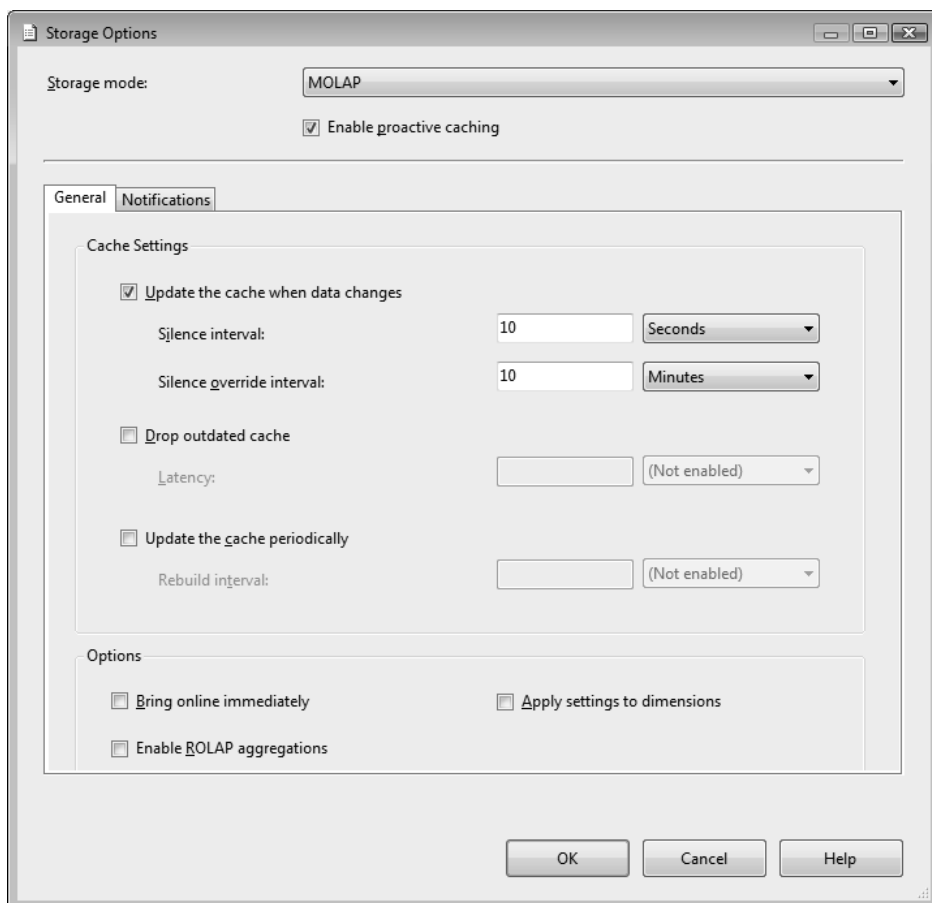
Druhá možnost je nastavit parametry přesně podle konkrétních požadavků pomocí dialogu aktivovaného tlačítkem „Options“ (viz obrázek 2.7).

Nástroje pro práci s analytickými službami

První předpoklad pro jednotný přístup k modelování je unifikace technologií. Tento předpoklad je splněný integrací databázových analytických a reportovacích služeb do jednoho balíku. Druhým předpokladem je unifikované uživatelské rozhraní, což v případě databázového serveru představují nástroje pro administraci, dotazování a návrh struktur a dílčích modelů. Také tento požadavek je v případě SQL Serveru 2008 splněný. K dispozici máme dva komfortní nástroje:

- ◆ SQL Server Management Studio
- ◆ Business Intelligence Development Studio

Kromě těchto nástrojů můžeme s výhodou využívat těsnou interakci mezi SQL Serverem 2008 a vývojovým prostředím Visual Studio 2008.

