

**Česká zemědělská univerzita v Praze**

**Provozně ekonomická fakulta**

**Katedra informačních technologií**



**Diplomová práce**

**Implementace Business Intelligence**

**Radek Lebr**

© 2015 ČZU v Praze

# ČESKÁ ZEMĚDĚLSKÁ UNIVERZITA V PRAZE

Katedra informačních technologií

Provozně ekonomická fakulta

## ZADÁNÍ DIPLOMOVÉ PRÁCE

Radek Lebr

Systemové inženýrství a informatika

Název práce

**Implementace Business Intelligence**

Název anglicky

**Implementation of Business Intelligence**

---

### Cíle práce

Prvním cílem diplomové práce je vytvořit ucelenou rešerši o aplikacích typu Business Intelligence (dále jen BI). Při studiu literárních pramenů bude kladen důraz na kategorie BI systémů, dále na typické funkce BI aplikací a jejich význam pro soudobé organizace. Poznatky načerpané v rámci zpracování literární rešerše budou následně využity při tvorbě vlastní přidané hodnoty práce. Druhým cílem této diplomové práce bude zpracování případové studie, jejíž výstupem bude návrh optimalizace BI řešení ve zvoleném podniku. V úvodu případové studie student popíše výchozí situaci, provede analýzu informačních potřeb uživatelů BI na různých úrovních řízení a na základě této analýzy vyvodí výchozí neoptimální vlastnosti BI řešení a navrhne cesty k jejich odstranění. Přidanou hodnotou této případové studie tak bude posun od výchozího neoptimálního stavu do stavu cílového optimálního.

### Metodika

Při zpracování literární rešerše bude využito metody studia literárních pramenů. V rámci případové studie budou informace o výchozím stavu zjištěny metodou nestrukturovaných rozhovorů s klíčovými uživateli, dále bude provedena analýza materiálů poskytnutých společností, v rámci které probíhá optimalizace využití BI. Při zpracování případové studie bude kladen důraz na ochranu osobních dat a na ochranu citlivých obchodních informací poskytnutých předmětnou společností. Vyvození neoptimálních charakteristik BI řešení provede student na základě analýzy materiálů poskytnutých předmětnou společností a dále rozpořem výstupů z nestrukturovaných rozhovorů. Syntézou takto získaných poznatků získá autor podklady pro logickou dedukci návrhů na změny, které budou směřovat k optimalizaci BI řešení.

---

**Doporučený rozsah práce**

60 – 80 stran

---

**Doporučené zdroje informací**

Hubbard D. W.: How to Measure Anything: Finding the Value of Intangibles in Business. Wiley: 2 edition, April 19, 2010. 320 p. ISBN-10: 0470539399

Sabherwal R.: Business Intelligence. Wiley: 1 edition, March 1, 2010. 304 p. ISBN-10: 0470461705

Turban E., Sharda R., Dursun D., King D.: Business Intelligence. A Managerial Approach. Prentice Hall, 2 edition, July 28, 2010. 312 p. ISBN-10: 013610066X

---

**Předběžný termín obhajoby**

2015/06 (červen)

**Vedoucí práce**

Ing. Tomáš Rain, Ph.D.

---

Elektronicky schváleno dne 31. 10. 2014

**Ing. Jiří Vaněk, Ph.D.**

Vedoucí katedry

---

Elektronicky schváleno dne 11. 11. 2014

**Ing. Martin Pelikán, Ph.D.**

Děkan

V Praze dne 27. 02. 2015

### Čestné prohlášení

Prohlašuji, že svou diplomovou práci " Implementace Business Intelligence" jsem vypracoval samostatně pod vedením vedoucího diplomové práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou citovány v práci a uvedeny v seznamu literatury na konci práce. Jako autor uvedené diplomové práce dále prohlašuji, že jsem v souvislosti s jejím vytvořením neporušil autorská práva třetích osob.

V Praze dne 1. 3. 2015



---

## Poděkování

Rád bych touto cestou poděkoval Ing. Tomáši Rainovi, Ph.D. za odborné vedení při vypracování diplomové práce a Ing. Josefu Jurákovi za neúnavnou podporu při zpracování případové studie.

# Implementace Business Intelligence

---

## Implementation of Business Intelligence

### Souhrn

Tato diplomová práce se věnuje implementaci Business Intelligence (dále jen BI) v podnikovém prostředí. Úvodní část práce obsahuje cíle práce a metodiku, pomocí níž bude cílů dosaženo. Práce ve své teoretické části charakterizuje BI a postupně všechny jeho součásti a procesy. Nejdříve obecně vymezí pojem BI jako podpůrný nástroj pro klíčová rozhodování managementu, vysvětlí vztah Data – Informace – Znalost. Odliší OLAP jako součást BI od operativních procesů OLTP. Seznámí s hlavními přínosy BI. Dále se již detailněji věnuje jednotlivým komponentám - primárním zdrojům dat, datovému skladu, datovým tržištím, analytické a prezentační vrstvě, ETL procesu, který zajistí optimální přesun dat z primárních systémů do datového skladu. Práce detailně analyzuje reporting, jako součást BI a rozdělí reportovací nástroje do skupin podle různých kritérií. Následně navazuje s popisem datové kostky a OLAP modelu. Seznámí s dimenzemi, fakty, mírami a KPI. Dále popisuje termín Business Performance Management a jeho prezentační nástroje (např. Dashboard). Práce v teoretické kapitole implementace BI charakterizuje jednotlivé kroky procesu od úvodní studie a analýzy, návrhu přes samotnou implementaci až po zhodnocení celé implementace. Případová studie v druhé části práce pak aplikuje teoretické kroky implementace na firmu GTS Czech nad jejími IT prostředky, procesy a uživateli. Výstupem studie je funkční datová kostka.

### Summary

The master's thesis focuses on an implementation of Business Intelligence (hereinafter BI) within a business environment. Opening of the thesis includes an objective of the thesis and methodologies, through which the objectives will be achieved. The theoretical part of the thesis characterizes BI and one by one all of its parts and processes. Firstly, the thesis defines generally the term BI as a decision-supporting tool for managers, explains relation Data-Information-Knowledge. It distinguishes OLAP as a part of BI from

operational processes OLTP. It familiarize with main benefits of BI. Further, it deals with BI components in detail - primary data sources, data warehouse, data marts, analytical and presentation layer, ETL process, which ensures optimal data transport between primary systems and data warehouse. The thesis in detail analyses reporting as a component of BI and divides report tools into groups, according different criteria. Then it continues with a description of a data cube and OLAP model. It makes dimensions, facts, measures and KPI's familiar. Next step is description of the term Business Performance Management and its presentation tool (e.g. Dashboard). The thesis during a theoretical chapter of an implementation of BI characterizes every single step of the process – from introductory study and analysis, design through actual implementation to evaluation of the whole implementation. Case study in the second part applies the theoretical implementation steps to GTS Czech Company over its IT resources, processes and users. The result of the case study is the functional data cube.

**Klíčová slova:** Business Intelligence, podpora pro rozhodování, Reporting Services, Analysis Services, OLAP, metadata, datová kostka, DWH, ETL, Dashboard

**Keywords:** Business Intelligence, Decision Support, Reporting Services, Analysis Services, OLAP, metadata, data cube, DWH, ETL, Dashboard

## Seznam zkratek:

ACRM	Analytical CRM
API	Application Programming Interface
BI	Business Intelligence
BIRT	Business Intelligence and Reporting Tools
BPM	Business Performance Management
CAPEX	Capital Expenditures
CDR	Call Detail Records
CPM	Corporate Performance Management
CPU	Central Processing Unit
CRM	Customer Relationship Management
CSR	Central Service Repository
ČTÚ	Český telekomunikační úřad
DAX	Data Analysis Expressions
DB	Database
DMA	Data Marts
DSA	Data Staging Area
DWH	Data Warehouse
EAI	Enterprise Application Integration
ETL	Extract, Transform, Load
EBITDA	Earn Before Interest, Taxes, Depreciation and Amortization
HW	Hardware
HDD	Hard Disc Drive
IČ	identifikační číslo
IS	Information Services
IVR	Interactive voice response
KPI	Key Performance Indicator
LE	Large Enterprise
LTS	Long Term Support
MD	Manday
MDX	Multidimensional Expressions
NACE	Nomenclature statistique des activités économiques dans la Communauté européenne
OLAP	Online Analytical Processing
OPEX	Operating Expense
OS	operační systém
PPS	Performance Point Services
ROI	Return of Investments
RFID	Radio-frequency identification devices
SCM	Supply Chain Management
SLA	Service Level Agreement
SME	Small and Medium Enterprise



SQL	Structured Query Language
SŘBD	system řízení báze dat
SSAS	SQL Server Analysis Services
SSIS	SQL Server Integration Services
SSRS	SQL Server Reporting Services
SW	Software
WHSL	Wholesale
VoIP	Voice over Internet Protocol

## Obsah

1	Úvod.....	12
2	Cíl práce a metodika .....	13
2.1	Cíl práce .....	13
2.2	Metodika .....	13
3	Teoretická východiska .....	14
3.1	Úvod do BI.....	14
3.1.1	Co je BI? .....	14
3.1.2	Přínosy BI .....	15
3.1.3	Podstata BI.....	16
3.1.4	Komponenty a členění BI .....	17
3.2	Primární systémy .....	20
3.3	Datový sklad .....	20
3.3.1	Charakteristika datového skladu.....	20
3.3.2	Tvorba DWH .....	23
3.4	ETL .....	26
3.5	Metadata.....	28
3.6	Reporting .....	30
3.6.1	Charakteristika Reportingu .....	30
3.6.2	Statický a dynamický Reporting.....	30
3.6.3	Pravidelnost Reportingu a Analýz .....	30
3.7	Reportovací, dotazovací a analytické nástroje BI.....	32
3.7.1	Podle typu použití .....	33
3.7.2	Licenční dělení.....	34
3.7.3	Umístění a přístup.....	36
3.7.4	Segmentové dělení BI nástrojů .....	38
3.7.5	Shrnutí dělení BI nástrojů .....	40
3.8	Multidimenzionální modelování dat a OLAP.....	43
3.8.1	Měření dat .....	43
3.8.2	Dimenzionální modelování.....	44
3.8.3	OLAP kostky .....	47
3.8.4	Druhy OLAP.....	48
3.9	Analytické aplikace BI.....	49
3.9.1	Analytické CRM (ACRM).....	49
3.9.2	Web Analytics.....	50
3.9.3	Competitive Intelligence (CI) .....	50
3.9.4	Business Activity Monitoring (BAM) .....	50
3.10	BPM (CPM).....	52
3.10.1	Úvod do BPM .....	52
3.10.2	Scorecard a Dashboard .....	53
3.11	Základní procesy implementace BI .....	55
3.11.1	Zpracování úvodní studie.....	58
3.11.2	Analýza .....	59
3.11.3	Design BI řešení.....	59
3.11.4	Návrh technologické platformy .....	60
3.11.5	Návrh ETL .....	61
3.11.6	Implementace a konsolidace .....	61
3.11.7	Migrace, DWH, infrastruktura.....	62

3.11.8	Hodnocení implementace BI .....	62
4	Případová studie.....	63
4.1	Představení společnosti GTS Czech .....	63
4.1.1	Poskytované služby.....	63
4.1.2	Profil společnosti .....	63
4.1.3	Situace na trhu .....	64
4.2	Plánování a příprava studie.....	64
4.3	Nestrukturované pohovory s klíčovými uživateli a jejich vyhodnocení.....	66
4.3.1	Anketa.....	66
4.3.2	Nestrukturovaný pohovor .....	68
4.3.3	Vyhodnocení pohovoru a ankety .....	69
4.4	Analýza stávajícího stavu BI v GTS Czech.....	72
4.4.1	Analýza DWH a ETL .....	72
4.4.2	Analýza reportovacích nástrojů .....	76
4.4.3	Analýza technologické platformy .....	80
4.5	Výstup z analýzy a návrhy řešení .....	82
4.5.1	Návrh řešení pro DWH a ETL.....	82
4.5.2	Návrh řešení pro reportovací nástroje.....	82
4.5.3	Návrh řešení pro technologickou platformu .....	83
4.5.4	Náklady pro DWH, ETL, technologie a reportovací nástroje .....	84
4.5.5	Souhrnný návrh řešení .....	85
4.6	Design BI řešení.....	88
4.6.1	Analýza dostupných datových zdrojů v relačních databázích .....	88
4.6.2	Dimenze .....	89
4.6.3	Fakta.....	90
4.6.4	KPI.....	91
4.6.5	Příprava podkladové relační databáze .....	92
4.7	Implementace do SSAS .....	97
4.8	Návrh ETL .....	100
4.8.1	SSIS řešení .....	100
4.8.2	Řešení s uloženými procedurami .....	101
4.8.3	Popis ETL procesu.....	101
4.9	Autorizace a autentizace .....	103
4.10	Školení uživatelů, testy a předání BI řešení do produkce.....	103
4.11	Zhodnocení výsledků implementace a další doporučení .....	104
5	Závěr .....	105
	Seznam literatury a použitých zdrojů .....	107
	Seznam obrázků.....	110
	Seznam tabulek .....	110
	Seznam příloh .....	111
	Přílohy.....	112

# 1 Úvod

Manažeři se špatně rozhodují v klíčových situacích. Firmy přichází o peníze a zákazníky, kteří jsou nespokojeni. Toto je stav, který nastává ve společnostech při nedostatečném nebo žádném zdroji informací o zákaznících, obchodu, trhu. Nebo ten zdroj i existuje, ale firma z něj neumí čerpat. Přitom v dnešní době je většina běžných trhů (například trh fixních telekomunikací, o kterém bude řeč v případové studii) již vysoce saturována konkurencí a bojuje se o každého zákazníka. A každá konkurenční výhoda může rozhodovat o přežití na trhu.

Řešení existují a jedno z nich je vhodné využívání Business Intelligence (dále jen BI) nad podnikovými IT systémy a datovými zdroji. Běžné reporty (např. MS Excel) jsou při velkém množství zdrojových dat zdlouhavé, omezené na úzké skupiny uživatelů, bez možnosti pokročilých filtrů. Každé oddělení si navíc vykládá získané informace jinak a datům nerozumí a při jejich prezentaci před kolegy pak dochází k nedorozuměním. Výroba každého nového reportu také zabere nějaký čas pro IT zaměstnance. BI všechny tyto nedostatky řeší. Management ví, že nějaké BI existuje, ale... vždy si najdou nějaké ale, proč BI nepotřebují. Je to drahé, s nejistým a neměřitelným výsledkem, nebo prostě třeba nemají zrovna čas.

Tématem této diplomové práce je Implementace BI. Vysvětlím pojem BI, jeho komponenty, jeho využití a hlavně jeho výhody. Praktická část práce je zvláště zaměřena na stěžejní část BI – datovou kostku. Očekávám, že většina uživatelů reportů netuší, co BI je, co dělá, k čemu je dobré a v čem by zrovna jim mohlo usnadnit práci.

Proto jsem si vybral toto téma BI Implementace – v této práci chci pomoci změnit neznalost uživatele na znalost tedy na jeho (konkurenční) výhodu díky možnostem BI.

## **2 Cíl práce a metodika**

### **2.1 Cíl práce**

Prvním cílem práce je vytvořit ucelenou rešerši o aplikacích typu BI. Při studiu literárních pramenů bude kladen důraz na kategorie BI systémů, dále na typické funkce BI aplikací a jejich význam pro soudobé organizace. Poznatky, načerpané v rámci zpracování literární rešerše budou následně využity při tvorbě přidané hodnoty práce.

Druhým cílem této diplomové práce bude zpracování případové studie, jejímž výstupem bude návrh optimalizace BI řešení ve zvoleném podniku. V úvodu případové studie popíšu výchozí situaci, provedu analýzu informačních potřeb uživatelů BI na různých úrovních řízení a na základě této analýzy vyvodím výchozí neoptimální vlastnosti BI řešení a navrhne cesty k jejich odstranění. Přidanou hodnotou této případové studie tak bude posun od výchozího, neoptimálního, stavu do stavu cílového – optimálního.

### **2.2 Metodika**

Při zpracování literární rešerše bude využito metody studia literárních pramenů, S ohledem na zkoumanou problematiku bude využito jak odborných publikací, tak i elektronických zdrojů. Literární rešerše bude teoretickým východiskem, na základě kterého bude zpracovaná samotná případová studie.

V rámci případové studie budou informace o výchozím stavu zjištěny metodou nestrukturovaných rozhovorů s klíčovými uživateli, dále bude provedena analýza materiálů poskytnutých společností, v rámci které probíhá optimalizace využití BI. Při zpracování případové studie bude kladen důraz na ochranu osobních dat a na ochranu citlivých obchodních informací, poskytnutých předmětnou společností.

Vyvození neoptimálních charakteristik BI řešení provedu na základě analýzy materiálů, poskytnutých předmětnou společností a dále rozpořem výstupů z nestrukturovaných rozhovorů. Syntézou takto získaných poznatků získám podklady pro logickou dedukci návrhů na změny, které budou směřovat k optimalizaci BI řešení.

## 3 Teoretická východiska

Pojem BI obsáhne (kromě typických prvků) velmi širokou oblast IT nástrojů a standardů.

V této teoretické části se soustředím více na ty části BI, které následně v části praktické budou analyzovány, hodnoceny a nad nimiž se bude implementace provádět.

### 3.1 Úvod do BI

#### 3.1.1 Co je BI?

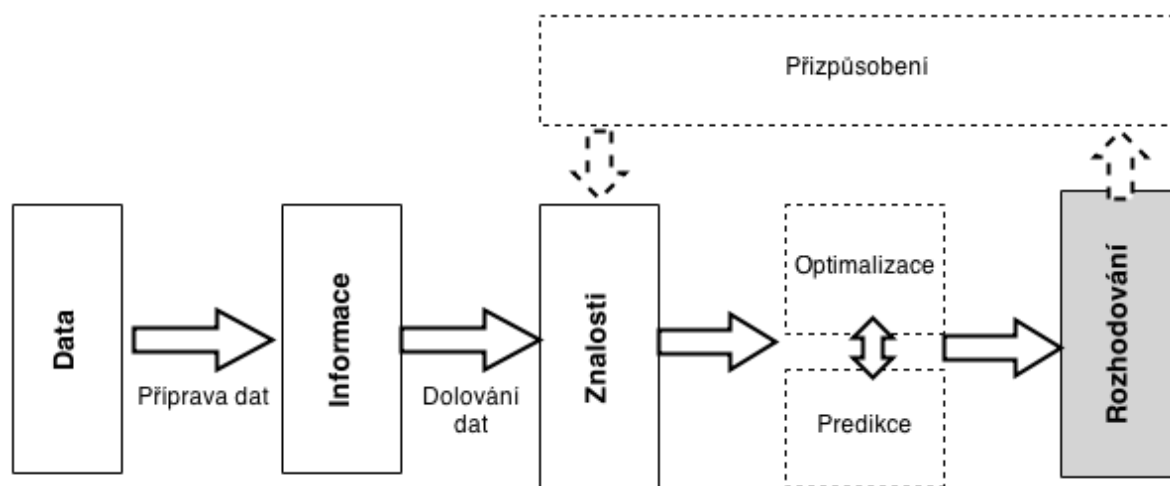
Kapitola představí BI, jeho základní komponenty, členění a nástroje. Vysvětlí k čemu je BI vlastně dobré a kdy je potřeba implementovat BI.

*„Termín Business Intelligence poprvé definoval Howard Dresner ze společnosti Gartner Group jako množinu konceptů a metodik, které zlepšují rozhodovací proces za použití metrik, nebo systémů založených na metrikách. Je to proces transformace údajů na informace a převod těchto informací na poznatky prostřednictvím objevování. Účelem procesu je konvertovat velké objemy údajů na poznatky, které jsou potřebné pro koncové uživatele. Tyto požadavky lze potom efektivně použít například v procesu rozhodování[1].“*

Tedy cílem BI je pomoci rozhodnout manažerům, zaměstnancům, akcionářům, běžným uživatelům nějaký problém na základě kvalitně zpracovaných a správně zobrazených dat z databází či jiných dalších zdrojů. Obrázek 3.1 popisuje základní vztah mezi (na začátku) hrubými, nezpracovanými daty a (na konci) znalostí, poznáním a prostorem pro rozhodnutí.

- *„Data jsou shromažďována na určité časové bázi ve formě bitů, čísel, symbolů nebo třeba objektů.*
- *Informace jsou pak setříděná data, tj. předpřipravena, vyčištěna, strukturována, zbavena duplicit.*
- *Znalost lze vyložit jako sjednocené informace, kdy získáme fakta a vazby, které jsme pochopili, objevili, naučili se[2].“*

Dále je pak ještě možné učit se, proces před finálním rozhodováním zlepšovat a optimalizovat. Je to vlastně poměrně běžnou praxí managementu, po rozhodnutí znovu analyzovat a kontrolovat své předchozí kroky, tak, aby příští rozhodnutí bylo provedeno lépe a s větší jistotou správné volby.



**Obrázek 3.1 Adaptivní BI**

Zdroj: MICHALEWICZ, Zbigniew. *Adaptive business intelligence*. New York: Springer, c2007, xiii, 246 p. ISBN 978-354-0329-282.

### 3.1.2 Přínosy BI

Podle průzkumu BARC BI SURVEY 2013 se hlavní přínosy správně implementovaného BI<sup>1</sup> dají (v pořadí od nejdůležitějšího podle hodnocení dotazovaných) shrnout následovně:

- Rychlejší reporting, analýzy a plánování
- Přesnější reporting, analýzy a plánování
- Lepší podpora manažerského rozhodování (je rychlejší a přesnější)
- Zlepšení kvality dat
- Zvýšení spokojenosti zaměstnanců s nástroji
- Zvýšení spokojenosti zákazníků
- Nárůst tržeb
- Snížení nákladů na externí IT služby
- Úspora dalších nákladů mimo IT oblast
- Redukce stavu zaměstnanců společnosti
- Snížení stavu zaměstnanců IT podpory

<sup>1</sup> Zdroj: BUSINESS APPLICATION RESEARCH CENTER. The BI Survey 13 [online]. 2013 [cit. 2014-08-16]. Dostupné z: <http://www.bi-survey.com>

### 3.1.3 Podstata BI

V IT Existují dva hlavní typy datového zpracování nebo úloh datového charakteru, tedy ukládání a přenosu do a z databází[3]. První je OLTP – transakční procesy, kterým **BI není** a druhý OLAP – analytické a plánovací zpracování, kterým **BI přesně je**.

**OLTP** si lze představit jako procesy, které datově zajišťují každodenní operativu firmy[4].

- Okamžitý přístup k aktuálním datům (na službě, objednávce, účtu, kartě pacienta).
- Možnost dostupná data modifikovat a na základě těchto úprav budou automaticky aktualizovány návazná data v systémech (např. změna stavu kreditu zákazníka podle přijaté platby).
- Tvorba nových záznamů, dat nebo samostatné vytváření dat na základě souvisejících změn v systému (generování faktury, objednávky, zahájení procesu zřízení služby na základě uzavřené smlouvy atp.)

**OLAP** pak jsou analytické procesy, vycházející z datového zdroje<sup>2</sup> DWH[4].

- Umožní hodnotit KPI (firemní ukazatele) např. EBITDA, churn a intake<sup>3</sup> zákazníků, mezipodnikové (v případě více zemí) náklady a výnosy a samozřejmě celkové tržby, náklady, hrubý zisk, ROI atd.
- KPI musí být dostupné v pohledech podle různých hledisek a jejich kombinací (dimenzionální a faktový přístup viz kapitola o Multidimenzionálním modelování), tj. uživatel má možnost modelovat nad daty. Příkladem je měsíční relativní hodnocení tržby/churn pro jednotlivé obchodní týmy, mezi které jsou zákazníci rozdělení<sup>4</sup>.
- Vysoká pružnost odpovídající potřebám analytického managementu.
- Analýza, vývoj a predikce podnikových ukazatelů v čase s použitím hierarchie analyzovaných hodnot.

