

MASARYKOVA UNIVERZITA
FAKULTA INFORMATIKY



Manažerské informační systémy a jejich úloha v řízení podniku

DIPLOMOVÁ PRÁCE

Bc. Matúš Chamula

Brno, 2011

Prehlásenie

Prehlasujem, že táto diplomová práca je mojím pôvodným autorským dielom, ktoré som vypracoval samostatne. Všetky zdroje, pramene a literatúru, ktoré som pri vypracovaní používal alebo z nich čerpal, v práci riadne citujem s uvedením úplného odkazu na príslušný zdroj.

Vedúci práce: prof. RNDr. Jiří Hřebíček, CSc.

Pod'akovanie

Na tomto mieste by som rád pod'akoval prof. RNDr. Jiřímu Hřebíčkoví, CSc., za odborné vedenie diplomovej práce a pripomienky k danej téme. Zvláštne pod'akovanie patrí predovšetkým Kurtovi Millerovi za neoceniteľnú spoluprácu pri vývoji vytvoreného systému. Rovnako chcem vysloviť uznanie svojej mame za jej podporu a trpezlivosť a takisto aj priateľom, ktorí mi pri vypracovaní tejto práce pomohli.

Zhrnutie

Podnikanie sa v súčasnosti nezaobíde bez efektívneho rozhodovania a kvalitného vedenia zo strany manažmentu. Úlohu manažmentu v danej sfére značne uľahčuje nasadenie manažérskych informačných systémov, za pomoci ktorých je možné získavať kľúčové informácie potrebné pre strategické a obchodné rozhodovanie. Cieľom tejto práce je dôkladne preskúmať manažérske informačné systémy, ich prínosy ako aj možnosti využitia pri riadení podniku.

Teoretická časť je zameraná na architektúru manažérskych informačných systémov a rovnako sa venuje dostupným metódam v oblasti ich tvorby. Úlohou praktickej časti je takýto systém navrhnuť a implementovať ho vo vybranom podniku využitím patričných business intelligence nástrojov.

Kľúčové slová

Manažérsky informačný systém, MIS, business intelligence, BI, transformácia dát, ETL, dátový sklad, dátový trh, OLAP, reporting, Pentaho

OBSAH

1	Úvod.....	1
1.1	Využitie manažérskych informačných systémov	1
1.2	Štruktúra diplomovej práce	2
2	Podnikové informačné systémy	3
2.1	Architektúra informačných systémov v podniku	3
2.2	Strategické riadenie podniku	4
2.3	Taktické riadenie podniku.....	4
2.4	Operatívne riadenie podniku.....	4
2.4.1	Supply Chain Management	5
2.4.2	Customer Relationship Management	5
2.4.3	Enterprise Resource Planning	5
3	Manažérske informačné systémy a business intelligence.....	7
3.1	Vývoj manažérskych informačných systémov	7
3.2	Využitie systémov podpory rozhodovania	8
3.3	Business intelligence.....	8
3.4	Marketing a business intelligence	9
3.5	Corporate Performance Management	9
3.6	Architektúra business intelligence.....	10
3.7	Dátové kategórie procesu rozhodovania.....	12
3.8	Interná spolupráca podnikových systémov	12
3.9	Transakčné a analytické systémy	13
3.9.1	OLTP systémy.....	13
3.9.2	OLAP systémy.....	14
3.10	Analytické spracovanie dát	14
4	Transformácia dát a dátové sklady.....	16
4.1	Kvalita dát.....	16
4.2	Transformačný proces	17
4.2.1	ETL/ELT	17
4.2.2	Extrakcia.....	17
4.2.3	Transformácia	18
4.2.4	Import.....	19
4.3	Dátové sklady.....	19

4.4	Architektúra dátového skladu.....	20
4.4.1	Dátový sklad a dátové trhy	20
4.4.2	Centralizovaný dátový sklad	21
4.4.3	Porovnanie architektúr dátových skladov	21
4.5	Dimenzionálny dátový model.....	22
4.5.1	Schéma hviezdy	23
4.5.2	Schéma snehovej vločky	23
5	Analýza OLAP a dolovanie dát.....	25
5.1	OLAP analýza.....	25
5.2	Dátová kocka OLAP.....	25
5.3	Operácie OLAP analýzy	27
5.4	OLAP Server.....	28
5.4.1	XMLA.....	28
5.4.2	MDX	28
5.5	Architektúra OLAP serveru	29
5.5.1	MOLAP	29
5.5.2	ROLAP	30
5.5.3	HOLAP	30
5.5.4	DOLAP.....	31
5.6	Dolovanie dát	31
6	Reporting.....	32
6.1	Tvorba reportov.....	32
6.2	Účel reportov a výsledkov analýzy.....	33
7	komplexné BI riešenia.....	34
7.1	Open source BI.....	34
7.1.1	BIRT	35
7.1.2	Jaspersoft.....	36
7.1.3	Palo for Excel a Palo Suite	37
7.1.4	Pentaho BI Suite	37
7.1.5	SpagoBI	39
7.2	Komerčné BI riešenia.....	40
7.2.1	Microsoft BI.....	40
7.2.2	IBM Cognos.....	41
7.2.3	SAP BI.....	41
7.3	Postavenie open source BI v rámci trhu	42

8	Navrhovaný systém.....	44
8.1	Predstavenie spoločnosti	44
8.2	Špecifikácia zdrojového systému.....	45
8.2.1	Dátové entity.....	45
8.2.2	Užívateľské roly.....	45
8.3	Systémové požiadavky	45
8.4	Kritériá voľby BI balíka.....	46
8.5	Realizácia navrhovaného systému.....	47
8.5.1	ETL proces	47
8.5.2	Dátový sklad.....	47
8.5.3	OLAP kocka.....	48
8.5.4	Fyzická vrstva systému	49
8.5.5	Ukážka systému.....	49
9	Záver.....	51
10	Bibliografia.....	53
	Príloha A.....	1
	Príloha B.....	2
	Príloha C	4
	Príloha D.....	6

1 ÚVOD

Podnikateľské prostredie sa príchodom nových trendov a technológií neustále transformuje. Aktuálne prebiehajúce zmeny v podnikaní a obchodnej sfére spôsobili dva primárne faktory. Prvým je globalizácia, chápaná ako neustále rastúca vzájomná závislosť medzi ekonomikami rôznych krajín sveta. Jej konečným dôsledkom je rast ekonomiky globálnej. Druhým faktorom je rozvoj informačných technológií a internetu, ktoré umožnili vysokorýchlostne prenášať obrovské množstvá dát. Takýto dopad na ekonomické vzťahy viedol k tomu, že na trh a konkurenciu na trhu je nutné prihliadať v globálnom meradle.

Kvôli takýmto zásadným zmenám v prostredí obchodovania vzrástla dôležitosť získavania a výmeny informácií. V podnikoch začali rapídne stúpať investície spojené s modernizáciou informačných technológií a nasadzovaním informačných systémov.

1.1 Využitie manažérskych informačných systémov

Trendom v posledných dekádach sa stalo využívanie manažérskych informačných systémov, ktoré umožňujú zhromažďovať kvantá dát a následne ich transformovať na znalosti použiteľné v procese rozhodovania. Podniky sú vďaka manažérskym informačným systémom schopné vytvoriť účinnú obchodnú stratégiu a pružnejšie reagovať na dianie ekonomického trhu.

Spomínaný proces transformácie dát a informácií na znalosti je základom nástrojov *business intelligence*, ktoré sú charakterizované ako systémové celky na podporu rozhodovania. Na základe výsledkov získaných využitím týchto nástrojov môžu manažéri realizovať rozhodovacie kroky pohotovejšie a efektívnejšie. Manažérske informačné systémy ako súčasť komplexných business intelligence riešení s nimi úzko súvisia.

1.2 Štruktúra diplomovej práce

Práca je zameraná na architektúru manažérskych informačných systémov a ich využitie v praxi. Zahŕňa podrobný popis jednotlivých súčastí manažérskych informačných systémov a nástrojov business intelligence, spolu s princípmi ich fungovania. Práca zároveň predstavuje techniky a spôsoby využitia týchto nástrojov ako systémov podpory rozhodovania.

Druhá kapitola sa zaoberá informačnými systémami podnikovej sféry. V teoretickej rovine popisuje základnú štruktúru podnikových informačných systémov a podrobnejšie sa venuje jednotlivým typom informačných systémov v podniku. Samotné manažérske informačné systémy sú rozoberané v tretej kapitole. Na úvod je predstavená história týchto systémov a jednotlivé možnosti ich využitia. Súčasťou kapitoly je takisto popis štruktúry a detailnejšia analýza manažérskych informačných systémov a nástrojov business intelligence.

Dáta využívané v procesoch rozhodovania sú rozličného charakteru a nejednotnej štruktúry. Pre potreby business intelligence je nutné tieto dáta transformovať do požadovanej formy, zjednotiť ich štruktúru a umiestniť ich do dátového skladu v podobe dimenzionálneho modelu. Rozborom príslušných transformačných procesov a architektúr dátových skladov sa zaoberá štvrtá kapitola. Po etape spracovania a uchovania dát nasleduje ich štatistika a analýza. Piata kapitola interpretuje multidimenzionálnu dátovú kocku ako vhodný analytický dátový model, ktorý umožňuje so spracovávanými dátami efektívne manipulovať pre potreby analýzy. Kapitola zároveň predstavuje OLAP server, jednotlivé typy jeho architektúr a konkrétne operácie tvoriace jadro OLAP analýzy. Obsahom šiestej kapitoly je popis fázy reportingu analyzovaných dát spolu s účelom jeho využitia.

Záverečné kapitoly sa venujú praktickej časti práce a sú zamerané na využitie nástrojov business intelligence v praxi. Okrem všeobecného prehľadu dostupných open source a komerčných business intelligence riešení obsahujú aj popis použitého balíka nástrojov spoločnosti Pentaho a detailne sa venujú architektúram a postupom realizovaným pri tvorbe výsledného manažérského informačného systému.

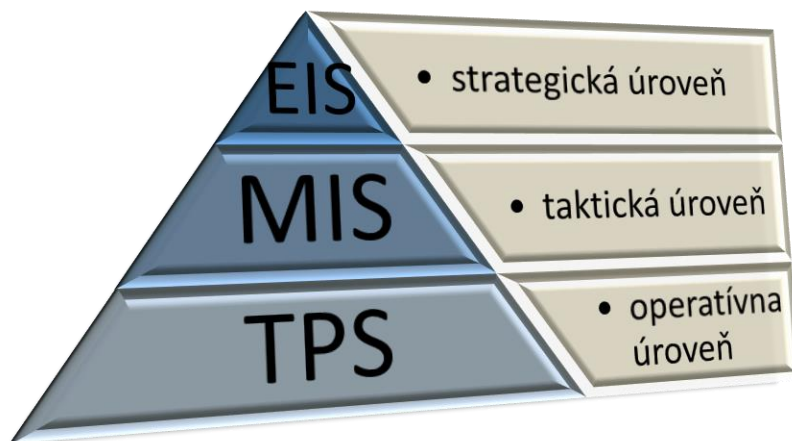
2 PODNIKOVÉ INFORMAČNÉ SYSTÉMY

Objem informácií a súvisiacich dát vo svete neustále rastie a rovnako dochádza k zvyšovaniu komplexnosti dostupných technológií. Tieto faktory majú za následok akvizíciu využívania informačných systémov v podnikoch. S ich pomocou sú podniky schopné maximalizovať mieru využitia a spracovania informácií. Kvalitné informačné systémy sú tak najmä pre väčšie podniky nevyhnutnosťou.

Podnikový informačný systém predstavuje softwarové riešenie na spracovanie dát a riadenie procesov. Jeho funkciou je zabezpečiť komunikáciu medzi jednotlivými procesmi podniku a poskytovať čo najkvalitnejšie a najaktuálnejšie informácie [1].

2.1 Architektúra informačných systémov v podniku

Každý typ podnikového informačného systému slúži konkrétnej funkčnej oblasti, pre ktorú je navrhnutý a pre ktorú spĺňa určité kritériá. Jednotlivé skupiny týchto informačných systémov sú priamo určené adekvátnym užívateľským skupinám. Primárne typy podnikových IS ilustruje obrázok 2.1. Pyramídová hierarchia rovnako znázorňuje úroveň postavenia špecifických skupín informačných systémov v rámci organizačnej štruktúry podniku. Štruktúra podnikového subjektu sa obvykle rozdeľuje na tri základné úrovne – *strategickú, taktickú a operatívnu* [2].



Obrázok 2.1: Architektúra podnikových informačných systémov [2]

2.2 Strategické riadenie podniku

Na najvyššie postavenej strategickej úrovni riadenia podniku vystupujú EIS (*Executive Information Systems*) pozostávajúce z aplikácií a nástrojov na podporu vedenia podnikov a inštitúcií. Sú koncipované priamo pre potreby vrcholového manažmentu, ktorý sa zaoberá tvorbou dlhodobého operačného plánu, koordináciou a monitorovaním obratu či výsledkov hospodárenia podniku. EIS ako vstupný zdroj dát využívajú informačné systémy nižších úrovní.

2.3 Taktické riadenie podniku

Poslaním stredného manažmentu je riadiť podnik po taktickej stránke, kde je hlavnou úlohou spravovanie rozpočtu a plánovanie produkcie. Pre taktickú úroveň riadenia podniku je príznačné využívanie MIS (*Management Information Systems*), ktorých relevantnú podmnožinu tvoria systémy podpory rozhodovania DSS (*Decision Support Systems*).

Systémy typu MIS vychádzajú z ekonomických a účtovných systémov. Poskytujú analýzy obchodných záležitostí týkajúcich sa podniku z rôznych uhlov pohľadu, pričom analyzujú určitý objem dát a pomocou rozhodovacích metód poskytujú manažérovi možnosti riešenia, prípadne usmerňujú ich postup.

2.4 Operatívne riadenie podniku

Predmetom operatívneho riadenia podniku je spracovávanie objednávok, realizácia produkcie a vedenie evidencie o samotných zamestnancoch. Užívatelia na tento účel typicky využívajú systémy transakčného charakteru TPS (*Transaction Processing Systems*).

Systémy operatívnej úrovne integrujú interné obchodné procesy podniku do samostatného celku a slúžia na mechanizáciu agendových úloh ako sú fakturácie, inventarizácie a mzdy [3]. Pod TPS spadá viacero systémov, z ktorých každý pokrýva špecifickú oblasť bežných a rutinných podnikových procesov.

2.4.1 Supply Chain Management

Jedným z dôležitých organizačných procesov je riadenie a spravovanie kompletného reťazca dodávateľov, o ktoré sa starajú SCM (*Supply Chain Management*) systémy. Riadenie dodávateľských subjektov príslušným systémom prináša na jednej strane optimalizáciu efektivity spracovania jednotlivých objednávok a na strane druhej zvyšuje spokojnosť zákazníka.

SCM systémy pomáhajú riadiť a koordinovať aktivity spojené s výrobou, nákupom a distribúciou produktu. Spravujú všetky náležitosti spojené s úvodnými fázami výroby až po dodanie produktu koncovému spotrebiteľovi [4]. Využitím systémov SCM je možné efektívne riadiť dodávky materiálov a služieb, redukovať tak potrebné náklady a tým doceliť pokles výslednej ceny produktu.

2.4.2 Customer Relationship Management

Do skupiny operatívnych cieľov podnikateľských subjektov spadá takisto zlepšovanie spolupráce so zákazníckymi entitami. Primárne sa jedná o získavanie nových zákazníkov a efektívnu komunikáciu s existujúcimi zákazníkmi. Cieľom týchto aktivít je porozumenie požiadaviek zákazníka a predvídanie jeho potrieb, čo prináša celkové skvalitnenie poskytovaných služieb.

Administráciou prvkov daného odvetvia sa zaoberajú systémy typu CRM (*Customer Relationship Management*). Úlohou systémov riadenia vzťahov so zákazníkmi je zhromažďovať a spracovávať detailné informácie o zákazníkoch, ako aj archivovať podrobnosti ich komunikácie s podnikom.

2.4.3 Enterprise Resource Planning

Základným predpokladom úspešného podnikania je finančný profit a jeho maximalizácia. Tento cieľ je najčastejším prvkom obchodnej stratégie každého podniku. K nadobúdaniu zisku je v prvom rade nutné, aby podnik efektívne narábal so zdrojmi, ktoré má k dispozícii, a zvyšoval tým úspory v investíciách. Riešenie tejto problematiky pokrývajú ERP (*Enterprise Resource Planning*) systémy.

ERP systémy sú orientované na plánovanie a riadenie rôznorodých podnikových procesov, akými sú produkcia, predaj, správa financií či logistika [4]. Implementácia týchto systémov v konečnom dôsledku prináša zefektívnenie priebehu ekonomických procesov a prípadné urýchlenie celkového rastu podniku.

3 MANAŽÉRSKE INFORMAČNÉ SYSTÉMY A BUSINESS INTELLIGENCE

Mohutné podniky a korporácie sa v súčasnosti nezaobídu bez moderných technológií podporujúcich strategické plánovanie, medzi ktoré patria práve manažérske informačné systémy. Tieto systémy sú primárne založené na informačných technológiách, rôznych štatistických postupoch a analýzach. Dokážu tak poskytnúť manažérovi potrebné informácie o dianí v podniku, analyzovať jednotlivé podnikové procesy a náležitou formou výsledky týchto analýz prezentovať.

3.1 Vývoj manažérskych informačných systémov

Počítače a informačné technológie prišli na scénu v štyridsiatych rokoch 20. storočia a spočiatku ich podniky využívali prevažne na spracovávanie dát a automatizáciu triviálnych úloh. S neskorším vývojom trhu a ekonomiky sa však objavovali nové požiadavky na softwarové technológie podpory rozhodovania a súčasne s nimi sa prvýkrát objavil pojem *manažérsky informačný systém*. Pokusy takýto systém vyvinúť však stroskotali na úrovni vtedajších informačných technológií, ktoré vďaka absencii požadovaného grafického výstupu a nedostatočnej úrovni interakcie s užívateľom neboli schopné prevádzky komplexných systémov.

S príchodom osobných počítačov v osemdesiatych rokoch, ktoré disponovali operačným systémom a pokročilejšími možnosťami využitia, sa však situácia značne zmenila. Podniky boli vďaka nim schopné vyvíjať vlastné aplikačné a databázové riešenia. Začali tak vznikať napríklad tabuľkové editory, ktoré sa považujú za predchodcov dnešných business intelligence architektúr [5]. Zároveň s nimi sa v značnej miere rozvíjali aj samotné informačné systémy a ich využívanie v

podnikovej sfére rapídne rástlo. Pokrok v tejto oblasti bol spečatený nástupom prvých systémov podpory rozhodovania, primárne určených pre stredný a vrcholový manažment.

3.2 Využitie systémov podpory rozhodovania

Činnosť podnikov ako ekonomických subjektov vystupujúcich na trhu sa riadi procesom strategického plánovania. Základom tohto procesu je stanovenie primárnych cieľov podniku, medzi ktoré v prvom rade patrí maximalizácia zisku a ekonomického rastu. Primárnym cieľom podniku pritom priamo zodpovedá štýl rozhodovania manažmentu.