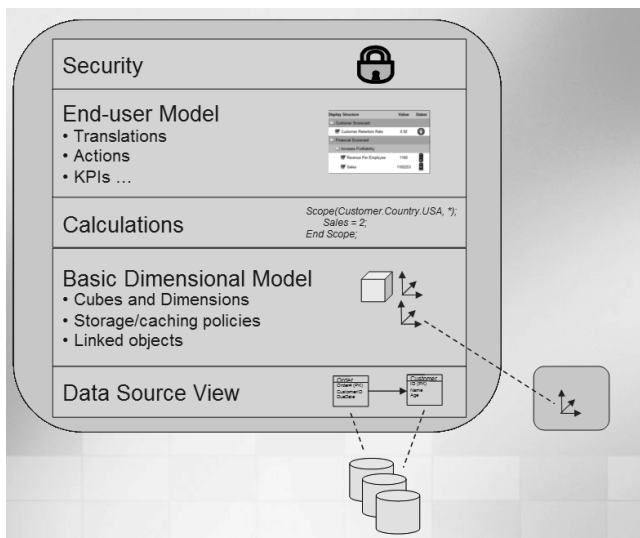


Obrázek 2.7. Vlastní nastavení parametrů

Implementace a přístup k UDM přes Business Intelligence Development Studio

Pochopit filozofii modelování nám pomůže hierarchické schéma. Celá vrstva sloužící pro vytvoření výstupu (reportu) je z architektonického hlediska postavená nad údaji v databázích. Samozřejmě je výhodné, aby tato data byla co nejvíce vyčištěná, případně do rozumné míry normalizovaná bez zbytečných redundancí a nepřehledných komplikovaných relačních vztahů. Samozřejmě ne každý den jsou Vánoce, a tak často při budování BI vrstvy musíme vycházet z dat, která jsou k dispozici, případně jaká vyhovují logice předmětu podnikání, nebo k jaké struktuře dat se příslušná firma dopracovala postupným budováním informačních systémů.

Samotná BI nadstavba nad databází je rozdělená do čtyř vrstev (viz obrázek architektury 2.8). Nad **databázemi** je vrstva **datových pohledů**, ve které jsou už přesně specifikované tabulky, pohledy, a to až na úroveň jednotlivých atributů a relačních vazeb. Modelování na úrovni vrstvy pohledů na data (v anglické terminologii Data Source View) do určité míry přispěje ani ne tak k fyzickému vyčištění dat, jako spíše s nadefinování přímějšího přístupu k nim.



Obrázek 2.8. Schéma architektury s využitím Unified Dimensional Modelu



Východisko

Na této úrovni se vytvářejí relační schémata, na základě kterých budou vytvářené dimenze a kostky. Velmi důležitou úlohou architektonické úrovně pohledů na data je sjednocení dat z heterogenních zdrojů.

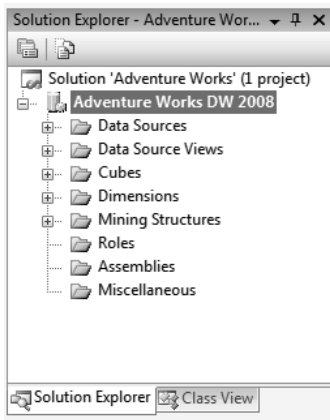
Následuje **úroveň dimenzí**, kde specifikujeme, které atributy se vážou k měrným jednotkám obchodování. Na úrovni základního dimenzionálního modelu do značné míry dochází i k ovlivňování způsobu ukládání a cacheování multidimenzionálních dat. Kromě faktů a dimenzí vstupují do hry i kalkulace, to znamená data vypočítaná podle jednoduchých, případně i složitých vztahů z hodnot jiných atributů. Například když je v některé tabulce údajů o osobách jedním z atributů rodné číslo, dokážete z něj určit datum narození, a tedy i věk příslušné osoby.

Vrstva **End-user Model** definuje manipulace s daty a výsledky souhrnů na nejvyšší úrovni. Uvědomte si, že výsledky analýz a agregovaná data obsahují mnohem citlivější informace než surová data. Proto je na úrovni výsledků analýz mnohem důležitější zabezpečení cílového přístupu k informacím.

Unifikovaný přístup k BI úlohám podle tohoto vícevrstvého schématu je naznačený i složkami v okně vývojového prostředí Solution Explorer v pravém horním rohu (viz obrázek 2.9).

- ◆ Data Sources
- ◆ Data Source Views
- ◆ Cubes
- ◆ Dimension
- ◆ Mining Models
- ◆ Roles

- ◆ Assemblies
- ◆ Miscellaneous



Obrázek 2.9. Business Intelligence Development Studio – okno Solution Explorer



Poznámka

Při vytváření modelu vlastně s větší nebo menší pomocí interaktivních průvodců budujeme v záložkách jednotlivé hierarchické vrstvy modelu.