Transakční procesy data vlastně vytvářejí a v databázích je zpřístupňují[4]. Kdežto analytické procesy s vytvořenými daty a databázemi pouze pracují a převádějí do potřebných struktur analytických nástrojů a aplikací. Je třeba zmínit, že pro OLAP není

---

<sup>2</sup> Datový sklad je charakterizován v kapitole 3.3

<sup>3</sup> Churn neboli odliv, odchod zákazníků. Intake neboli nová pravidelná tržba

<sup>4</sup> Toto je skutečný příklad a slouží jako jeden z ukazatelů pro vyplacení provizí obchodníkům



realtime<sup>5</sup> pohled na data potřebný, ale tato situace se poslední době mění. Management vyžaduje tzv. near-to-realtime BI – např. do patnácti minut od zápisu do primárního systému tuto informaci vidět v reportovací a analytické aplikaci. Toto samozřejmě přináší nové výzvy a postupy pro BI developery a správce.

### 3.1.4 Komponenty a členění BI

Na BI lze nahlížet z několika různých úhlů pohledu. Níže představím běžné dva, podle kterých zhruba budou následovat další kapitoly.

1. **Uživatelský přístup** – BI se skládá z komponent, které spravují a zpracovávají data do požadovaných výstupů[3]:

**Prostředí datového skladu** – slouží jako zdroj dat a úložiště. O jeho chod a výstavbu a vývoj se starají technické složky firmy (ať už běžní technici při instalaci HW v datovém centru/serverové místnosti, tak databázoví specialisté).

**Analytické aplikace** – koncoví uživatelé mají k dispozici data a provádí analýzy a získávají informace pomocí různých nástrojů. Sem potom patří vytváření a správa reportů (dále jen reporting), jednorázové dotazy nad databázemi (Query), multidimenzionální analýzy OLAP a dolování dat (Data mining).

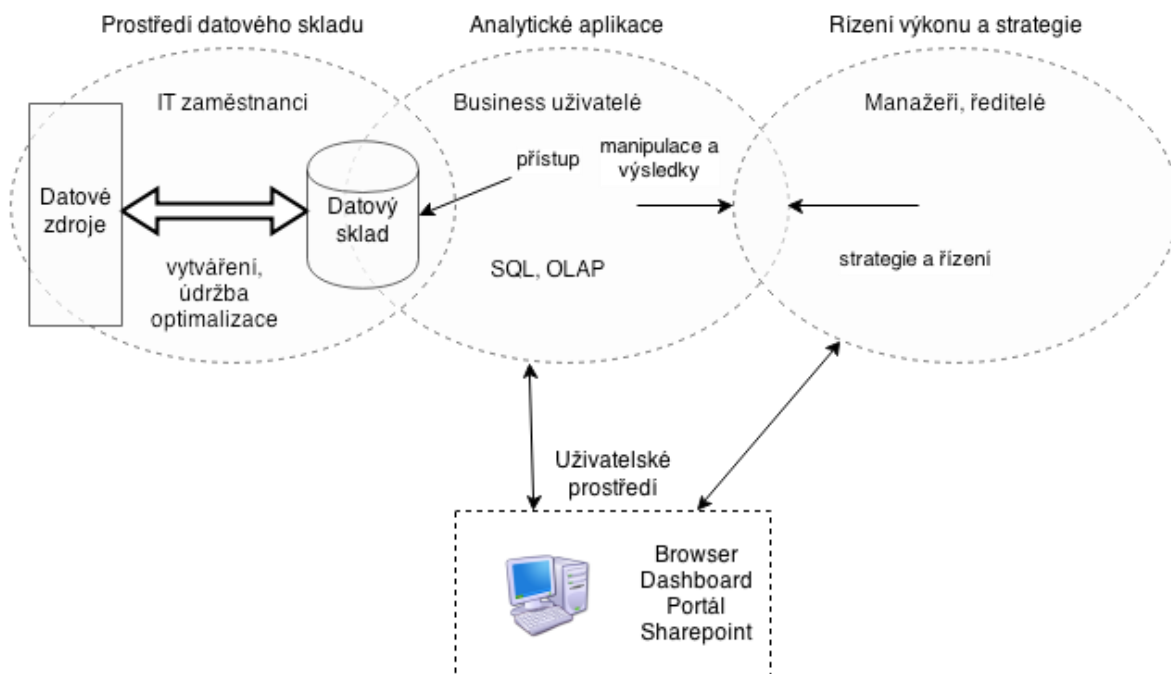
**Řízení výkonu a strategie** – také označována jako BPM nebo CPM. Jde o propojení a jistou nadstavbu nad běžnými BI nástroji a architekturou. Využívá k tomu BI aplikace, které se specializovaly na jednotlivé oblasti KPI a integruje je. Zahrnuje procesy, technologie, metodiky a měření a obohacuje je o řídicí (management) pohled[4].

**Uživatelské prostředí** – prezentační vrstva pro koncové uživatele a to na úrovni analýzy i řízení. Typickým médiem je dnes webový prohlížeč. Dále pak specializované analytické

---

<sup>5</sup> Realtime neboli aktuální data, nejvíce se blíží skutečnosti a je možné je s okamžitým efektem měnit.

nástroje jak Scorecardy, Dashboardy<sup>6</sup>. Patří sem také firemní Portály, Intranet a například SharePoint jako úložiště firemních a týmových informací.



**Obrázek 3.2** Uživatelská architektura BI

Zdroj: TURBAN, Efraim. *Business intelligence: a managerial approach*. 2nd ed. Boston: Prentice Hall, c2011, xx, 292 p. ISBN 01-361-0066-X.

## 2. Přístup dle vazeb BI komponent – lze dělit do těchto vrstev[5]:

**Zdrojové systémy** – primární systémy a zdroje dat ze kterých BI aplikace čerpají informace.

**Datová transformace** – ETL a EAI komponenty pro sběr a přenos dat z primárních systémů do další vrstvy BI.

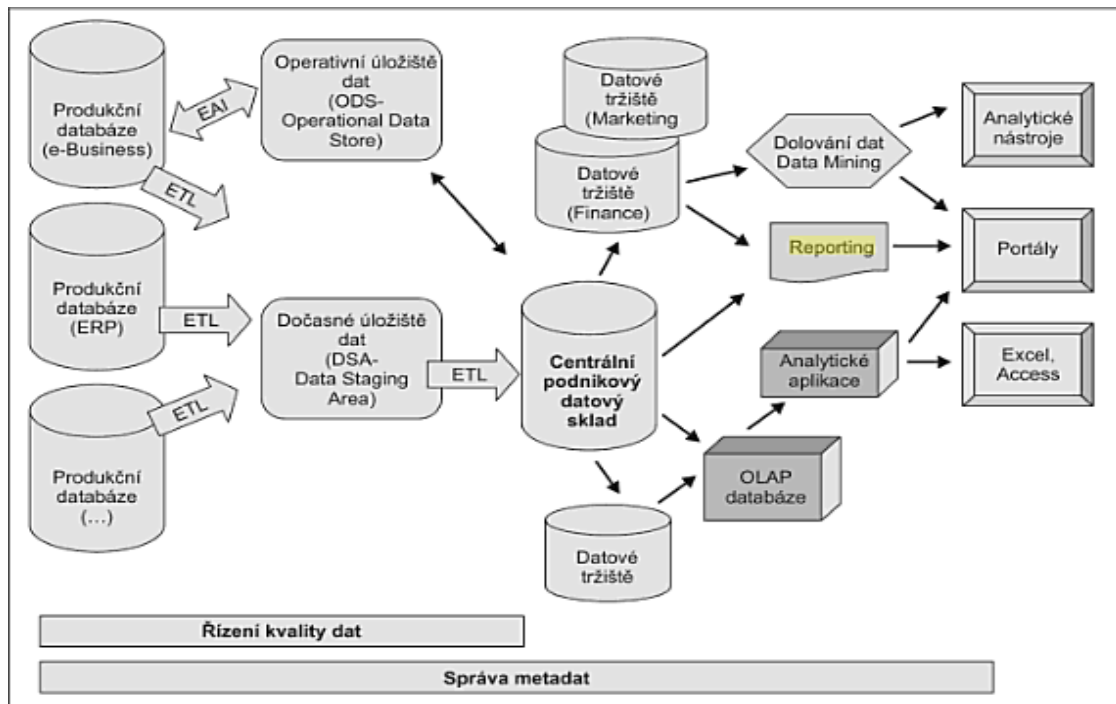
**Ukládání dat** – obsahuje různé typy databází. Zajišťuje tedy správu uložených dat. Patří sem datové sklady, datová tržiště operativní a dočasná úložiště dat.

**Analýza dat** – komponenty pro analytickou a reportovací činnost. Zahrnuje Reporting, OLAP a dolování dat.

**Prezentační vrstva** – slouží ke komunikaci s uživateli (obdobně jako uživatelské prostředí v předchozím modelu).

<sup>6</sup> Bude charakterizováno v kapitole 3.10

Následující obrázek popisuje vazby mezi jednotlivými vrstvami.



**Obrázek 3.3 BI schema**

Zdroj: GÁLA, Libor, Jan POUR a Zuzana ŠEDIVÁ. Podniková informatika. 2., přeprac. a aktualiz. vyd. Praha: Grada, 2009, 496 s. Expert (Grada). ISBN 978-80-247-2615-1

K tomuto přístupu je ještě zmínit doplnit tři oblasti pro samotnou správu dat manipulaci s nimi[5]:

- Nástroje pro zajištění datové kvality a datové čistoty např. Microsoft SSIS
- Nástroje pro správu metadat např. Microsoft SQL Server a jeho Management Studio.
- Technické zkušenosti jednotlivců a týmů, potřebné k implementaci komplexního BI. Opět to zahrnuje jak oblasti fyzické instalace, tak programování a schopnosti pracovat s vývojovými nástroji.

## 3.2 Primární systémy

Primární nebo zdrojové systémy jsou databáze, ze kterých pochází veškerá původní data [4]. Jejich charakteristikou je realtime práce s daty. Většinou jde tedy o OLTP databáze a typickými představiteli jsou ERP<sup>7</sup>, SCM<sup>8</sup> a další systémy, které podporují operativu divizí pro dané oblasti a pomáhají nižšímu managementu řídit tyto oblasti. Primární zdroje lze rozdělit na interní (zdroje vlastního podniku nebo podniků v rámci skupiny) a externí (databáze analytických společností, Excel nebo csv soubory statistického úřadu, veřejné správy a další). Data jsou v primárních systémech téměř vždy různorodá (struktura, formát dat) a často bývá náročné i získat metadata (tedy pochopit správný význam dat) k těmto zdrojům, zvláště pokud se jedná o externí zdroje.

## 3.3 Datový sklad

V této kapitole bude charakterizován Datový sklad (Data Warehouse), dále DWH. Jeho možné členění podle konkrétních potřeb firmy a také podle charakteru dat.

### 3.3.1 Charakteristika datového skladu

Datový sklad je logické místo v IT infrastruktuře firmy. Jsou zde uchovávána současná a historická data z oblastí podnikání firmy, která jsou potřebná pro klíčová manažerská rozhodování a reporting[3]. Data jsou obvykle uchovávána v takové struktuře, aby byla dostupná pro následné analytické činnosti, OLAP, data mining, databázové dotazy a reporting. Pro DWH by měly být fundamentální následující charakteristiky[6]:

- **Subjektově vymezený** – data budou detailně organizována podle klíče (subjektu), jako jsou produktové portfolio, obchod, zákazníci, poruchy. Naopak nerelevantní data (data, která nenapomáhají získání potřebných informací) vůbec nebudou zahrnuta. Tím se DWH liší od operační databáze ODS, která bývá zaměřena na produkt takový a zajištění včasné aktualizace transakcí. DWH je tedy mnohem více komplexní v náhledu na firemní fungování.

---

<sup>7</sup> Podnikový IS integrující správu procesů a dat z různých oblastí a nástrojů firemní činnosti, jako finance, nákup, logistika, produkt atd[1]. Na ERP se specializuje např. německá firma SAP.

<sup>8</sup> Řízení vztahu s dodavateli, umožní kontrolu celého nákupního procesu[3]

- **Integrovaný** – u DWH se očekává konsolidování dat z různých datových zdrojů. Je tedy třeba zajistit konsistentní formát jednotek (například jednotky času, stejnou měnu apod.) a také vyřešit konflikty v názvech (V jedné db bude nazván produkt „GTS Internet“, ve druhé pak „Novera Internet“, což je původní název produktu).
- **Historizující (časové řady)** – čas je jedna z důležitých dimenzí, které musí DWH podporovat. DWH uchovává historii a díky tomu lze pomocí analýzy získat informace o trendech, odchylkách, predikovat vývoj. Časové otisky se současně sbírají v různých intervalech (dny, týdny, měsíce), tak aby analýzy byly co nejpřesnější pro daný účel.
- **Stálý** – Jakmile jsou data vložena do DWH, uživatelé ani procedury je nemohou změnit. V případě, že dojde k zapsání nového otisku dat, starší (předchozí časový otisk) data jsou v DWH označena jako neplatná/zastaralá.

DWH lze logicky rozdělit do čtyř základních typů[3]: Datové tržiště, Centrální podnikový datový sklad, Operativní úložiště dat a Dočasné úložiště dat. Zatímco první dva jsou charakteristikou dlouhodobého uložení dat skutečně datové sklady, zbývající dva slouží jako přechodné úložiště.

**Datové tržiště** (Data Mart, DMA) – je na rozdíl od DWH zaměřený na konkrétní oblast firemního působení (Obchod, Zřizování služeb, Marketing, Customer Care), nebo celou určitou divizi, pobočku[4]. Je tedy přirozeně menší než DWH. V některých firmách může být vnímán jako součást DWH.

Tři hlavní výhody datových tržišť oproti DWH jsou menší náklady na fyzické vybudování databáze (server, menší prostor, paměť), stejně tak i menší časová náročnost BI specialisty a klíčových uživatelů tržiště na definování struktury tržiště. Třetí výhodou je pak snížení rizika z globálního na lokální při zavádění a následně i při provozu. Pokud dojde k výpadku jednoho tržiště (databáze), ostatní datové sklady jsou stále v provozu a dostupné uživatelům BI.

Datové tržiště je možné dále rozdělit na závislé a nezávislé[3].

Závislé přímo vychází z DWH a při tvorbě tržiště musí být nejdříve existovat DWH. Toto je typický případ ve velkých firmách, které již náklady na vybudování DWH vynaložily.

Výhodou je, že uživatel vždy vidí stejná a konsistentní data v tržišti i v DWH. Pokud tedy dvě různé divize reportují nad společným produktem různá čísla a každý používá svoje tržiště, rozdíl bude způsoben spíše různými náhledy na informace jejich výkladem než datovými zdroji.

Opačným typem je pak nezávislé datové tržiště[3]. Takové se používá v malých firmách, které nevládní DWH (tedy jej vlastně nahrazuje) nebo se jedná o zcela specifickou oblast (např. strategicky významnou).

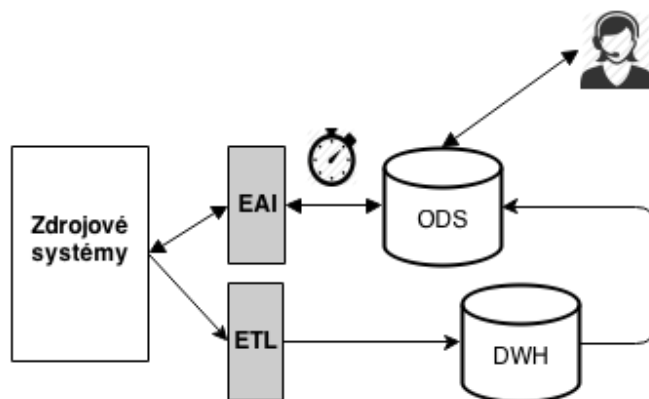
**Centrální podnikový datový sklad** (Enterprise Data Warehouse, EDW) – svým zaměřením a rozsahem pokrývá většinu oblastí firmy a to jak působením, tak i BI[3]. EDW pak bývá také důležitým zastřešujícím BI prvkem pro nadnárodní společnosti. Běžně pak pokrývá různé aplikační oblasti BI např.<sup>9</sup>: CRM (customer relationship management), BPM (business performance management), oblast lidských zdrojů, vztahy s dodavateli, logistiku, finance a kontroling, výrobu, marketing, BAM (business activity monitoring).

**Operativní úložiště dat** (Operation Data Sources, ODS) – jedná se o datové uložení, kde jsou data z primárních systémů ukládána pomocí nástrojů EAI<sup>10</sup> v téměř reálném čase a uživatelé tedy získávají vždy aktuální informace[1]. Minimální doba odezvy je díky velikosti databáze, protože v ODS se udržuje vždy pouze poslední, aktuální verze dat, ostatní data jsou smazána, resp. přesunuta do DWH. Typickým příkladem použití je systém podpory on-line komunikace se zákazníkem, kdy potřebuje operátor okamžitě reagovat (Call Centra, Řízení letového provozu apod.). Zároveň ale lze aktuální, ale omezené informace z ODS zkombinovat s DWH zdroji, kde bude komplexní doplněk, např. o zákazníkovi.

---

<sup>9</sup> Aplikační oblasti a systémy pro podporu rozhodování budou přiblíženy v následujících podkapitolách

<sup>10</sup> Enterprise Application Integration platforma sjednocuje rozhraní různých primárních zdrojů a snižuje odezvy při obousměrných přenosech mezi ODS a EAI[1].



**Obrázek 3.4 ODS funkce**

Zdroj: GÁLA, Libor, Jan POUR a Zuzana ŠEDIVÁ. Podniková informatika. 2., přeprac. a aktualiz. vyd. Praha: Grada, 2009, 496 s. Expert (Grada). ISBN 978-80-247-2615-1

**Dočasné úložiště dat** (Data Staging Area, DSA) – slouží k dočasnému uložení extrahovaných dat z primárních systémů, tedy předtím, než jsou data ve finální struktuře a formátu uložena do DWH nebo do DMA[1]. DSA, stejně jako ODS jsou nepovinné komponenty, je možné mít BI řešení pouze s primárními databázemi a pak DWH. Data jsou do DSA ukládána po ETL<sup>11</sup> procesu existují pouze v aktuální verzi.

### 3.3.2 Tvorba DWH

Při stavbě nového DWH je možné využít tři základní přístupy[3]:

**Inmonův model** – jedná se o původní přístup k vytváření DWH a předpokládá se, že již při vytvoření centrálního datového skladu bude toto řešení natolik komplexní, že nebude třeba vytvářet žádné další DMA. DMA, vytvořená na počátku, jsou závislá na centrálním DWH. Hierarchie toho modelu je shora-dolů.

**Kimballův model** – centrální datový sklad se skládá z více DMA a je postupně rozšiřován o další tržiště (doplňován). Každé DMA je ale samostatně schopné fungovat. Hierarchie toho modelu je zdola-nahoru.

<sup>11</sup> ETL proces popisuje kapitola 3.4

**Tabulka 3.1 Srovnání hierarchických přístupů**

	Inmon	Kimball
Hierarchie	Shora-dolů	Zdola-nahoru
Architektura	Data v EDW jsou atomická a mohou být šířena do lokálních databází	Přes všechny DMA je zajištěna konsistence EDW pomocí společné datové sběrnice a jednotného procesu pro datové modelování.
Složitost přístupu	Velká	Spíše jednoduchá
Metodika vývoje SW	Spirálový model <sup>12</sup>	Základní <sup>13</sup> objektově orientovaný model, odklon od relačního databázového systému
Příprava db designu	Důkladná a náročná	Lehká – mírně obtížná
Duplicitní komponenty	Nejsou vzhledem k centrálnímu řešení	Existují
Datová orientace	Subjektová orientace	Zaměření na procesy
Databázové nástroje	E-R diagramy, data-flow diagramy	Dimenzionální modelování
Dostupnost pro koncové uživatele	Minimální	Vysoká
Primární uživatelé	IT profesionálové	Koncoví uživatelé
Finanční zátěž	CAPEX (investice)	OPEX (provoz)
Firemní zařazení	Hlavní firemní informační zdroj	Datový sluha s vhodnou transformací dat z primárních zdrojů
Cíl	Solidní a robustní řešení, založené na ověřených databázových metodách a postupech	Řešení, které je lehce dostupné a snadno použitelné pro koncového uživatele a zároveň při nižší odezvě umožní uživatelům přímé databázové dotazy.
Rizika	Před dokončením implementace se změní primární zdroje/ požadavky uživatelů	U velkých organizací (více zemí) nelze sdílet jednotné číselníky a pravidla (příliš unikátní). Zaniká tím výhoda konsolidace.
Rizika 2	Drill-down je náročnější, přechází se mezi různými vrstvami(komponentami)	Jednotná reportingová vrstva nelze vybudovat

Zdroj: TURBAN, Efraim. Business intelligence: a managerial approach. 2nd ed. Boston: Prentice Hall, c2011, xx, 292 p. ISBN 01-361-0066-X. a NOVOTNÝ, Ota. Business intelligence: jak využít bohatství ve vašich datech. 1. vyd. Praha: Grada, 2005, 254 s. ISBN 80-247-1094-3.

<sup>12</sup> **Spirálový model** náleží do skupiny tzv. přístupů řízených riziky. To znamená, že postup do další fáze závisí na důsledně provedené analýze všech rizik a možných problémů. Zpočátku se vývoj provádí na základě hrubé specifikace požadavků, v pozdějších fázích je tato specifikace i po konzultaci se zákazníkem postupně upřesňována[20].

<sup>13</sup> Koncept, Analýza, Design, Implementace

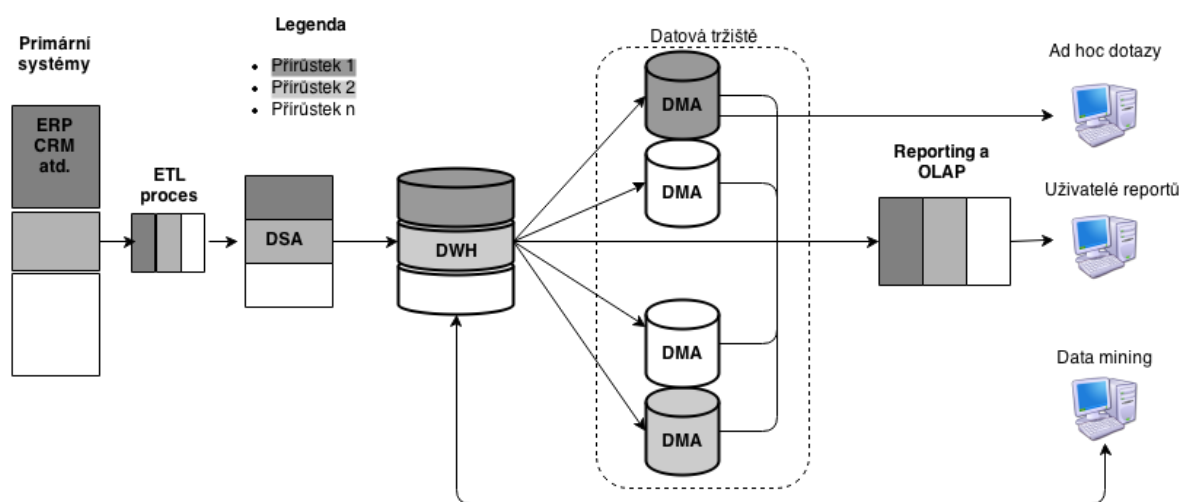


Třetí způsob se nazývá **Přírůstkový přístup**[6]. U tohoto přístupu se předpokládá, že před vytvořením DWH existuje již celková koncepce BI řešení. Všechny uživatelské požadavky jsou zaznamenány, architektura je rozvržena, jsou stanoveny priority a časový harmonogram. Následně je tato koncepce implementována postupně v časově a finančně omezených krocích. Každý krok (podobně jako u Kimballova modelu) je komplexní a zasáhne všechny potřebné oblasti BI (Kimball model ale změny aplikuje nad právě jedním DMA). Obsáhlost kroků je pak dána finančními a časovými možnostmi firmy [5]. Po kompletní implementaci pak vznikne Inmonův model, kde jsou DMA plně závislá na centrálním DWH.