Manažérske informačné systémy ako systémy podpory rozhodovania sa uplatňujú pri tvorbe podnikovej stratégie a dosahovaní podnikových cieľov. Pomocou zostavených analýz a prognóz podporujú manažérov v strategických a obchodných rozhodnutiach. Za predpokladu adekvátneho využívania dokážu byť tieto systémy prínosom v oblasti organizačnej efektivity a poskytovania informácií, ktoré následne pomáhajú manažérom konať korektne v súlade s danou situáciou či stanovenými cieľmi podniku.

3.3 Business intelligence

Termín *business intelligence* možno chápať ako samostatné odvetvie či disciplínu vo sfére podnikových informačných systémov, ktorého cieľom je získavanie obchodne dôležitých informácií za pomoci analytických nástrojov. Podľa [2] sa všeobecne za business intelligence považuje súbor procesov, aplikácií a technológií, ktorých cieľom je účinne a účelne podporovať rozhodovacie procesy v podniku. Podporujú analytické a plánovacie činnosti podnikov či organizácií a sú postavené na špecifických OLAP technológiách a ich modifikáciách, teda na princípoch dimenzionálnych pohľadov na podnikové dáta.

Manažérske informačné systémy, ktoré sa radia medzi systémy spracovávajúce obrovské množstvo dát a ktoré zároveň poskytujú rôznorodé aplikácie a nástroje spracovávajúce tieto dáta, poskytujú údaje potrebné pre obchodné rozhodovanie manažmentu podnikov a spadajú tak priamo do sféry business intelligence.

Princíp business intelligence je založený na čerpaní *know-how* zo všetkých dostupných dát, ktoré má podnik k dispozícii. Tieto dáta, získané primárne extrakciou z podnikových informačných systémov prípadne externých zdrojov, sa pomocou business intelligence nástrojov transformujú na informácie a znalosti, z ktorých je následne možné tvoriť podklady pre rozhodovacie kroky týkajúce sa obchodných plánov a stratégie podniku. Analýzou dostupných dát je rovnako možné

monitorovať javy a fakty týkajúce sa podniku a podrobiť tak výkon a efektivitu jednotlivých podnikových procesov podrobnému prieskumu. Business intelligence využíva na tento účel súbor matematických modelov a analytických metód, ktorých úlohou je systematicky zužitkovávať dostupné dáta a pomocou štatistických postupov z nich extrahovať patričné znalosti využiteľné v procese rozhodovania [5].

3.4 Marketing a business intelligence

Business intelligence vďaka analýze dát, vzťahujúcich sa k zákazníkovi, úzko súvisí s marketingom. Výsledky business intelligence analýz sú pre podnik a jeho marketingové oddelenie ideálnou referenciou. Vďaka nim možno porozumieť zákazníckym potrebám, monitorovať dopyt po jednotlivých produktoch a tvoriť efektívnu marketingovú kampaň.

Nástroje business intelligence však poskytujú manažérom podporu aj v širšej sfére marketingového rozvoja. Typickými požiadavkami stredného a vrcholového manažmentu sú odpovede na otázky typu:

- Má pri zvyšovaní investičných nákladov vzniknúť nová pobočka vo vybranom regióne alebo má byť rozšírená súčasná pobočka?
- Ako vyzerá vývoj nákupu konkrétnej skupiny tovaru za určité obdobie a aké množstvo nadobudnutého tovaru bolo následne efektívne využité?
- Počas ktorých kvartálov kalendárneho obdobia bola najvyššia miera predaja konkrétnych produktov a aká skupina zákazníkov nakupuje tieto produkty najviac?
- Do akej miery sú zákazníci spokojní s poskytovanými službami a ako sa čo najefektívnejšie prispôbiť vývojovým trendom spojenými s požiadavkami zákazníkov?

Analýzy vyhotovené pomocou business intelligence systémov dokážu poskytnúť indície k podobným otázkam.

3.5 Corporate Performance Management

S business intelligence úzko súvisí aj problematika *Corporate Performance Management* (CPM) popisujúca celopodnikové riadenie výkonnosti. Spoločnosť Gartner definuje CPM ako súhrnný pojem, ktorý popisuje metodiky, metriky, podnikové procesy a systémy, ktoré sa používajú pre monitorovanie a riadenie výkonnosti celého podniku [6].

Metodika CPM ako nadstavba business intelligence je rovnako určená pre vrcholový manažment, no obidva prístupy oddeľuje viacero odlišností. Zatiaľ čo business intelligence rozvíja u podnikového manažmentu schopnosť v danej situácii adekvátne reagovať, CPM sa zameriava na proaktívny prístup a riadenie business procesov organizácie. Namiesto spracovania a analýzy historických dát je predmetom CPM riadenie smerovania podniku ako celku a manažment za pomoci CPM nástrojov stanovuje smerovanie podniku do budúcnosti a monitoruje podnikové ciele a vízie.

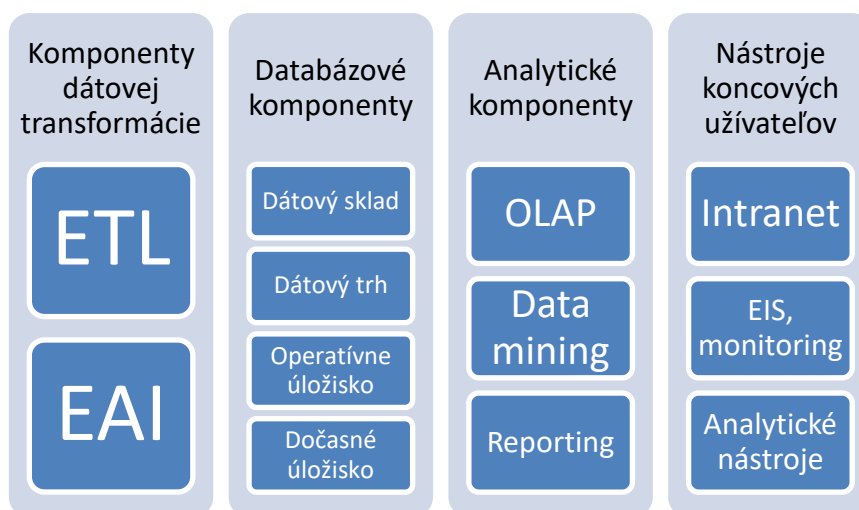
Okrem iného je pomocou jednotlivých služieb CPM možné monitorovať výkonnostné aspekty nasledovných odvetví podniku:

- Riadenie rozpočtu a plánovanie, s pomocou ktorého je možné určiť primárne ciele podniku a k nim prislúchajúce priority. Ich vývoj je priebežne monitorovaný v čase a na základe výsledkov tohto procesu manažment plánuje kroky do budúcnosti.
- Strategické riadenie podniku zaoberajúce sa podnikovými víziami. Jeho jadro spočíva v dekompozícii strategického plánu podniku na samostatné celky, ktoré je možné následne analyzovať. K tomuto účelu sa obvykle implementujú *balanced scoreboards*, pomocou ktorých je zvolenými metodikami možné určiť tzv. kľúčové indikátory podniku (*Key Performance Indicators*), analyzovať ich z viacerých perspektív a na základe výsledkov týchto analýz následne vypracovať čo najefektívnejšiu stratégiu.
- Konsolidáciu predstavujúcu globálne spravovanie podnikových dát týkajúcich sa financií a umožňujúcu efektívne vykazovať finančnú situáciu podniku a jeho jednotlivých celkov. Výstupom konsolidačného procesu je súhrn finančných informácií pre investorov a interne využiteľné finančné správy.

So zameraním CPM úzko súvisí problematika BPM (*Business Process Management*), nakoľko sa oba prístupy snažia globálne maximalizovať výkonnosť podniku a efektivitu jeho interných business procesov.

3.6 Architektúra business intelligence

Business intelligence ako súbor nástrojov analýzy dát a podpory rozhodovania pozostáva z viacerých interakčných komponentov. Obrázok 3.1 ilustruje architektúru typického business intelligence systému a štruktúru spolupráce jeho jednotlivých súčastí.



Obrázok 3.1: Diagram business intelligence [7]

- Dáta získavané z interných podnikových systémov či externých zdrojov je nutné pre potreby business intelligence transformovať do patričnej podoby. Prvú časť business intelligence procesu tak tvoria aplikačné nástroje získavania a transformácie dát *ETL* a *EAI*. Tieto nástroje umožňujú vstupné dáta extrahovať, transformovať do požadovanej formy a následne ich presunúť do dátového skladu (viď podkapitola 4.2).
- *Dátový sklad* spolu s *dátovými trhmi* predstavujú primárne úložisko dát a možno ich charakterizovať ako databázu, ktorá priamo vyhovuje potrebám business intelligence. Dáta sú v dátovom sklade uložené vo forme dimenzionálneho modelu, ktorého štruktúra bola navrhnutá pre analytické procesy business intelligence. Obidve úložiská zvykne navyše dopĺňať operatívne a dočasné úložisko dát (viď podkapitola 4.3).
- Nástroje uskutočňujúce nad dátami samotné analýzy tvoria ďalší separátny celok business intelligence. Zaradujú sa medzi ne *OLAP analýza* a proces *dolovania dát*. *OLAP analýza* umožňuje užívateľom v reálnom čase sledovať rozbor dostupných dát na rôznych stupňoch agregácie podľa zvolených kritérií (dimenzií), zatiaľ čo komplexnejšie analýzy na predikciu javov a tvorbu konkrétnych vzorov rieši sféra dolovania dát (viď podkapitola 5.1 a podkapitola 5.6). *Reporting* predstavuje finálnu fázu činnosti analytických komponent v business intelligence procese. Jeho úlohou je poskytovať užívateľom výsledky analýz v čo najzrozumiteľnejšej forme. Primárne ide o grafy, tabuľky a ďalšie prostriedky, ktoré dokážu efektívne a zrozumiteľne oboznámiť užívateľa s detailmi výsledkov jednotlivých analýz (viď kapitola 6).

3.7 Dátové kategórie procesu rozhodovania

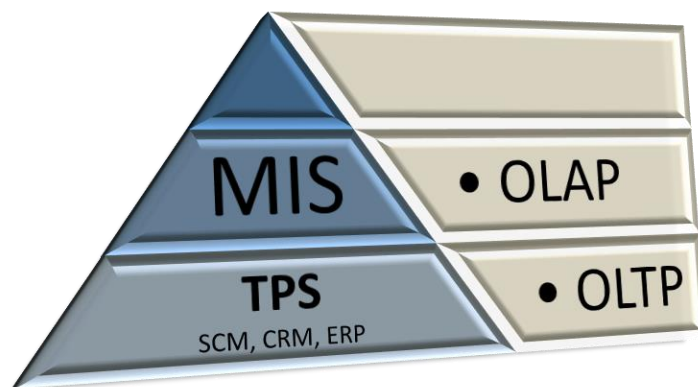
Údaje, nad ktorými sa v procesoch rozhodovania vykonávajú analýzy a ktoré rovnakou mierou ovplyvňujú výsledné prognózy, pochádzajú z heterogénnych zdrojov, ako sú podnikové informačné systémy operatívnej úrovne a špecifické externé systémy. V procese rozhodovania sa tým pádom narába s rozličnými typmi dát, ktoré spadajú do troch základných kategórií [5]:

- **Interné dáta** – ako zdroj interných dát slúžia najmä transakčné systémy operatívnej podnikovej úrovne, ktoré ich uchovávajú vo vlastných databázach tvoriacich jadro TPS. Tieto dáta nesú detailné informácie o podniku najmä z oblasti administratívy, účtovníctva, výroby a logistiky. Interné dáta dokážu poskytnúť rôznorodé fakty o každodennom chode podniku, ktoré sú následne využiteľné v procese rozhodovania. Ide napríklad o záznamy týkajúce sa objednávok, numerické údaje, ktoré sú súčasťou faktúr alebo výsledky interných činností podniku a marketingových kampaní.
- **Externé dáta** – zdrojom externých informácií bývajú najmä nezávislé spoločnosti monitorujúce dianie ekonomického trhu, ktoré na základe určených indikátorov dokážu odhadnúť trendy a predpovedať vývoj trhu pre špecifické odvetvia podnikania. Externé dáta však podnik môže získavať aj vlastným prieskumom trhu, prípadne zisťovaním spotrebiteľskej mienky formou ankiet a dotazníkov.
- **Osobné dáta** – v mnohých prípadoch vykonávajú manažéri rozhodovacie kroky na základe vlastných informácií a odhadov nadobudnutých skúsenosťami, ktoré sa zaraďujú medzi osobné dáta rozhodovacieho procesu.

3.8 Interná spolupráca podnikových systémov

Manažérske informačné systémy ako súčasť business intelligence pracujú predovšetkým s dátami podnikových informačných systémov operatívnej úrovne. Transakčné systémy obsahujú množstvo informácií využiteľných v procese rozhodovania a pre MIS predstavujú primárny zdroj dát, čo naznačuje obrázok 3.2.

Dáta extrahované z viacerých typov transakčných systémov (SCM, CRM, ERP) sú heterogénneho charakteru. Analýza takýchto dát by bola značne komplikovaná, preto je potrebné dáta transformovať. Samotná extrakcia a analýza dát manažérskymi informačnými systémami pritom prebieha bez akejkoľvek záťaže systémov transakčného typu.



Obrázok 3.2: Podnikové systémy taktickej a operatívnej úrovne

Schéma spolupráce informačných systémov taktickej a operatívnej úrovne zároveň naznačuje ich vzájomné diferencie. TPS spadajú do skupiny OLTP (*Online Transaction Processing*) systémov, ktoré narábajú s dynamickými dátami a sú primárne optimalizované na vysoký výkon. MIS sa zaraďujú medzi systémy typu OLAP (*Online Analytical Processing*), ktoré spracovávajú jednotlivé dáta analyticky a sú optimalizované na vykonávanie analýz a vysokú flexibilitu.

3.9 Transakčné a analytické systémy

Všeobecne sú transakčné systémy operatívnej úrovne charakterizované ako systémy pre radových užívateľov, zatiaľ čo manažérske informačné systémy ako súčasť business intelligence riešení slúžia na tvorbu analýz a podporu rozhodovania, pričom ich využívajú výhradne manažéri a analytici. Najzásadnejším rozdielom medzi oboma typmi systémov je však odlišný spôsob spracovania dát.

3.9.1 OLTP systémy

Transakčné typy podnikových systémov sú postavené nad štandardnou relačnou databázou a možno ich prirovnať ku klasickému databázovému riešeniu, v ktorom obyčajne prebieha veľké množstvo transakcií. Hlavný dôraz sa kladie na okamžité spracovanie transakcií a krátku dobu odozvy jednotlivých dotazov. OLTP systémy musia zaisťovať dátovú integritu a zaručovať aktuálnosť uchovávaných dát.

3.9.2 OLAP systémy

Manažérske informačné systémy sú členené do skupiny systémov typu OLAP, u ktorých jadro funkcionality spočíva v analytickom spracovávaní dát. OLAP systémy centralizujú jednotlivé dáta do dátového skladu, ktorý na ich uchovávanie využíva dimenzionálny dátový model. Dáta tak môžu byť pre potreby efektívnej analýzy reprezentované multidimenzionálnou dátovou kockou (viď podkapitola 5.2).

OLAP systémy sa vyznačujú relatívne malým počtom transakcií – dátový sklad je oproti relačnej databázy určený výhradne na čítanie zhromaždených dát za účelom ich následného rozboru. Eventuálne dotazy nad dátami sú však komplexnejšie, nakoľko je pre procesy rozhodovania nutné zostavovať detailné analýzy. Viac ako samotný výkon a doba odozvy je u týchto systémov podstatnejšia efektivita spracovania dát, od ktorej sa odvíja kvalita výsledných analýz a prognóz.

3.10 Analytické spracovanie dát

Počas tvorby výsledných reportov manažérskymi informačnými systémami je pre potreby procesu rozhodovania nevyhnutné spracovávať uchovávané dáta analyticky. Transakčné OLTP systémy kvôli svojim vlastnostiam nie sú vhodné na realizáciu analýz nad uchovávanými dátami a oproti OLAP systémom majú v spôsobe spracovávania dát viacero nevýhod.

- Veľké podniky obyčajne prevádzkujú niekoľko informačných systémov, z ktorých každý disponuje separátnou databázou. Pri tvorbe komplexných analytických dotazov nad jednotlivými relačnými databázami by tak bolo nutné spájať väčšie množstvo tabuliek.
- Relačný databázový model je navrhnutý na uchovávanie dát, efektívny priebeh transakcií a na účely analýzy je nevhodný. Dimenzionálny model dátového skladu je adaptáciou relačného modelu s dimenziami a faktami a predstavuje dátovú štruktúru špeciálne prispôbenú analytickým dotazovacím technikám.
- Manažéri pri tvorbe a realizácii strategických krokov potrebujú sledovať trendy za dlhšie časové obdobie a často vychádzajú z faktov minulosti, napríklad z analýz predaja skupiny výrobkov v danej oblasti za posledných 5 rokov. Dáta uložené v databázach transakčných systémov sa však po stanovenej dobe odstraňujú, prípadne zálohujú na externé médium a neboli by tak v prípade potreby priamo dostupné. Dátový sklad je prispôbený na uchovávanie historických dát, ktoré sú súčasťou manažérskych informačných systémov.
- Ďalšou nevyhnutnou požiadavkou systémov analytického typu je uchovávanie dát v agregovanej forme, čo relačný model OLTP systémov nespĺňa. Pri potrebe manipulovať

s úrovňou detailu uchovávaných dát je dátový sklad schopný poskytovať dáta s rôznym stupňom granularity.

Rozdiely medzi transakčnými a analytickými systémami existujú vo viacerých podstatných aspektoch, no vyplývajú najmä z ich odlišného využitia. Celkový súhrn vlastností oboch typov systémov je zhrnutý v nasledujúcej tabuľke.

	OLTP	OLAP
Dáta	Dynamické, výhradne aktuálne	Statické, aktuálne a historické
Granularita	Detailné dáta	Agregované a konsolidované dáta
Aktualizácia	Nepretržitá	Periodická
Operácie	Jednoduché dotazy a transakcie	Komplexné dotazy
Doba odozvy	Krátka, rádovo v sekundách na dotaz	Dlhšia kvôli komplexným dotazom
Požiadavky	Vysoký výkon, kvalitné spracovanie transakcií	Flexibilita, efektivita odozvy
Zameranie	Uchovávanie potrebných dát	Získavanie znalostí a poskytovanie analýz
Užívatelia	Rádoví zamestnanci	Manažéri a analytici
Veľkosť DB	MB až GB	GB až TB

Tabuľka 3.1: Porovnanie systémov OLTP a OLAP [5]

4 TRANSFORMÁCIA DÁT A DÁTOVÉ SKLADY

Jednotlivé dáta prúdiace zo zdrojových systémov sú heterogénneho charakteru. Je preto potrebné ich transformovať do patričnej podoby tak, aby s nimi dátový sklad dokázal narábať korektne. Na tento účel slúži proces dátovej transformácie zaisťujúci kvalitu dát pred ich naplnením do dátového skladu.