Pohledy na datové zdroje – Data Source Views

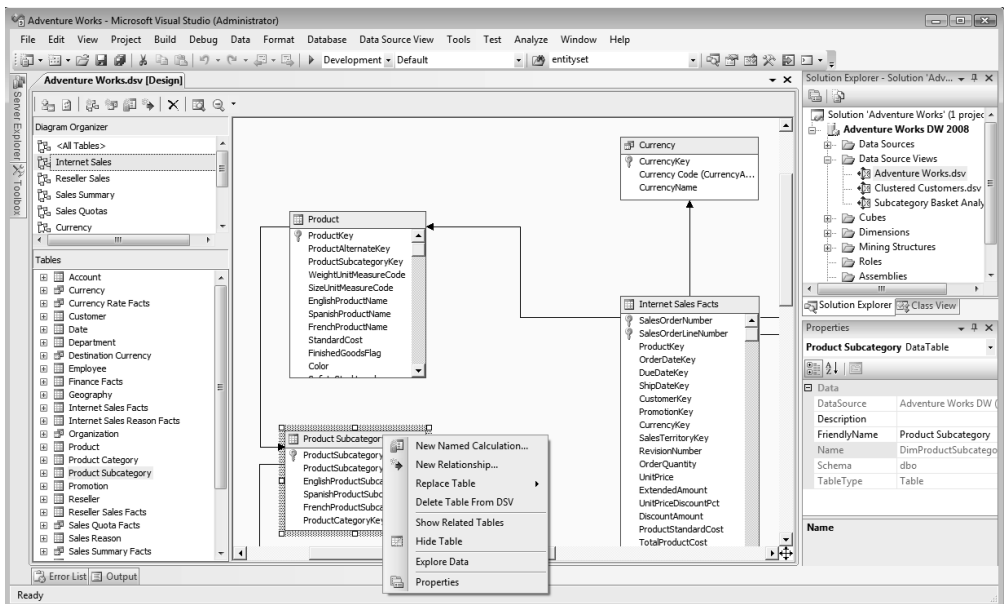
Tuto vrstvu si můžeme představit jako jakousi logickou virtuální databázi, ve které se nacházejí databázové tabulky, sloupce, primární klíče a podobně. Zpravidla je jedná o jakousi podmnožinu dat jedné nebo několika databází.

Pohledy na datové zdroje umožňují virtuálně měnit strukturu databáze.

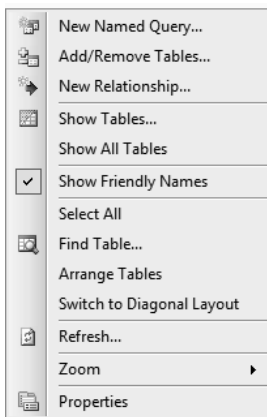
Na úrovni této vrstvy je možné spojovat více databázových zdrojů, používat vypočítané sloupce, vytvářet vlastní dotazy. Hlavní úlohou této vrstvy je oddělit analytické služby od fyzické struktury dat v databázi. Taktéž abstrahuje model od fyzické struktury dat. Při změně fyzické struktury dat stačí udělat jen změnu mapování na úrovni Data Source Views.

S pohledy na datové zdroje můžeme velmi efektivně pracovat v prostředí Visual Studio 2008 nebo v BI Development Studiu (viz obrázek 2.10). K dispozici máme místní nabídku (viz obrázek 2.11) pro jednotlivé tabulky nebo pro pohledy na datové zdroje jako celek. Nabídka pro konkrétní tabulku aktivujeme klepnutím pravého tlačítka myši na symbol příslušné tabulky. Všeobecnou nabídku aktivujeme klepnutím na volnou plochu diagramu.

Takto můžeme modelovat strukturu dat včetně relačních vztahů mezi tabulkami bez toho, abychom zasahovali do návrhových struktur zdrojových databází.