Výhody tohoto řešení[5]:

- V případě potřeby změny je vždy možnost provést úpravy v následujícím chystaném kroku. Tedy vysoká pružnost.
- Menší a kratší přírůstky se lépe kontrolují. Finančně i z hlediska provedené práce.
- Nevznikají duplicity komponent BI

Mezi nevýhody lze řadit počáteční prodlevu před vznikem prvních přírůstků a nutnost důkladné tvorby celé koncepce, časové a finanční.



**Obrázek 3.5 Přírůstkové řešení**

Zdroj: NOVOTNÝ, Ota. *Business intelligence: jak využít bohatství ve vašich datech*. 1. vyd. Praha: Grada, 2005, 254 s. ISBN 80-247-1094-3.

### 3.4 ETL

Zkratka ETL znamená „Extract, Transform, Load“<sup>14</sup>. Proces, který se takto označuje, je základní integrační komponentou BI modelu. Skládá se tedy z[4]:

- **Výběr** (Extract) – získání dat z jedné nebo více databází (OLAP i OLTP) a jiných datových zdrojů (MS Excel, MS Access, csv soubory).
- **Uspořádání** (Transform) – převedení vybraných dat z původní formy do potřebné podoby v cílovém datovém skladu.
- **Nahrání** (Load) – umístění uspořádaných dat do stávající struktury DWH nebo DMA (případně i DSA)

Vhodné ETL nástroje by pak měly mít minimálně tyto vlastnosti[4]:

- Vybrat pouze ta data nebo datové oblasti, která jsou potřebná pro následující části BI (Analytické nebo BPM).
- Konsolidace dat, tj. vyloučení duplicit a multiplicit zdrojových dat
- Čistota nebo kvalita dat – odstranění chyb a zavádějících nepřesností
- Schopnost číst nebo zapisovat data neomezeného počtu datových architektur[3].
- Automaticky v procesu zpracovávat a ukládat i metadata zdrojových dat
- Podpora Open Standardů a schopnost update a upgrade
- Srozumitelné a použitelné uživatelské rozhraní ETL nástroje
- Optimalizace výkonu v závislosti na HW a možnost modifikace výkonu
- Dlouhodobá podpora dodavatele SW

ETL nástroje bývají součástí balíku BI instalace a všichni velcí poskytovatelé BI je dodávají (pro Microsoft např. MS SQL Server 2012 Integration Services, pro Oracle pak Oracle Data Integrator Enterprise Edition 12c).

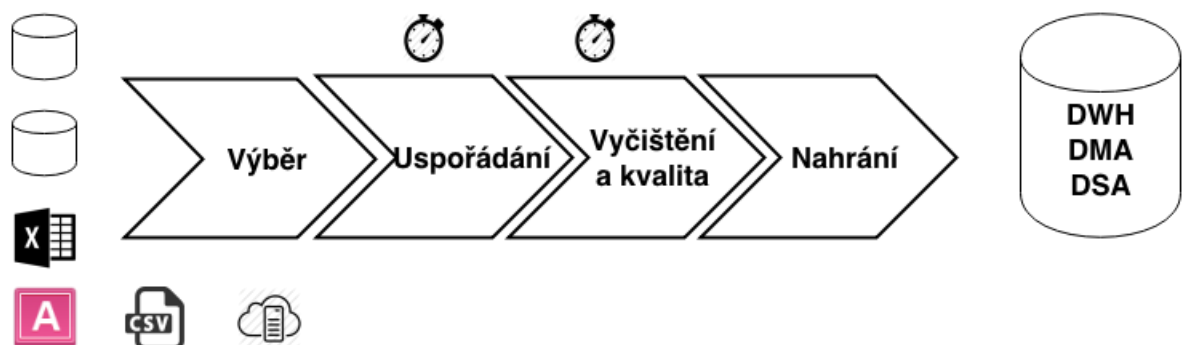
Na procesu ETL jsou nejvíce náročné právě uspořádání a podílí se 60% na časovém, finančním a pracovním využití IT zaměstnanců[4]. Bohužel ETL proces v souvislosti s BI implementací má tři velké nevýhody, které management firmy vidí[3]:

---

<sup>14</sup> Rovněž se označuje jako Datová pumpa[1]

- Nástroj ETL má velké vstupní investice (CAPEX) a nelze pak nástroj ve finální podobě použít někde jinde, většinou je přizpůsoben na míru právě daným potřebám podniku.
- Nasazení ETL trvá velmi dlouho (záleží na velikosti a množství primárních zdrojů, množství a kvalitě metadat a požadavkům managementu na Analytické a BPM data). Po celou dobu je třeba financovat vývoj.
- Výsledky dokončeného ETL nepůjdou nijak změřit nebo ohodnotit, takže management neví, zda IT odvedlo dobrou práci. Kvalita nebo nedostatky ETL se projeví až po nějakém čase, kromě zásadních provozních problémů.

Pokud je ale ETL kompletně a dle požadavků a očekávání implementován, výrazně ušetří firmě čas, finance, údržbu dat.



**Obrázek 3.6 ETL proces**

Zdroj: TURBAN, Efraim. *Business intelligence: a managerial approach*. 2nd ed. Boston: Prentice Hall, c2011, xx, 292 p. ISBN 01-361-0066-X.

### 3.5 Metadata

Nejjednodušší definice metadat je že se jedná o **data o datech**[7]. Slouží tedy k dokumentaci, popisu, rozlišení, lokalizaci konkrétních dat i celků IS podniku. BI Metadata jsou v podniku nezbytnou podmínkou pro zachování znalostí o datových zdrojích. Pokud bychom je nepoužili, nepopsali datové zdroje, s odchodem klíčového zaměstnance, který např. datové tržiště nebo ETL vytvářel, veškeré znalosti ztratíme a bude třeba začít budovat danou komponentu BI téměř od začátku. S nadsázkou lze připodobnit potřebu metadat legendě o Pražském orloji, který potom, co jej mistr Hanuš zastavil, nedokázal žádný hodinář dlouhá desetiletí zprovoznit.

Metadata mohou být dělena na[7]:

- Popisná metadata (Descriptive) – popisují zdroje tak abychom jim porozuměli a dokázali je identifikovat. Bude sem patřit popis dimenze v OLAP, autor, abstrakt a klíčová slova u knihy.
- Strukturální metadata (Structural) – určují, jakou vzájemnou vazbu mají zkoumané komponenty objektů i celé objekty, v knize to bude např. řazení stránek, v DWH pak u tabulek popis jejích atributů, primární a cizí klíče, řazení dat, datové typy, omezení.
- Administrativní metadata (Administrative) - ukládají informace pro základní správu dat, jako kdy byla tabulka vytvořena a jak (import, ad-hoc skriptem atd.), obecně všechny informace potřebné pro uchovávání a archivaci v daném typu SŘBD (např. Oracle MS SQL Server), dále pak autentizaci a autorizaci<sup>15</sup>.

Metadata se vyskytují na různých úrovních agregace od nejnižší entity, domény jednotlivé procedury po popis datového tržiště, komplexního popisu DWH, procesu ETL jako celku[7]. V BI řešení bývají uložena společně s daty v databázích (Strukturální a

---

<sup>15</sup> „Autentizací rozumíme techniku, s jejíž pomocí ověřujeme proklamovanou identitu nějakého subjektu a jejím cílem je přesvědčit protistranu o své identitě a zajistit ochranu před jejím falšováním“[1]. Autorizace je proces, který dle definovaných pravidel (oprávnění) umožní přistoupit ověřeným uživatelům k souborům a nakládat s nimi[1].

Administrativní) a navíc by měl být k dispozici i datový slovník, který pak abstraktem jednotlivá potřebná data detailně popíše. Typickým příkladem z podnikové praxe je atribut ID zákazníka. Pokud se zeptám tří různých oddělení, dostanu tři i různé odpovědi: Pro finanční oddělení je to fakturační účet, pro segmentový marketing je to IČ a pro oddělení výstavby je číslo občanského průkazu IT správce z lokality. Proto je třeba bezpochybně definovat význam každého nejasného atributu (v praxi u větších firem s více divizemi tedy většiny z nich).

## **3.6 Reporting**

Pro popis Reportingu a analytických nástrojů použiji z větší části popis ze své bakalářské práce, protože tuto oblast jsem v podobnějším rozsahu zkoumal[8].

### **3.6.1 Charakteristika Reportingu**

Reporting lze definovat jako jednosměrnou činnost, spojenou s dotazováním nad databázemi, strukturovanými soubory (např. CSV) nebo jinými datovými zdroji pomocí jím vlastním dotazovacím jazykem (např. MS-SQL). Získaným výsledkem by měl být přehled, lze také říci informace, potřebná k rozhodování a statistice nad zkoumanou oblastí. Mohou existovat reporty pravidelné např. na měsíční bázi se stejně strukturovaným výstupem a vstupními filtry nebo ad-hoc reporty vytvořené pro jednorázové použití a většinou s definovaným zadáním koncového uživatele[1]. V rámci BI jsou vstupem (zdrojovými daty) pro reporting a jeho nástroje datové sklady DWH a DMA. Tato data jsou již pomocí ETL očištěna a strukturována.

### **3.6.2 Statický a dynamický Reporting**

Reporting lze dělit dle jeho pružnosti na statický a dynamický způsob užívání.

Statický způsob je charakteristický neměnnými vstupními parametry a bývá spojována s pravidelným Reportingem[14]. V rámci firmy bude mít standardní strukturu i vzhled. Jeho výhodou je, že je vnímán uživateli jako standard a je jednoduchý na obsluhu.

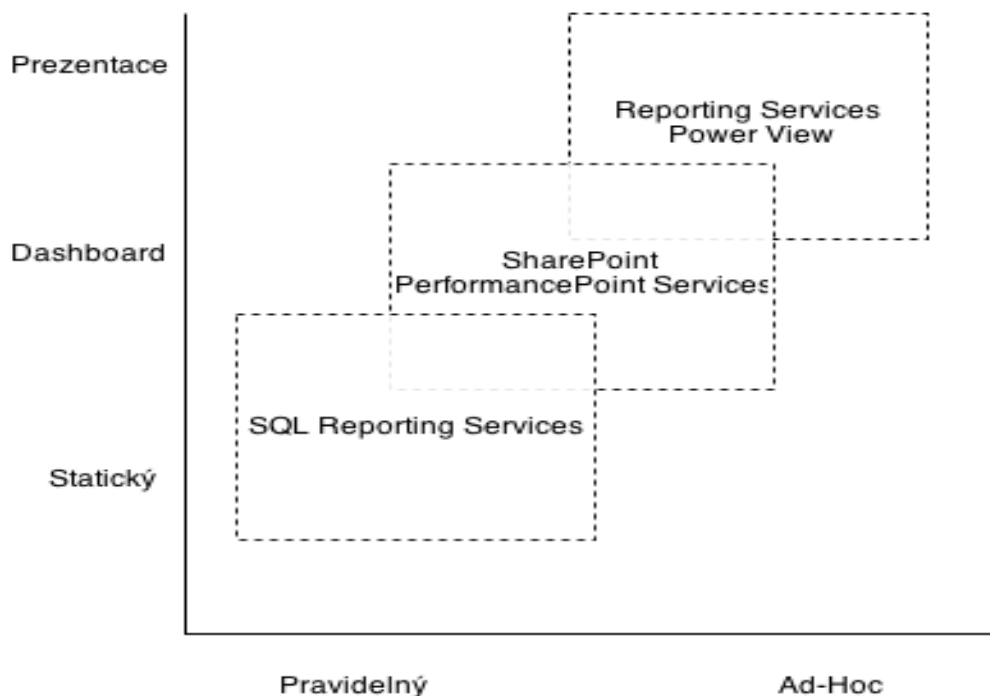
Dynamický způsob naopak umožňuje ovlivňovat obsah, případně i formu očekávaného reportu[14]. Uživatel mění vstupní parametry, upravuje vstupní filtry někdy i částečně design reportu. Výhodou dynamického Reportingu oproti statickému je, že uživatel získá výstup alespoň částečně (např. oproti OLAP) dle svých představ. Nevýhodou může být někdy přílišná složitost dynamického Reportingu, uživatel pak ztrácí orientaci ve vstupních parametrech a výstup nesplní jeho očekávání. Toto platí zvláště při užívání jednoho Reportingu různými skupinami uživatelů.

### **3.6.3 Pravidelnost Reportingu a Analýz**

Z hlediska periodicity se dělí Reporting na dvě základní skupiny[5]:

- Reporting běžný (pravidelný) – reporty jsou vytvářené v určených intervalech a dané struktuře. Příkladem může být uzávěrka za účetní období (měsíc, rok) nebo výroční zpráva pro externí odběratele. Pro interní uživatele např. počet nových zákazníků a jejich tržba.
- Jednorázové dotazy – neboli ad hoc query. Reporty jsou vypracovány na požádání či v případě okamžité potřeby. Mimořádný Reporting je výhradně používán interními uživateli. Uživatelé dopředu přesně nevědí, jaký obsah a formát má výstupní report splňovat a teprve po vytvoření reportu vyhodnotí, zda splnil jejich poptávku[14]. Pokud ano, lze z mimořádného reportu dodatečně vytvořit report pravidelný.

Příkladem může být charakter BI nástrojů společnosti Microsoft na následujícím obrázku. Uživatelé budou preferovat velkou volnost v úpravách pro ad-hoc prezentační a analytické nástroje<sup>16</sup>, zatímco čím více se bude report (analýza) stávat pravidelným, očekává se spíše rychlost získání informací než možnost modelování a úprav.



**Obrázek 3.7 Microsoft Reporting a OLAP nástroje dle přístupu a pravidelnosti**  
Zdroj:vlastní

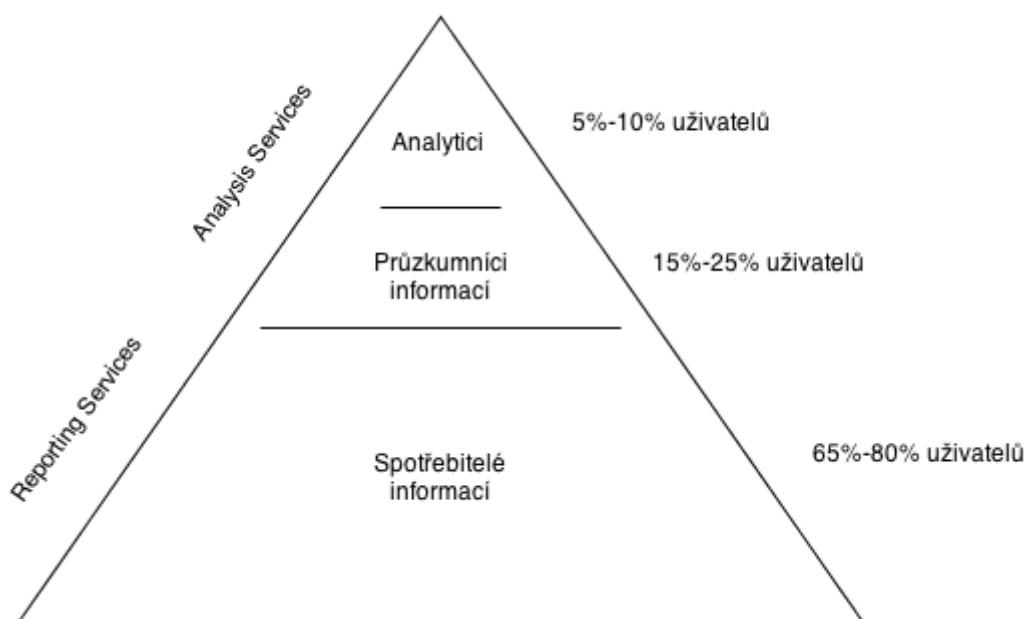
<sup>16</sup> V tomto případě nástroj Power View a SharePoint PPS

### 3.7 Reportovací, dotazovací a analytické nástroje BI

Reporting i analytické nástroje jsou SW součástí BI a jejich účelem je usnadnit přístup a analýzu vstupních dat, jejich přeměnu do informací, a prezentaci výsledků[9]. V teoretické rovině lze oblast reportingu oddělit od oblasti analytické. Z praxe ale vyplývá, že uživatelé preferují komplexnost a tedy většina dnešních BI nástrojů obsahuje prvky reportingu a také analytických funkcí. Proto není jednoduše možné říci, že některý nástroj je pouze analytický nebo pouze reportovací.

V dnešní době je možností reportování a analýzy dat velké množství. Uživatelé se s analytickými nástroji a reporty setkávají někdy i nevědomky a vnímají je jako běžnou funkcionalitu nebo součást nějaké jiné aplikace. Velké podniky pak často používají jednotky až desítky reportovacích a analytických nástrojů, zvláště pokud nevznikaly „na zelené louce“ a rostly různými fúzemi, akvizicemi.

Již první odkaz po zadání do internetového vyhledávače pojmu „reporting tools“ vede na adresu Wikipedie, kde je pod tímto pojmem zobrazeno přes 40 názvů SW a odkazů na společnosti, zabývajícími se reportovacími nástroji. Další nově vznikají, mění se nebo se slučují do jednoho komplexního řešení. Obrázek 3.8 charakterizuje odhadované procentuální dělení odběratelů informací. Zde se dělí na pouhé konzumenty, průzkumníky a analytiku.



Obrázek 3.8 Dělení BI spotřebitelů

Zdroj: LACKO, Luboslav. Business Intelligence v SQL Serveru 2005: reportovací, analytické a další datové služby. Vyd. 1. Brno: Computer Press, 2006, 391 s. ISBN 80-251-1110-5.



### 3.7.1 Podle typu použití

Základním typem dělení bude způsob použití, resp. BI oblast použití nástroje. Zde lze ještě členit na čtyři skupiny[9]:

#### **Nástroje podporující strukturovanou paměť**

Do této skupiny patří nástroje, které umožňují systému organizovat a zapamatovat si data[9].

- Správa datových úložišť
- ERP nástroje
- Správa znalostní báze
- Správa digitálního obsahu (Digital content management)
- Nástroje pro správu dokumentace

#### **Nástroje podporující pochopení dat**

Tato skupina se zaměřuje na použití získaných dat. Spíše ne na pravidelnosti analýzy je založena na její nepřetržitosti[9]. Typické nástroje pak jsou:

- BI analytické nástroje
- Nástroje pro dolování dat
- Real-time podpora rozhodování

#### **Nástroje podporující informační integraci**

Informační integrace je schopnost spojit (sloučit) dříve strukturovaný a nestrukturovaný obsah z různých zdrojů, která obohacuje strukturovanou paměť o nový obsah (data a informace)[9].

- Dolování z textu
- Dolování z webu
- Průzkum a analýza prostředí a okolí společnosti (Scanning)
- RFID<sup>17</sup> analýza

---

<sup>17</sup> Zařízení s radiovou technologií, podávající informaci o pozici.

## **Nástroje podporující prezentaci informací**

Tato skupina nástrojů se zaměřuje na podporu prezentaci informací, jedná se o rozhraní mezi BI systémem a uživatelem[9].

- OLAP nástroje
- Vizualizační nástroje
- Dashboardy
- Scorecardy
- BPM nástroje

### **3.7.2 Licenční dělení**

Na trhu dochází, stejně jako u jakéhokoliv typu SW, k rozdělení reportovacích nástrojů dle licenční politiky neboli také distribuce licence. Typů licencí je poměrně velké množství, pro účely této práce budou členěny pouze do dvou skupin. Jedna skupina, kdy je třeba za SW výrobcí zaplatit jednorázově nebo platit pravidelně a druhá skupina se všemi ostatními distribucemi, které umožní používat SW bezplatně.

#### **Bezplatné užívání SW**

Skupina bezplatného užívání reportovacích nástrojů bude zahrnovat takové typy SW, které jsou k dispozici zdarma, ale většinou existují nějaké podmínky, které uživatel musí před užíváním SW akceptovat a potvrdit. Mezi tyto podmínky může patřit například dohoda o nešíření SW bez udání jména autora, zákaz používání SW pro firemní účely, omezení instalací na počet PC, povinná registrace na stránkách autora, poskytnutí zdrojového kódu SW atd.

Do této skupiny patří např. tyto typy licencí: open-source a freeware. V některých případech se v SW zobrazuje v určité části grafického rozhraní reklama, jinak omezení, ten se nazývá adware. Dalším typem SW licence je shareware. SW rozsah funkcionalit je pro bezplatnou verzi omezen, případně je užívání mírně ztíženo přihlašovací oknem při spuštění programu. Po zaplacení uživatel obdrží licenční klíč, který vloží do programu a tento se stane plnou verzí bez omezení. Ještě je třeba zmínit

licence typu demo nebo také trial, kdy je SW plně funkční po nějaký časový úsek a po jeho vypršení je uživatel SW nucen buď SW koupit, nebo přestane fungovat či se zásadním způsobem omezí jeho funkcionality (např. nepůjde ukládat výsledek).

Výhoda bezplatného SW je hlavně v nulové ceně, nevýhoda bývá většinou ve slabší technické podpoře produktu a omezeného vývoje. Příkladem může být BIRT Project komunity Eclipse[15]. Je vytvořen v jazyce JAVA a podporuje instalaci na OS Windows a Linux.

### **Komerční užívání**

V drtivé většině firem je používán SW s licencí pro komerční provoz. Jedná se o plné verze SW bez omezení a hlavní výhodou je plná podpora výrobce a také tzv. LTS podpora. To znamená, že výrobce SW garantuje, že se o SW bude starat po určitou deklarovanou dobu, bude vydávat opravné balíčky a reagovat na potřeby uživatelů. A to i přesto, že již vydá novou verzi SW a původní stáhne z prodeje. Další výhodou může být například lokalizace SW (tedy SW bude v jazyce uživatele). U neplacených licencí bývá obsaženo jen pár běžných jazyků. Nevýhodou je někdy odmítavý postoj výrobce implementovat novější funkcionality, byť by to nebylo náročné. Toto potvrzují kolegové v GTS Czech z IS oddělení. Jako příklad je SharePoint 2010 a jeho wikistránky. Novinky (někdy se ale jedná vlastně o opravy nedodělků) jsou nasazeny až do verze 2013. Zákazník je pak nucen zakoupit novější verzi SW, který danou funkcionalitu obsahuje. Příklad placené plné verze (s možností vyzkoušet produkt ve verzi trial) je JReport[16] společnosti Jinfonet Software, který se skládá ze tří, samostatných komponent: Report Designeru, Server a Dashboardu. Dashboard je vlastně grafické rozhraní s předdefinovanými výstupy z reportů. Obsahuje jen základní nástroje na filtrování či změny. Slouží pro rychlý manažerský přehled.

### **3.7.3 Umístění a přístup**

Reportovací nástroje jsou většinou dostupné ze třech základních umístění: lokálně, on-line a na sdíleném firemní prostoru (Serveru). Mezi nástroji z různých oblastí existuje možnost synchronizace a sdíleného přístupu.