4.1 Kvalita dát

Zaistenie príslušnej úrovne kvality dát predstavuje dôležitú úlohu. Výsledky zostavených analýz budú presné jedine v prípade, že dátový sklad uchováva čo najdokonalejšie dáta a od úrovne kvality dát sa tak odvíja úvodná fáza business intelligence procesu. Medzi primárne faktory ovplyvňujúce úroveň kvality uchovávaných dát patria najmä [5]:

- **Presnosť a správnosť** – za účelom získavania čo najexaktnejších výsledkov musia byť dáta náležite presné a pravdivé. Je preto nevyhnutné overiť, či sú hodnoty jednotlivých entít spracovávaných analýzou reprezentované korektne.
- **Kompletnosť** – dáta by nemali obsahovať príliš mnoho prázdnych hodnôt, aby nedochádzalo k zníženiu presnosti výsledných analýz. Proces transformácie dát sa snaží túto nežiaducu vlastnosť redukovať v čo najväčšej miere.
- **Konzistencia** – obsah a podoba výsledných dát po integračnom procese musí byť konzistentná vzhľadom na rozličnosť zdrojov, z ktorých tieto dáta pochádzajú. Dáta nadobúdajú požadovanú konzistenciu napríklad unifikáciou menových jednotiek, nakoľko sa v procese business intelligence často spracovávajú heterogénne dáta v podobe rôznorodých finančných údajov, prípadne jednotiek miery.

- **Aktuálnosť** – dáta by mali byť často aktualizované vzhľadom na frekvenciu tvorby jednotlivých analýz a potreby čo najaktuálnejších výsledkov. Aktualizácia dátového skladu obvykle prebieha pravidelne v denných alebo týždenných intervaloch.
- **Relevancia** – výsledkom zberu a vyhodnocovania dát poskytovaných dátovým sklado- musia byť relevantné informácie. Zásluhou množstva dostupných informácií sú business intelligence systémy schopné poskytnúť dostatok precíznych faktov nevyhnutných v procese rozhodovania.

4.2 Transformačný proces

Po zaručení dostatočnej úrovne kvality vstupných dát nasleduje proces ich transformácie, ktorého účelom je upraviť dáta do podoby vhodnej na ich analýzu. Samotný proces transformácie pozostáva z troch základných operácií a je označovaný skratkou ETL (*Extract, Transform, Load*). Úspešným vykonaním všetkých operácií sa dáta z dostupných zdrojov extrahujú, transformujú podľa vybraných požiadaviek a následne prenesú na cieľový dátový sklad.

4.2.1 ETL/ELT

Proces transformácie ETL predmetné dáta po extrahovaní zo zdrojových systémov transformuje na samostatnej platforme či ETL serveri a následne ich predá dátovému skladu. Priebeh transformačného procesu tak nijakým spôsobom nezaťažuje dátový sklad. Naproti tomu sa v ELT (*Extract, Load, Transform*) procese dáta presúvajú priamo na dátový sklad, ktorý sa stará o ich transformáciu. Celá transakcia sa procesne zjednoduší a nie je tak potrebná prítomnosť transformačnej entity, avšak za cenu zvýšenia záťaže dátového skladu.

Vyššiu úroveň abstrakcie v transformačnom procese predstavuje EAI (*Enterprise Application Integration*), ktorý nepracuje priamo s dátami, ale s aplikáciami a rieši tak integráciu na aplikačnej úrovni.

4.2.2 Extrakcia

Extrakcia, ako úvodná operácia ETL/ELT, má za úlohu extrahovať dáta z dostupných systémov do formátu vhodného na ich transformáciu. Miera akosti extrakcie, ako aj spôsob výberu samotných dát sa odrazí na kvalite dát výsledných. Tejto operácii je tak nutné prikladať vyššiu dôležitost. Proces extrakcia dát obvyčajne prebieha formou vzdialeného prístupu ku zdrojovým databázam za pomoci API metód ODBC/JDBC.

Všeobecne existujú dva základné typy extrakcie – iniciálna a inkrementálna (následná). Pri iniciálnej extrakcii sa do prázdneho dátového skladu importujú všetky relevantné a dostupné dáta, zatiaľ čo v prípade inkrementálnej extrakcie sa dátový sklad aktualizuje iba novo vzniknutými dátami.

4.2.3 Transformácia

Operácia transformácie sa riadi konkrétnymi predpismi a využitím rôznorodých funkcií transformuje dáta do adekvátnej podoby. Primárnym cieľom transformačného procesu je odhaliť v dátach akékoľvek nepresnosti či inkonzistencie a následne ich korigovať. Nakoľko pôvodné dáta pochádzajú z rozličných zdrojov a majú odlišnú štruktúru, musí transformácia zaručovať, že výsledné dáta budú homogénne a integrovateľné do dátového skladu.

Jednotlivé dáta sa za účelom efektívnejšieho narábania v priebehu transformačného procesu agregujú a konsolidujú. V konečnom dôsledku je tak dátový sklad schopný realizovať podrobnejšie analýzy nad dátami v kratších časových intervaloch a efektívnejšie. Medzi najčastejšie vykonávané transformačné operácie nad vybranými dátami patria [8]:

- Redukovanie určitých položiek databázy, respektíve položiek neobsahujúcich žiadne dáta (hodnoty typu null a pod.). V prípade databázy obsahujúcej stĺpce `order_no`, `customer_name` a `sum` pri monitorovaní tržieb podľa zákazníkov budú čísla jednotlivých objednávok predstavovať irelevantný údaj. Všetky položky `order_no` by boli v takom prípade ignorované a nedošlo by k ich umiestneniu do dátového skladu.
- Konvertovanie kódových hodnôt – potrebné, ak sa v zdrojových dátach používa napríklad hodnota 1 na označenie muža a hodnota 2 na označenie ženy, no dátový sklad na tento účel využíva odlišné hodnoty M a F.
- Unifikácia hodnôt označujúcich rovnaký predmet – hodnoty typu male, 1 a Mr označovať jednotne ako M.
- Odvodenie nových a zároveň potrebných hodnôt výpočtom – pri zisťovaní tržieb podľa jednotlivých produktov by išlo o položku `sale_amount = qty * unit_price`.
- Odstránenie duplicitných a redundantných dát, ktoré zvyšujú nároky na úložnú kapacitu dátového skladu.
- Filtrovanie a zoradovanie transformovaných dát.
- Rozdelenie stĺpca databázy do viacerých nezávislých stĺpcov – týmto spôsobom sa obyčajne rozdeľujú stĺpce obsahujúce niekoľko separátnych hodnôt.

4.2.4 Import

Import predstavuje finálnu fázu dátovej transformácie. Jeho úlohou je importovať dáta, ktoré prešli procesmi extrakcie a transformácie, do dátového skladu. Priebeh tohto procesu a najmä frekvencia importu sa líši v závislosti od charakteristiky stanovenej politiky systému. Dátové sklady môžu nahradzovať pôvodné dáta aktuálnymi raz za týždeň, zatiaľ čo iné pridávajú dáta k už existujúcim v denných intervaloch.

4.3 Dátové sklady

Dátový sklad predstavuje v architektúre business intelligence primárne úložisko dát. Pojem dátový sklad v sebe indikuje množinu vzájomne prepojených aktivít týkajúcich sa návrhu, implementácie a využívania takéhoto dátového úložiska. Dátový sklad fyzicky vystupuje na rozhraní medzi systémami OLTP a analytickými nástrojmi business intelligence, pričom hlavnou úlohou dátového skladu je poskytovať vstupné dáta OLAP aplikáciám.

Podľa [5] existuje viacero charakteristických prvkov, ktoré dátovému skladu predurčujú úlohu primárneho dátového úložiska business intelligence systémov:

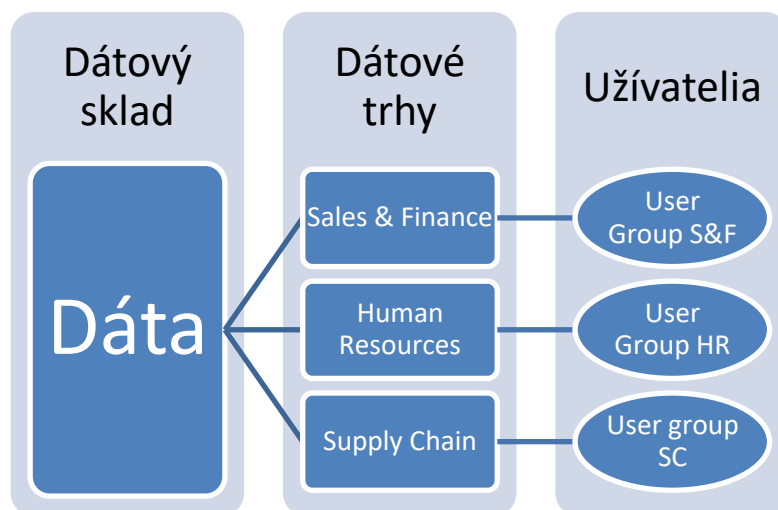
- **Integrácia** – business intelligence systémy pracujú so skupinami rôznorodých dát. Pri takejto heterogénnosti je nutná integrácia a unifikácia dát v rámci precedentného transformačného procesu prípadne na úrovni dátového skladu, aby bolo možné k výsledným informáciám pristupovať efektívnejšie a v rámci jediného dátového úložiska.
- **Kvalita** – dáta prenesené do dátového skladu sú vopred korigované. Týmto spôsobom sa predchádza akýmkoľvek chybám a nezrovnalostiam vo výsledných analýzach.
- **Efektivita** – vyhotovenie komplexných databázových dotazov s účasťou dátového skladu, ktoré sú zamerané na extrakciu informácií pre účely business intelligence, by bolo časovo neefektívne. Administrácia dotazov smerujúcich do dátového skladu je tak u business intelligence systémov prenechaná OLAP analýze.
- **Rozšíriteľnosť** – objem dát uchovávaných transakčnými systémami časom narastá a kvôli obmedzeniam týkajúcich sa dostupnej veľkosti úložiska, je po určitej dobe žiaduce dáta z OLTP systému odobrať a zálohovať na externé médium. Business intelligence systémy však potrebujú mať prístup ku všetkým dostupným dátam. Túto funkcionality zabezpečuje dátový sklad uchovávaním historických dát.

4.4 Architektúra dátového skladu

Súčasnité architektúry dátových skladov vychádzajú z dvoch pôvodných návrhov, ktoré priniesli Bill Inmon a Ralph Kimball, pričom obidva návrhy sú navzájom protichodné. Zatiaľ čo Kimball predstavoval dátový sklad ako zhľuk nezávislých dátových trhov, Inmon presadzoval centralizovanú a normalizovanú architektúru.

4.4.1 Dátový sklad a dátové trhy

Podľa Kimballovej architektúry je dátový sklad zložený z viacerých separovaných dátových trhov. Dátový trh možno definovať ako separátne dátové sklad obsahujúci špecifickú podmnožinu dát určenú obmedzenému okruhu užívateľov. Každý z dátových trhov obsahuje výhradne dáta prislúchajúce k určitej oblasti podniku (napr. oddelenie predaja alebo dáta o zamestnancoch). Jednotlivé dátové trhy sú budované postupne a môžu k nim pristupovať jedine adekvátni užívatelia. Architektúru dátového skladu ako zjednotenia dátových trhov ilustruje obrázok 4.1.



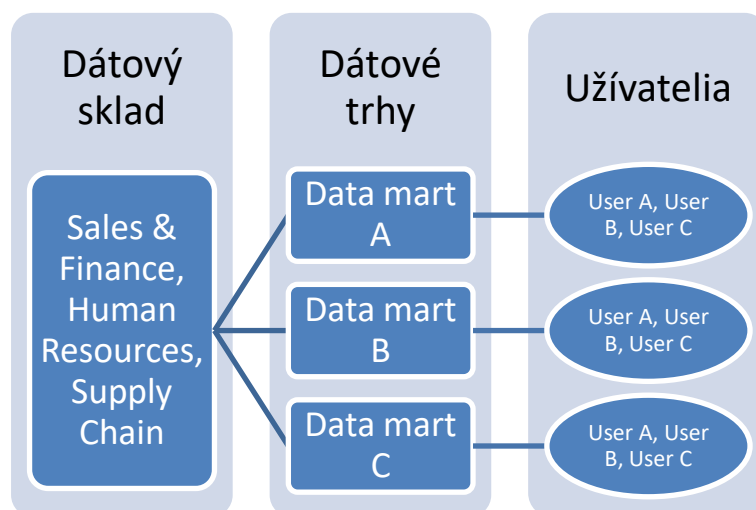
Obrázok 4.1: Architektúra podľa Kimballa

Užívatelia majú k uchovávaným dátam limitovaný prístup, nakoľko každý dátový trh je určený výhradne konkrétnej užívateľskej skupine. Pracovník marketingového oddelenia je schopný z dátového trhu získať napríklad podrobnejšie informácie o tržbách prebiehajúcich projektov. K sprístupneniu údajov o zamestnancoch by však bolo nutné adresovať dotaz na oddelený dátový trh obsahujúci patričné dáta, ku ktorému tento užívateľ nemusí mať prístup [9].

4.4.2 Centralizovaný dátový sklad

Bill Inmon priniesol opačnú koncepciu architektúry, ktorou je centralizovaný dátový sklad. Podstatou centralizovanej architektúry je fakt, že dátový sklad predstavuje jediné úložisko určené pre všetky dáta potrebné v procese business intelligence a odstraňuje separáciu informácií dátovými trhmi, ktoré v tomto prípade neobsahujú žiadne dáta.

Podobne ako v predošlej architektúre tu vystupujú dátové trhy, ale jedine ako sprostredkovateľ dát formou rozhrania a slúžia tak na zníženie záťaže dátového skladu. Pomocou nich pristupujú koncoví užívatelia k dátam uchovávaným v centralizovanom dátovom sklade. Pri centralizovanej architektúre má každý užívateľ systému možnosť pristupovať k akýmkoľvek dátam zhromaždeným v dátovom sklade. Obrázok 4.2 poskytuje celkový pohľad na architektúru centralizovaného dátového skladu.



Obrázok 4.2: Architektúra podľa Inmona

4.4.3 Porovnanie architektúr dátových skladov

Základný rozdiel medzi prístupmi Kimballa a Inmona je v koncepcii budovania dátového skladu. Architektúra podľa Inmona charakterizuje dátový sklad ako úložisko dát v globálnom zmysle, zatiaľ čo Kimball prezentoval dátový sklad založený na decentralizovanom princípe. Existuje však viacero ďalších charakteristík, ktorými sa obe architektúry odlišujú.

- Výhodou zavedenia decentralizovanej architektúry je jednoduchá implementácia dátových trhov a s tým spojené zníženie počiatočných investícií do business intelligence sys-

tému. Táto architektúra navyše umožňuje efektívne pripojenie dodatočných dátových trhov v budúcnosti. Vytvorenie centralizovaného dátového skladu je značne komplikovanejšie a jeho následné rozširovanie predstavuje pre podnik netriviálnu úlohu.

- Dátové trhy podľa Kimballa obsahujú dáta pre špecifické účely, no centralizovaný dátový sklad v sebe integruje všetky dostupné dáta. V prípade centralizovanej architektúry majú koncoví užívatelia prístup ku všetkým dátam, zatiaľ čo s využitím dátových trhov boli schopní narábať iba s ich podmnožinou. V dôsledku toho môže u Kimballovej architektúry nastať situácia, keď sa výsledky užívateľských analýz líšia.
- Zavedenie separátnych dátových trhov ako samostatných úložísk so sebou prináša riziko, že dátové trhy budú obsahovať o rovnakých podnikových entitách odlišné údaje. Táto inkonzistencia medzi uchovávanými dátami môže spôsobiť, že výsledné analýzy sa budú podľa jednotlivých dátových trhov líšiť. Dátový sklad ako centrálné úložisko dát toto riziko kompletne eliminuje.
- Pri postupnom budovaní dátových trhov dochádza k redundancii uchovávaných dát a mnohé dáta sa stávajú nadbytočnými. Centralizovaný dátový sklad redundanciu dát vylučuje, nakoľko je jediným dátovým úložiskom pre potreby business intelligence.

4.5 Dimenzionálny dátový model

Bežné relačné modely sú primárne optimalizované na vykonávanie transakcií. Business intelligence systémy však pre svoj chod využívajú alternatívny spôsob uloženia dát v podobe *dimenzionálneho modelu*, ktorý je vhodný na realizáciu analýz nad uchovávanými dátami. Dimenzionálny dátový model je vhodný formát pre generovanie multidimenzionálnej dátovej kocky.

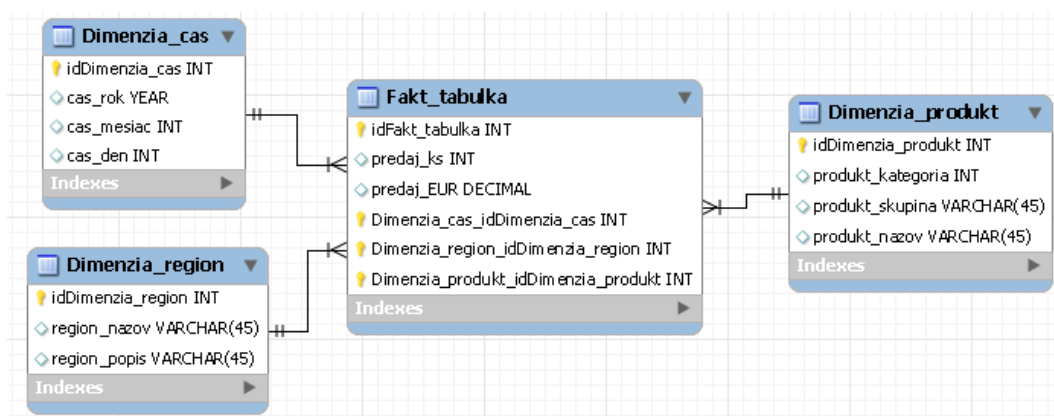
Dimenzionálny model pozostáva z dvoch dátových kategórií – *dimenzií* a *faktov*. Dimenzie reprezentujú typy analyzovaných dát a popisujú aspekty podnikových procesov. Fakty predstavujú konkrétne hodnoty týchto procesov. Typickými príkladmi dimenzií sú čas, územie alebo pohlavie, zatiaľ čo príslušnými faktami budú rok, štát a muž.

Dátový sklad predstavuje typ relačnej databázy navrhnutý špeciálne pre potreby business intelligence. Tvorbe business intelligence riešení musí byť prispôsobená koncepcia uchovávania dát v dátovom sklade. Nakoľko analytické OLAP nástroje pracujú s dimenzionálnymi dátami, je nutné, aby dátový sklad takýto dátový model simuloval relačnou databázou. Na dosiahnutie tejto vlastnosti slúžia dve základné relačné schémy – schéma hviezdy (*star schema*) a jej variant schéma snehovej vločky (*snowflake schema*).