Obrázek 2.10. Práce s Data Source View v prostředí Business Intelligence Development Studio



Obrázek 2.11. Místní nabídka možností Data Source View v prostředí Business Intelligence Development Studio

Administrace databází a dotazování přes SQL Server Management Studio

Analytické služby, podobně jako databázový server, jsou koncipované jako služba na pozadí. Navíc s databázovými službami analytické služby velmi úzce spolupracují. Taktéž při přípravě dat nebo testování integračních služeb se neobejdete bez efektivního přístupu k databázovému serveru. Některé příklady v této publikaci předpokládají, že některé cvičné databáze jsou už nainstalované, případně uvádíme SQL příkazy pro vytvoření vlastních databázových tabulek přizpůsobených problematice daného příkladu. To všechno jsou úlohy pro klientskou konzolovou aplikaci.

Pomocí klientské konzolové aplikace zadáváme databázovému serveru příkazy v jazyce SQL a v jejím okně jsou následně výstupy, které databázový server vygeneruje jako odezvu na zadané příkazy.

Kromě klientské konzolové aplikace budete potřebovat také nástroj pro správu databáze. Prostřednictvím této aplikace můžete vytvářet nové databáze, vytvářet a spravovat účty jednotlivých uživatelů, přidělovat a rušit jejich oprávnění pro práci s jednotlivými objekty databáze. Pomocí nástroje pro správu databáze můžete nastavit i strategii údržby a zálohování dat v databázi a podobně. Pro práci s databázovým serverem Microsoft SQL Server 2008 je pro oba zmiňované účely k dispozici nástroj **SQL Server Management Studio**. Tento nástroj umožňuje i přístup k multidimenzionálním databázím.

SQL Server Management Studio je komplexní integrované prostředí pro správu databázového serveru SQL Server 2008. Podobně jako Business Intelligence Development Studio nebo Visual Studio i tento nástroj vychází z unifikovaného vývojového prostředí Microsoft Development Environment. Nástroj se spouští standardním způsobem z nabídky Start operačního systému Windows ve skupině aplikací a nástrojů složky Microsoft SQL Server 2008.

Připojení Management Studia k analytickým službám

Po spuštění nástroje se zobrazí dialog pro připojení k databázovému serveru. Jestliže děláme první pokusy na lokálním nebo virtuálním počítači, můžeme nastavit název serveru na localhost a jako metodu přihlášení použijeme Windows Authentication. Přihlašovací dialog je koncipovaný tak, že se lze ke každé službě připojit izolovaně. Můžeme se připojit k těmto typům služeb:

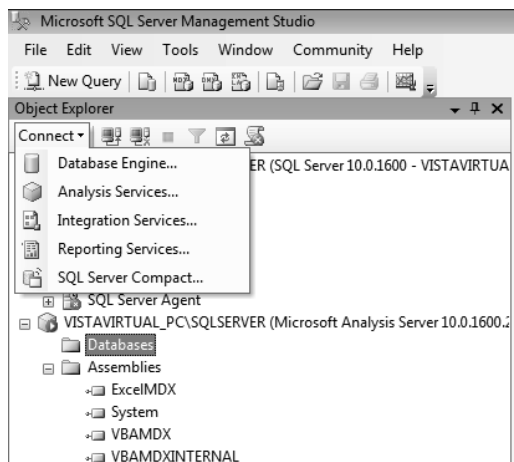
- ◆ Database Engine
- ◆ Analysis Services
- ◆ Reporting Services
- ◆ SQL Server Mobile
- ◆ Integration Services