#### **Lokální nástroje**

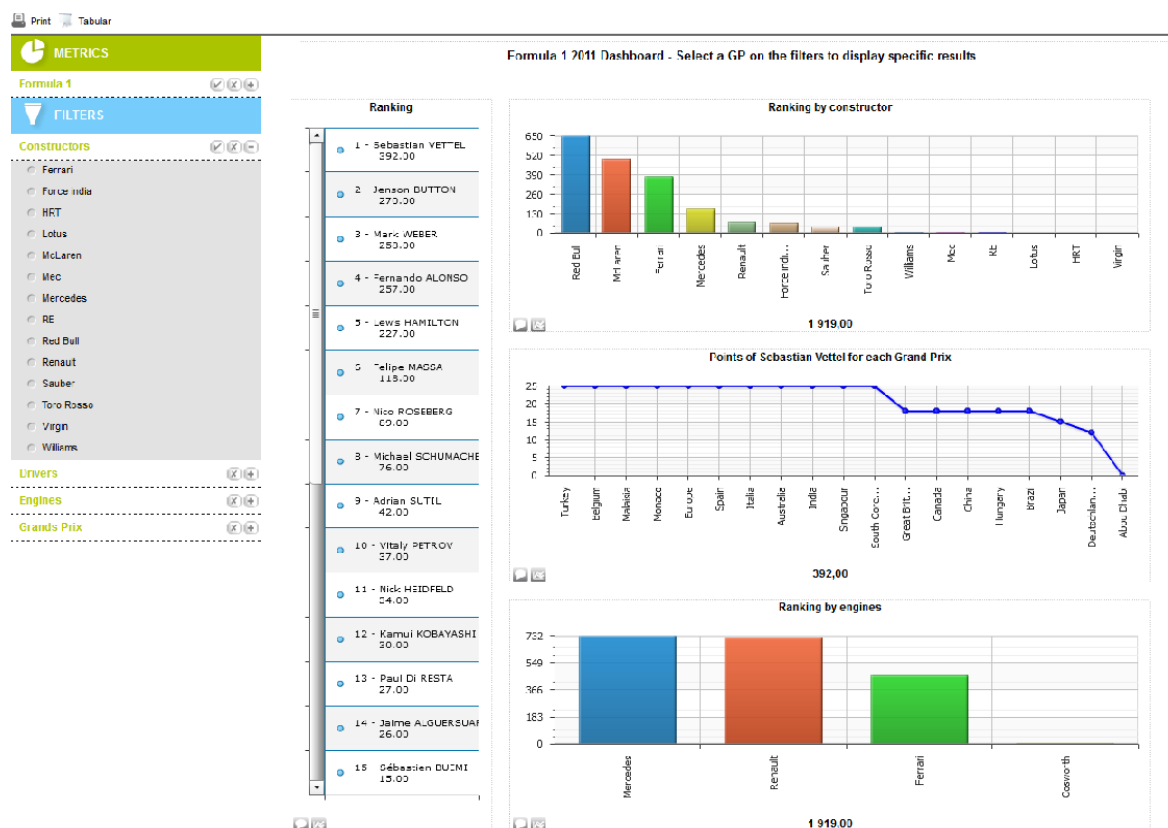
SW se instaluje přímo na PC uživatele, odkud je pak spouštěn. Typickými SW v této oblasti jsou některé součásti kancelářských balíků MS Office (MS Excel nebo MS Access) a OpenOffice (Calc a Base). Jedná se o nástroje pro jednoduchý, ale hlavně rychlý způsob reportingu a analýzy nad daty. Pomocí doplňků od výrobce (add-in) lze rozšiřovat jejich funkcionality. U těchto nástrojů se podmínkou, že výstup lze někde sdílet ať už jako samostatný prvek (MS Excel list s kontingenční tabulkou, MS Access databázi), převést na prezentační aplikaci typu MS PowerPoint nebo publikovat na firemním sdíleném prostoru (SharePoint). Výhodou lokálních nástrojů je známé prostředí kancelářských aplikací pro uživatele a nevýhodou je omezení reportovacích funkcionalit (např. počtem řádků) a v případě velkého množství dat i náročnost na lokální HW. Bohužel speciálně BI PowerPivot doplněk pro Excel po doinstalování start této aplikace značně (dle mého odhadu 2x-3x) zpomalí a je tedy na zvážení, kdy jej používat a zda jej nechat nainstalován stále.

#### **On-line (Cloud) nástroje**

SW je nainstalován na serveru poskytovatele nebo výrobce nástroje, odkud je pak pomocí internetu poskytován přes grafické rozhraní. Znamená to, že si uživatelé nemusí instalovat žádného reportovacího klienta na PC a s nástrojem pracují přes webový prohlížeč. Tento způsob práce se vzdáleným SW se někdy nazývá tenký klient (thin client). Jako příklad jsem zvolil francouzskou společnost Bittle solutions a její on-line BI nástroj[13]. Umožňuje import dat např. formátů CSV, XLS, Google sešitu, interní či externí databáze a následnou práci s daty ve webovém prohlížeči a export zpět do běžných formátů. Rozhraní lze vidět na Obrázek 3.9.

Výhodou je, že odpadá potřeba instalovat lokální server nebo nějaké samostatné BI nástroje. Další výhodou je, že výpočetní výkon v podstatě řeší poskytovatel a uživatel již

jenom pracuje s daty, které zobrazuje v prohlížeči. Nevýhodou je, že citlivá firemní data jsou svěřena do rukou třetí osoby a může dojít k narušení bezpečnosti.

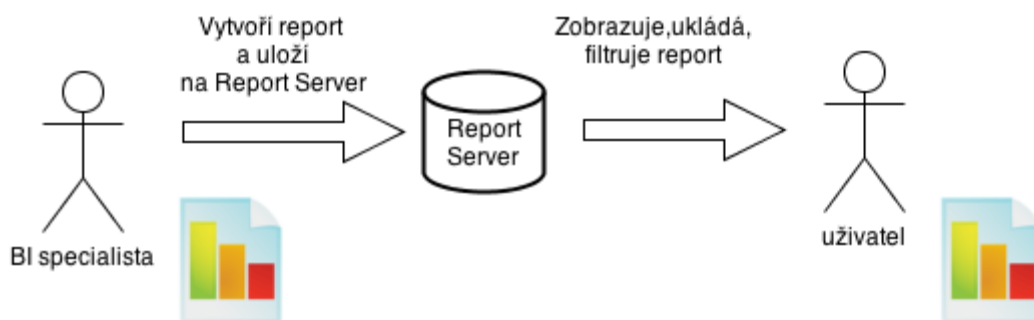


Obrázek 3.9 Bittle thin client

Zdroj: Lebr, Radek. Reportovací nástroje MSSQL Server 2012 ve společnosti GTS Czech. Praha : Vysoká škola manažerské informatiky, ekonomiky a práva, 2013. Bakalářská práce.

## Klient-Serverové nástroje

Třetí skupinou nástrojů je ta, která je nasazena na severu umístěného lokálně v rámci firmy. Může se jednat o komplexní řešení jako je JReport nebo třeba MS Reporting Services[18]. Reporty jsou vytvářené například pomocí lokálních nástrojů na PC (Report Designer), následně jednosměrně publikovány na serveru a dále s nimi lze pracovat pomocí modelovacích nástrojů jako je třeba Report Builder[19], které nemodifikují datové zdroje vystaveného reportu na serveru, ale vytváří v podstatě pouze vlastní náhled nad publikovaným reportem (například změnou typu grafu ze sloupcového na koláčový), který si pak mohou uložit.



**Obrázek 3.10** Vystavení reportu pomocí Report Builderu

Zdroj: Lebr, Radek. Reportovací nástroje MSSQL Server 2012 ve společnosti GTS Czech. Praha : Vysoká škola manažerské informatiky, ekonomiky a práva, 2013. Bakalářská práce.

Hotové reporty pak mohou být zobrazeny přímo pomocí Report Builderu. Report Designer se od Report Builderu liší ve způsobu vytváření reportu. Designer umí pouze vytvořit nový, ale znovu ten stejný nedokáže použít jako podklad pro vytvoření jiného reportu. Builder toto dokáže. Druhou možností je vystavení reportu na SharePoint serveru, kdy uživatelé již vidí report v klasickém webovém prohlížeči a nepotřebují speciální aplikaci. Výhodou je, že je zajištěna bezpečnost dat a reporty jsou umístěny na jednom místě, tedy zálohované a se správou verzí. Nevýhodou mohou být vyšší náklady.

### 3.7.4 Segmentové dělení BI nástrojů

Zákazníci neboli uživatelé reportů se dělí na základní dva až tři typy: na společnosti nebo firmy (Enterprise a Business) a přímé koncové uživatele včetně domácností neboli osobní.

#### Osobní

Typickým příkladem kategorie personal BI je MS Excel nebo MS Access. Uživatel si analýzu dat řeší lokálně s minimálními nároky na síť a není třeba žádného datového skladu nebo BI vrstev. Vše je řešeno tak, aby vzniknul minimální náklad na další BI komponenty. Všechny vrstvy jsou (v tomto případě prezenční, analytická i datové úložiště) bývají součástí jednoho nástroje<sup>18</sup>.

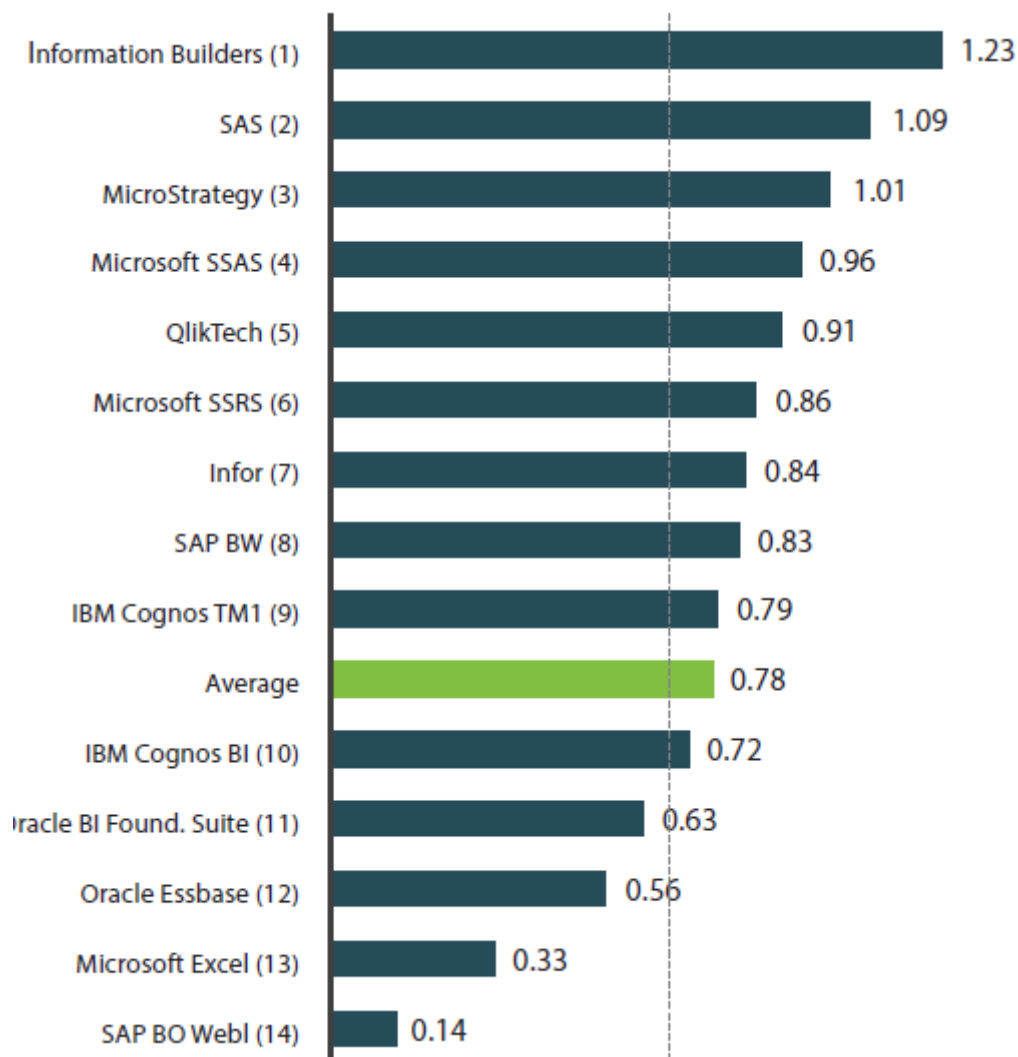
<sup>18</sup> Pozn. BI nástroje asi nejsou typickou osobní (nebo dokonce domácí) aplikací, a proto se výrobci zaměřují na segment firemní. Nicméně jsou zde výjimky, které je třeba zmínit. Jedná se třeba o Google Analytics – nástroj pro sledování návštěv definované webové stránky a konkrétního pohybu návštěvníků po stránce. Uživateli Google Analytics je pak generován report se zaznamenanými údaji – počty návštěv, IP adresami, dobou strávenou na stránkách atd.

## Podnikové

Podnikové BI nástroje jsou dělitelné na dvě skupiny a to dle velikosti podniku. Velké podniky (Large Enterprise) mívají stovky až tisíce zaměstnanců a desetitisíce zákazníků, zatímco firmy (Small Enterprise) jednotky, desítky. Podle velikosti společnosti (počtu zaměstnanců, oddělení, zákazníků) je adekvátně nasazen IT/IS základ. Od toho se odvíjejí i požadavky na reportovací nástroje.

U Large Enterprise bude jistě třeba, aby byly reportovací nástroje provázány s IS společnosti, a bývá tedy zvykem, že jsou již součástí (komponentou) většího, celofiremního IS řešení. Je také ale běžné mít nasazeno několik různých reportovacích nástrojů od různých výrobců, které mezi sebou musí umět spolupracovat. Tyto podniky mají většinou jednotky až desítky databází (agregovaných prostřednictvím mechanismu získání dat ETL [6] do datových skladů DWH), ze kterých reportovací nástroje získávají data. Výstupy (samotné reporty) jsou zprostředkovány uživatelům přes Intranet, SharePoint nebo jiný firemní prostor (např. firemní WIKI). SW má být připraven na velkou zátěž, zvládat souběžné operace, interagovat s IS a mít možnost záloh. Výrobce SW někdy přímo modifikuje nástroj dle potřeb společnosti pro snazší implementaci do specifického prostředí.

Spokojenost s BI produkty pro skupinu Large Enterprise v provedeném průzkumu v roce 2013 společností Business Application Research Center ukazuje následující obrázek. Lze předpokládat, že v uvedeném průzkumu jsou zahrnuty klíčové produkty, které pokrývají většinu trhu, a používá je majoritní skupina firem. Z toho plyne, že výčet uvedených SW představuje současný stav na trhu.



**Obrázek 3.11 BARC průzkum spokojenosti**

Zdroj: BUSINESS APPLICATION RESEARCH CENTER. The BI Survey 13 [online]. 2013 [cit. 2014-08-16].

Dostupné z: <http://www.bi-survey.com>

Firmy typu Small Enterprise očekávají od reportovacího nástroje malou náročnost na HW a případně na OS (Unix, Linux), možnost zálohování, exportování reportů do různých formátů. Reporting a analytická činnost bývá většinou jednodušší i vzhledem k menšímu počtu datových zdrojů. Většinou je rozhodující cena.

### 3.7.5 Shrnutí dělení BI nástrojů

V následující tabulce jsem zmapoval typické prvky pro jednotlivé skupiny a typy reportovacích nástrojů. Zmiňuji zde výhody a nevýhody a příklad SW. Je třeba si uvědomit, že toto dělení je vícerozměrové, tedy se nevylučuje kombinace prakticky žádného typu mezi skupinami (např. Bezplatné-Online nebo Společnosti-Lokální). Atypická je zde pouze kombinace Domácnosti-Servery.



Nejčastěji zmiňovaná je cena a náročnost na HW, důležitá je také podpora výrobce a řešení „šité na míru“. V některých případech může dojít zcela ke zrušení SW bez náhrady (tak to praktikuje např. Google), kdy uživatel přijde o veškerá data. SW pro společnosti bude určitě splňovat vyšší kritéria na dostupnost a stabilitu (SLA).

Tabulka 3.2 Porovnání BI nástrojů

dělení BI nástrojů	členění	Největší výhody	Největší nevýhody	Příklad SW
<b>Licenční</b>	Bezplatné	<ul style="list-style-type: none"> <li>• Cena</li> </ul>	<ul style="list-style-type: none"> <li>• bez podpory výrobce</li> <li>• adware</li> <li>• demo</li> </ul>	BIRT
	Komerční	<ul style="list-style-type: none"> <li>• Podpora výrobce,</li> <li>• možnost reklamace,</li> <li>• více funkcionalit</li> </ul>	<ul style="list-style-type: none"> <li>• Placený SW</li> </ul>	JReport
<b>Umístění a přístup</b>	Lokální	<ul style="list-style-type: none"> <li>• Práce offline</li> <li>• Známé prostředí (MS Office)</li> </ul>	<ul style="list-style-type: none"> <li>• Zatížení HW PC</li> </ul>	MS Excel
	Online	<ul style="list-style-type: none"> <li>• Silný HW serveru</li> <li>• Bez instalace</li> <li>• Vše v browseru</li> <li>• Synchronizace s lokálními daty</li> </ul>	<ul style="list-style-type: none"> <li>• Kompromitace dat</li> <li>• Bez zálohy</li> <li>• Možnost terminace SW bez náhrady</li> </ul>	Bittle
	Serverové	<ul style="list-style-type: none"> <li>• Silný HW serveru</li> <li>• Zálohy dat</li> <li>• Vysoká bezpečnost</li> </ul>	<ul style="list-style-type: none"> <li>• Cena</li> <li>• Náklady na Server</li> <li>• Údržba Serveru</li> </ul>	SSAS Share-Point
<b>Segmentové</b>	Společnosti a firmy	<ul style="list-style-type: none"> <li>• SLA řešení</li> <li>• Dostupná podpora</li> <li>• Školení</li> <li>• Komplexní řešení</li> </ul>	<ul style="list-style-type: none"> <li>• Cena (licence)</li> <li>• HW náročnost</li> </ul>	SSRS Oracle SAP
	Osobní	<ul style="list-style-type: none"> <li>• Cena</li> <li>• Jednoduchost</li> </ul>	<ul style="list-style-type: none"> <li>• Není samostatný SW</li> <li>• Omezené možnosti</li> <li>• Možnost terminace SW bez náhrady</li> </ul>	Linked-In Google Analytics Libre Office

Zdroj: Lebr, Radek. Reportovací nástroje MSSQL Server 2012 ve společnosti GTS Czech. Praha : Vysoká škola manažerské informatiky, ekonomiky a práva, 2013. Bakalářská práce.

Z neznámějších<sup>19</sup> dodavatelů BI řešení jmenujme společnosti [4]:

- Oracle s db systémem Oracle
- Microsoft s db systémem MS SQL Server
- IBM s db systémem DB2
- SAP
- SAS Institute
- Microstrategy
- Actuate
- Information Builders
- Jaspersoft
- Pentaho

---

<sup>19</sup> Detailnější výpis dodavatelů na trhu naleznete v příloze v tabulce BI produktů

## 3.8 Multidimenzionální modelování dat a OLAP

Jak již bylo zmíněno v předchozích kapitolách, datové modelování pomocí tabulek dimenzí a faktů jsou typické pro BI analytické aplikace. OLAP kostky pak právě představují „stavebnici“ pro BI analytika, se kterou data modeluje a získává potřebné informace.

### 3.8.1 Měření dat

Než vůbec BI analytik začne modelovat a vytvářet OLAP kostky, reporty, musí pochopit, co vlastně za informace hledá a jak je má získat. Měl by tedy rozumět pojmu měření.

V zásadě existují tři základní aspekty problémů při měření<sup>20</sup>, nejenom dat[10]:

- Pojetí (Concept) měření – je třeba si definovat jakou chybu, odchylku jsme ochotni tolerovat v naměřených hodnotách nebo při opakování měření pro získání věrohodné informace. Naopak pokud je měření vysoce přesné díky omezujícím vstupním podmínkám, může dojít k zásadnímu zkreslení získané informace. Proto je třeba hledat kompromis a dívat se „nad měřené hodnoty“. K tomu je třeba také zvolit vhodné stupnice a jednotky měření. Existují dva přístupy k měření; ordinální a nominální. Příkladem jednotek **ordinálního** měření může být hodnocení filmů podle oblíbenosti. Čtyři „jednohvězdičkové“ filmy nemají určitě stejnou váhu jako jeden „čtyřhvězdičkový“, víme jen, že film s jednou hvězdou je horší než se čtyřmi, ale nevíme o kolik. Homogenní jednotky **nominálního** měření jsou pak např. kila, litry, koruny. Tam je vazba jednoznačná a 5 kilových závaží = 1 pětikilové. S těmito jednotkami lze dále matematicky pracovat (násobit, sčítat, dělit).
- Předmět měření (Object) – pokud se předmět měření vhodně popíše a odstraní abstrakce, např. namísto „Jak změřit IT bezpečnost firmy“ se zeptat na počet DDoS útoků a virových napadení, lze získat nějaké výsledky. Pro usnadnění rozhodování, zda lze daný předmět měřit je možné použít následující tříbodové hodnocení:
  - 1 Lze to vůbec zjistit a pozorovat?
  - 2 Pokud to lze pozorovat, lze určit množství?
  - 3 Pokud lze určit rozsah množství, je to měřitelné.

---

<sup>20</sup> Kvantitativně vyjádřené omezení neurčitosti, založené na jednom či více pozorování[10].

- Metody měření – základním předpokladem je u analytika elementární znalost měřících technik. Mluvíme o popisné statistice, algebře, případně ekonometrii a geometrii. Existují čtyři užitečné předpoklady pro volbu měřících metod[10]:

- 1 Stejná věc už se někdy dříve měřila.
- 2 K dispozici je víc podkladových dat, než se zdá.
- 3 K získání výsledku měření je potřeba méně dat, než se zdá; nedostatek lze řešit vhodnými analytickými schopnostmi.
- 4 Dodatečná data jsou dostupnější, než se jeví na první pohled.

Kromě komplikací při a před měřením lze očekávat tři typické námitky k měření a prezentaci závěrů: ekonomické, obecné a etické. Bohužel zde je již na vhodné argumentaci (analytika vůči vedení, vedení vůči akcionářům, PR vůči novinářům atd.), zda budou zamítnuty nebo ne.

- Ekonomickým základem (a zároveň námitkou) je, že náklad na měření nesmí přesáhnout jenom přínos. Zároveň je také vhodné náklad na měření mít s čím porovnat.
- Obecný pohled zase nevěří statistickým závěrům. Statistika je vnímána negativně, protože pracuje s určitou pravděpodobností.
- Etické námitky by měly zajistit diskuzi a odmítnout závěry, byť statisticky a odborně správné. Příkladem budiž využití statistiky ČSÚ počtu sebevražd žen mezi 15-60 v Praze pro prodej a reklamu na antidepresiva.

### 3.8.2 Dimenzionální modelování

Dimenzionální modely se skládají ze dvou typů komponent. Komponenty dimenzí a faktů. Zatímco dimenze umožní data klasifikovat (popisovat), fakta obsahují pak samotnou skutečnost, skutečně získané hodnoty.