4.5.1 Schéma hviezd

Jadrom schémy hviezd je centrálna tabuľka uchovávajúca všetky fakty. Pre každú dimenziu existuje separátna tabuľka, ktorá je na tabuľku faktov odkazovaná kľúčom. Zložením jednotlivých kľúčov tabuliek dimenzií vzniká primárny kľúč tabuľky faktov.

Schéma hviezd jasne a zreteľne definuje dimenzionálny model uchovávaní dát a zároveň umožňuje k dátam rýchlo pristupovať. Predpokladom k jej využitiu je uchovávanie statických dát, nakoľko schéma hviezd stráca svoju efektivitu u frekventovane sa meniacich dát. Ďalšou z nevýhod tohto riešenia sú možné problémy s postupným rastom tabuliek dimenzií pri hlbších dimenzionálnych hierarchiách [10].

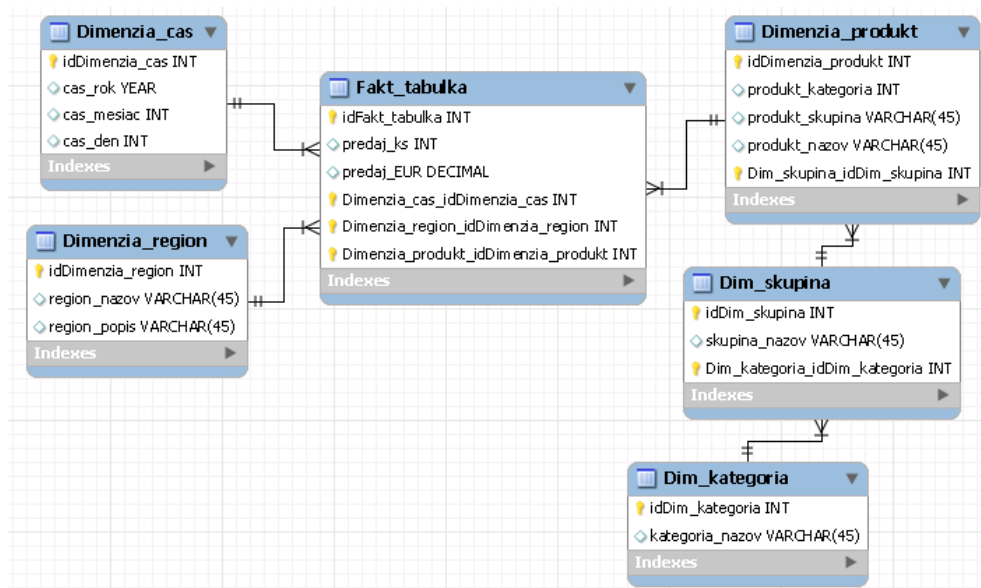


Obrázok 4.3: Schéma hviezd

4.5.2 Schéma snehovej vločky

Schéma snehovej vločky je typom hviezdicovej schémy, v ktorej sú jednotlivé tabuľky dimenzií normalizované. Výsledkom normalizácie je rozdelenie dimenzií do viacerých tabuliek za účelom odstránenia redundancií v uchovávaných dátach. Pre každú úroveň dimenzie tak existuje samostatná tabuľka, ktorá je so susednými úrovňami prepojená cudzím kľúčom.

Výhodou využitia schémy snehovej vločky je úspora úložného priestoru vďaka menšiemu objemu výsledných dát. Pri tvorbe potrebných dotazov je však nutné spájať viacero tabuliek, čo môže negatívne ovplyvniť časovú efektivitu analýzy.



Obrázok 4.4: Schéma snehovej vložky

5 ANALÝZA OLAP A DOLOVANIE DÁT

Transformáciu dát na *know-how* riadia v rámci business intelligence systémov analytické OLAP aplikácie, ktorým dátové sklady poskytujú potrebné vstupné dáta. Objem dát uchovávaných v dátovom sklade však nemusí znamenať zodpovedajúce množstvo znalostí. S dátami je počas analýzy nutné narábať efektívne a vytážiť z nich čo najviac potrebných faktov. Analytické aplikácie sú jadrom celého business intelligence procesu a od ich efektivity závisí kvalita a výpovedná hodnota výsledných reportov.

5.1 OLAP analýza

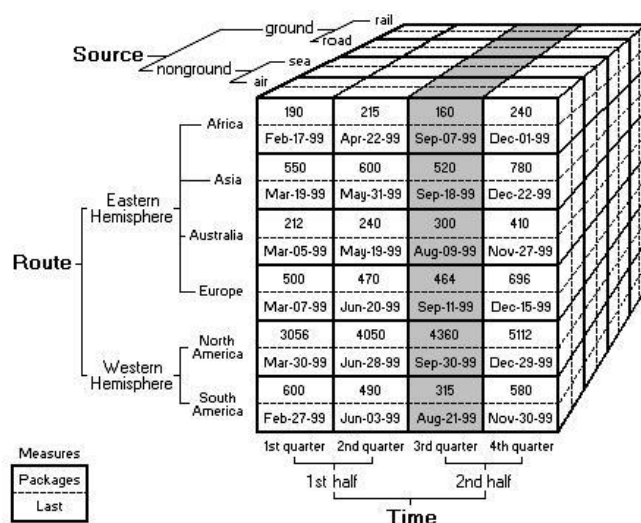
Realizovať dátovú analýzu spojenú s procesom business intelligence nad oddelenými transakčnými systémami je značne neefektívne. Pre potreby analýzy je nutné využiť dátový sklad, v ktorom sú dáta v patričnej podobe zoskupené do jediného úložiska.

Business intelligence systémy využívajú na riešenie problematiky analýz uchovávaných dát OLAP (*On-Line Analytical Processing*) technológie. Nástroje OLAP analýzy patria medzi komponenty business intelligence, ktorých úlohou je podrobne analyzovať vstupné dáta a ako výstup produkovať podklady pre tvorbu reportov. Reportovacie nástroje z nich následne dokážu generovať detailné výsledky.

5.2 Dátová kocka OLAP

Špecifikom business intelligence systémov je uchovávanie dát v dátovom sklade vo forme dimenzionálneho modelu. OLAP nástroje narábajú s dimenzionálnymi dátami spôsobom vhodným na ich analýzu. Na tento konkrétny účel slúži *OLAP kocka*, ktorú ilustruje obrázok 5.1.

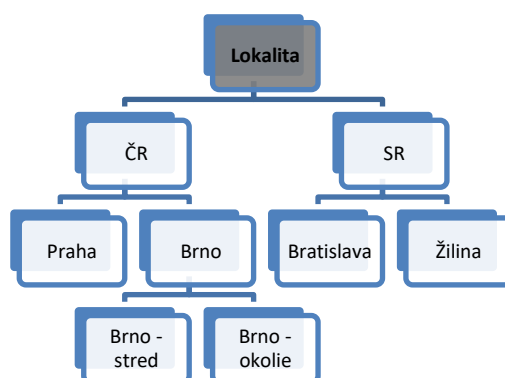
Dátová kocka OLAP nemusí nutne pozostávať z troch dimenzií, všeobecne sú OLAP nástroje schopné analyzovať dáta prislúchajúce k n dimenziám.



Obrázok 5.1: Dátová kocka OLAP [11]

Dimenzionálna reprezentácia uchovávaných dát má väčšiu vyjadrovaciu silu ako dvojrozmerné tabuľky, nakoľko jednotlivé dáta OLAP kocky možno sledovať z viacerých uhlov pohľadu. Samotné bunky OLAP kocky reprezentujú fakty, ktoré sú kategorizované podľa jej dimenzií.

Každý dimenzii OLAP kocky náleží skupina hodnôt, v uvedenom prípade ide o produkt, lokalitu a čas. Fakty (nazývané aj miery alebo metriky) predstavujú konkrétne číselné ohodnotenia daného aspektu, napr. množstvo, špecifický štát a definovaný časový interval. OLAP kocka uchováva jednotlivé dáta v rôznych úrovniach agregácie. Dimenzie a príslušné fakty tak možno kategorizovať podľa stanovených skupinových celkov. Vzorovú hierarchickú štruktúru OLAP kocky demonštruje nasledujúci obrázok.



Obrázok 5.2: Hierarchia geografickej dimenzie

S využitím dátovej kocky OLAP je možné vykonávať analýzy nad dátami z množstva rôznych perspektív – napríklad vyextrahovať štatistiku predaja skupiny produktov v danej oblasti za vymedzený štvrťrok. K modelu OLAP kocky možno pridávať dodatočné dimenzie, ako sú pohlavie či vek zákazníka, a analyzovať tak popularitu jednotlivých produktov v rámci konkrétne sledovaných spotrebiteľských skupín.

5.3 Operácie OLAP analýzy

Analytické nástroje OLAP dokážu efektívne analyzovať dáta využívaním viacerých operácií nad dátovou OLAP kockou. Táto skupina operácií umožňuje realizovať nad dátami špecifické variácie a modelovať tak štruktúru výslednej analýzy do požadovanej formy [5].

- *Roll-down (drill-down)* je operácia využívaná pri potrebe detailnejšieho pohľadu na daný aspekt. Pri operácii roll-down je konkrétna položka dimenzie znížená o jednu či viacero úrovní alebo je k predmetu analýzy pridaná dodatočná dimenzia. V uvedenom prípade dátovej kocky by sa miera dimenzie *lokalita* posunula v rámci úrovňovej stupnice zo štátu nižšie na mesto alebo ulicu. Naopak pridaním dimenzie *zákazník* by sa zoskupenie dát odstránilo a vo výsledku by boli k dispozícii údaje o predaji produktov podľa zákazníckych skupín.
- Operácia *roll-up (drill-up)* je presným opakom roll-down. Jej využitím je možné generovať štatistiky na základe agregovanejších dát. Prevedením operácie roll-up sa položka dimenzie zvýši o jednu či viacero úrovní, prípadne sa celá dimenzia odstráni. Zvýšením úrovne miery dimenzie *lokalita* by ako výsledná miera vystupovala oblasť strednej Európy, prípadne Európy ako kontinentu. Ak by sa na druhej strane odstránila dimenzia *čas*, boli by vo výslednej analýze k dispozícii štatistiky predaja produktov za kompletne časové obdobie.
- Prevedením operácie *slice* sa vybraný prvok dimenzie pevne zafixuje a zobrazia sa iba údaje, ktoré priamo súvisia s touto fixnou hodnotou. Ak je takýmto spôsobom zafixovaných viacero dimenzií, danú operáciu nazývame *dice*. Napríklad zafixovaním atribútu dimenzie *čas* na rok 2009 dostávame štatistiky predaja výrobkov v jednotlivých regiónoch za konkrétne ročné obdobie.
- Operácia *pivot* dokáže s jednotlivými dimenziami rotovať a zameniť ich zobrazenie. Dimenzie zobrazované v stĺpcoch sa tak premiestnia do riadkov a naopak, čím je možné získať rôzne pohľady na dátovú kocku.

5.4 OLAP Server

Jadro analytického procesu tvorí OLAP server, ktorý stojí medzi nástrojmi OLAP analýzy a dátovým sklado. OLAP server predstavuje prostriedok, pomocou ktorého majú jednotlivé analytické nástroje možnosť pristupovať k dimenzionálnym dátam určeným na analýzu.

Primárnou úlohou OLAP serveru je generovanie dimenzionálnych dát z dát uložených v dátovom sklade. OLAP server pomocou špecifických procesov generuje z relačného modelu hviezdy či snehovej vločky dátový kocku OLAP a uchováva ju vo vlastnom dátovom úložisku, prípadne ju vytvára v reálnom čase.

5.4.1 XMLA

Ako štandard pre prístup k dátam v analytických OLAP systémoch vznikol XMLA (*XML for Analysis*). XMLA bol vyvinutý Microsoftom v spolupráci so spoločnosťou Hyperion s cieľom vytvoriť štandardizované API v oblasti business intelligence.

Klientske analytické aplikácie komunikujú s OLAP serverom práve prostredníctvom XMLA, pričom pri komunikácii sú využívané štandardy HTTP, SOAP a XML. XMLA tvorí špecifickú množinu XML správ zasielaných pomocou protokolu SOAP, vďaka ktorým majú klientske aplikácie prístup k dimenzionálnym dátam generovaných z dátových zdrojov.

5.4.2 MDX

OLAP nástroje narábajú s dimenzionálnymi dátami a pre prístup k nim je nutné využiť odlišný dotazovací jazyk než SQL, nakoľko SQL nedisponuje dostatočnou vyjadrovacou silou pre dáta s viacerými dimenziami. MDX (*Multi-dimensional Expressions*) predstavuje v súčasnosti najbežnejší dotazovací jazyk pre prístup k dimenzionálnym dátam, ktorý je v XMLA špecifikovaný ako mdXML.

MDX štruktúrou svojich dotazov pripomína SQL. Zatiaľ čo SQL klauzulami SELECT a WHERE vymedzoval stĺpce a riadky tabuľky relačnej databázy, MDX využíva klauzulu SELECT na vymedzenie dimenzií a faktov dimenzionálnych dát a klauzulu WHERE na ohraničenie a definovanie podmienok vybraných dimenzií a faktov. Štruktúra základného MDX dotazu je nasledovná [12]:

```
SELECT [<axis_specification>[, <axis_specification>...]]  
FROM [<cube_specification>]  
[WHERE [< slicer_specification>]]
```

Klauzula **SELECT** špecifikuje podmnožinu dimenzionálnych dát, respektíve skupinu sledovaných dimenzií, ktoré majú byť zahrnuté vo výsledku. Pomocou **FROM** klauzuly sa vyberie konkrétna dátová kocka, z ktorej budú jednotlivé dáta extrahované. **WHERE** klauzula stanovuje os obmedzujúcu množinu výsledkov na dáta, ktoré sa nachádzajú výhradne na tejto osi. Konkrétna podoba výsledného MDX dotazu môže vyzeráť napríklad takto [12]:

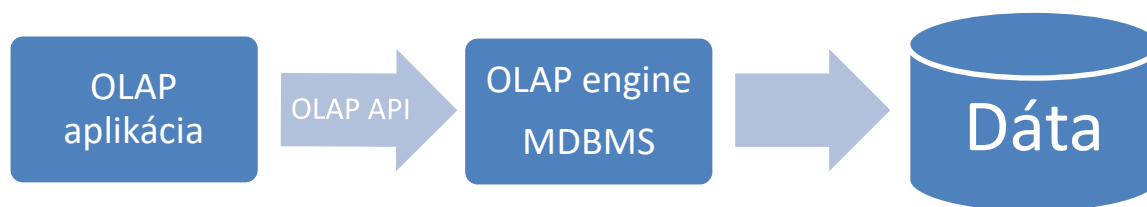
```
SELECT
{ [Measures].[Unit Sales], [Measures].[Store Sales] } ON COLUMNS,
{ [Time].[1997], [Time].[1998] } ON ROWS
FROM Sales
WHERE ( [Store].[USA].[CA] )
```

5.5 Architektúra OLAP serveru

OLAP server ako primárna entita analytického procesu môže mať v závislosti od spôsobu spracovania dimenzionálnych dát odlišnú štruktúru. Jeho architektúra špecifikuje metódy, pomocou ktorých sa konštruuje mutlidimenzionalita dátovej kocky v rámci internej OLAP databázy. OLAP server tak môže byť realizovaný štyrmi základnými spôsobmi: MOLAP (*Multidimensional OLAP*), ROLAP (*Relational OLAP*), HOLAP (*Hybrid OLAP*) a DOLAP (*Desktop OLAP*).

5.5.1 MOLAP

Jedná sa o pôvodný formát OLAP architektúry, v ktorom je obsah dátového skladu uchovávaný v príslušnej dimenzionálnej forme reprezentovanej *n*-rozmernou dátovou kockou. Dátová kocka je umiestnená v multidimenzionálnom databázovom systéme (MDBMS), ktorý je súčasťou OLAP engine realizujúceho analýzy nad uchovávanými dátami. Schému architektúry MOLAP ilustruje nasledujúci obrázok.



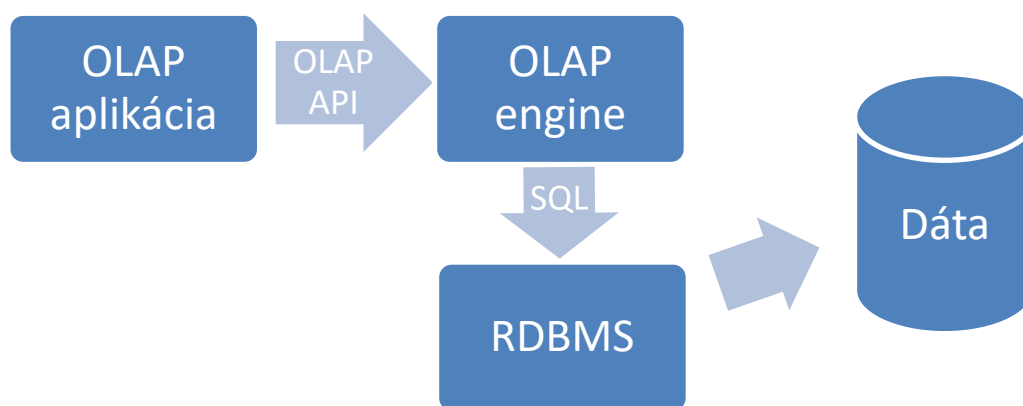
Obrázok 5.3: Architektúra MOLAP

Architektúra MOLAP je vďaka prítomnosti dátovej kocky v MDBMS optimalizovaná pre potreby frekventovaného dotazovania a realizácia analýz je tak na rozdiel od ostatných prístupov časovo

efektívna. Nevýhodou tohto prístupu je však nutnosť administrácie MDBMS, ktorý manipuluje s komplexnými dimenzionálnymi dátami, a súvisiace zvýšené nároky na hardvér a potrebný úložný priestor.

5.5.2 ROLAP

V architektúre ROLAP pristupuje OLAP engine k dimenzionálnym dátam relačnej databázy (RDBMS), ktorá je prevzatá z dátového skladu. Všetky MDX dotazy smerujúce do databázy musia byť nutne transformované do optimalizovanej formy SQL, pričom samotnú transformáciu spracováva oddelená vrstva OLAP engine. OLAP engine počas procesu analýzy schematicky prevádza relačné dáta do multidimenzionálnej podoby a vytvára dátovú kocku OLAP v reálnom čase. Nasledujúci obrázok demonštruje architektúru ROLAP.



Obrázok 5.4: Architektúra ROLAP

Charakteristikou relačnej databázy je skutočnosť, že nie je optimalizovaná pre potreby OLAP. Dátovú kocku je nutné vytvárať v priebehu analýzy, čo má negatívny dopad na jej časovú náročnosť. Výhodou architektúry ROLAP sú však menšie nároky na úložný priestor, vysoká flexibilita uložených dát a jednoduchšia správa relačnej databázy oproti koncepcii MDBMS.