Možnosti výběru pomocí prvku combo box představuje obrázek 2.12



Obrázek 2.12. Přihlašovací dialog SQL Server Management Studia

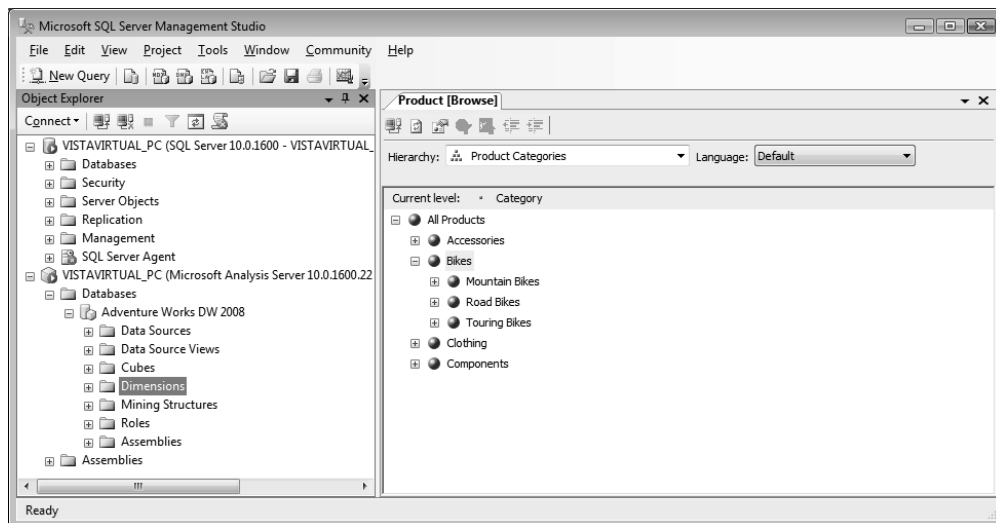
Když je SQL Server Management Studio připojené k nějaké službě, například k analytickým službám, a nastane potřeba operativního přepnutí na jiný druh služby, třeba na integrační služby, použijeme tlačítko „Connect“ na toolbaru okna Object Explorer (viz obrázek 2.13).



Obrázek 2.13. Možnosti přepínání připojení SQL Server Management Studia na jednotlivé druhy služeb

Práce s BI strukturami

Pracovní plocha nástroje SQL Management Studio je podle režimu práce rozdělena na několik částí. Nejtypičtější rozmístění je okno Object Exploreru v levé části a hlavní okno ve středu. Okno Object Explorer poskytuje grafický, přehledný, hierarchický pohled na jednotlivé komponenty až na nejnižší úroveň sloupců a indexů. V levé části je možné zobrazit i okno Registered Servers, které obsahuje seznam serverů a jejich služeb, na něž se můžeme pomocí Management Studia připojit.



Obrázek 2.14. Typické rozvržení oken nástroje SQL Server Management Studio při práci s BI strukturami

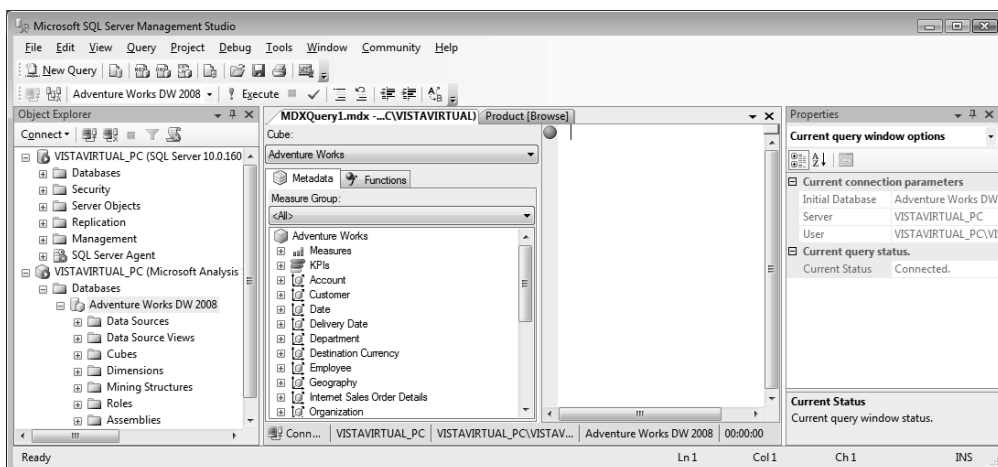
Pravé svislé okno můžeme přepínat záložkami do módu Solution Explorer nebo Template Explorer. Pomocí Solution Exploreru můžeme seskupovat skripty v databázových jazycích do projektů.

Dotazy v prostředí SQL Server Management Studio

Pomocí tlačítka New Query aktivujeme dotazovací mód Management Studia. Kromě klasických SQL a T-SQL dotazů můžeme pomocí tlačítek na toolbaru přepnout Management Studio do jiných dotazovacích režimů. Pro dotazování do BI struktur slouží:

- ◆ MDX Query
- ◆ DMX Query
- ◆ XMLA Query

Dotazovací režim zvolíme podle toho, zda budeme pracovat s relačními, multidimenzionálními, nebo data miningovými strukturami.



Obrázek 2.15. Typické rozvržení oken nástroje SQL Server Management Studio při dotazování v jazyku MDX