*„Dimenzí se z pohledu jejího užití rozumí analytické hledisko pro hodnocení sledovaných ukazatelů, z informatického hlediska se jeví jako struktura dat, případně jako databázová tabulka obsahující záznamy o jednotlivých prvcích dimenze (např. o položkách zboží).[4]“*

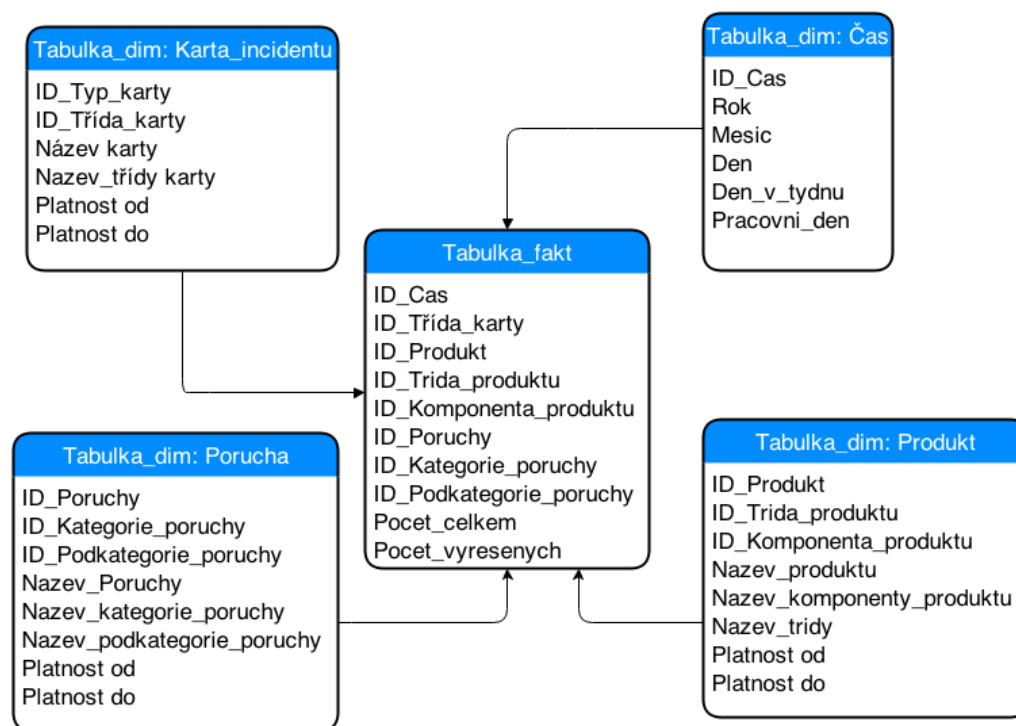
Prvky dimenzí pak mohou obsahovat několik hierarchických úrovní (tříd, skupin, podskupin)[4]. Data jsou pak ukládána do databáze právě v takové míře detailu, do jaké sahají hierarchické úrovně. Tento přístup umožňuje provést analýzu na libovolné úrovni a většinou jsou data obohacena o agregované hodnoty a míry<sup>21</sup>.

Multidimenzionální modelování lze rozdělit do dvou skupin:

- Data jsou uložena v relačních databázích
- Data jsou uložena v OLAP kostce

Pro relační databáze existují dvě základní schémata datového modelu[4]. Schéma STAR a model SNOWFLAKE. Tabulka fakt je pak u obou propojena pomocí cizích klíčů s primárními klíči tabulek dimenzí. Zde fungují dimenze i jako číselníky. Tabulky dimenzí jsou ve vazbě s tabulkami faktu 1:N.

Model STAR (schéma do hvězdy) je jednodušší a běžnější. Bývá to jedna faktová tabulka obklopená různými dimenzemi. Tyto dimenze mají celou svou hierarchii v jedné tabulce.

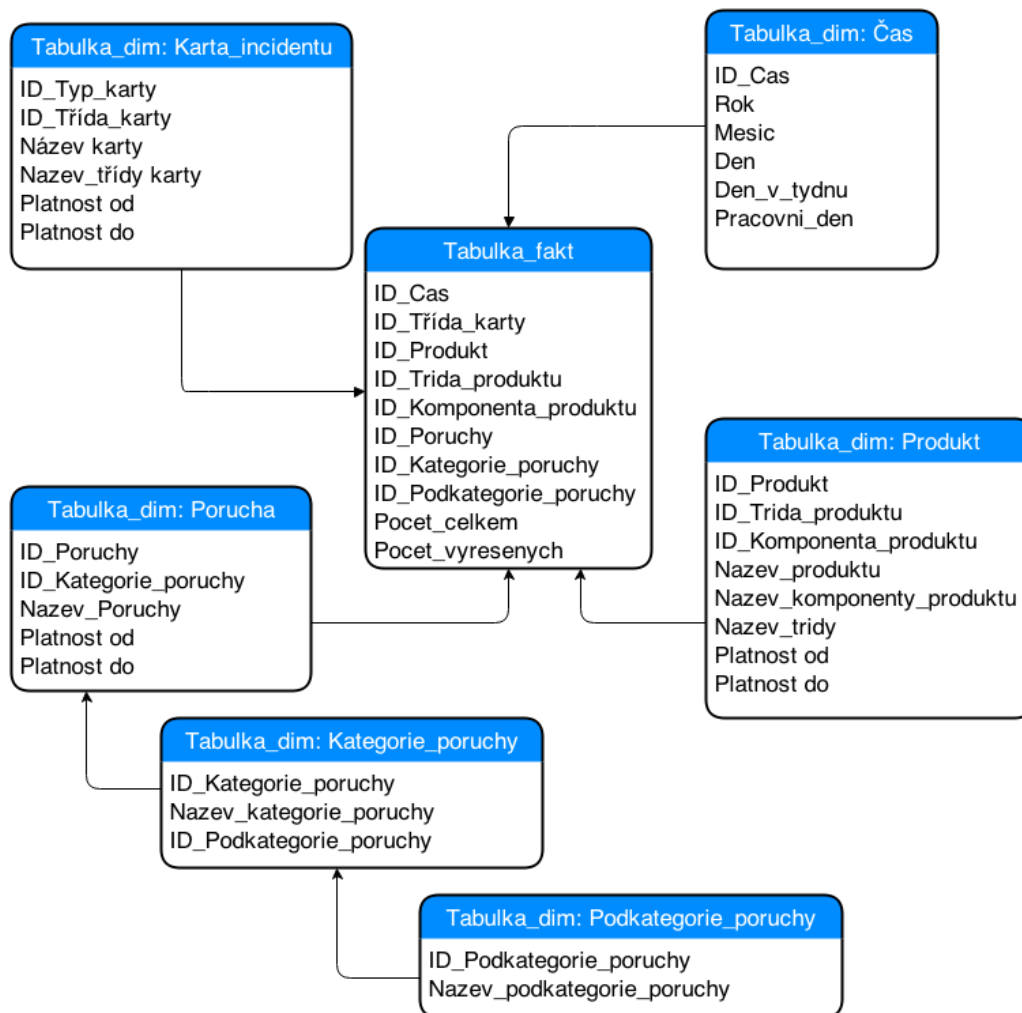


**Obrázek 3.12 STAR schéma**

Zdroj: TURBAN, Efraim. *Business intelligence: a managerial approach*. 2nd ed. Boston: Prentice Hall, c2011, xx, 292 p. ISBN 01-361-0066-X.

<sup>21</sup> Míry (measures) jsou vypočtené nebo dopočtené hodnoty ze známých dat. Např. pro danou službu vidíme tržbu, ale ještě se nám dopočítá jako další informace průměrná tržba služby za posledních 12 měsíců, aniž bychom museli dopočít provádnět dodatečně sami.

SNOWFLAKE schéma již oproti STAR řeší normalizaci, tedy aby se hodnoty vyšších úrovní hierarchie (zde např. ID\_Poruchy, Nazev\_Poruchy) v jedné tabulce dimenze neopakovala i pro nižší úrovně. Vše je odděleno do samostatných tabulek.



**Obrázek 3.13 SNOWFLAKE schema**

Zdroj: TURBAN, Efraim. *Business intelligence: a managerial approach*. 2nd ed. Boston: Prentice Hall, c2011, xx, 292 p. ISBN 01-361-0066-X.

V praxi je pak ale běžné obě schémata kombinovat, tabulek fakt se používá současně více a dimenzí jsou stovky. Tedy jedna dimenze je použita současně pro více tabulek faktů. Zde hrozí (při nedostatečném popisu dat), že bude nějaká dimenze použita nevhodně i z důvodu usnadnění si práce při implementaci BI. Bude například bez dostatečného detailu, či nebude popisovat správně hodnoty ve faktových tabulkách.

Tabulka 3.3 Porovnání SNOWFLAKE a STAR modelu

	STAR	SNOWFLAKE
<b>Hierarchie</b>	Jednotabulková	Vícetabulková
<b>Normalizace</b>	NE	ANO
<b>Doba zavedení modelu</b>	Rychlá	Průměrná
<b>Změny v modelu</b>	Nenáročné (méně tabulek)	Nutná pečlivost (Náročnější)

Zdroj: vlastní

### 3.8.3 OLAP kostky

Základní strukturou pro získání informací z dat v OLAP je kostka, která se skládá z dimenzí a faktů[4]. Z tabulek faktů se vybírají takové atributy, které jsou potřebné pro získání sledovaných ukazatelů. Dále se zde definují typy agregací pro případné vyšší hierarchie (MAX, MIN, AVG, SUM, COUNT). Z tabulek dimenzí pak vybereme takové, které pokryjí potřebu uživatelů, pro které je OLAP kostka připravována.



Obrázek 3.14 OLAP kostka  
Zdroj: vlastní

Získání konkrétní informace z kostky se alternativně nazývá Slice&Dice. Nabízí se totiž přirovnání ke „krájení na plátky a na kostičky“ celé OLAP kostky. Vzniklá „ukrojená“ jednotka pak již obsahuje výslednou hodnotu, číslo a tedy je informací.

### 3.8.4 Druhy OLAP

Existuje několik druhů OLAP databází, toho jsou tři základní (nejčastěji používané)[22]:

- ROLAP – relační OLAP, faktové i dimenzionální tabulky jsou uloženy v klasických relačních databázích. Data jsou pak z databáze získávána běžným dotazovacím jazykem (např. MS SQL) a výběr se provádí pomocí klauzule WHERE v dotazu. Oproti ostatním typům OLAP je ROLAP vhodný pro databáze s velkým množstvím dat a je nejsnáze rozšiřitelný (díky relační struktuře). Další výhodou oproti ostatním typům je schopnost zpracovat číselná i textová data. Také umožní rychlé rozložení dat na nejnižší úroveň. Nevýhodou jsou pak nižší výkon zobrazení informací z důvodu výpočtů přímo v databázovém serveru. Další nevýhodou je nutnost používání SQL pro manipulaci s daty. Některé specifické výpočty nemusí být přeložitelné do SQL dotazu.
- MOLAP – multidimenzionální OLAP je vlastně typický OLAP. Tento typ je založen na OLAP kostkách. Data jsou již dopředu agregována (na rozdíl od ROLAP). MOLAP díky agregaci a struktuře dat v kostce je více optimalizovaný a tedy výstupy (ukládání) jsou rychlejší než u ROLAP a také méně náročný na HW. Nevýhodou je pak menší schopnost rozšiřitelnosti (pro MOLAP existuje limit objemu dat), někdy dokonce ani MOLAP rozšířit nejde. Dále pak jsou data redundantní z podstaty datové kostky.
- HOLAP – hybridní OLAP kombinuje ROLAP i MOLAP do jedné architektury. HOLAP ukládá detailní data do relačních databází, ale agregované hodnoty do připravených OLAP kostek. HOLAP umí rozklad (drill) z kostek až na jednotlivé datové entity v relačních databázích. Výhody HOLAP jsou dobrá rozšiřitelnost, rychlé zpracování dat a pružnost v přístupu k datům.

Další typy jenom krátce zmíním: WOLAP (Web OLAP pro spolupráci OLAP-webový browser), DOLAP (Desktop OLAP pro práci na lokálním PC s vybranou částí OLAP), Mobile OLAP (pro spolupráci OLAP-mobilní telefon či tablet) a SOLAP (Spatial OLAP pro spolupráci mezi prostorovými a běžnými daty)[22].



### 3.9 Analytické aplikace BI

V následující části představím kromě OLAP, které již bylo zmiňováno v předchozích kapitolách, další oblasti a možnosti aplikace analytických prvků BI v kombinaci s dalšími systémy.

#### 3.9.1 Analytické CRM (ACRM)

*„Analytické CRM představuje komplex aplikací IS/IT, zaměřených na poznání zákazníka, jeho hodnoty, preferencí rizikovosti nebo pravděpodobnosti odchodu ke konkurenci. Za účelem splnění tohoto cíle využívají tato řešení spojení systémů BI a CRM.[4]“*

Základem pro ACRM je vyhrazený datový sklad nebo tržiště [4], ve kterém jsou již vhodně předpřipravená data (např. agregací), vztahující se k zákazníkovi v různých pohledech a zákaznických i řídicích procesech. Hlavními aktivitami v rámci ACRM jsou pak:

- Segmentace zákazníka (potenciální tržba, skutečná tržba, počet zaměstnanců, počet poboček, oblast podnikání atd.)
- Marketingové kampaně – vyhodnocení (počet kampaní za rok, úspěšnost kampaně pro daný region, pokles nákladů díky migracím na jiný produkt atd.)
- Predikce chování zákazníka - jeho reakce na změny cen, reakce na kampaně, na konkurenční nabídky apod.
- Individuální přístup k zákazníkům a personalizace služeb.
- Retenční a reklamační aktivity (zákazník chce odejít ke konkurenci, ale díky ACRM lze vysledováním aktivit zákazníka identifikovat důvod a jeho rozhodnutí změnit)

ACRM je tedy velmi závislé na kvalitě dat a jejich shromažďování z interakcí se zákazníkem. Následně pak umožňuje provádět analýzy nebo třeba připojit analytické nadstavby na ACRM. Navíc je možné určit hodnotu zákazníka např. V následujících oblastech[4]:

- Ziskovost zákazníka – určení hrubé marže (výnosy-náklady /výnosy).
- Odliv zákazníků (Churn Management) – snížení odchodu zákazníků ke konkurenci.
- Celoživotní hodnota zákazníka (Customer Life-Time Value) – jedná se o sumu absolutní marže ze služeb zákazníka v uplynulém i odhadnutém budoucím období.

### 3.9.2 Web Analytics

Jedná se o BI analytické aplikace spojené s webovými stránkami a provozem nad nimi. Příkladem může být Google Analytics<sup>22</sup>. Umožňují získat například následující informace a převést je do datového skladu, kde je lze dále po vhodné transformaci analyzovat[4]:

- Popisné statistické údaje – počty přístupů, časy, typy prohlížečů
- Chování uživatelů na stránkách – jejich pohyb, klikání na záložky a odkazy.
- Identifikace – IP adresy, cookies informace, data z formulářů.
- Analýza webového kanálu – na stránky lze nahlížet jako na jeden z prodejních kanálů a tím pádem je i stejně vyhodnocovat (náklady, výnosy, analýza kampaní)

### 3.9.3 Competitive Intelligence (CI)

Úkolem Competitive Intelligence (dále CI) je zmapovat do detailů trh a okolí podniku se zaměřením na konkurenci. Za tímto účelem využívá jak interních zdrojů, tak i externích analýz. Získávaná data mohou být jak strukturovaná (veřejné zdroje) tak nestrukturovaná formou rozhovorů[4]. Následně pak po získání dat dochází k porovnávání (benchmark) s konkurencí v různých aspektech a ukazatelích. Data jsou ukládána např. do znalostních bází, indexována pro snazší možnost vyhledávání a dále analyzována nebo prezentována.

### 3.9.4 Business Activity Monitoring (BAM)

BAM umožňuje získat klíčové ukazatele výkonu podniku v reálném čase na rozdíl od běžného BI. Nepoužívá se pro strategické rozhodování ale pro nižší úrovně (nejvíce

<sup>22</sup> Dostupné na <http://www.google.com/analytics/>

operativní) rozhodování. Vedoucí zaměstnanec musí mít možnost okamžité kontroly stavu a případně jeho nápravy tak, aby byly plány podniku dodrženy. Typický příklad je výroba v podniku (porovnání výroby s plánem, případně posílení výrobních kapacit) nebo technická podpora (průměrná doba čekání na řešení poruchy a opět např. navýšení operátorů). Nástroje BAM často využívají Dashboardy s ukazateli okamžitého stavu (např. na kontaktním centru počet volajících ve frontě s rozdělením dle IVR<sup>23</sup>). Další funkcionalitou nástrojů pro BAM jsou spouštěče (triggery), které na základě nějaké události (např. překročení určitého stavu nebo délky) reaguje definovaným způsobem (odešle email, zápis do logu, přesměruje volající na záložní kontaktní centrum atd.).

---

<sup>23</sup> Hlasový strom, kdy si volající volí pomocí stisku čísla svoji frontu

### **3.10 BPM (CPM)**

V této části představím BI komponenty typu Business Performance Management (dále BPM), také nazývané Corporate Performance Management (CPM), Enterprise Performance Management (EPM) nebo Strategic Enterprise Management (SEM).

#### **3.10.1 Úvod do BPM**

Pojem BPM je termínem, který pokrývá procesy, metodiky, měření a technologie potřebné k měření a řízení výkonnosti firmy[23]. Toto zahrnuje tři hlavní oblasti[24]:

- Sadu integrovaných a provázaných řídicích a analytických procesů, které jsou zaměřeny na finanční i operativní oblasti podnikání. Vše je postaveno na odpovídající IT.
- Nástroje pro přípravu strategického plánování a následného měření a porovnání skutečných výsledků oproti těmto plánům.
- Základní sadu firemních procesů, zahrnujících finanční a operační plánování, reporting, datové modelování, analýzy a sledování klíčových ukazatelů (KPI) potřebných k hodnocení firemní strategie.

BPM není to samé jako BI. Zatímco BI se zaměřuje na měření tržeb, kvality, zkoumání zákazníka, BPM tyto nástroje obohacuje o celý manažerský cyklus (strategie, plán, analýza, rozhodování)[3]. Další rozdíl od BI je že pohled BI na komponenty a společnost je zdola-nahoru a BPM je shora-dolů.

BPM využívá některé typické metodiky pro řízení výkonnosti, jako jsou Balanced Scorecard, Six Sigma, Activity Based Costing, Enterprise Value Added a další[4]. K tomu, aby tyto metodiky mohly být efektivně uplatněny, je třeba využít BI řešení s jejich podporou.

Nasazení BPM je možné buď společně s BI formou modulů na ERP (Oracle SAP) nebo dedikovanými aplikacemi BI/BPM, které se integrují s primárními zdroji samostatně[4]. Typickými oblastmi užití BPM pak jsou[21]:

- Strategický management
- Rozpočtování (Budgeting), plánování a předpovídání (Forecasting)

- Konsolidace finančních dat
- Optimalizační modelování a analýzy profitability
- Finanční, manažerský, regulovaný (státem nařízený) reporting

### 3.10.2 Scorecard a Dashboard

Pro vizualizaci BPM se používají dva hlavní typy komponent nebo nástrojů: Scorecardy a Dashboards, souhrnným názvem Pracovní panely[4]. V praxi bývají často zaměnitelné, mají hodně společného. Rozdílem může být oblast použití. Zatímco Scorecard se využívá pro strategické sledování výkonnosti, Dashboard spíše pro taktické a operativní, viz kapitola 3.9.4.

Charakteristické vlastnosti pro Pracovní panely pak jsou[11]:

- Monitoring – abstrahovaná data v grafické podobě k sledování KPI
- Analýza – data v souhrnné podobě pro analýzu hlavních příčin problémů
- Řízení – detailní operační data pro zjištění, jak problém efektivně vyřešit

Velmi typické a uživateli vyžadované je zobrazení veškerých požadovaných informací na jedné obrazovce (webové stránce nebo prezentační BPM SW).

Dobře navržené Pracovní panely pak obsahují a splňují následující[25]:

- Používají vizuální prvky (grafy koláčové, grafy sloupcové, řádkové minigrafy (sparklines), měřidla typu tachometr) pro zviditelnění či zvýraznění dat na první pohled, zvláště výjimek a extrémů.
- Uživatelsky přívětivé, jednoduché a srozumitelné.
- Kombinují data z různých zdrojů v jeden agregovaný a konsolidovaný pohled
- Umožňují Drill-down a drill-through<sup>24</sup>, tzn. dostat se ke zdrojovým, neagregovaným hodnotám nebo reportům.
- Zobrazují změny s minimálním zpožděním proti skutečné události.
- Vyžadují minimální zásahy programátora pro přípravu, nasazení do provozu a správu při provozu.

---

<sup>24</sup> Drill-down umožní postupný rozklad sumarizovaných dat na vysoký detail (stát, město, čtvrť, ulice) a Drill-through rozkládá data z agregované podoby rovnou na vysoký detail.



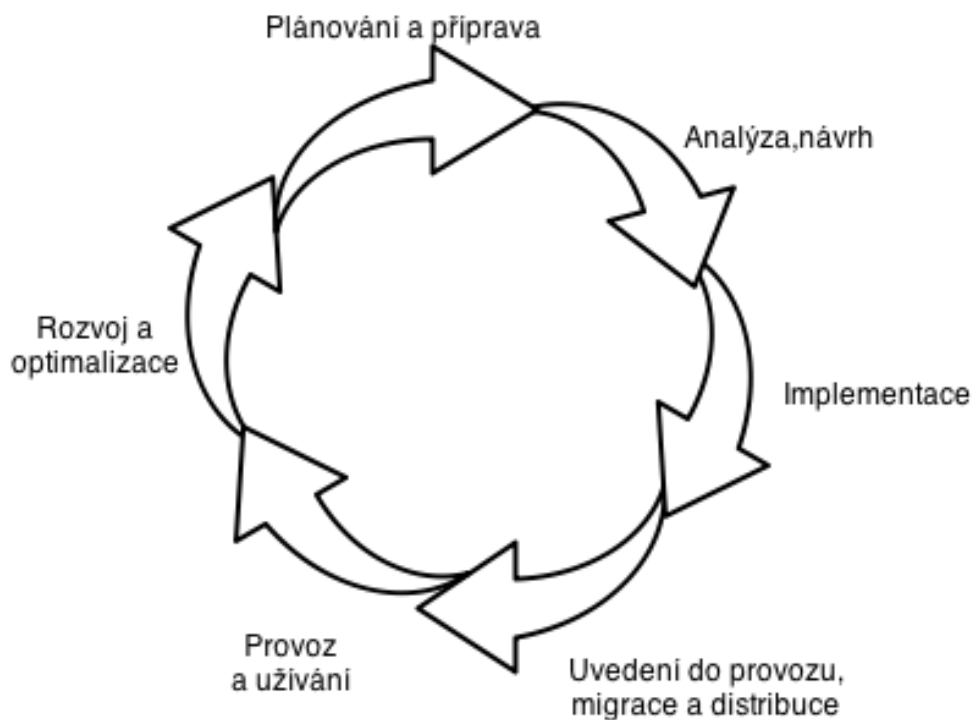
**Obrázek 3.15 Dashboard Operativní**

Zdroj: <http://www.analytikdat.cz/dashboard/2013/02/15/co-je-dashboard.html>

### 3.11 Základní procesy implementace BI

Proces implementace BI je natolik komplexní a specifický pro jednotlivé společnosti, že se v této práci zaměřím pouze na hlavní a typické fáze procesu.

Celý proces implementace vychází z charakteristiky životního cyklu SW, viz Obrázek 3.16. Jednotlivé fáze jsou pak v BI pojetí obohaceny o konkrétní prvky, které budou popsány v následujících stránkách.