5.5.3 HOLAP

Hybridný OLAP je špecifickým typom architektúry, ktorá kombinuje prístupy použité v ROLAP a MOLAP. Podstatou HOLAPu je fakt, že agregované dáta sú uložené v štruktúre multidimenzionálnej dátovej kocky a detailné dáta sa uchovávajú v podobe relačnej databázy. Hybridný OLAP tak využíva výhody efektivity dotazovania dimenzionálnej architektúry a nenáročnosti spracovania relačného prístupu.

5.5.4 DOLAP

Architektúra DOLAP je postavená na princípe lokálnej separácie OLAP kocky. Pred inicializáciou procesu analýzy sa podmnožina dimenzionálnej databázy alebo extrakcia dátovej kocky prenese na úložisko lokálneho klienta (desktop). Klient má vďaka tejto metóde k dispozícii dáta v podobe lokálnej dátovej kocky, pričom realizácia samotných analýz nie je závislá od pripojenia k OLAP serveru. Táto architektúra je obzvlášť vhodná na využitie v mobilných zariadeniach [13].

5.6 Dolovanie dát

Úlohou OLAP nástrojov je analyzovať dáta spôsobom, ktorý je špecifikovaný užívateľom. Manažérsku analýzu však možno odvodiť zo samotného obsahu informácií, a to pomocou dolovania dát. Dolovanie dát predstavuje špeciálny typ analýzy, ktorá využíva postupy na objavovanie strategických informácií v uchovávaných dátach. Pri tvorbe analýz pomocou OLAP nástrojov je cieľom nájsť odpovede na vopred známe otázky týkajúce sa podniku. Dolovanie dát naopak evokuje tvorbu týchto otázok a následné odvodzovanie predikcií. Proces dolovania dát napomáha manažérom a analytikom objavovať nové skutočnosti a testovať vlastné hypotézy.

Jednou z najpoužívanějších techník dolovania dát sú rozhodovacie stromy, ktoré systematicky delia dáta do skupín a vetvia ich podľa úrovni. Získavanie nových poznatkov dolovaním dát môže byť realizované aj za pomoci odvetví umelej inteligencie, akými sú neurónové siete. Tieto metódy sú vhodné najmä pre úlohy, v ktorých je potrebné nadobúdať predikcie týkajúce sa obchodnej stratégie podniku.

6 REPORTING

Po spracovaní dát analytickými nástrojmi a získaní výsledkov jednotlivých analýz nasleduje proces prezentovania týchto výsledkov koncovému užívateľovi. Na tento účel sa využívajú reportovacie nástroje, ktoré tvoria finálnu fázu business intelligence procesu.

Spoločne s dátovou analýzou je reporting ďalšou dôležitou súčasťou business intelligence. Aj v prípade využitia dokonalých analytických nástrojov a procesov dolovania dát môžu byť získané informácie nepoužiteľné, ak v nich manažéri ako koncoví užívatelia nedokážu nájsť odpovede na svoje otázky.

6.1 Tvorba reportov

Reporty sú prostriedkom prezentácie výsledkov analýz koncovým užívateľom. Tvorba reportov predstavuje pre business intelligence systém aktivitu spojenú s nadmernou réžiou. V typickom business intelligence systéme pozostáva 75-80 % záťaže systému z reportingu, ďalších 15-20 % z využívania analytických OLAP nástrojov a iba 0-5 % záťaže tvoria nástroje dolovania dát [9].

Výsledné reporty najčastejšie vznikajú vykonávaním štandardných dotazov nad analyzovanými dátami v pravidelných časových intervaloch. Report v tomto prípade predstavuje preddefinovanú šablónu, ktorá sa opakovane vyplňa aktuálnymi dátami. Uvedeným spôsobom bývajú generované reporty kvartálnych či celoročných správ.

V prípade potreby pokročilých a menej štandardných analýz má užívateľ možnosť tvoriť reporty na základe vlastných požiadaviek. Táto metóda sa nazýva *ad-hoc reporting*, pričom vytvorené

metódy majú obvykle jednorazový charakter a užívatelia ich štandardne generujú pomocou tzv. *wizardov*.

6.2 Účel reportov a výsledkov analýzy

Výsledkom práce reportovacích nástrojov je súbor dokumentov obsahujúcich detailné analytické údaje a štatistiky využiteľné v procesoch rozhodovania. Samotné reporty sa najčastejšie distribuujú v podobe HTML stránok, formátu Adobe PDF, Microsoft Excel alebo čistého textu tak, aby boli pre koncových užívateľov čo najzrozumiteľnejšie. Reporty bývajú často interaktívne s pokročilými grafickými prvkami a s možnosťou prispôsobenia špecifických atribútov reportu podľa vlastných užívateľských kritérií, napríklad v podobe tzv. *dashboardov*.

Obsahom reportu sú najmä tabuľky, grafy, diagramy a grafické nákresy, ktoré vhodnou a prehľadnou formou detaily analyzovaných aspektov podniku prezentujú. Pokročilejšie reporty môžu okrem iného obsahovať aj mapy spájajúce jednotlivé výsledky s geografickými dátami. Všeobecnou požiadavkou zo strany koncových užívateľov je možnosť finálne reporty a ich prvky vecne prezentovať. Výsledné reporty môžu manažéri a analytici využiť napríklad na prezentáciu dát týkajúcich sa anuitných výsledkov prevádzky jednotlivých oddelení podniku, prípadne pre vlastnú potrebu v rámci monitorovania konkrétnych aspektov spoločnosti.

7 KOMPLEXNÉ BI RIEŠENIA

Odvetvie business intelligence zažíva v poslednom období značný rozmach a veľké množstvo popredných spoločností prejavuje iniciatívu zaviesť do svojej infraštruktúry business intelligence aplikácie. V súčasnosti sú dostupné viaceré komplexné riešenia pre podnikové subjekty, ktoré sa snažia pokryť čo najviac aspektov business intelligence a s ich využitím tak vytvoriť plnohodnotný manažérsky informačný systém s pokročilou funkcionalitou.

V komerčnej sfére produkcie business intelligence riešení aktuálne dominujú softwarový giganti Microsoft, IBM, SAP a Oracle ponúkajúci robustné business intelligence systémy so širokým zázemím a podporou. Naproti tomu existujú na trhu open source alternatívy ponúkajúce dostatok možností pre menšie podniky.

7.1 Open source BI

Hlavnou prednosťou open source produktov sú nulové poplatky týkajúce sa licencie a samotného používania produktu. Zavádzanie business intelligence do podniku je však komplexný proces a implementácia tohto typu systému je časovo a finančne náročná. Úvodné investície týkajúce sa tohto procesu tak rozhodne nie sú zanedbateľné.

Prednosťou open source business intelligence systémov je, že podnik vďaka nim môže s nízkymi prevádzkovými nákladmi využívať všetky základné business intelligence komponenty spolu s možnosťou editácie ich zdrojového kódu. Oproti komerčným riešeniam sú primárne zamerané na malé podniky a poskytujú im technicky jednoduchšie riešenia. Open source business intelligence systémy v rámci svojho portfólia často ponúkajú aj platené alternatívy

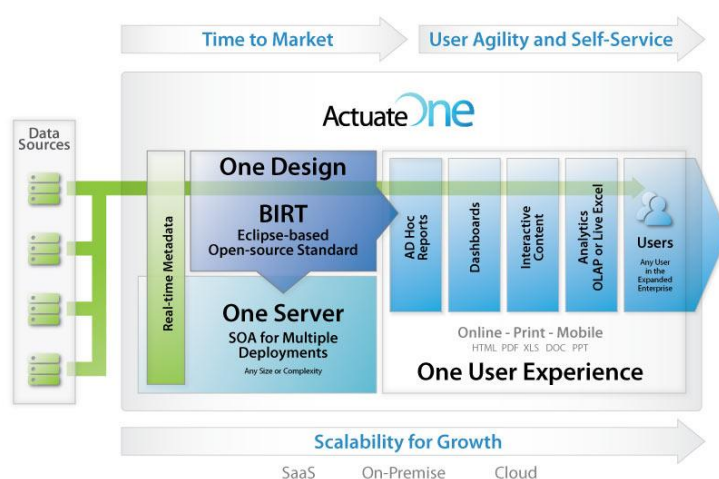
s pokročilejšou funkcionalitou, kde býva v rámci licencie zahrnutá aj profesionálna technická podpora.

7.1.1 BIRT

Projekt Eclipse BIRT (*Business Intelligence Reporting Tools*) vznikol ako výsledok spolupráce spoločností Actuate a Eclipse Foundation. Ide o pomerne mladý a stále sa rozvíjajúci projekt s dobrou víziou do budúcnosti.

BIRT je reportovací systém založený na vývojovom prostredí Eclipse a pozostáva z dvoch základných súčastí. Prvým primárnym komponentom je *BIRT Designer*, ktorého využitím je možné špecifikovať jednotlivé dátové zdroje pre potreby reportingu a rovnako tak aj definovať spôsoby vizualizácie finálnych dát. Ďalšou súčasťou je *BIRT Engine*, ktorý tvorí množinu Java tried a aplikačných rozhraní slúžiacich na generovanie reportov do širokej škály formátov. Eclipse BIRT je zameraný výhradne na reportovaciu časť business intelligence procesu a jeho možnosti sa dajú na aplikačnej úrovni integrovať do existujúcich produktov a webových aplikácií.

BIRT zároveň tvorí základ komerčného produktového balíka ActuateOne, pomocou ktorého je možné vyvíjať tzv. RIA (*Rich Internet Application*) aplikácie. Viaceré zdroje definujú RIA ako webové aplikácie s užívateľským komfortom blížiacim sa bežnému desktopovému softwaru – svojím spôsobom ide o kombináciu webovej aplikácie a tučného klienta. V súčasnosti existuje niekoľko konceptov pre efektívnu tvorbu RIA – AJAX, Flash, Silverlight atď.

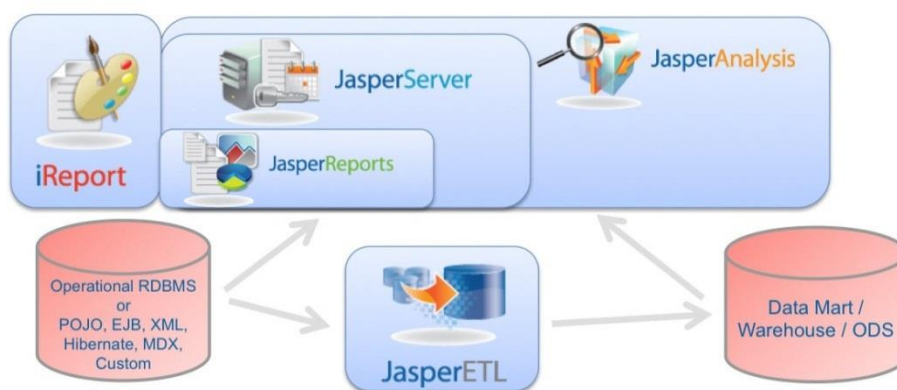


Obrázok 7.1: Architektúra ActuateOne [14]

ActuateOne predstavuje plnohodnotný business intelligence systém. Samotný BIRT je v ActuateOne rozšírený o aplikačný server *iServer*, ktorý ponúka možnosť distribúcie v tzv. *cloud* prostredí formou SaaS (*Software as a Service*) prípadne PaaS (*Platform as a Service*). Jednotlivé elementy ActuateOne sú uložené na strane poskytovateľa a ponúkané ako služby. Vďaka týmto vzdialeným službám si podnik môže prenajať funkčnú infraštruktúru a značne uľahčiť prvé zavádzanie business intelligence riešenia. ActuateOne takisto dokáže generovať pokročilé a interaktívne analytické výstupy s možnosťou ich integrácie do mobilných zariadení.

7.1.2 Jaspersoft

Ďalším open source business intelligence produktom na trhu je Jaspersoft od rovnomennej softvérovej firmy. Ide o plne komplexný balík aplikácií, ktorý pokrýva všetky základné aspekty business intelligence procesu. Jaspersoft sa svojím pokročilým partnerským programom poskytovania služieb snaží primárne orientovať na cloud business intelligence a je rovnako poskytovaný aj v komerčných verziách Professional a Enterprise Edition.



Obrázok 7.2: Architektúra Jaspersoft [15]

Softvérový balík Jaspersoft pozostáva z nasledovných aplikácií:

- **JasperETL** – nástroj disponujúci funkcionalitou na modelovanie ETL procesu, ktorý zaisťuje transformáciu zdrojových dát a ich následnú agregáciu do dátového skladu.
- **JasperAnalysis** – analytický nástroj, ktorý v sebe kombinuje OLAP server spolu s príslušným webovým klientom, databázu MySQL a servlet Tomcat tvoriaci webový server pre beh Java kódu.
- **JasperServer** – reportovací server podporujúci in-memory analýzu, tvorbu ad hoc dotazov a finálne generovanie reportov do širokej škály formátov.

- **JasperReports** – tvorí nadstavbu pre JasperServer a slúži na tvorbu pokročilých reportov za použitia moderných vizualizačných metód s využitím platformy Flash. Jeho jadrom je WYSIWYG nástroj iReport, pomocou ktorého je možné s jednotlivými reportami detailne manipulovať.

Open source verzia Jaspersoft však ponúka iba základnú funkcionálnosť a v rámci nej nie sú poskytované dodatočné návrhové nástroje s pokročilejšími možnosťami spracovania a reportingu dát. Jaspersoft je takisto šírený pod licenciou GPL a neumožňuje modifikáciu zdrojového kódu, čo z neho robí skôr komerčne zameraný produkt.

7.1.3 Palo for Excel a Palo Suite

Palo for Excel je zaujímavým open source produktom na poli business intelligence od spoločnosti Jedox, ktorý predstavuje analytickú a reportovaciu nadstavbu pre aplikáciu Microsoft Excel (existujúca alternatíva pre Open office Calc nesie názov PalOOCa). Jeho súčasťou sú OLAP Server a Excel/Calc plugin, za pomoci ktorých je možné vytvoriť funkčné business intelligence riešenie nad firemnými dátami priamo v tabuľkovom editore.

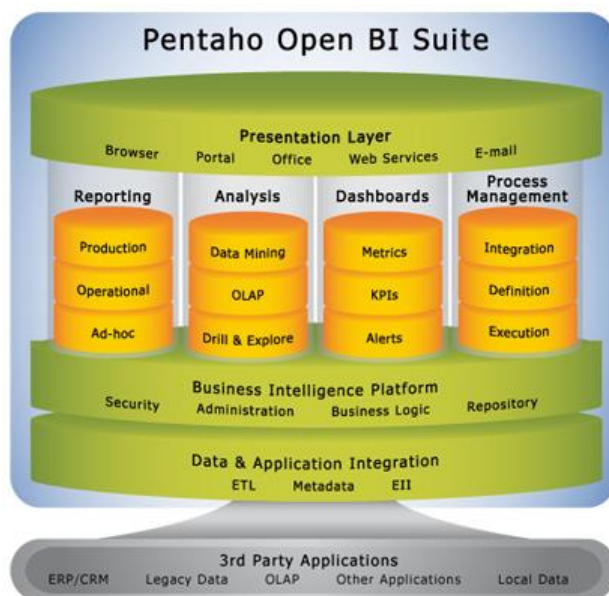
Excel ako široko používaný tabuľkový editor zaručuje, že mnoho malých podnikov má svoje firemné dáta využiteľné v procese business intelligence práve v Excel formáte. Palo for Server stavia na jednoduchosti svojich modulov, ich hardvérovej nenáročnosti a celkovej kompatibilitate. Zabudovaný OLAP Server uchováva jednotlivé Excel dáta vo forme dátovej kocky, umožňuje ich zdieľanie v centrálnom dátovom úložisku. Nechýba ani podpora XMLA a MDX a s tým súvisiaca možnosť vykonávať nad dátami dimenzionálne dotazy. Palo for Excel je zakomponované do rozličných klientskych knižníc (Java, PHP, C/C++, .NET) a eventuálne je možné integrovať ho do rôznorodých systémov a platforiem [16].

Naproti tomu je komerčná verzia Palo Suite kompletným business intelligence riešením kombinujúcim všetky hlavné BI aplikácie Jedoxu – Palo Web, Palo ETL Server, Palo for Excel a Palo OLAP Server. Palo Web ako webové rozhranie pre prístup k jednotlivým modulom a ich administráciu tvorí jadro celej platformy.

7.1.4 Pentaho BI Suite

Pentaho je aktuálne jedným z lídrov na trhu open source business intelligence, ktorý poskytuje široké spektrum možností postačujúce pre nároky malých a stredných podnikov. V rámci open source edície aplikačnej platformy Pentaho BI Platform šírenej pod licenciou GNU GPL je možné spravovať jednotlivé nástroje balíka Pentaho BI Suite.

Zdrojový kód Pentaho BI Suite je vyvíjaný v jazyku Java a jednotlivé business intelligence nástroje sú tak nezávislé od cieľovej platformy. Pentaho BI Suite je serverovo založené riešenie, pričom jadro celej platformy tvorí aplikačný server JBoss. Užívateľ k nemu pristupuje pomocou internetového prehliadača, ktorý vystupuje ako klient. Podrobnejšiu schému architektúry business intelligence balíka spoločnosti Pentaho ilustruje obrázok 7.3.



Obrázok 7.3: Architektúra Pentaho BI Suite [17]

Štruktúru samotného balíka Pentaho BI Suite tvorí viacero komponentov. Každý z nich predstavuje samostatný klientsky nástroj riešiaci konkrétnu časť business intelligence procesu.

- **Pentaho Data Integration** predstavuje sadu ETL nástrojov, ktoré zo zdrojových systémov transformujú a integrujú vybrané dáta využiteľné v procese business intelligence. Pomocou Data Integration je rovnako možné celý transformačný proces automatizovať a plniť tak dátový sklad v definovaných časových intervaloch.
- **Mondrian Schema Workbench** je modelovací nástroj, pomocou ktorého možno vytvárať schémy dátovej kocky OLAP priamo pre potreby analýzy. Tento nástroj užívateľovi takisto umožňuje upravovať jednotlivé schémy mutlidimenzionálnej databázy podľa vlastných potrieb a pomocou vizuálnych nástrojov vytvárať príslušné MDX dotazy. Schema Workbench úzko spolupracuje s OLAP serverom business intelligence balíka Mondrian.

- **Pentaho Metadata Editor** obstaráva pripojenie dodatočnej vrstvy metadát k spracovávaným dátam zdrojového systému. Metadáta obyčajne dopĺňujú informácie výsledných analýz dodatočnými údajmi určenými pre špecifické potreby reportingu.
- **Pentaho Report Designer** (pôvodom reportovací engine jFreeReport) slúži ako modelovací nástroj na jednoduché vytváranie detailných reportov súvisiacich s analýzou a procesom rozhodovania. Pomocou Report Designeru je užívateľ schopný prispôbovať jadro a štruktúru reportu podľa vlastných kritérií a takisto upravovať vizuálnu stránku výsledných analýz.
- **Pentaho Design Studio** je založený na vývojovej platforme Eclipse a predstavuje sadu multifunkčných nástrojov poskytujúcich grafické rozhranie, ktoré užívateľovi umožňuje vytvárať interaktívne dokumenty a pokročilejšie formy reportov.
- **Pentaho Data Mining** pozostáva z Java nástrojov strojového učenia, ktorých postupy a algoritmy sú určené na dolovanie dát a na účely plánovania poskytujú výsledky prediktívnej analýzy.