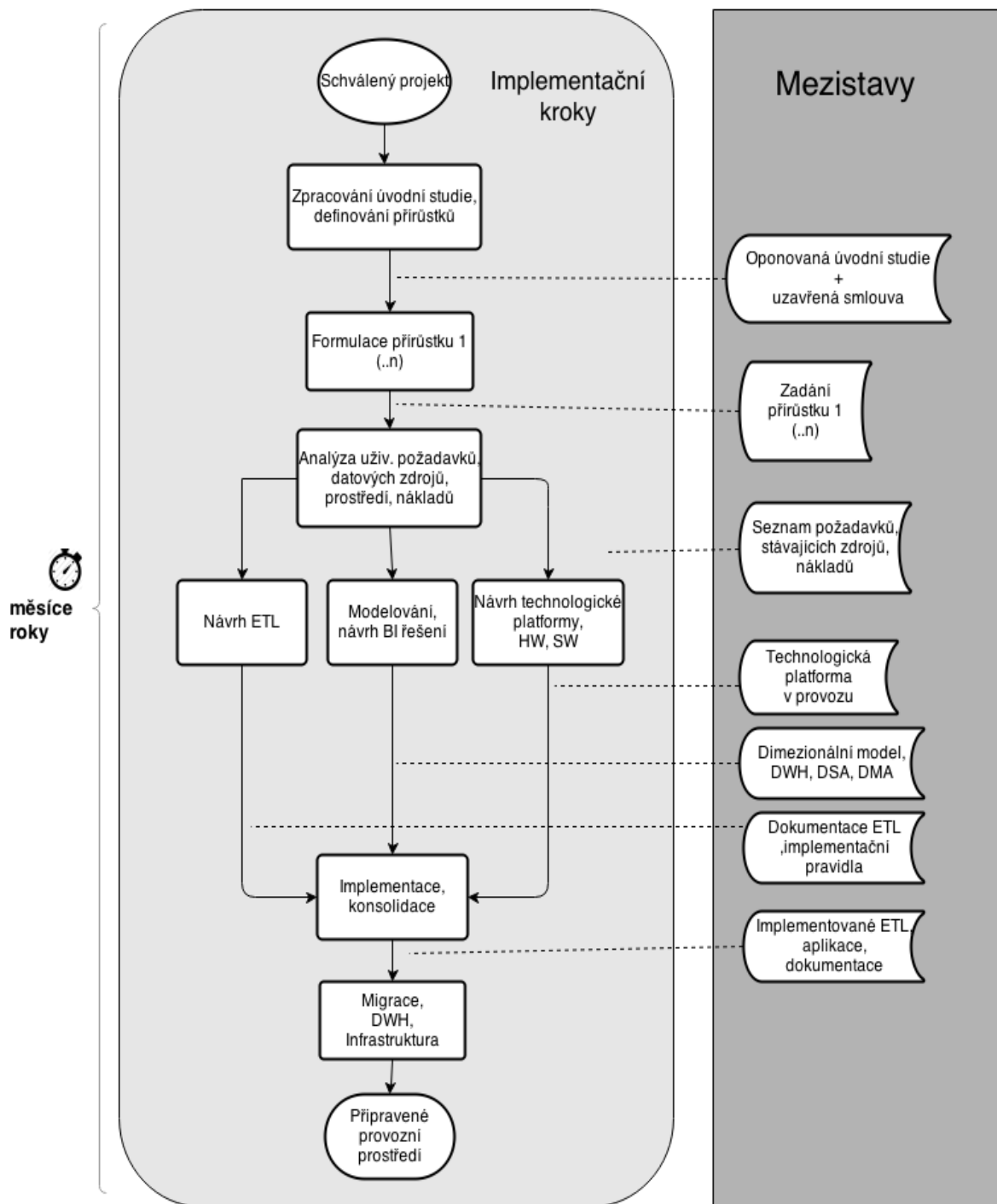
**Obrázek 3.16** Životní cyklus SW

Zdroj: POUR, Jan, Miloš MARYŠKA a Ota NOVOTNÝ. *Business intelligence v podnikové praxi*. 1. vyd. Praha: Professional Publishing, 2012, 276 s. ISBN 978-80-7431-065-2.

Následující Obrázek 3.17 tedy již popisuje celý BI proces implementace. Skládá se z výchozího stavu, což je schválený projekt implementace, jednotlivých implementačních kroků a tzv. mezistavů, což jsou naplánované, očekávané a dokončené skutečnosti (např. vytvořená dokumentace) v průběhu implementace. Výsledný stav diagramu je pak připravené BI řešení k předání do provozu nebo již existující řešení obohacené o dohodnutý přírůstek na různých BR vrstvách. V případě celého nového řešení se implementace BI pohybuje v řádech měsíců, spíše let. U přírůstkového řešení BI bych odhadoval čas spíše na týdny až měsíce (technologická platforma je zavedená, ETL a ostatní komponenty funkční).

V porovnání s životním cyklem SW by tedy do Analýzy a návrhu patřilo zpracování úvodní studie (to částečně i do Plánování), formulace přírůstku, Analýza uživatelských požadavků, modelování BI, návrh ETL a návrh technologické platformy. Další fáze, Implementace a Migrace mají své samostatné označení v obou pohledech. Fáze, týkající se provozu a užívání BI v této práci nebude zkoumána a pro fázi optimalizace a rozvoje je vyhrazena případová studie v druhé části práce.





**Obrázek 3.17 BI implementace**

Zdroj: POUR, Jan, Miloš MARYŠKA a Ota NOVOTNÝ. *Business intelligence v podnikové praxi*. 1. vyd. Praha: Professional Publishing, 2012, 276 s. ISBN 978-80-7431-065-2.

### 3.11.1 Zpracování úvodní studie

Kvalitní a detailní plánování a zpracování úvodní studie je základem celé implementace. Při plánování je důležité zaměřit se na informace (výstupy), které BI řešení může zprostředkovat a jak tyto nově získané informace pomohou firmě na trhu[9]. Vstupem tohoto implementačního kroku je například schválený plán firmou s termíny. Výstupem pak zadokumentovaný schválený návrh plánu (úvodní studie). V tomto úvodním implementačním kroku se tedy implementační tým seznamuje s prostředím firmy a provádí při tom následující dokumentační aktivity [4][9]:

- Definice cílů projektu a jejich soulad s podnikovou strategií
- Vytvoření seznamu klíčových uživatelů (specialistů), zastupující jednotlivé úseky nebo oddělení.
- Specifikace všech hlavních požadavků na BI, rámcově zachycené do mapy procesů, činností a postupů. Požadavky obsahují analytické, reportovací a plánovací potřeby. V této fázi jde o zachycení pouze rámcové
- Přiřazení priorit jednotlivým skupinám/požadavkům uživatelů v souvislosti s domluveným rozsahem projektu.
- Analýza stávající IT infrastruktury, její slabé a silné stránky. Analyzuje se jak HW, tak i SW.
- Analýza nákladů a benefitů. A to jak ekonomických, tak i nefinančních (např. posílení pozice na trhu bez konkrétního procentního bodu o kolik)
- Návrh BI architektury tj. určení hlavních komponent BI řešení (aplikace, DWH, DMA, ETL) a vazeb mezi nimi.
- Harmonogram projektu rozepsán na jednotlivé přírůstky a implementační kroky
- Organizační záležitosti, tj. ověření kompetencí a nastavení zodpovědností jednotlivých týmů a manažerů.
- Zajištění firemního garanta projektu z nejvyššího vedení, který zajistí případně potřebná klíčová rozhodnutí.
- Pečlivá dokumentace

V případě specifikace přírůstku BI řešení je pak situace jednodušší, protože platforma již existuje. Zde pak jsou aktivity BI týmu následující[4]:

- Analýza aktuálního stavu, kde jsou zahrnuté nové nebo upravené uživatelské požadavky oproti Úvodní studii. Také sem spadá analýza stávajících komponent BI řešení (DWH, ETL atd.)
- Definice přírůstku, tj. cíl, efekty, obecná funkcionalita, garant, interní zákazník řešení.
- Harmonogram a finanční hlediska (náklady a přínosy)

### 3.11.2 Analýza

V rámci analýzy se soustředí implementační tým na detailní rozbor funkčních a technických uživatelských požadavků, které pak společně s detailní analýzou datových zdrojů a IT infrastruktury vyhodnotí a navrhne obsah přírůstku[4].

- Nestrukturované pohovory s klíčovými uživateli pomohou upřesnit uživatelské požadavky, zadané v úvodní studii. Jedná se o i o verifikaci požadavků.
- Analýza požadavků s cílem odstranit duplicitní požadavky, sloučit podobné. Během této analýzy pak probíhá posouzení realizovatelnosti požadavků ve vazbě na dostupné zdroje (finanční, IT, HR, data).
- Analýza datových zdrojů, jejich čistoty a kvality, dostupnosti. Vyzdvihnutí problémových oblastí nebo úzkých míst, odhad předpokládaných nároků na nové ETL procesy a revize časového harmonogramu.
- Analýza dopadu změn implantace do primárních zdrojů. Je možné, že zadané uživatelské požadavky budou vyžadovat změny až v produkčních databázích, např. přidat nové atributy.

### 3.11.3 Design BI řešení

Cílem tohoto implementačního kroku je návrh datových modelů pro DWH, DMA nebo OLAP. K dosažení cíle se používá dimenzionální modelování, jehož součástí je i analýza dimenzí (stávajících a nových). V této fázi jsou třeba přinejmenším těchto činností[4]:

- Dimenzionální modelování potřebných a domluvených firemních ukazatelů do formy dimenzí a faktů, případně provedení potřebných změn pro přírůstkové řešení.

- Definice základního obsahu dat v datových úložištích (DSA, ODS, DWH a DMA) a jejich vazeb. Obsahuje i případný přírůstkový návrh změn.
- Definice relační databáze pro datová úložiště (logický model). Zahrnuje i odhadovanou velikost datového skladu a očekávané absolutní nebo relativní přírůstky.
- Design OLAP kostek, jejich atributy. Zde se také určují zdrojové tabulky z datových úložišť. Všechny potřebné OLAP atributy se pro kostky určují v této fázi, tedy i KPI, dimenze, hierarchie, míry atd.
- Definice chování BI analytických aplikací pro jednotlivé uživatelské role/oprávnění. Toto může zahrnovat obecné určení uživatelských rolí pro přístupy do datových úložišť, tak i přístupy pro aplikace na vyšších vrstvách. Tedy nastavení autentizace i autorizace.
- Návrh reportů, jejich formátu a struktury, omezující filtry. Volba atributů, typ zobrazení (web, SharePoint, další aplikace) a umístění. Doplňující možnosti (export reportu do domluvených formátů).
- Příprava „Semaforů“, tj. zadání jednotlivých limitních hodnot pro klíčové ukazatele (např. zákaznická tržba červená <5000Kč, oranžová >5000, zelená > 20000)
- Vytvoření demo prostředí pro testování a prezentaci připravovaných nástrojů, případně testovacích scénářů. Po připomínkách uživatelů lze schválené nástroje nasadit do produkce. Toto testování probíhá paralelně s dalšími implementačními kroky.
- Příprava školicích materiálů pro uživatele.

### 3.11.4 Návrh technologické platformy

V tomto implementačním kroku je třeba co nepřesněji odhadnout nároky na potřebný SW a HW a tyto zadokumentovat[4]:

- Vymezení podstatných komponent BI řešení (nebo jejich upgrade). V rámci SW se řeší OS, SŘBD, ETL aplikace. U HW se pak řeší komponenty serverů (paměť, prostor, výkon), případně třeba virtualizace front-end aplikační vrstvy<sup>25</sup>.

---

<sup>25</sup> Virtualizační nástroje jako např. VMware

- Fyzický datový model tabulek v databázích (datové typy), včetně jejich atributů a vazeb.
- Řešení samotného uložení dat (RAID pole), výkony databází, provozní charakteristiky.
- Konkrétní řešení očekávaného nárůstu objemu uložených dat (rozšíření diskového pole za provozu) a možnosti optimalizace.
- Školení uživatelů, vytvoření komunikačního kanálu pro helpdesk jako podporu uživatelů, certifikace uživatelů.
- Definice konkrétních uživatelských skupin a nastavení přístupových práv pro jednotlivé aplikace, databáze, OLAP. Přiřazení uživatelů do skupin.
- Seznam nároků na zajištění provozu BI řešení (personální nároky, zabezpečená konektivita do firmy pro vzdálenou správu, sdílené úložiště pro dokumenty, další bezpečnostní prvky).

### 3.11.5 Návrh ETL

U návrhu ETL se předpokládají takové kroky, které detailně vymezí potřebné procesy[4]:

- Detailní pravidla ETL, nastavení parametrů pro kontroly, určení periodicity.
- Nastavení zodpovědností pro jednotlivé ETL procesy.
- Specifikace ETL mezi různými BI vrstvami a komponentami pro aktuální přírůstek.
- Otestování navržených ETL procedur na demo prostředí a vyřešení vzniklých problémů, např. chybějících dat a NULL hodnot.
- Vyhledání a vyřešení možných problémů u integritních vazeb, tj. vazby mezi záznamy v různých tabulkách (primární vs. cizí klíč), kdy jeden odstraněný záznam způsobí chybovou situaci v jiné tabulce s vazbou na něj.

### 3.11.6 Implementace a konsolidace

Implementační fáze realizuje tři předchozí, na sebe paralelní, kroky návrhů (Tech. Platformy, ETL a Designu)[4]. Výstupem implementace je funkční BI řešení, připravená kompletní dokumentace (uživatelská i provozní). Konsolidace je zde zmíněna pro

přírůstkové BI řešení a je třeba mít celý funkční a otestovaný přírůstek na všech komponentách a vrstvách připraven k zavedení do provozu.

### **3.11.7 Migrace, DWH, infrastruktura**

Výstupem této fáze jsou plná datová úložiště v **provozním prostředí**, provedený upgrade nebo nasazení technologické platformy a aktualizovaná provozní dokumentace[4]. V praxi se provádí pod kontrolou vývojového týmu z důvodu možných problémů s kompatibilitou provozního a demo prostředí.

### **3.11.8 Hodnocení implementace BI**

Po kompletním nasazení do provozu všech komponent BI je implementace vyhodnocena [9]. Projektoví manažeři, klíčoví uživatelé, garanti projektu a další zúčastnění hodnotí v předem připraveném anketním formuláři, případně nestrukturovanými odpověďmi. Hodnotí se jednotlivé fáze implementace a také pohled na celek jako výstup.

Typické důvody neúspěšné implementace BI:

- Nevhodně postavený implementační tým a nedostatečně kvalifikovaní BI zaměstnanci
- Příliš mnoho dat a Nevhodná strategie
- Nízká kvalita dat
- Malá podpora/spolupráce klíčových uživatelů a koncových uživatelů
- Přílišná důvěra v dodavatele BI řešení
- Nepochopení účelu BI řešení
- Malá podpora a nezájem vedení firmy
- Nedostatečný výkon (pomalé dotazy, reporty, ETL atd.)
- Zbytečné doplňky
- Problémový Outsourcing
- Nedostatečná příprava

## **4 Případová studie**

Cílem případové studie je optimalizace stávajícího řešení reportingu a analýz pomocí implementace BI nástrojů.

Studie bude částečně charakterizovat a analyzovat situaci v GTS Czech s.r.o., firmě, ve které jsem zaměstnán jako specialista reportů. Součástí analýzy je i nestrukturovaný pohovor. Následně po vyhodnocení analýzy aplikuji prvky BI implementace v souladu s předchozí teoretickou částí. V závěru vyhodnotím celou optimalizaci.

### **4.1 Představení společnosti GTS Czech**

V krátkém popisu zde přiblížím společnost GTS Czech, jaká je oblast podnikání a segment zákazníků, na které se zaměřuje, dále zmíním krátce i současný stav na zacíleném trhu.

#### **4.1.1 Poskytované služby**

GTS Czech s.r.o. poskytuje telekomunikační (dále telco) a ICT služby pro uživatele i své velkoobchodní partnery. Mezi telekomunikační služby lze zahrnout VoIP i TDM telefonování, video i telekonference, příchozí volání na zelená, modrá, bílá čísla (900xxx), mobilní služby, připojení k internetu, VPN propoje, Ethernet propoje, pronájem datových okruhů, pronájem celé fyzické přenosové kapacity (dark fiber). Z ICT typických služeb lze zmínit Anti-DDOS řešení, cloudové služby jako je Virtual Hosting, Virtual server, back-up služby apod.

#### **4.1.2 Profil společnosti**

GTS Czech s.r.o. je českou částí mezinárodní skupiny GTS Central European Holding B.V., která působí na telekomunikačním trhu dále v Polsku, Rumunsku, Maďarsku a na Slovensku. GTS Czech s.r.o. vznikla jako GTS Novera sloučením dvou velkých tuzemských alternativních telco operátorů Aliatel a GTS CZECH v roce 2005. O rok později majoritní vlastníci GTS Novera koupili další tři telco operátory Contactel, Telenor Networks a Nextra. V roce 2010 došlo k přejmenování na současné GTS Czech s.r.o. V současné době má společnost přibližně 450 zaměstnanců a poskytuje služby pro

přibližně 16 000 zákazníků. Společnost se zaměřuje na tržní segment zákazníků a jejich potenciálem SOHO+ a výše, tj. na zákazníky s celkovým měsíčním výdajem v oblasti telco 2000 korun, i kdyby byla část tržby u jiného operátora. Tržby za rok 2013: 4,559 mld. Kč, z toho 55% tvoří hlasové služby, 29% datové, 10% připojení k internetu, 5% managed a ICT služby.

Významní a referenční zákazníci: Ministerstvo práce a sociálních věcí, TESCO<sup>26</sup>.

V roce 2014 se stává GTS Czech součástí Deutsche Telekom AG a od 1.1.2015 pak integruje do společnosti T-Mobile.

### **4.1.3 Situace na trhu**

V dnešní době je český telekomunikační fixní i mobilní již saturovaný a v podstatě dochází pouze k churnu(odlivu) zákazníků mezi jednotlivými majoritními operátory. Největším konkurentem pro GTS Czech s.r.o. je národní poskytovatel O2.

Kvalita telco i ICT služeb je na vysoké úrovni jak po stránce technické (např. vysoká rychlost datového a internetového připojení, nízké odezvy, dostatečné SLA), tak po stránce technické podpory, CRM a péče o zákazníka všeobecně. Operátoři se snaží držet posledních trendů a hledají neustále nové produkty, které by oslovily uživatele. Ovšem stále platí, že zákazník slyší ponejvíce na změny cen služeb. Situace mezi operátory je taková, že v mnoha případech dochází k nákladovým i dumpingovým cenám v rámci konkurenčního boje. Novým směrem v posledních letech jsou konvergované služby a paušální hlasové tarify. Proto se čím dál více nabízí doplňkové služby k hlavní službě a zvýhodněné balíčkové služby, tzv. Triple play, tedy pevné hlasové i internetové služby do kanceláře a ještě navíc mobilní hlasové služby. Jednodušším typem je double play, což je vlastně pouze kombinace hlasových a internetových služeb za jeden paušál.

## **4.2 Plánování a příprava studie**

Při plánování a přípravě studie vycházím z metodiky práce kdy při zachování bezpečnosti a ochraně citlivých dat musím zajistit informace o výchozím stavu ve společnosti.

---

<sup>26</sup> Údaje jsou dostupné ve výroční zprávě GTS Czech pro rok 2013, která je poskytována PR oddělením na požádání.



Analýza datových zdrojů a technologické platformy bude tedy abstrahována na běžnou úroveň informace (náklady budou uváděny orientačně v řádech tisíců korun nebo v MD). Nestrukturované pohovory s klíčovými uživateli budou anonymizovány. Známe bude pouze oddělen klíčových uživatelů.

V rámci případové studie musí být přinejmenším naplánovány a následně provedeny následující analytické kroky:

- určení klíčových uživatelů, které oslovím v rámci **analýzy stávajícího stavu reportovacích nástrojů a potřeb uživatelů.**
- **analýza technologické platformy, DWH, ETL** – pro tuto část je třeba spolupracovat s IT oddělením a získat potřebné technické informace, metadata.
- **Určení neoptimálních charakteristik aktuálního BI stavu** – z předchozích dvou bodů
- **Návrh na změny stávajícího stavu**
- **Implementace návrhu**

Při samotném nasazení vytvořeného řešení musí být dodržena domluvená pravidla s BI oddělením, které zajistilo prostor na virtuálním serveru pro celý projekt pod podmínkami, že:

- nebude plýtváno s RAM a úložným prostorem serverů
- nebude vytěžován transaction log produkčního DWH v pracovní době (omezení uživatelů)
- nebudu zasahovat do stávajících tabulek, DWH logiky
- nebudu rekonfigurovat virtuální server (servery)

## **4.3 Nestrukturované pohovory s klíčovými uživateli a jejich vyhodnocení**

Zde je třeba se zamyslet, jak určit klíčové uživatele. Kdo používá nebo by mohl používat nástroje BI ať už na základě této nebo dalších implementací BI. Vzhledem k tomu, že společnost má 450 zaměstnanců a jistě všichni nepotřebují ke své práci nástroje BI, bylo třeba zvolit nějaký klíč výběru.

Vybral jsem tedy tyto skupiny:

- Musí minimálně pracovat pokročile s MS Excel (kontingenční tabulka, kontingenční grafy)
- Manažer nebo produktový manažer
- Je ve firmě minimálně 2 roky
- Respondenta musím znát osobně

Povinné byla pouze délka zaměstnání a to, že respondenta znám. Délka zaměstnání proto, že po dvou letech ve firmě je uživatel zkušený, zná primární systémy, zná již IT prostředky a procesy firmy, rozumí informacím, získaných z dat a dokáže tedy určit, co je v rámci BI a reportingu v GTS Czech dobře a co se může ještě zlepšit. Oslovil jsem 25 kolegů jako výstup této klasifikace. Získal jsem z toho celkem 17 uživatelů, kteří byli ochotni a měli k dispozici svůj čas k poskytnutí rozhovoru. To, že respondenta osobně znám pracovně také pomohlo ověřit jeho schopnosti. Většinou jsme spolu již něco řešili a měl jsem možnost poznat jeho analytické schopnosti a zkušenosti, které jsou pro připravovaný pohovor důležité.

Celý pohovor se skládal ze dvou částí: ankety a samotného nestrukturovaného pohovoru. Očekával jsem, že zkušenosti uživatelů s BI budou velmi malé (nebo nulové). Předpokládal jsem z vlastní zkušenosti, že úroveň statického reportingu je na vysoké úrovni a dynamický reporting nebude využíván.

### **4.3.1 Anketa**

Anketa byla odeslána uživatelům emailem, je strukturovaná formou otázka - volná odpověď a ohodnocení naléhavosti potřeby v případě dotazu, týkající se nějakého nedostatku (nice to have – optional - critical). Otázky jsem připravoval sám dle nejlepšího vědomí a první verze formuláře konzultoval se dvěma kolegy. Jejich rady a doporučení

jsem se pokusil zapracovat a poslední verzi ankety odeslal vybraným respondentům. Jednotlivé odpovědi uživatelů pak naleznete v příloze této práce. Níže vidíte podobu ankety v původní textové podobě.

Název ankety

Pozice:

Divize/oddělení:

Kdo je pro Vás zákazník? Vyberte Interní|Externí|Oba (interní = jiné oddělení):

1. Jaké BI nástroje ve firmě používáte? (např. Excel, Access, Reporty na Sharepointu, ProClarity, MS SQL Management studio)

2. Používáte kontingenční tabulku v MS Excel při vaší práci? ANO | NE

3. a) Jaké informace nebo oblasti z GTS systémů vám chybí v současném stavu reportů a analýz?

(jednotlivé požadavky ohodnoťte volbou z Critical | Optional | Nice to have)

b) Pokud ano, vypište Vaše požadavky:

.....

Critical | Optional | Nice to have

4. Jsou dostatečně rychlé aktualizace dat? Které reporty/oblasti (pokud nějaké) je třeba aktualizovat častěji? (jednotlivé požadavky ohodnoťte)

.....

Critical | Optional | Nice to have

5. Je pro vás dostatečný detail stávajících reportů? Které oblasti byste potřebovali vidět detailnější a které naopak více agregované? (jednotlivé požadavky ohodnoťte)

.....

Critical | Optional | Nice to have

6. a) Využíváte grafy a vizualizační prvky BI nástrojů (grafy reagují na poklikání myši přímo uvnitř grafu, rozbalují se do detailu, data se přepočítávají při výběru nějaké oblasti grafu)?

b) Pokud ano, jsou pro vás tyto prvky dostatečné? (vyberte Ano|Ne|Nelze určit)

7. a) Jaké jsou vaše zkušenosti s rychlostí odezvy a zobrazením při práci s reporty a BI nástroji v GTS?

b) Jak by bylo pro vaši práci zrychlení důležité?