Pentaho BI Suite a jeho jednotlivé komponenty sú mocným nástrojom na tvorbu business intelligence systémov s pevným technickým zázemím, rozsiahlou komunitou a podrobnou dokumentáciou. Pentaho ako komerčná open source spoločnosť ponúka svoj balík aj v platenej verzii Enterprise Edition, ktorá navyše ponúka pravidelné aktualizácie produktov, množstvo pokročilých technických prvkov, školenia a profesionálnu technickú podporu.

7.1.5 SpagoBI

Jedným zo zaujímavých komplexných open source riešení sa stáva SpagoBI, ktorého developerom je Engineering Ingegneria Informatica. SpagoBI poskytuje všetky prvky business intelligence, no zároveň preberá časť funkcionality z dostupných open source a komerčných aplikácií, ktoré je možné do SpagoBI integrovať vo forme modulov. Ide tak o prispôbovateľný business intelligence balík ťažiaci z dostupnosti open source BI a analytických nástrojov.

Jadrom balíka je SpagoBI Server tvoriaci platformu, ktorá dovoľuje jednotlivé nástroje integrovať a modifikovať tak primárnu časť business intelligence systému na základe vlastných kritérií. SpagoBI Server vyžaduje ku svojmu behu aplikačný server JBoss umožňujúci modulom navzájom efektívne komunikovať. V rámci analytickej a reportovacej fázy je plne integrovateľných viacero z predošlých produktov – JPalo a JPivot na účely OLAP analýzy, BIRT a JasperReports/iReport zaoberajúce sa reportingom [18].

7.2 Komerčné BI riešenia

Využitie komerčných riešení môže mať mnoho výhod, no ich zavedenie do podniku predstavuje zložitejší proces, ktorý je finančne nákladný. Sú teda vhodné do prostredia veľkých korporácií, ktoré ich potenciál dokážu naplno využiť.

Komerčné riešenia sa oproti open source business intelligence vyskytujú na poli business intelligence dlhšiu dobu. Zastrešujú ich softwarové spoločnosti s obrovským rozpočtom a dokážu tak okrem kvality a pokročilosti riešenia poskytnúť najnovšie technologické trendy v oblasti business intelligence. V stručnosti si dovoľím spomenúť produkty troch najväčších hráčov v oblasti tvorby business intelligence systémov.

7.2.1 Microsoft BI

Microsoft je jedným z obrov pôsobiacich na trhu business intelligence, ktorý vyvíja vlastné analytické produkty podporujúce manažérske rozhodovanie. Balík Microsoft BI, ako komplexná business intelligence platforma, je spojením viacerých komerčných produktov Microsoftu – Office, Sharepoint a SQL Server.

Štruktúra Microsoft BI predstavuje trojvrstvovú architektúru. Business intelligence vrstva, ktorej jadrom je databázový systém SQL Server obsahujúci špecifické dátové a BI analytické nástroje, tvorí základ celej platformy. Prezentačná vrstva zahŕňajúca reporty, štruktúru jednotlivých analýz a zdieľanie výsledných informácií predstavuje výstupný portál business intelligence procesu. Aktivity spojené s prezentáciou potrebných údajov riadi Sharepoint spolu s Microsoft Office (primárne Excel). Poslednou súčasťou sú vývojové a doplnkové komponenty, ktoré ďalej rozširujú silu a možnosti business intelligence platformy ako celku.

SQL Server ako nosný pilier Microsoft BI uchováva dátový sklad spolu s dátovými trhmi a zabezpečuje chod ich primárnych databázových služieb. Okrem iného disponuje dodatočnou funkcionalitou, ktorá z neho robí plnohodnotnú business intelligence databázu [19]:

- SSIS (*SQL Server Integration Services*) – súbor nástrojov zaisťujúcich funkcionality ETL procesu a finálnu integráciu dát do Microsoft SQL Serveru,
- SSAS (*SQL Server Analysis Services*) – komponenty OLAP tvoriace špeciálnu databázu *Cube*, ktorá je určená pre analýzu dát,
- SSRS (*SQL Server Reporting Services*) – doplnkové aplikácie slúžiace na tvorbu reportov z dostupných dátových zdrojov.

Jadrom prezentačnej vrstvy je webová aplikácia Microsoft Sharepoint, pomocou ktorej možno zdieľať jednotlivé dokumenty a distribuovať potrebné dáta. K pokročilej úrovni prezentácie jednotlivých výsledkov môžu takisto poslúžiť aplikácie Microsoft Excel a Microsoft Visio, ktoré dokážu skompilovať finálne analýzy do formy tabuliek, grafov či interaktívnych diagramov.

Vďaka vývojovej vrstve možno pomocou technológií Microsoft .NET a Silverlight rozšíriť a prispôbovať funkcionality jednotlivých komponentov Microsoft BI, konkrétne využitím vývojových aplikácií Microsoft Visual Studio a Report Builder.

7.2.2 IBM Cognos

Cognos od spoločnosti IBM obsahuje kompletnú škálu špecializovaných aplikácií a analytických nástrojov na riešenie business intelligence problematiky v rozsiahlych podnikoch. Pre menšie a stredné podniky ponúka IBM alternatívu v podobe Cognos Express, ktorý disponuje výhradne primárnou BI funkcionalitou.

Architektúra Cognosu je založená na Kimballovom prístupe budovania dátového skladu s využitím separátnych dátových trhov. Koncovým užívateľom sú v systéme pridelené preddefinované užívateľské roly odvodené od organizačnej pozície, ktorú v danej spoločnosti vykonávajú. Na základe oprávnení a úrovni prístupu do systému určených samotnými rolami, pristupujú užívatelia k jednotlivým aplikáciám platformy Cognos. Funkčná logika dátového skladu je implementovaná pomocou databázového serveru IBM DB2.

Zaujímavosťou novej verzie Cognos 10 je, že okrem klasických business intelligence aplikácií v sebe integruje nástroje Lotus Connections pre sociálnu spoluprácu a vytvára tak BI sociálnu sieť. Štruktúru a poznatky jednotlivých analýz možno pomocou Cognosu zdieľať medzi užívateľmi v reálnom čase, pričom nechýba plná podpora mobilných zariadení [20].

Cognos je takisto úzko spojený s performance managementom a jeho pole pôsobnosti siaha do oblasti CPM a BPM. Cognos poskytuje širokú škálu procesne orientovaných business aplikácií umožňujúcich riadenie business procesov a strategického plánovania podniku a plné využitie potenciálu spoločnosti.

7.2.3 SAP BI

Posledným uvedeným veľkým hráčom na poli komerčných business intelligence riešení je spoločnosť SAP a jej balík NetWeaver, ktorý je komplexným a multifunkčným firemným produktom. NetWeaver predstavuje prepojenie viacerých SAP aplikácií, ktorých určité

komponenty tvoria plnohodnotný business intelligence systém SAP BI. Okrem toho ponúka NetWeaver ako celok širokú škálu ďalších funkcií a umožňuje spravovať firemné procesy a dáta, integrovať portálové riešenia či využiť aplikačnú platformu na vývoj informačného systému.

Najciteľnejšou výhodou SAP BI je odbúranie ETL procesu v prípade, že daná spoločnosť využíva ako zdroj dát informačné systémy v rámci portfólia SAP. Primárnym komponentom dátového skladu je SAP Business Information Warehouse, ktorý je vhodný aj pre distribuované prostredia, pričom Business Explorer rieši analytickú a reportovaciu fázu [21].

NetWeaver poskytuje business intelligence systému integrovaný portál prepojujúci jeho jednotlivé aplikácie. Vďaka nemu dokáže SAP BI spolupracovať s ostatnými komponentmi Netweaveru a možno tak vytvoriť rozsiahly a multifunkčný informačný systém.

7.3 Postavenie open source BI v rámci trhu

Komerčné business intelligence riešenia majú vďaka mohutným korporáciám, ktoré ich vyvíjajú, širokú technologickú základňu. Obyčajne ide o softwarových gigantov, ktorí aj vďaka nemalým investíciám do týchto produktov oproti open source riešeniam technologicky napredujú. Ich implementáciou a administráciou sa rovnako zaoberá široká sieť dodávateľov, čo ich predurčuje na využitie vo veľkých podnikoch.

Open source nedisponujú takým širokým a pokročilým spektrom funkcionality ako komerčné produkty, ale ponúkajú finančne výhodnú alternatívu poskytujúcu všetky základné prvky business intelligence. V súčasnosti už býva pravidlom, že open source BI produkty prechádzajú z formy vývoja v rámci open source komunity na komerčne zameranú produkciu a bývajú zastrešené renomovanejšími spoločnosťami [22]. Okrem open source licencovaných verzií tak vznikajú ich alternatívy v podobe komerčných riešení.

Napriek nulovým poplatkom vzťahujúcim sa k licenci zavedenie open source business intelligence riešení rozhodne nie je zdarma. Je nutné brať do úvahy fakt, že samotná implementácia open source business intelligence predstavuje proces, ktorého realizácia môže byť finančne náročnejšia. Dodatočne býva spoplatnená aj technická podpora spoločnosti, ktorá dané open source BI riešenie zastrešuje. Kvôli slabším možnostiam v oblasti dodávania týchto riešení môžu (nefinančné) náklady na ich údržbu narásť.

Nasadenie open source business intelligence systémov je zamerané na malé podniky požadujúce technologicky jednoduchšie riešenie, ktoré bude primárne využívané na účely reportingu a

monitorovania výsledkov podniku. Na základe výsledkov ankety publikovanej spoločnosťou BEyeNetwork bolo zistené, že približne tretina z oslovenej tisícky IT profesionálov v minulosti zavádzala open source business intelligence software, pričom ďalšia tretina oslovených uvažuje o jeho zavedení v blízkej budúcnosti [23]. Je možné predpokladať, že tento nárast záujmu o open source business intelligence produkty je dôsledkom dostatočne vysokej úrovne funkcionality týchto produktov a globálne nepriaznivej ekonomickej situácie.

Všeobecný záujem o problematiku business intelligence má rastúcu tendenciu. V roku 2009 uverejnilo IBM podrobnosti vlastného prieskumu, počas ktorého bolo oslovených 2600 osôb pracujúcich vo funkcii CIO (*Chief Information Officer*) v 78 rôznych krajinách a 19 rôznych odvetviach. Až 83% z oslovených respondentov uviedlo, že považujú zavedenie business intelligence za cieľ s najvyššou prioritou v rámci svojho funkčného obdobia [24].

Ako dobrá vízia do budúcnosti business intelligence sa ukazuje byť aj *cloud computing*, ktorého potenciál umožňuje implementovať business intelligence riešenia do širšej škály cieľových podnikov a poskytovať tak tieto systémy ako službu so vzdialeným prístupom. Viaceré open source business intelligence produkty už touto funkcionalitou disponujú.

8 NAVRHOVANÝ SYSTÉM

Ďalšou súčasťou etapy praktickej časti práce bolo pomocou zvoleného business intelligence riešenia vytvoriť pre vybraný podnikový subjekt funkčný manažérsky informačný systém. V kapitole bude predstavená spoločnosť, pre ktorú bol systém realizovaný ako aj podrobná architektúra MIS vrátane jeho funkčných komponentov.

8.1 Predstavenie spoločnosti

Business intelligence riešenie vytvorené v rámci praktickej časti práce bolo realizované v spolupráci s firmou pôsobiacou v súkromnom sektore českého trhu. Kurt Miller ako hlavný predstaviteľ rovnomennej spoločnosti založil svoje podnikanie v roku 1994. S počiatočným kapitálom do 500 000 Kč inicioval so svojím spoločníkom obchodnú činnosť získaním živnostenského listu v oblasti predaja zmiešaného tovaru a v súčasnosti pôsobí pod názvom Prodejny téměř všeho.

Spočiatku firma na predaj využívala dočasné obchodné priestory na trhovisku, jej postupná expanzia však vyústila do následnej kúpy permanentných priestorov v podobe kamenného obchodu. Do roku 1999 disponovala spoločnosť štyrmi obchodnými priestormi v rámci mesta Brna a jeho blízkeho okolia. Spoločnosť Kurta Millera sa tak z malého rodinného podniku vyvinula na sieť predajní zamestnávajúcu 9 osôb ako stály personál spoločne so 4 brigádnikmi.

V minulom roku spoločnosť úspešne zahájila proces integrácie interného informačného systému primárne určeného na koordináciu predajného procesu a administráciu finančných údajov. Spoločnosť rovnako plánuje zaviesť vlastnú divíziu v oblasti dodávania tovaru a logistiky a v budúcnosti rozšíriť svoje pole pôsobnosti aj na oblasť elektronického obchodovania. V rámci

realizácie manažérskeho informačného systému bolo prihladené k prevádzke e-shop portálu v budúcnosti.

8.2 Špecifikácia zdrojového systému

Ako samotný objekt a zdroj dát určených na analýzu poslužil interný informačný systém, nad ktorým bolo vytvorené oddelené business intelligence riešenie. Aspekty vyvinutého MIS boli prispôbené tak, aby spĺňali interné požiadavky spoločnosti.

Predmetom podnikania spoločnosti je sprostredkovanie širokej škály produktov, akými sú potraviny, drogeria, lieky a podobný zmiešaný tovar. Spoločnosť plánuje v blízkej budúcnosti prevádzkovať vlastný internetový portál spolu s integrovaným elektronickým obchodom. Databáza súčasného informačného systému tak bude následne rozšírená.

8.2.1 Dátové entity

Interný informačný systém podniku ako dátové úložisko uskutočnených obchodných transakcií uchováva dostatok dôležitých informácií, z ktorých je možné následne čerpať podklady pre predajné štatistiky a manažérske rozhodovanie:

- nákupné transakcie z pokladničných terminálov,
- účtovnícke dáta a sumárne informácie,
- zamestnanecké výkazy a dáta týkajúce sa predajných pobočiek,
- údaje o zákazníkoch a dodávateľskom procese (implementované v budúcnosti).

8.2.2 Užívateľské roly

Plný prístup k internému informačnému systému majú vlastníci spoločnosti a účtovnícke subjekty. Zamestnanci ako pracovníci fyzických pobočiek v systéme figurujú jedine ako predmetná užívateľská skupina, no nemajú k nemu žiadny prístup. Pre administrátorské entity externého dodávateľa, ktorý informačný systém spravuje, bola vytvorená špecializovaná užívateľská skupina umožňujúca jeho administráciu. Manažérsky informačný systém ako separátne business intelligence riešenie bude umožňovať prístup výhradne manažmentu (vlastníkom) spoločnosti.

8.3 Systémové požiadavky

Manažment spoločnosti ako skupina koncových užívateľov business intelligence riešenia disponoval konkrétnou predstavou o jeho možnostiach. Funkcionalita analytickej časti vytvoreného

manažérskeho informačného systému tak spĺňa viaceré funkčné požiadavky a možno pomocou neho vytvoriť výstupy, ktoré majú blízko k disciplíne marketingu:

- analýzy vzťahujúce sa na nákupy a obchodné transakcie uskutočnené na predajných miestach,
- zoskupenie jednotlivých nákupov a analýza ich vývoja v čase,
- kategorizácia produktov a štatistika predaja definovaných produktových skupín,
- sumarizovanie obchodných transakcií a obratu jednotlivých predajných miest,
- analýza zvolených zákazníckych skupín a ich chovania (v rámci budúceho e-shopu).

8.4 Kritériá voľby BI balíka

Cieľový podnik predstavuje malú spoločnosť na domácom trhu, ktorá rozhodne neplánuje vynaložiť dodatočné finančné prostriedky súvisiace s licenčnými poplatkami – open source BI riešenie je teda optimálnou voľbou. Na vytvorenie jednotlivých súčastí navrhovaného manažérskeho informačného systému bol použitý Pentaho BI Suite (viď podkapitola 7.1.4), ktorý predstavuje kompletný balík business intelligence nástrojov riešiacich ETL proces, analýzu dostupných dát a reporting výsledkov.

Open source alternatíva Pentaho BI Suite vo verzii Community Edition je oproti dostupným riešeniam uvedeným v podkapitole 7.1 najkomplexnejším a najvhodnejším business intelligence balíkom pre navrhovaný manažérsky informačný systém.

- BIRT a Palo for Excel svojou funkcionalitou nepokrývajú všetky aspekty business intelligence procesov. Sú zamerané výhradne na reporting, respektíve analytickú fázu a dátové transformácie ETL spoločne s návrhom a realizáciou OLAP kocky by tak museli byť riešené adekvátnymi produktmi tretích strán.
- Jaspersoft neponúka vo svojej open source verzii dostatočné množstvo užitočných nástrojov využiteľných pri tvorbe navrhovaného BI riešenia. Dostupnosť nástrojov OLAP Schema Workbench pre štrukturálny návrh dátovej kocky OLAP a Query/Report Designer až v rámci komerčnej Enterprise Edition je v tomto prípade kritický nedostatok.
- SpagoBI združuje rôznorodé aplikačné riešenia a kvôli ich variabilite ako aj nedostatočnej dokumentácii by bola implementácia značne komplikovanejšia.

8.5 Realizácia navrhovaného systému

Na základe definícií manažérskych požiadaviek bola navrhnutá koncepcia business Intelligence riešenia a jeho následná implementácia. Záverečné časti kapitoly praktickej časti práce popisujú architektúru hlavných prvkov realizovaného manažérského informačného systému a jeho fyzickú vrstvu.

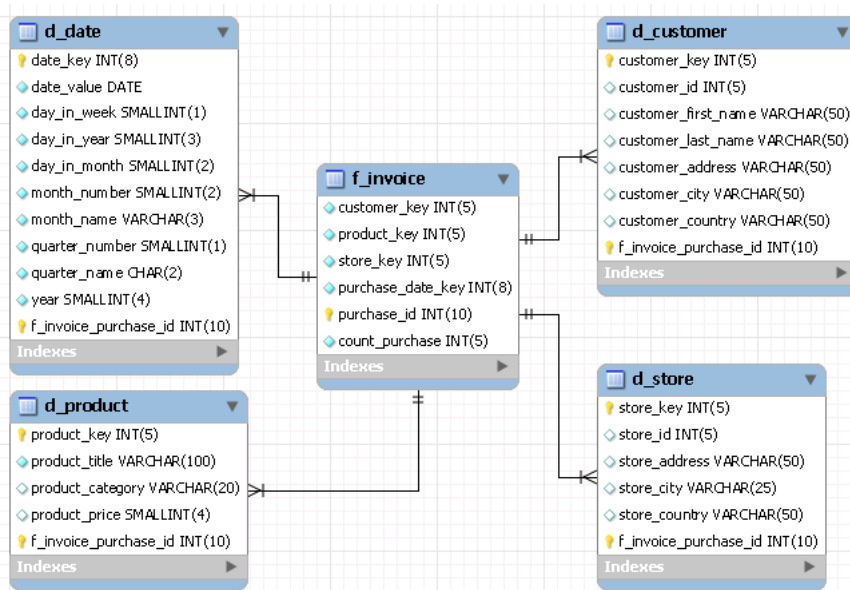
8.5.1 ETL proces

Kvôli pomernej jednoduchosti štruktúry dát interného informačného systému neboli vyžadované zásadné transformačné kroky slúžiace na ich konverziu. Dáta v rámci dátového skladu spĺňali podobu zohľadňujúcu manažérske požiadavky na ich analýzu. Vzhľadom na momentálnu slabú frekvenciu obsiahnejšej aktualizácie interných dát bol zvolený ich import do dátového skladu v mesačných intervaloch.