(důležitost zrychlení ohodnoťte Critical | Optional | Nice to have)

8. a) Jaké máte zkušenosti s datovými OLAP kostkami?  
b) Pomohlo by vám vytvoření datové kostky ve vašich analýzách?  
(vytvoření nové kostky ohodnoťte Critical | Optional | Nice to have)
9. a) Jste spokojeni s rychlostí a kvalitou tvorby nových reportů, pružností změn a úprav v reportech?  
b) Kolik času průměrně trvá vytvoření reportu? (vyberte níže)  
1-3 dny  
4-7 dní  
7-14 dní  
14dní - měsíc  
Více než jeden měsíc  
c) Jak důležité by pro vás bylo zkrácení doby vytvoření (změny) požadovaných reportů u oddělení IT nebo u dalších interních dodavatelů?  
(vyberte Critical | Optional | Nice to have)

Mým záměrem při vytváření ankety bylo oddělit „zrno od plev“. Tedy ověřit si, že oslovení respondenti jsou skutečně kompetentní. Otázky jsem volil jak s volnou odpovědí (s otevřeným koncem), tak uzavřené s nucenou odpovědí (výběrové)[12]. Tázal jsem se hlavně na uživatelské zkušenosti s BI a se současným stavem v GTS Czech a také na zkušenosti s IT stavem z pohledu běžného uživatele (rychlost odezvy systému, tvorba reportů atp.).

#### 4.3.2 Nestrukturovaný pohovor

Pohovor následoval po vyplněné anketě, kterou jsem si přečetl a následně podle odpovědí upřesňoval jejich význam. Některé odpovědi v anketě byly nejednoznačné a naopak některé otázky uživatelé pochopili v různých významech. Tedy sladění otázek a odpovědí byl jeden z cílů nestrukturovaného pohovoru. Dalším cílem pak bylo získat od uživatelů „soft“ informace, tedy jejich zkušenosti nikde neměřené a nevyčíslitelné přesně. Tyto pak doplnily výstupy z ankety a podaly celkem komplexní představu o současném vnímání BI z pohledu uživatelů. Zároveň se tím pádem již začal rýsovat prostor pro zlepšení současného stavu.

### 4.3.3 Vyhodnocení pohovoru a ankety

Jak jsem očekával od výsledků, reporting GTS dle hodnocení uživatelů je na vysoké úrovni a to ze dvou důvodů: Dobře postavený datový sklad a velikost firmy.

- Správně a vhodně postavené DWH s kvalitními metadaty umožní vytvoření nového reportu poměrně rychle a snadno.
- Velikost GTS pak stále dovolí pružnost od zadání požadavku na nový report, tak změny stávajícího. Není to ještě příliš byrokratický moloch, jakým se stávají společnosti nad tisíc zaměstnanců (v našem případě např. O2 a T-Mobile) – tam je procesní stránka pak již nutností a to s sebou přináší i podstatně delší čas na tvorbu reportů. Než celý proces proběhne od zadavatele k tvůrci reportu, všechny úrovně manažerů, bezpečnosti, IT schválí a okomentují, uběhne většinou několik týdnů<sup>27</sup>.

Anketa ukázala, že technická oddělení (network management center- dohled sítě a řešení poruch se zákazníky) BI vůbec neznají nebo nepoužívají. Naopak velkoobchodní divize BI používá denně a to datové kostky wholesale (dále WHSL) hlasového provozu a nad nimi analytický nástroj ProClarity, který sice ukončil svůj vývoj (společnost koupena firmou Microsoft a v 2007 vydána poslední verze ProClarity[26]). Nicméně uživatelům WHSL kostky poslední verze ProClarity 6.3 vyhovuje, změny by preferovali pouze drobné, kosmetické úpravy („nice to have“).

Obchodní oddělení, oddělení vymáhání pohledávek a oddělení péče o zákazníka jsou spokojeni se stávajícím stavem, vyjma častějších aktualizací. Oddělení financí a mezinárodní produkt by preferovali častější aktualizace a kolegové by rádi objevili BI možnosti datových kostek, ale pouze jako „nice to have“ .

Oddělení segmentového marketingu je spokojeno, reporty si vytváří sami v MS SQL s přímým přístupem do DWH, stejně tak oddělení financí. Opět ale „nice to have“ zájem o OLAP.

Víceméně většina uživatelů ve své oblasti zájmu (pracovní náplně) preferuje častější aktualizaci dat reportů. Ale stav reportingu jim naprosto vyhovuje, datové kostky nepotřebují.

---

<sup>27</sup> Vycházím zde ze zkušeností kolegů, kteří pro společnosti typu Enterprise pracovali a reporty požadovali.

Co mě překvapilo jako výstup ankety a pohovoru bylo, že i střední management<sup>28</sup> používá kontingenční tabulky a pracuje v MS Excelu, nezaměřuje se pouze na vizualizační prvek informace (např. grafy), pracuje i se syrovými daty. Nicméně právě vyšší management požaduje chybějící OLAP a BI nástroje, kromě stávajících reportů. Je zde tedy znát již odklon od běžného reportingu k analýze a BPM nástrojům (Scorecard, Dashboard).

Střednímu managementu stačí aktualizace dat 1 denně (možná i méně často), naopak klíčoví uživatelé z nižších pozic u některých typů reportů preferují častější aktualizace (2x a více/den).

Pouze jeden z klíčových uživatelů vyhodnotil svou potřebu BI jako kritickou a na toho se nyní blíže zaměříme. Jedná se o produktové oddělení, konkrétně hlasové služby. Z pohovoru vyplynulo, že by oddělení velmi uvítalo možnost analýz nad datovými kostkami, tak jak to mají jejich kolegové na WHSL. Prosím představte si, že hlasový provoz v takto velké firmě tvoří obrovské množství záznamů do databází a následná práce nad relačními tabulkami je velmi pomalá a zdlouhavá. Toto je jedna z hlavních výhod OLAP kostek. Podstatně **urychlí práci** (hledání informací) nad ohromným množstvím dat. Pokud by uživatel použil například MS Excel, tak z vlastní zkušenosti vím, že práce nad 10 000 řádků při počtu sloupců 5 a více je již pomalá a nekomfortní. Pokud se tedy podívám na současný stav do DWH, tak například jedna relační tabulka má za jeden den více jak 330000 záznamů (jednotlivých hovorů) jenom v odchozím směru! A v rámci analýzy se uživatelé hlasového produktu dívají spíše pohledem TOP-DOWN po měsících, kdy v případě potřeby rozbalují (drill down) potřebné segmenty informací a klesají hlouběji k větší granularitě na týdny, dny hodiny. Zde tedy vidím velký prostor pro zlepšení, tj. návrh BI řešení se všemi potřebnými komponentami.

---

<sup>28</sup> bohužel top management se mi nepodařilo do ankety zajistit

Níže příkládám zjednodušenou tabulku pro porovnání odpovědí uživatelů jednotlivých oddělení.

**Tabulka 4.1 Seskupené a zjednodušené odpovědi uživatelů**

<b>Oddělení</b>	<b>Jsou aktualizace dat dostatečně rychlé?</b>	<b>Používají BI</b>	<b>Chtějí používat BI</b>	<b>Naléhavost potřeby BI</b>
<b>Wholesale</b>	Ano	Ano	-	-
<b>Finance</b>	Ne	Ne	Ano	Nice to have
<b>Segment marketing</b>	Spíše ano	Ne	Ano	Nice to have
<b>Obchod</b>	Ne	Ne	Ne	-
<b>Dohled sítě</b>	Ano	Ne	Ne	-
<b>Produkt</b>	Ano	Ne	Ano	Optional/Critical
<b>Customer care</b>	Ne	Ne	Ne	-
<b>Vymáhání</b>	Ne	Ne	Ne	-
<b>Střední Mgmt</b>	Ano	Ne	Ano	Optional

## 4.4 Analýza stávajícího stavu BI v GTS Czech

Společnost GTS Czech je má v základu zákaznicky rozdělené na maloobchodní (retail) a velkoobchodní (wholesale). Ve své studii budu zkoumat pouze svět retailu, protože oba světy jsou poměrně rozdílné a tedy i jejich BI budou celkem odlišné, přinejmenším separátní. Postupně prozkoumám současný stav jednotlivých komponent, potřebných pro BI řešení, navrhnu vhodné úpravy a další kroky.

### 4.4.1 Analýza DWH a ETL

#### DWH

Hlavní databáze DWH v GTS Czech se označuje jako DSA, což je trošku zvláštní, protože DSA v logice DWH obecně je pouze mezistupeň, kdežto zde je to již koncový prvek pro většinu dat z primárních systémů po provedení ETL. Tedy do DSA se zdroje z primárních systémů nahrají a pak nad nimi provedou potřebné transformační a čistící operace a případné historizace<sup>29</sup> dat. Jednotlivé oblasti, seříděné podle primárních systémů se nazývají schémata, logicky se tím oddělí od ostatních tabulek. Například tabulky, které se týkají primárního zdroje DUR mají celou cestu jako název databáze/schéma /název tabulky tedy například tabulka DSA.DUR.INVOICE zaznamenává faktury.

Bezpečnost přístupu a práce nad DWH je řešena autentizací uživatelů a v rámci autorizace potom nastavením omezení např. mazání tabulek, nepovolené úpravy vytvořených view<sup>30</sup> a funkcí, zakázané vytváření procedur a triggerů atd.).

Pro přímý přístup do relačních tabulek DWH se používají v podstatě pouze dva nástroje: SQL Management Studio a MS Access.

Velikost DWH je v současnosti přes 6TB (včetně wholesale dat), obsahuje přes 12000 tabulek<sup>31</sup>. Data jsou komprimována, v nekomprimované mají dvojnásobnou velikost. BI oddělení odhaduje, že v DWH se nachází přes 60 miliard záznamů.

Základní vztah mezi DWH a přehled primárních systémů v GTS Czech viz Obrázek 4.1. Stručný popis jednotlivých primárních systémů, uvedených na obrázku pak naleznete v příloze k diplomové práci. Nicméně ty hlavní primární zdroje, ze kterých budu čerpat

---

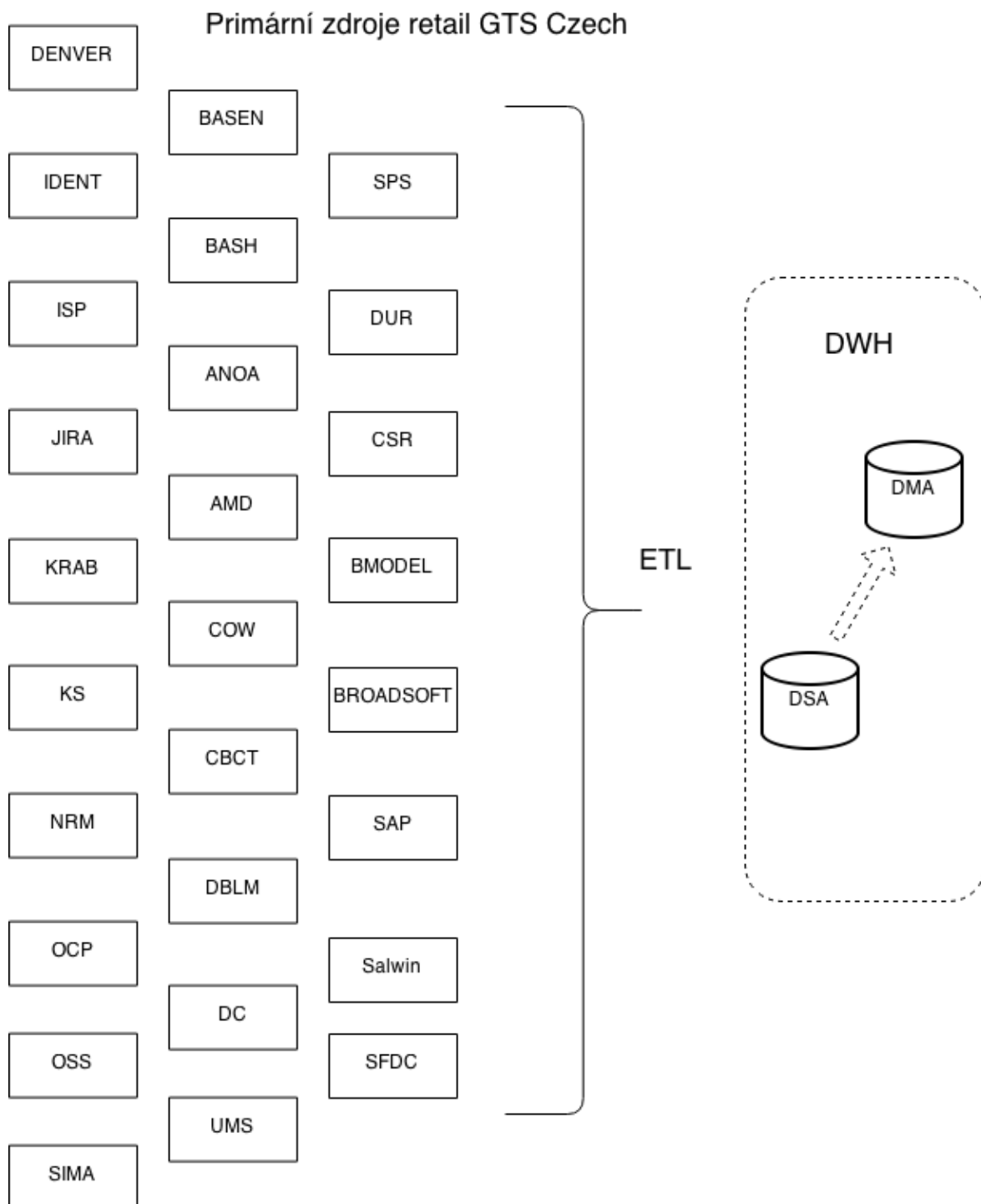
<sup>29</sup> Označení otisku dat časovou nálepkou, tj. umožní uživateli dívat se na data i historicky, nejenom poslední a aktuální verzi.

<sup>30</sup> Jedná se o virtuální tabulku, která je výsledkem nějakého SQL dotazu. Uživatel už ale nemusí znát samotný dotaz, vidí pouze tabulku.

<sup>31</sup> Vycházím z interní prezentace BI oddělení



data pro OLAP kostku zde popíšu trochu více. Bude se jednat o tři základní primární zdroje dat: DUR, CSR a Salwin.



**Obrázek 4.1 Primární zdroje GTS Czech v DWH**  
Zdroj: vlastní

**DUR** je označení pro Durian. Jedná se o účetní systém a zpracovává kompletní fakturaci v rámci služeb, účtů a zákazníků. Jsou tedy zde vazby účetní období (měsíc zdanitelného plnění), číslo služby, zákazník, částka a samozřejmě další atributy. Veškerá fakturační nastavení ke službám (ať už novým nebo změnám) se provádí a udržují v tomto systému. Důležitou součástí Durianu je v rámci DWH správné přiřazení ceny za jednotlivé uskutečněné hovory (hovorné) a to podle délky hovoru, přiřazeného hlasového tarifu k telefonnímu číslu, případné individuální ceně pro daný volaný směr (destinaci). V Durianu nalezneme i číselníky produktů, jejich komponent a ceníkových položek na komponentách. Zákazníkem se v Durianu rozumí jedinečný fakturační účet zákazníka (ano, jedno IČ jich může mít i více).

**CSR (Central Service Repository)** ukládá informace o jednotlivých samotných službách, zejména pak aktuální i historickou konfiguraci služby, adresu, komponenty služby, vazbu na přístupové okruhy poskytovatele, tedy kompletní informace o konkrétní službě z pohledu koncového bodu. Například u služby telefonní přípojky zde tedy bude uvedeno číslo služby, telefonní číslo/čísla, typ přípojky (Analog, ISDN), číslo přípojného okruhu, adresa, port na ústředně, jméno ústředny apod.

**Salwin** - jedná se o CRM. Uchováváme zde všechny potřebná data ve vztahu k zákazníkovi (nikoliv na úrovni služby ale skutečně až na úrovni zákazníka). Adresy sídel, kontaktní osoby s popisem jejich pozice, průměrné tržby za určitá období, v jakých kampaních byl zákazník zahrnut, a dále který obchodník (obchodníci) se o zákazníka stará. Zákazníkem se v Salwinu rozumí jedinečné IČ subjektu.

## **ETL**

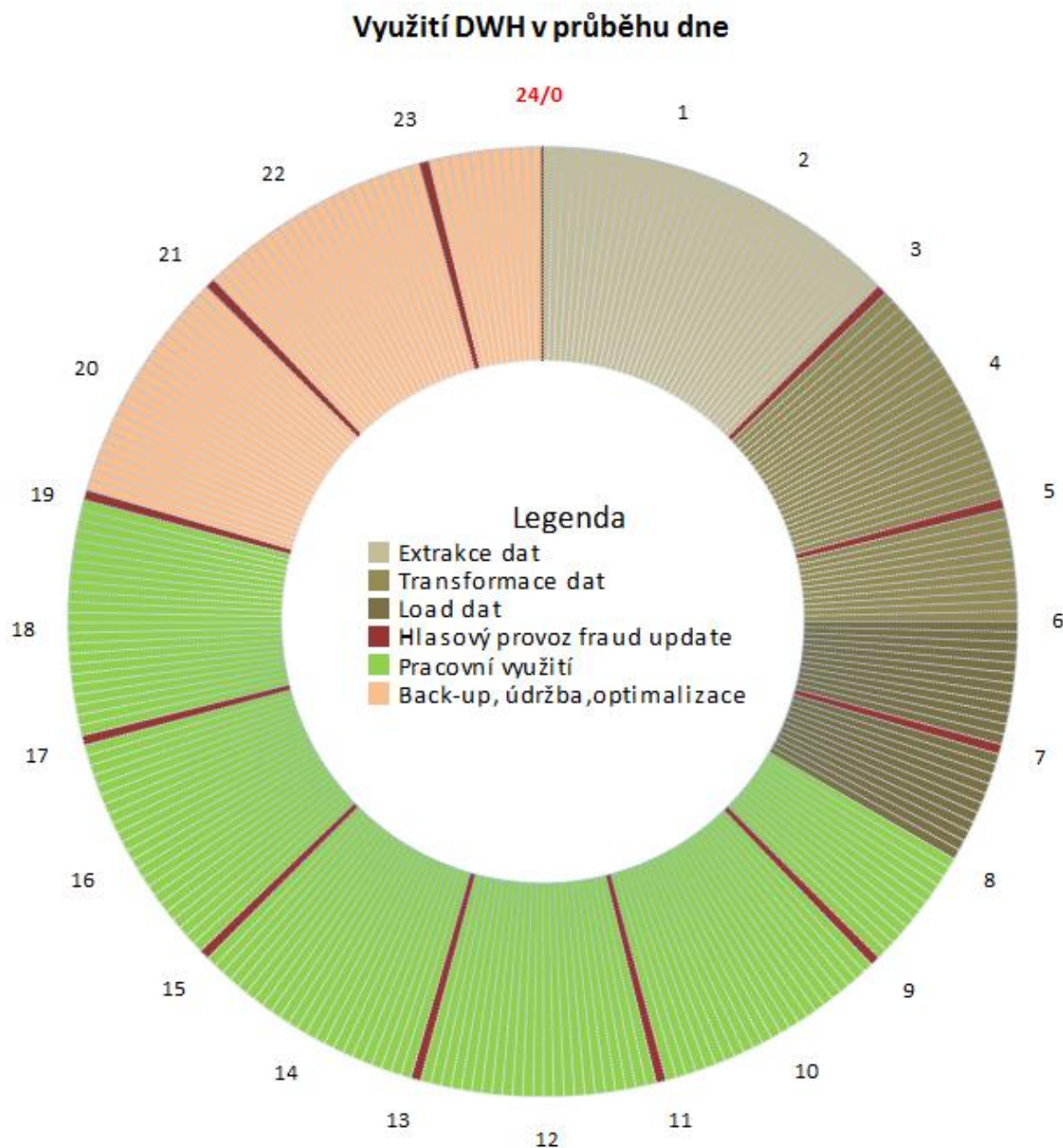
Většina tabulek v DWH je nahrávána z primárních systémů denně. Tyto tabulky jsou nahrávány celé (tedy smazat původní a znovu naplnit celou). Některé tabulky pak jsou nahrávány přírůstkovou metodou (tedy doplní se pouze o rozdíl mezi původním stavem a novým stavem). Tento přístup se používá, pokud je zmíněná tabulka velmi velká a celé přehrání by trvalo příliš dlouho, nebo když je spojení na primární systém příliš pomalé, případně nespolehlivé. Další tabulky se nahrávají vícekrát za den. Jedná se o podklady pro Fraud system<sup>32</sup> nebo pro WHSL služby.

---

<sup>32</sup> Ochrana proti podvodnému volání přes netušícího zákazníka. Může se stát, že hacker napadne ústřednu zákazníka a pak na jeho účet provolává velké částky/ dlouhé hovory. GTS se tímto proaktivním nástrojem snaží tyto útoky odhalit dříve, než dojde k milionovým škodám, které jsou pak vymáhány po zákazníkovi.

V jedné denní (respektive noční) dávce se přenese okolo 500GB dat v 2 miliardách záznamů. V měsíčním přírůstku do DWH pak jde o 100GB dat. BI odhaduje maximální propustnost na 2 miliardy záznamů / hodinu.

S ETL pro Fraud system a WHSL analytické potřeby se jedná celkem o 3 miliardy záznamů za den.



**Obrázek 4.2 Využití DWH v průběhu dne**  
Zdroj: vlastní

Z obrázku je patrné, že pro uživatele je vyhrazen prostor mezi osmou a devatenáctou hodinou. To znamená, že v DWH neprobíhají žádné ETL, které by vytížily databázi natolik, aby byli uživatelé omezeni, či dokonce zcela bez možnosti SQL dotazů do databází. Tento „čas DWH k dispozici“ je narušován v dvouhodinových intervalech nahráním dat pro potřeby Fraud kontroly a v pětiminutových intervalech pak WHSL provoz, který ale nevytíží DWH natolik, aby uživatelé pocítili nějaký nekomfort.

Od devatenácté hodiny do půlnoci probíhají zálohovací a optimalizační práce nad DWH.

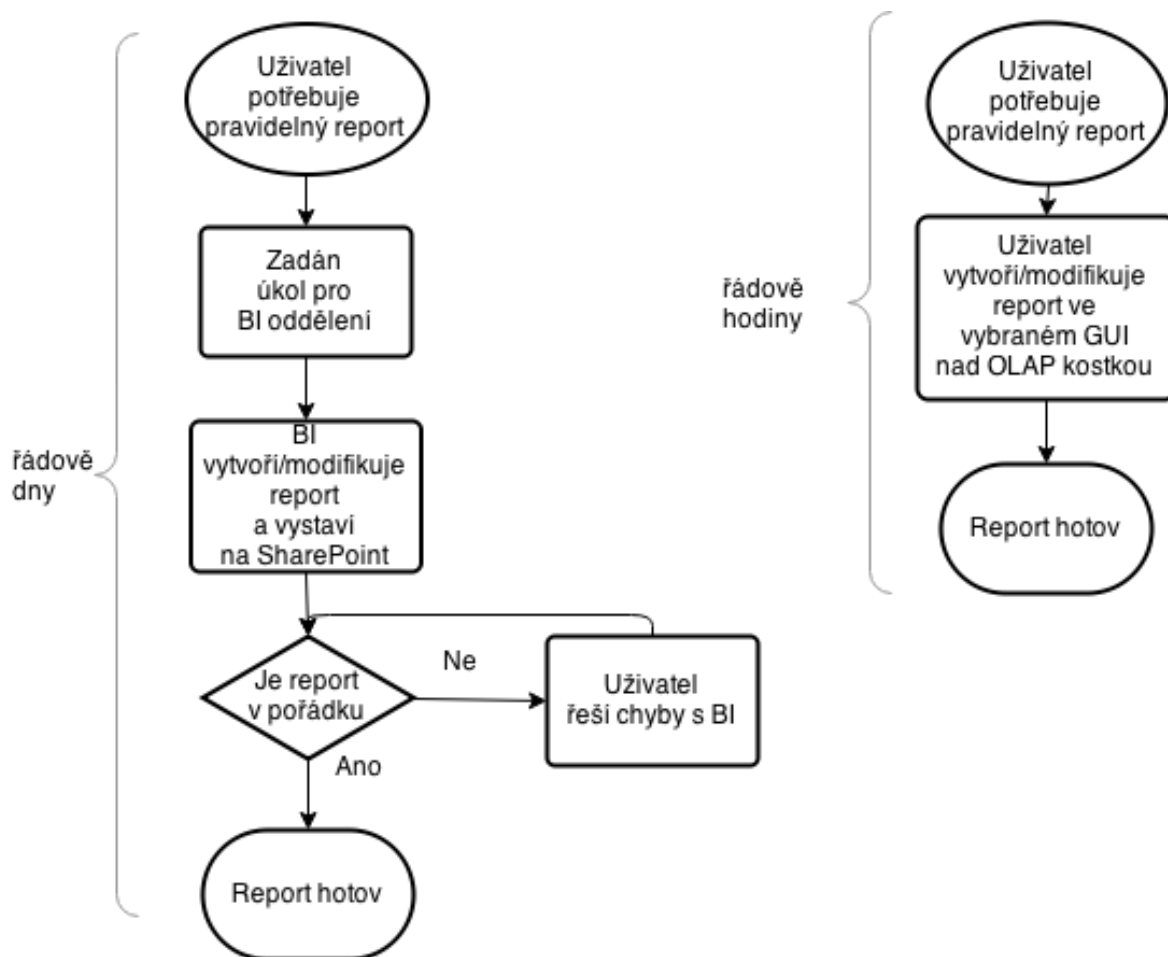
Mezi půlnocí a osmou hodinou ráno je prostor pro celý cyklus ETL. Většinou končí dříve, ale v případě problémů, kdy je třeba proces opakovat je prostor až do 8:00. Samozřejmě to neznamená, že by uživatel po sedmé večer nemohl pracovat a pouštět si dotazy a skripty, ale musí očekávat značné omezení v rychlosti vrácení výsledku, protože bude mít nejnižší prioritu.