8.5.2 Dátový sklad

Dátový sklad je postavený nad relačnou databázou interného informačného systému. V rámci návrhovej fázy dátového skladu bolo počítané s budúcou implementáciou e-shopu a so spracovávaním údajov o zákazníkoch.

Logický model dátového skladu bol realizovaný databázovým návrhom relačnej schémy hviezda. Jadro schémy dátového skladu tvorí centrálna tabuľka faktov (`f_rental`) obsahujúca špecifické detaily všetkých uskutočnených obchodných transakcií v rámci spoločnosti. Model dátového skladu rovnako obsahuje štyri tabuľky dimenzií s metrikami vzťahujúcimi sa ku kardinálnym entitám systému, ktoré boli vopred definované na základe požiadaviek na BI riešenie. Obrázok 8.1 ilustruje schému hviezdy použitú pri návrhu architektúry dátového skladu.



Obrázok 8.1: Model dátového skladu

Uchovávané dáta tak možno monitorovať a spracovávať z viacerých perspektív na základe aktuálnych dimenzií:

- dimenzia času (d_date) modelujúca časové body a intervaly prislúchajúce k jednotlivým obchodným transakciám a (užívateľským) operáciám uskutočneným v systéme,
- dimenzia produktov (d_product), ktorá uchováva detaily popisujúce portfólio ponúkaných produktov ako predmetov predaja,
- dimenzia zákazníka (d_customer) pozostávajúca z osobných a geografických údajov všetkých registrovaných zákazníkov e-shop portálu,
- dimenzia pobočiek (d_store), ktorá obsahuje podrobnejšie informácie popisujúce súčasné pobočky a obchodné priestory spoločnosti.

8.5.3 OLAP kocka

Spolu s dátovým skladoom bola nad uchovávanými dátami vytvorená príslušná dátová kocka OLAP, ktorá je založená na hviezdicovej schéme dátového skladu. Hierarchická štruktúra dimenzií OLAP kocky priamo zodpovedá analytickým potrebám a obsahuje potrebné úrovne detailov dát určených na dimenzionálnu analýzu. Modely dimenzií majú nasledovný charakter:

- Rok -> Kvartál -> Mesiac -> Deň
Hierarchia dimenzie času je separovaná do kontinuálnych časových intervalov odvodených od priebehu kalendárneho roka. V závislosti od manažérskych potrieb detailov jednotlivých analýz je možné určovať úrovne skúmaných časových období.

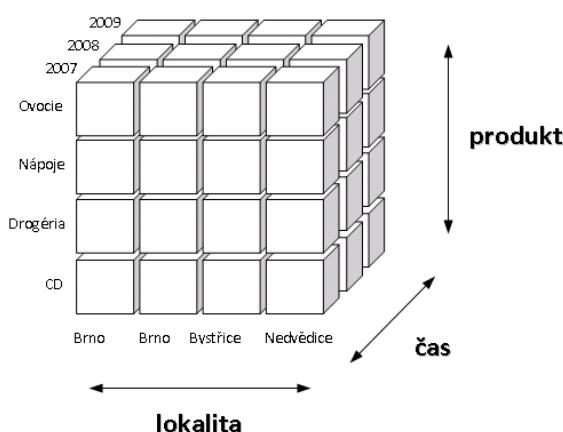
- Štát -> Mesto

Dimenzie zákazníka a pobočky definujú identickú hierarchiu týkajúcu sa geografickej úrovne detailu, ktorá popisuje fyzickú lokalitu zákazníka, respektíve pobočky.

- Kategória

Hierarchia dimenzie produktov pozostávajúca z jedinej úrovne odlišuje jednotlivé prvky dimenzie podľa ich kategórie.

Obrázok 8.2 obsahuje náčrt modelovej dátovej kocky OLAP v rámci vytvoreného manažérského informačného systému.



Obrázok 8.2: Dátová kocka OLAP

8.5.4 Fyzická vrstva systému

K vytvorenému manažérskemu informačnému systému budú pristupovať užívatelia manažérskej roly pomocou internetového prehliadača ako tenkého klienta. Samotný systém využíva na svoj beh integrovaný aplikačný server JBoss tvoriaci jadro business intelligence balíka aplikácií Pentaho. Ako databázové riešenie nad dátovým skladom vytvoreného business intelligence systému bol využitý open source databázový server MySQL Community Server šírený pod licenciou GNU GPL. Platforma Java, na ktorej je postavený balík business intelligence aplikácií Pentaho, zaručuje nezávislosť od architektúry či operačného systému.

8.5.5 Ukážka systému

Kompletné inštalačné a zdrojové súbory potrebné na chod vytvoreného business intelligence systému možno nájsť na priloženom DVD. Užívateľskú príručku spolu s adresárovou štruktúrou priloženého DVD popisujú prílohy A a B, pričom návod predpokladá, že na cieľovej stanici je

nainštalovaný systém Windows. Ďalšie časti príloh obsahujú snímky z behu systému a úryvky kódu z konfiguračných súborov.

Jednotlivé komponenty vytvoreného systému a ich prehľadane je sprístupnené prostredníctvom klientskych nástrojoch Pentaho BI Suite, ktoré sú vrátane dokumentácie dostupné na oficiálnych stránkach výrobcu. Na priloženom DVD sa nachádzajú prednastavené verzie týchto komponentov, ktoré umožňujú dodatočnú modifikáciu týchto komponentov.

9 ZÁVER

Aktuálna hospodárska situácia a vývoj trhu kladie rozvíjajúcim sa podnikom neľahké podmienky a sféra manažmentu musí čeliť náročnejším rozhodnutiam. Pokiaľ je však obchodná stratégia podniku a investovanie dostupných prostriedkov podložené správnymi opatreniami, podnik má väčšiu šancu prosperovať a zaujať dominantnejšie postavenie voči konkurencii. Využitie business intelligence metodík môže byť teda kľúčovým krokom zvlášť v dnešnom období recesie, kedy je potrebné zvyšovať marketingovo zamerané aktivity.

Business intelligence nástroje sa stali neoddeliteľnou súčasťou manažérskeho rozhodovania a tvorby analýz. Manažérske informačné systémy a nástroje business intelligence dokážu byť v tomto smere pre manažérov markantným prínosom, nakoľko pomocou vyhotovených analýz a prognóz usmerňujú ich postup a poskytujú informačnú podporu v procese rozhodovania. Vďaka konsolidovaným dátam, ktoré sú výsledkom business intelligence procesu, dokážu manažéri urobiť rozhodovacie kroky na základe exaktných faktov bez toho, aby vedeli aký postup k finálnemu výsledku viedol.

Okrem poskytovania informácií a znalostí pre kľúčové rozhodnutia týkajúce sa obchodnej stratégie podniku je takisto možné využitím business intelligence monitorovať určité podnikové aspekty a sledovať jeho výkonnostné analýzy. Manažéri tak majú navyše k dispozícii komplexné popisy využívania podnikových zdrojov a môžu tak efektívne hľadať spôsoby ich následnej úspory.

Súčasný open source produkty na trhu business intelligence sa snažia konkurovať komerčným alternatívam, no nedokážu užívateľovi poskytnúť takú škálu súvisiacich riešení ako komerčne

distribúované produkty. Open source balíky business intelligence sféry konkurujú hlavne cenou a absenciou akýchkoľvek poplatkov vzťahujúcich sa k licencií. Počiatočné investície na zavedenie open source business intelligence systému a jeho neskoršia administrácia však so sebou prináša riziko zvýšených nákladov na jeho administráciu a údržbu. Produkty v oblasti open source business intelligence sú každopádne vhodné pre segment malých a začínajúcich podnikov. Pre stredné a väčšie podniky môžu poslúžiť ako skúšobný manažérsky informačný systém a podniky tak na základe vlastných skúseností učinia prípadnú konverziu na komerčné riešenie.

Podkapitola 7.3 uvádza, že open source business intelligence systémy sú momentálne na vzostupe a v súčasnosti dokážu na trhu malých podnikov konkurovať komerčným produktom. Balík open source business intelligence aplikácií Pentaho BI Suite, ktorý bol využitý v rámci praktickej časti práce, plne vyhovoval potrebám a požiadavkám cieľového podniku. Využitie open source business intelligence riešenie spoločnosti Pentaho tvorí solídny základ tvorby samostatných manažérskych informačných systémov a systémov na podporu rozhodovania v malých organizáciách. Pre komplexné a robustné systémy stredne veľkých podnikov a mohutných korporácií sú však viac vhodné komerčné business intelligence riešenia.

Počas vývoja business intelligence riešenia nad zdrojovým systémom nedošlo ku značnejším komplikáciám, ktoré by boli spôsobené nedostatočnou funkcionalitou Pentaho BI Suite. Dátovým entitám interného informačného systému cieľového podniku bola následne prispôbena funkcionalita výsledného business intelligence riešenia na základne manažérskych požiadaviek. Na základe evidovaných dát je možné vo vytvorenom systéme sledovať špecifické trendy užívateľských nárokov, ako napríklad vývoj predaja konkrétnej kategórie produktov v zvolenom období kalendárneho roku, či štatistika nárastu dopytu v jednotlivých lokalitách. Interný informačný systém bude ako celok v budúcnosti rozšírený o e-shop a bude súčasťou nového internetového portálu. Do budúcnosti sa teda naskytuje možnosť existujúce business intelligence riešenie patrične rozšíriť o potrebnú funkcionalitu monitorovania zákazníckych entít.

10 BIBLIOGRAFIA

1. **Šmíd, V.** *Management informačního systému* [online]. 2005 [cit. 10. február 2010]. <<http://www.fi.muni.cz/~smid/managis.html>>.
2. **Rábová, I.** *Manažerské informační systémy* [online]. 2005 [cit. 25. február 2010]. <<https://is.mendelu.cz/eknihovna/opory/index.pl?opora=198>>.
3. **Tvrdíková, M.** *Aplikace typu EIS a přínosy z jejich užívání* [online]. 1999 [cit. 10. február 2010]. <<http://www.osu.cz/katedry/kip/aktuality/sbornik99/tvrdikova>>.
4. **Laudon, K., Laudon, J.** *Management Information Systems: Managing the Digital Firm*. 9th Edition. Upper Saddle River : Prentice Hall, 2005. 978-0-13-153841-2.
5. **Vercellis, C.** *Business Intelligence: Data Mining and Optimization for Decision Making*. Milan : John Wiley & Sons, 2009. 978-0-470-51139-8.
6. **Rayner, N.** *CPM: A Strategic Deployment of BI Applications*. Stamford : Gartner, 2002.
7. **Novotný, O., Pour, J., Slánský, D.** *Business Intelligence - Jak využít bohatství ve vašich datech*. Praha : Grada, 2005. 978-80-247-1094-3.
8. **Wikipedia.** *Extract, Transform, Load* [online]. 2010 [cit. 23. apríl 2010]. <http://en.wikipedia.org/wiki/Extract,_transform,_load>.
9. **Bouman, R., van Dongen, J.** *Pentaho Solutions: Business Intelligence and Data Warehousing*. Indianapolis : John Wiley & Sons, 2009. 978-0-470-48432-6.
10. **Berka, P.** *Dobývání znalostí z databází*. Praha : Academia, 2005. 978-80-200-1062-9.
11. **Sourceforge.** *BIOLAP* [online]. 2005 [cit. 15. október 2010]. <<http://biolap.sourceforge.net/>>.
12. **Microsoft Corporation.** *Analysis Services: The Basic MDX Query* [online]. 2010 [cit. 1. máj 2010]. <[http://msdn.microsoft.com/en-us/library/aa216770\(SQL.80\).aspx](http://msdn.microsoft.com/en-us/library/aa216770(SQL.80).aspx)>.

13. **Kupčík, J.** *OLAP technologie a mobilní klienti* [online]. 24. december 2007 [cit. 14. jún 2010]. <<http://www.fit.vutbr.cz/study/courses/TJD/public/0708TJD-Kupcik.pdf>>.
14. **Actuate Corporation.** *ActuateOne* [online]. 2010 [cit. 1. november 2010]. <<http://www.actuate.com/products/>>.
15. **Jaspersoft Corporation.** *Jaspersoft Business Intelligence Suite* [online]. 2010 [cit. 1. november 2010]. <<http://www.jaspersoft.com/jaspersoft-business-intelligence-suite>>.
16. **Jedox AG.** *Palo Suite* [online]. 2010 [cit. 5. november 2010]. <<http://www.jedox.com/en/products/Palo-Suite.html>>.
17. **Pentaho Corporation.** *Pentaho Community User Guide*. San Francisco : Pentaho Corporation, 2008.
18. **Engineering Ingegneria Informatica S.p.A.** *SpagoBI* [online]. 2010 [cit. 5. november 2010]. <<http://www.spagoworld.org/xwiki/bin/view/SpagoBI/>>.
19. **Microsoft Corporation.** *Microsoft Business Intelligence* [online]. 2010 [cit. 24. október 2010]. <<http://www.microsoft.com/bi/>>.
20. **IBM.** *Cognos Business Intelligence and Financial Performance Management* [online]. 2010 [cit. 26. október 2010]. <<http://www-01.ibm.com/software/data/cognos/>>.
21. **SAP Corporation.** *SAP NetWeaver: Business Intelligence* [online]. 2010 [cit. 27. október 2010]. <<http://www.sap.com/services/education/catalog/netweaver/bi.epx>>.
22. **Hanuš, J.** *Mýty kolem open source pro řešení business intelligence* [online]. 1. jún 2009 [cit. 22. september 2010]. <<http://www.systemonline.cz/business-intelligence/myty-kolem-open-source-pro-reseni-business-intelligence.htm>>.
23. **McKay, L.** *Open-Source Business Intelligence: 5 Times Bigger by 2012* [online]. 2010 [cit. 15. október 2010]. <<http://www.destinationcrm.com/Articles/CRM-News/Daily-News/Open-Source-Business-Intelligence-5-Times-Bigger-by-2012-60247.aspx>>.
24. **Schoenek, H.** *BI Solutions For the Midsize Market* [online]. 2009 [cit. 15. október 2010]. <<http://ezinearticles.com/?Business-Intelligence-Software-Review---BI-Solutions-For-the-Midsize-Market&id=3210675>>.
25. **Inmon, W.** *Building the Data Warehouse*. 4th. San Francisco : John Wiley & Sons, 2005. 978-0-7645-9944-6.

PRÍLOHA A

Obsah priloženého DVD

Priložené súbory obsahujú prednastavené súčasti balíka business intelligence aplikácií Pentaho potrebné na prevádzku vytvoreného systému.

Popis adresárovej štruktúry:

administration-console – administratorská konzola Pentaho BI Servera
bi-server-ce – Pentaho BI Server
report-designer – Pentaho Report Designer
schema-workbench – Pentaho Schema Workbench
ptv.mondrian.xml – schéma vytvorenej dátovej kocky OLAP
ptv-schema.sql – skript so schémou dátového skladu
ptv-data.sql – skript so vzorovými dátami dátového skladu

PRÍLOHA B

Návod na inštaláciu priloženého systému

Základnou požiadavkou pre beh systému je dispozícia nainštalovaného rozhrania Java (JRE, JVM). Pred spustením Pentaho BI Servera je rovnako potrebné mať na pracovnej stanici nainštalovaný databázový server MySQL, ktorý je dostupný na <http://dev.mysql.com/downloads/mysql/>. Inštalačný súbor JRE a databázy MySQL možno nájsť aj na priloženom DVD.

Po úspešnej inštalácii a prihlásení sa užívateľa k databázovému serveru je nutné importovať vytvorený dátový sklad spolu s vzorovými zdrojovými dátami do databázy pomocou príkazového riadku. Postupnosťou príkazov

```
SOURCE %path%/ptv-schema.sql;  
SOURCE %path%/ptv-data.sql;  
USE ptv;
```

zavedie dátový sklad ako databázu ptv a importuje do nej príslušné dáta.

Spustením súboru `bi-server-ce/start-pentaho.bat` sa inicializuje Pentaho BI Server. K serveru sa následne pripojuje pomocou užívateľskej konzoly, ktoré je prístupná na adrese <http://localhost:8080/pentaho/Login>. Kliknutím na 'Pentaho User Console Login' a vybraním administrátorského užívateľského konta Joe, dôjde k úspešnému prihláseniu do systému.

Na hlavnej stránke užívateľskej konzoly je na prehliadanie vytvorenej dátovej kocky nutné kliknúť na položku 'New Analysis View' a z ponúkaných schém vybrať 'PTV'. Po kliknutí na tlačidlo OK sa zobrazí interaktívna schéma dátovej kocky, ktorá užívateľovi umožňuje dáta uchovávané v dátovom sklade prehliadať podľa jednotlivých dimenzií.

Pre prehliadanie a modifikáciu štruktúry dátovej kocky OLAP je potrebné spustiť Pentaho Schema Workbench súborom `schema-workbench/workbench.bat` a následne otvoriť priložený súbor `ptv.mondrian.xml`.

Na vytváranie dodatočných reportov slúži Pentaho Report Designer, ktorý sa spúšťa súborom `report-designer/report-designer.bat`, kde je pomocou Report Wizardu a po vybratí príslušných zdrojových dát možné pomocou vizuálnych nástrojov tvoriť vlastné reporty.

PRÍLOHA C

Ukážka súboru dátovej kocky OLAP ptv.mondrian.xml

```
<Schema name="Prodejny temer vseho">
  <Dimension type="TimeDimension" name="Date">
    <Hierarchy name="time_hierarchy" hasAll="true" allMemberName="All Dates"
primaryKey="date_key">
      <Table name="d_date">
        </Table>
        <Level name="Year" column="year" type="Integer" uniqueMembers="true"
levelType="TimeYears" hideMemberIf="Never">
          </Level>
          <Level name="Quarter" column="quarter_number" type="String"
uniqueMembers="false" levelType="TimeQuarters" hideMemberIf="Never"
captionColumn="quarter_name">
            </Level>
            <Level name="Month" column="month_number" type="String" uniqueMembers="false"
levelType="TimeMonths" hideMemberIf="Never" captionColumn="month_name">
              </Level>
              <Level name="Day" column="day_in_month" type="Integer" uniqueMembers="false"
levelType="TimeDays" hideMemberIf="Never">
                </Level>
              </Hierarchy>
            </Dimension>

    <Dimension type="StandardDimension" name="Product">
      <Hierarchy name="product_hierarchy" hasAll="true" allMemberName="All Products"
primaryKey="product_key">
        <Table name="d_product">
          </Table>
          <Level name="Category" column="product_category" type="String"
uniqueMembers="false" levelType="Regular" hideMemberIf="Never"
captionColumn="product_category">
            </Level>
          </Hierarchy>
        </Dimension>

    <Dimension type="StandardDimension" name="Customer">
      <Hierarchy name="customer_hierarchy" hasAll="true" allMemberName="All
Customers" primaryKey="customer_key">
        <Table name="d_customer">
          </Table>
          <Level name="Country" column="customer_country" type="String"
uniqueMembers="true" levelType="Regular" hideMemberIf="Never"
captionColumn="customer_country">
            </Level>
```

```

        <Level name="City" column="customer_city" type="String" uniqueMembers="false"
levelType="Regular" hideMemberIf="Never" captionColumn="customer_city">
        </Level>
    </Hierarchy>
</Dimension>

    <Dimension type="StandardDimension" name="Store">
        <Hierarchy name="store_hierarchy" hasAll="true" allMemberName="All Stores"
primaryKey="store_key">
            <Table name="d_store">
            </Table>
            <Level name="Country" column="store_country" type="String"
uniqueMembers="true" levelType="Regular" hideMemberIf="Never"
captionColumn="store_country">
            </Level>
            <Level name="City" column="store_city" type="String" uniqueMembers="false"
levelType="Regular" hideMemberIf="Never" captionColumn="store_city">
            </Level>
        </Hierarchy>
    </Dimension>

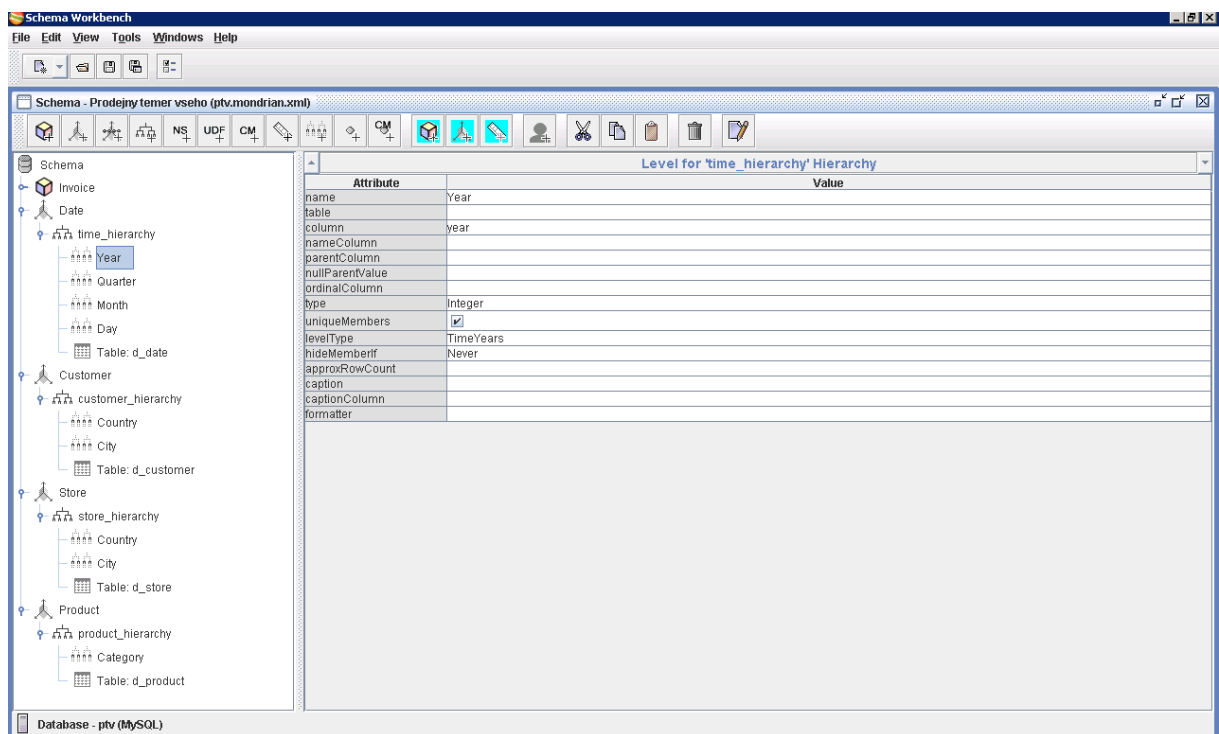
<Cube name="Invoice" cache="true" enabled="true">
    <Table name="f_invoice">
    </Table>
    <DimensionUsage source="Date" name="Purchase Date"
foreignKey="purchase_date_key">
    </DimensionUsage>
    <DimensionUsage source="Product" name="Product" foreignKey="product_key">
    </DimensionUsage>
    <Measure name="purchase" column="count_purchase" formatString="Standard"
aggregator="sum">
    </Measure>
</Cube>
</Schema>

```

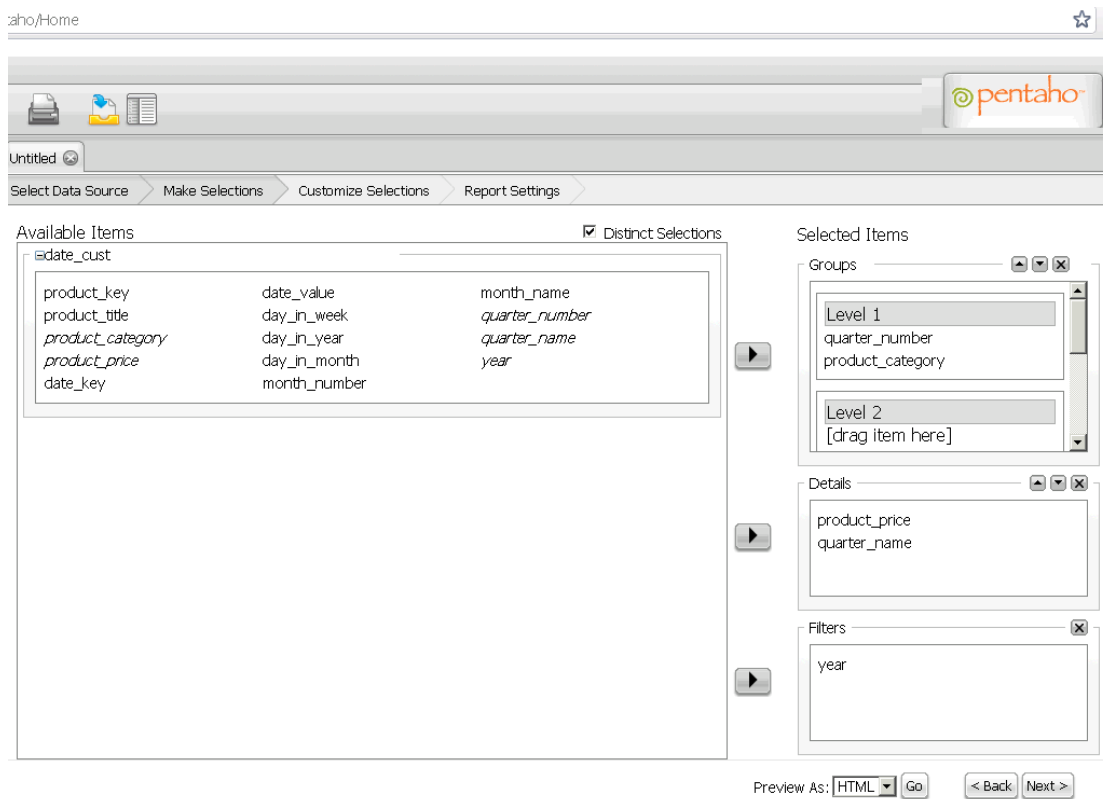
Ukážka súboru `ptv.mondrian.xml` demonštruje schému dátovej kocky OLAP vytvorenej v rámci business intelligence riešenia.

PRÍLOHA D

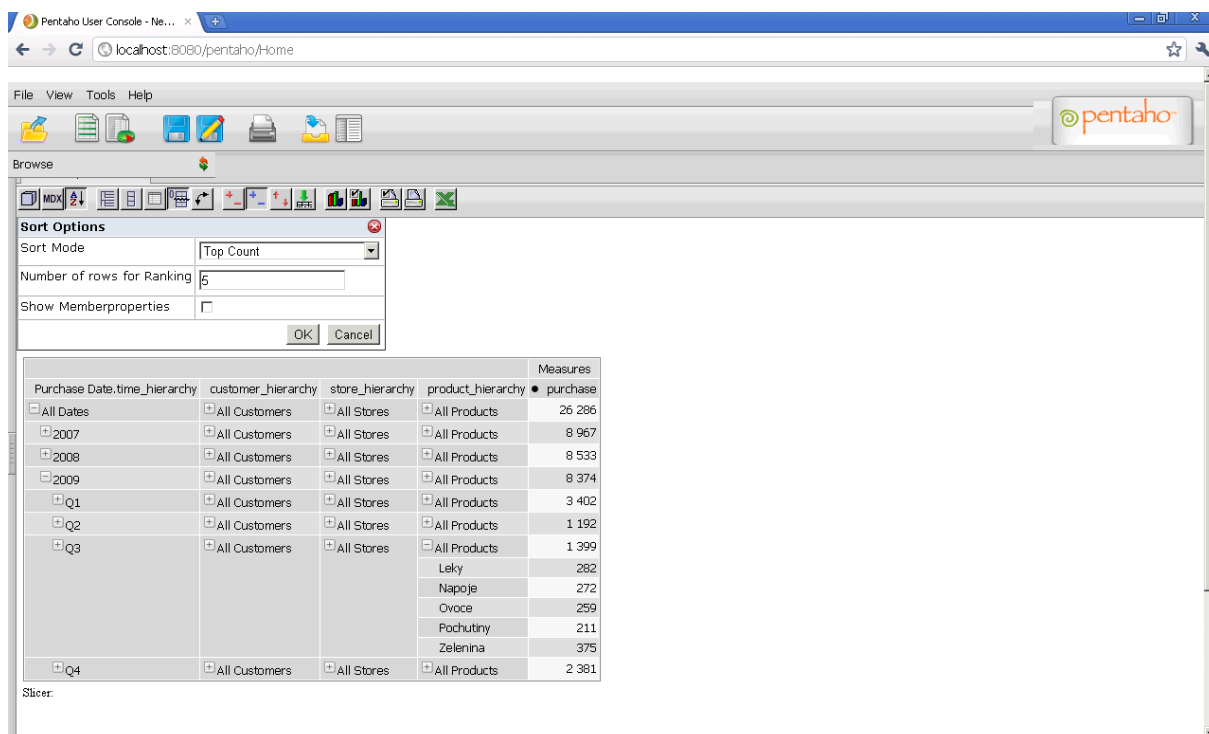
Snímky vytvoreného systému



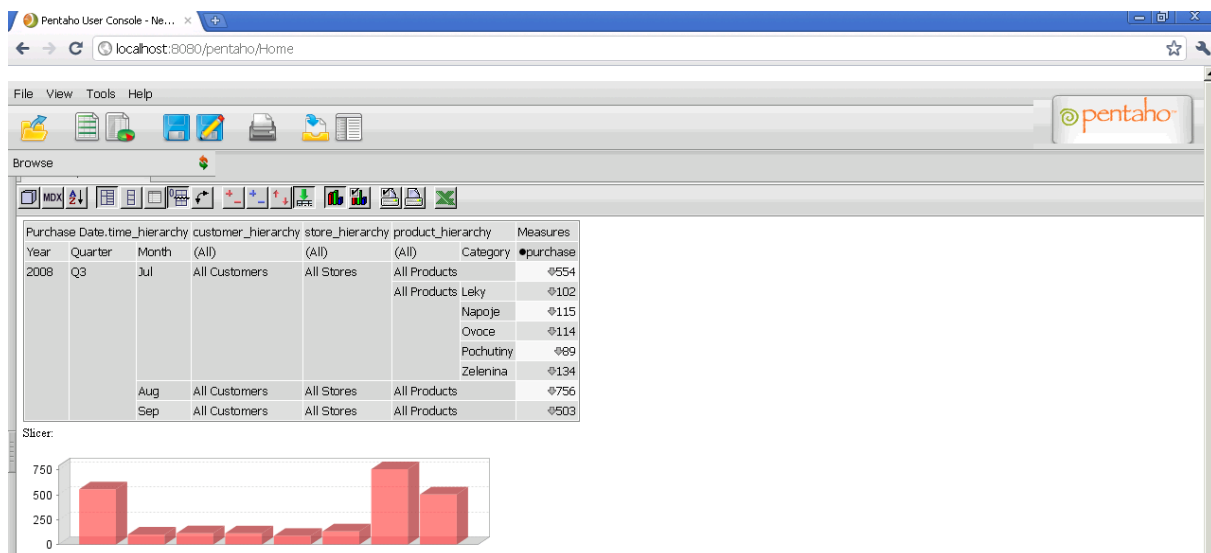
Obrázok D.1: Návrh schémy dátovej kocky OLAP v nástroji Schema Workbench



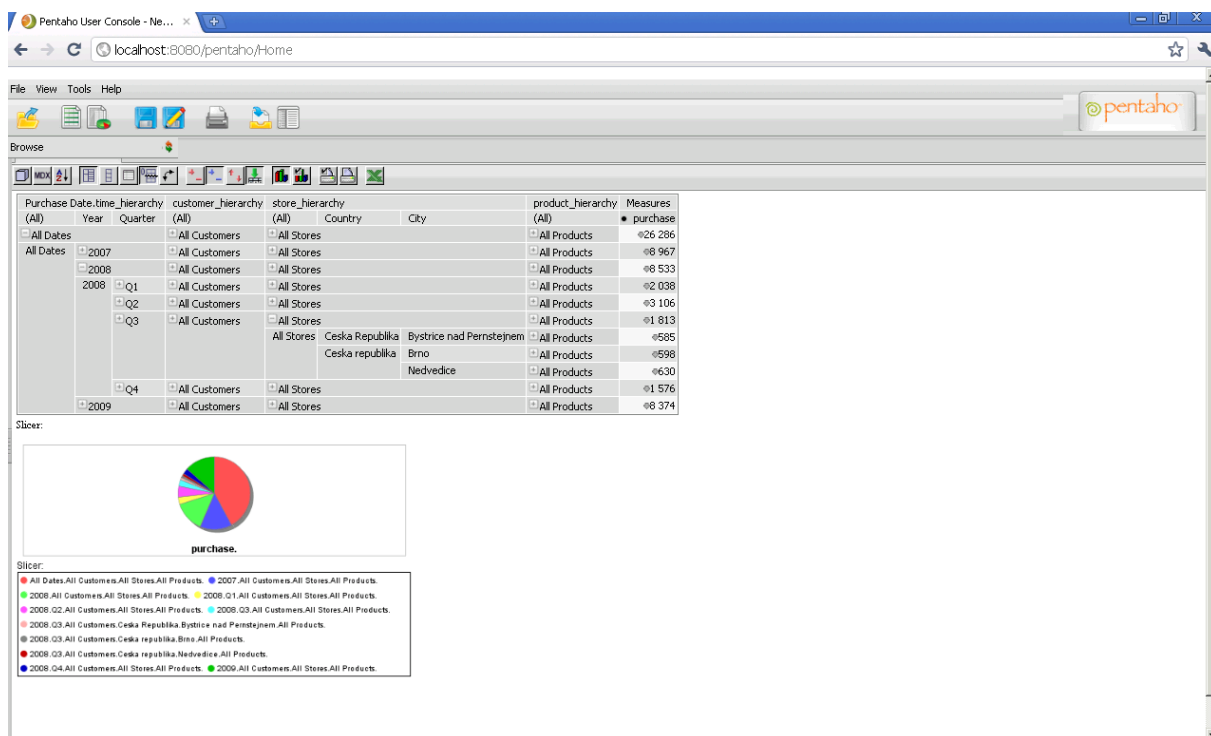
Obrázok D.2: Priebek tvorby reportu pomocou integrovaného wizaru



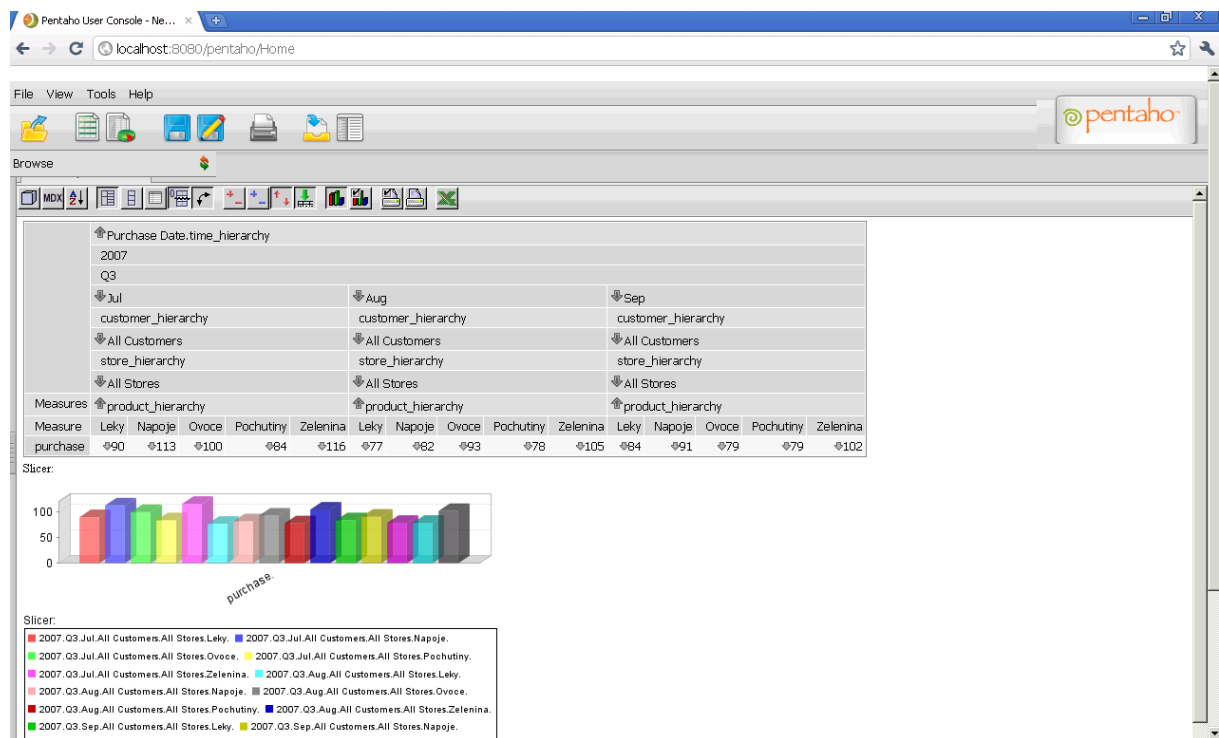
Obrázok D.3: Prostredie užívateľskej konzoly a prieskumníka dát



Obrázok D.4: Tvorba jednoduchéj analýzy interných dát



Obrázok D.5: Tvorba jednoduchéj analýzy interných dát



Obrázok D.6: Tvorba jednoduchéj analýzy interných dát