#### **4.4.2 Analýza reportovacích nástrojů**

Z pohledu uživatele je výčet reportovacích nástrojů velmi jednoduchý. Jedná se o MS Excel 2010 (případně Excel 2013), MS Access 2003 nebo již zmiňované ProClarity Desktop Professional 6. Byť někteří uživatelé dokáží pracovat s nástrojem SQL Management studio, provedou pouze samotné dotazy nad databází a následně stejně výstup dotazu kopírují do MS Excel, kde dále s daty pracují a mění je v informace. Tento způsob, kdy pracují se známým nástrojem (MS Office) je i vlastně pro uživatele reportingu komfortnější.

Z pohledu správce IT (BI) pak je nutné pohled na reportovací nástroje rozšířit. V GTS Czech jsou využívány SQL Server Reporting Services (SSRS) integrované do MS SharePoint. Toto řešení umožňuje vytváření běžných reportů tak, že uživateli stačí pro zobrazení a filtraci pouze webový prohlížeč a libovolný OS. Takový report pak umožní základní filtraci pomocí předpřipravených filtrů, viz Obrázek 4.4. Tento typ reportu má ale jednu nevýhodu. Pokud chce uživatel upravit nějaký filtr nebo přidat atribut, musí požádat BI oddělení formou helpdesku o úpravu a následně (jak bylo zřejmé z ankety a pohovorů) většinou dny, než byla změna provedena. Tedy možnost výhody úprav samotným

uživatelé oproti čekání na podporu BI oddělení v časovém měřítku se snažím naznačit, viz Obrázek 4.3. Jedná se vlastně tedy o jakousi dynamičnost reportu.



**Obrázek 4.3 Porovnání přístupu k vytváření reportů**  
Zdroj: vlastní

Pro zobrazení reportů a dat pomocí webového prohlížeče se v GTS Czech používají tři přístupy- všechny vlastně souvisí s platformou SharePoint: SharePoint Performance Point Service (PPS), PowerView<sup>33</sup> a již zmiňované SSRS. Porovnání těchto tří nástrojů jsem analyzoval v bakalářské práci[8] s výstupem viz Tabulka 4.2. V této práci pak budu

<sup>33</sup> „Nástroj Power View je nástroj pro interaktivní zkoumání, vizualizaci a prezentaci dat, který rozšiřuje možnosti vytváření interaktivních ad-hoc sestav. Nástroj Power View je funkce Microsoft Excelu 2013 a Microsoft SharePoint Serveru 2010 a 2013 v rámci doplňku SQL Server 2012 Service Pack 1 Reporting Services pro Microsoft SharePoint Server Enterprise Edition“ [27]

provádět testování implementované OLAP kostky a práci nad ní v PPS nástroji. Z předchozí zkušenosti se mi jeví vhodný pro jeho nejbohatší na možnosti práce s OLAP a věřím, že i klíčoví uživatelé ocení tyto rozšířené možnosti.

Tabulka 4.2 Porovnání web GUI SharePoint BI nástrojů

	SharePoint Insights (PPS)	SSRS	Power View(RS)
<b>Dynamický</b>	Ano	Ne	Ano
<b>Náročnost editace</b>	Mírně(DashBoard Designer)	Ano(Report Builder)	Ne
<b>Export</b>	Ano	Ano	pouze PowerPoint
<b>GUI</b>	Ano	Ano	Ano
<b>KPI</b>	Ano	Ano	Ano
<b>Externí editace</b>	Ano	Ano	Ne
<b>Dashboard</b>	Nativně	Ano	Ano
<b>Kompatibilita</b>	RS ano, PW ne	Ne	Ne
<b>Interakce s MS Office</b>	Ne	Ne	Ano(PowerPoint)
<b>Drill</b>	Plně(up, down through)	Omezeně	Omezeně(down,up)
<b>Nutná znalost SQL</b>	Ne - Částečně	Částečně	Ne
<b>Myš - pravé tlačítko</b>	Ano	Ne	Ne
<b>Největší konkurenční výhoda</b>	Drill a rozkladný strom	Bohaté možnosti nastavení,možnost implementace do .NET aplikací	Vizualizace, prezentace, přenositelnost, filtrování v jednom náhledu

Zdroj: vlastní

Shrnu-li analýzu reportingu v retailové části GTS Czech, tak uživatelé nevyužívají žádné OLAP kostky, pracují výhradně se statickými reporty. Chybí metadata k reportům (uživatel odejde z firmy a změny v reportu jsou obtížné až nemožné, často je snazší vytvořit report nový) a to jak na úrovni SQL dotazu, tak i na úrovni datových slovníků a celého pochopení reportu jako celku. Výhodou je, že je ve firmě pouze jedno oddělení, které celý reporting zastřešuje, tedy uživatel vždy ví, na koho se má obracet v případě žádostí i problému. Ten jeho žádost přímo vyřeší (tedy není zde žádný první stupeň helpdesku, kromě případných klíčových uživatelů).

Address report

Actions 1 of 2 ? Find Next

COUNTRY	REGION	REGION EN	TOWN	TOWN EN	STREET	STREET EN	STREET N
CZ	Hlavní město Praha	Hlavní město Praha	Praha	Prague	Hvozďanská	Hvozďanská	2237/5
CZ	Hlavní město Praha	Hlavní město Praha	Praha	Prague	Ořetrovo náměstí	Ořetrovo náměstí	542/16
CZ	Hlavní město Praha	Hlavní město Praha	Praha	Prague	Na družce	Na družce	2049/11
CZ	Hlavní město Praha	Hlavní město Praha	Praha	Prague	Prášíá	Prášíá	1335/63
CZ	Hlavní město Praha	Hlavní město Praha	Praha	Prague	Vešerova	Vešerova	1090/1
CZ	Paroubický	Paroubický	Chrudim	Chrudim	Čs. partyzáni	Čs. partyzánu	26
CZ	Paroubický	Paroubický	Paroubice	Paroubice	Jiráskova	Jiráskova	20
CZ	Střeckeský	Střeckeský	Příbram	Příbram	V Lukách	V Lukách	296
CZ	Hlavní město Praha	Hlavní město Praha	Praha	Prague	Žitná	Žitná	1575/49
CZ	Hlavní město Praha	Hlavní město Praha	Praha	Prague	Žitná	Žitná	592/10
CZ	Hlavní město Praha	Hlavní město Praha	Praha	Prague	Mostecká	Mostecká	40/26
CZ	Hlavní město Praha	Hlavní město Praha	Praha	Prague	Černokostelecká	Černokostelecká	270/137
CZ	Hlavní město Praha	Hlavní město Praha	Praha	Prague	Školná	Školná	1267/30
CZ	Hlavní město Praha	Hlavní město Praha	Praha	Prague	V jáně	V jáně	167/17
CZ	Hlavní město Praha	Hlavní město Praha	Praha	Prague	V jáně	V jáně	637/10
CZ	Hlavní město Praha	Hlavní město Praha	Praha	Prague	Řemická	Řemická	661/19
CZ	Hlavní město Praha	Hlavní město Praha	Praha	Prague	Kolárova	Kolárova	2278/24
CZ	Hlavní město Praha	Hlavní město Praha	Praha	Prague	Komárkova	Komárkova	1193/17
CZ	Hlavní město Praha	Hlavní město Praha	Praha	Prague	Čmická	Čmická	780/61
CZ	Hlavní město Praha	Hlavní město Praha	Praha	Prague	Trojská	Trojská	1997/13a
CZ	Hlavní město Praha	Hlavní město Praha	Praha	Prague	Točlá	Točlá	1726/5
CZ	Jihomoravský	Jihomoravský	Vyskov	Vyskov	Vila Nejedlého	Vila Nejedlého	691
CZ	Jihomoravský	Jihomoravský	České Budějovice	České Budějovice	Škuberského	Škuberského	1478/14
CZ	Jihomoravský	Jihomoravský	Veselí nad Moravou	Veselí nad Moravou	Kolárova	Kolárova	1694
CZ	Vysočina	Vysočina	Třebíč	Třebíč	Kpt. Jaroše	Kpt. Jaroše	638/7

Parameters

Country: CZ

Building type: . Access to incumbent and OLO

Access technology: . Fibre, Wire, Radio, Non Fibre OI

Location category: Off-Net, 3rd party-connected, On-Net

Obrázek 4.4 Příklad uživatelského reportu v prostředí SharePoint (web)  
Zdroj: vlastní

### 4.4.3 Analýza technologické platformy

#### HW část

Pro DWH, databázovou část (SSRS) je vyhrazen server s následující konfigurací:

- 2x osmijádrový Intel Xeon
- 128GB RAM
- 8TB prostoru využito v 36TB vyhrazeném HDD poli
- propustnost serveru/pole je 12TB/hodinu

Pro DWH SSAS část je pak použit server:

- 2x šestijádrový Intel Xeon
- 48GB RAM
- 1,5TB prostoru využito

ETL pak zajišťuje server s konfigurací:

- Virtuální server na platformě Hyper-V<sup>34</sup>, 16x virtuálních jader
- 16GB RAM

LAN v GTS Czech je řešena pomocí 1Gbps propojů mezi servery a některými primárními systémy a 100Mbps full-duplex infrastrukturou pro připojení koncových uživatelů. V současnosti jsou již přes LAN zapojeny i stolní telefony (VoIP) a na přepínačích (switch, druhá vrstva OSI modelu) oddělen jejich provoz do VoIP ústředny výrobce Broadsoft.

#### SW část

SSRS jsou v současnosti spuštěny na MS SQL Server 2012 Enterprise edition. SSAS mají stejnou konfiguraci. ETL je řešeno na MS SQL Server 2012 Standard edition pomocí SSIS.

---

<sup>34</sup> Hyper-V je virtualizační serverový systém od Microsoft, tzn. na jednom skutečném HW běží jedna a více virtuálních instancí serverů s plnou funkcionalitou skutečného serveru. Společně sdílí prostředky (RAM, HDD, CPU)



Uživatelé pak využívají firemně instalovaný OS Microsoft Windows 7, případně libovolnou Linux distribuci (toto se týká pouze oddělení vývoje). Z webových prohlížečů je v rámci firmy doporučován Firefox a Chrome – pouze u těchto dvou garantuje IT oddělení kompatibilitu s místními aplikacemi.

## 4.5 Výstup z analýzy a návrhy řešení

V následující části jsem rozdělil návrhy řešení podle oblastí analýzy a odhadnul náklady na řešení. V poslední podkapitole jsem navrhnul řešení, které dále bude práce rozvíjet.

### 4.5.1 Návrh řešení pro DWH a ETL

V rámci DWH a ETL, jak bylo zjištěno analýzou v kapitole 4.4.1, není třeba žádných zásadních změn. Po diskuzi s BI oddělením a objasnění si některých historických vazeb<sup>35</sup> zde **doporučuji tyto kroky:**

- Revize metadat jednotlivých datových zdrojů a atributů – jak jsem pochopil, spousta databází zde v DSA dožívá v původní podobě, ale původní tvůrci ve firmě již nejsou a informace o datech jsou neúplné. Týká se to i datových slovníků.
- Jednoznačně oddělit DSA od datových marketů, raději přidat více marketů podle potřeb uživatelů než jedna halda dat v DSA. DSA pak ponechat k dispozici pouze BI oddělení a rozhraní k uživatelům budou pouze DMA.
- Pro ETL platí hlavně důkladná dokumentace
- Věřím, že v DWH bude uloženo hodně zbytečných procedur a triggerů, které jsou již neplatné, nepoužívané. Revize těchto procesů by měla určit vlastníka procedury, případně (pokud se nepodaří) ověřit podle její funkce, kterému oddělení patří a ověřit zda je potřeba ještě aktuální.

**Důvodem** navrhovaných kroků je zvýšení přehlednosti a srozumitelnosti, které umožní efektivní a časově méně náročnou správu DWH a ETL procesů. Také tím usnadní případná budoucí rozšíření DWH, komponenty se budou snadněji doplňovat.

### 4.5.2 Návrh řešení pro reportovací nástroje

Reporting nástroje vychází z analýzy v kapitole 4.4.2 a jsou dostatečně flexibilní, ale nevyužité. Uživatelé potvrzují v pohovorech, že možnosti reportovacích nástrojů v GTS znají pouze okrajově. **Doporučuji tyto kroky:**

---

<sup>35</sup> Případy, kdy existují běžící procesy, o kterých neexistují informace, proč jsou právě tak nastavené, kdo je tak nastavil a pro koho slouží. A hlavně co se stane, pokud se odstraní.

- proškolit klíčové uživatele na možnosti nástrojů BI (SharePoint, MS Office, ProClarity)
- připravit wikistránky s Q&A částí
- prezentační a výukové materiály (Powerpoint, SharePoint)
- demo prostředí, kde si budou moci možnosti vyzkoušet. Vlastně by to byla jakási cvičná kostka<sup>36</sup>

**Důvodem** výše uvedených kroků (návrhů) je požadavek uživatelů na rozšíření jejich schopností (znalostí) práce se stávajícími BI nástroji. Tím umožnit lepší práci s informacemi a důsledkem bude přesnější business rozhodování.

#### 4.5.3 Návrh řešení pro technologickou platformu

Technologická platforma se mi jeví poddimenzovaná podle zjištěných technických parametrů, uvedených v kapitole 4.4.3. Z vlastní zkušenosti uživatele vím, že některé dotazy ve „špičce“ trvají skutečně dlouho (dotaz, který mimo špičku trvá 1-2 min trvá 10-30min) a **doporučuji** posílit SQL server pro DWH, přinejmenším alokovat více procesorového času a paměti pro uživatele. Z pohledu LAN pak provést revizi a ujistit se, že úzké hrdlo po cestě mezi uživatelem a databází nejsou síťové prvky (tzn. monitorovat vytižení jednotlivých switchů v LAN, případně posílit, spoje mezi switchi mít 1Gbps síť).

**Důvodem** posílení technologické platformy je požadavek na větší uživatelský komfort (rychlejší získání informací ze systémů i kratší odezvy při práci se systémy).

---

<sup>36</sup> Například Microsoft na svém SQL Serveru pro analytické potřeby má cvičnou databázi Adventureworks, kterou si zdarma uživatelé mohou instalovat a zkusit jak relační tak analytické (OLAP) modelování.

#### 4.5.4 Náklady pro DWH, ETL, technologie a reportovací nástroje

V Tabulka 4.3 jsem srovnal odhadované náklady na jednotlivá optimalizační doporučení.

Tabulka 4.3 Náklady na optimalizaci BI

	HR náklad jednorázový v MD	CAPEX (investiční náklad jednorázový) V tisících Kč	OPEX (Provozní náklad měsíční nový) v MD
Revize metadat	80-120	-	-
Oddělení DSA a DMA	40-60	Nákup HDD – 10-100	
ETL dokumentace			0,2
Revize procedur a triggerů	Max 10 MD	-	-
Školení uživatelů	2	-	0,5
Interní Wikistránky	2	-	0,1
Prezentační a výukové materiály	2	-	-
Vytvoření dema	1-2	-	-
Upgrade DWH	1	50-200	0,1
Revize/monitoring switchů	2		0,5

#### 4.5.5 Souhrnný návrh řešení

Syntézou jednotlivých analýz (vyhodnocení analýzy ankety, pohovorů v kapitole 4.3.3, analýzou aktuálního stavu ETL, DWH a technologické platformy) jsem dospěl k závěru, že v rámci BI GTS Czech jsou dvě oblasti, ve kterých lze situaci pro uživatele zlepšit. Jedná se o:

1. častější aktualizace vybraných reportů (optional) – **důvodem** je to, že uživatelé potřebují mít vždy čerstvé informace pro svoje rozhodování a návazné manažerské kroky. Informace zastarává již po několika hodinách od jejího získání.
2. vytvoření BI s potřebnými komponentami k analýze nad hlasovým provozem retailových služeb. (critical) – první **důvod** je obrovské množství dat (miliony řádků v tabulce), které běžný reportovací nástroj<sup>37</sup> nedokáže rychle zpracovat a filtrovat. Druhý pak je ten, že BI umožní pracovat s daty nad rámec reportu (modelovat dimenze, porovnávat historii nebo na trendy atd.). Souhrnně tedy uživatel požadoval vytvoření komplexního a rychlého nástroje pro podporu rozhodování v oblasti retailového hlasového provozu.

Aktualizace dat vybraných stávajících reportů bude stanovena na 2x denně (jedná se cca o 10 různých reportů), u OLAP hlasového provozu bude update také 2x denně.

---

<sup>37</sup> Např. MS Excel

Tabulka 4.4 Uživatelské požadavky na změny BI

Častější aktualizace dat		Náklady	Nový OLAP Hlasového provozu	Náklady
<b>ETL balíčků</b>	úpravy Pouze změna update	1 MD	Nový ETL proces	2MD
<b>SharePoint úpravy</b>	Beze změn	-	Vytvoření PPS prostředí	1 MD
<b>Server HW</b>	Navýšení HDD + alokace prostředků pro ETL	10 000 – 40 000Kč	Navýšení HDD – alokace prostředků pro ETL + prostředky na SSAS (celý server)	50 000 – 200 000 Kč
<b>Server SW</b>	Beze změny	-	Instalace nového SSAS, nový DMA	2 MD
<b>Implementace řešení</b>		-	5 MD analýza + 3MD implementace + 3 MD testování a nasazení + 1 MD školení uživatelů	12 MD
<b>Uživatel</b>	Beze změn	-	Nutné školení	1MD
<b>Dokumentace</b>	Beze změn	-	Ano, zcela nová /	2 MD
<b>BI model</b>	-	-	Přírůstkový (hybridní)	-

V obou případech (tabulky Tabulka 4.3 Tabulka 4.4) opomím komplikace ze strany managementu, možné zamítnutí navrhovaných kroků nebo omezení na uvažovaných nákladech. Návrhem řešení je tedy **splnit** oba **požadavky** uživatelů.

V případě častějších aktualizací je řešení poměrně snadné. Nevýhodou určitě budou dodatečné náklady na prostor databáze, zvláště pokud by se data historizovala. Pokud se historizovat nebudou, dodatečný náklad bude podstatně menší. Takže vlastně bude stačit upravit ETL proces (package) a aktualizace provádět častěji. Bohužel vše se všim souvisí, tedy bude nutné revidovat ETL těch primárních zdrojů, které do zmíněných deseti reportů dodávají data. Proto jeden celý MD (Manday) na pracovníka BI oddělení (IT). Neuvažuji zde žádné nové provozní náklady.

Pro vytvoření celého nového OLAP řešení s ETL je již potřeba nových nákladů více. Jedná se o kompletní proces BI implementace se všemi komponentami. Na druhou stranu se jedná stále pouze o přírůstkový model (viz Obrázek 3.5), a tedy již větší část platformy existuje, pouze se přidává do jednotlivých komponent dodatečná část[6]. Výjimku ale tvoří neexistující OLAP část (SSAS), zde tedy očekávám největší investici finanční i časovou. Neuvažuji zde také žádné nové provozní náklady. Samotné implementaci BI řešení se věnuji v následující kapitole.

## 4.6 Design BI řešení

V této praktické ukázce se nastíním řešení vytvoření OLAP kostky pro hlasový provoz nad reálnými daty a DWH v GTS Czech. Kolegy z GTS Czech BI oddělení jsem byl požádán, abych tento model skutečně připravil s tím, že jej následně v praxi uplatní, ať už jako demo prostředí pro výuku a předvedení možností BI nástrojů (novému managementu a interním zákazníkům), tak i pro reálné analýzy produktových manažerů v oblasti hlasu.

### 4.6.1 Analýza dostupných datových zdrojů v relačních databázích

Abych mohl připravit podklady z relačních databází pro vytvoření OLAP, je třeba nejprve dostupná data pochopit. Cílem analýzy je najít závislosti mezi tabulkami, pochopit atributy v tabulkách a strukturu dat. Díky tomu, že již tuto znalost mám díky působení ve firmě, analýza byla jednodušší. Postup by ale byl stejný, jen více rozsáhlý. Postup, kterou jsem zvolil při analýze, byl tento:

- Najdi si libovolnou skutečnou fakturaci zákazníkovi za konkrétní období a vybranou službu – použij primární systém Durian.
- Rozeber jednotlivé položky na faktuře a označ ID faktury a ID ceníkových položek
- Začni hledat ID faktury a položek v jednotlivých tabulkách v DB schématu pro fakturaci (DUR).
- ID faktury je jedinečné, jakmile nalezeš správnou tabulku, začni hledat závislosti podle názvů atributů v dalších tabulkách. Průběžně kontroluj s údaji v primárním zdroji (částky musí být stejné, počty komponent také i data vzniku i ukončení platností komponent).
- V případě nepochopení vazby mezi komponentami na faktuře se obrať na klíčové uživatele primárního systému Durian (fakturačního systému).

Je zřejmé, že se jedná o metodu shora-dolů (loupání cibule). V případě vytváření zcela nového BI tato analýza trvá několik týdnů až měsíců, zvláště pokud analytik primární systémy sám nezná (externista) a spolupracuje s klíčovými uživateli.



Tato metoda narazila na problém ve chvíli, kdy jsem hledal rozpad hovorného u služeb s hlasovým paušálem, tzv. flat tarifem<sup>38</sup>. V těchto případech nejsou jednotlivé směry hovorů (národní fix, mobilní, mezinárodní) rozpadlé, ale vše je na faktuře pod jednou částkou. Po konzultaci s BI oddělením jsem zjistil, že oceňování všech hovorů probíhá ve stejném databázovém schématu DUR, ale do fakturace pak již vstupují pouze agregovaně. V odkazovaných tabulkách jsem našel všechny uskutečněné hovory (takzvané CDR), byť byly oceněny za 0 Kč () a to jak příchozí, tak i odchozí.

V rámci analýzy dostupných zdrojů jsem zjistil, že z primárních systémů již žádné dodatečné informace nepotřebuji získat - potřebná data v dostatečném rozsahu jsou k dispozici v DWH a získávat je mohu přímo.

#### 4.6.2 Dimenze

Požadované dimenze a jejich části (hierarchie) pro OLAP kostku jsou následující:

##### Datum

- den v týdnu (číslo 1-7, jméno dne, je pracovní den/je víkend, je svátek)
- den v měsíci (1-31)
- měsíc (číslo 1-12, jméno měsíce)
- kvartál (1-4)
- rok (2014- xxx)
- kombinované formáty (ROK + Měsíc např. 2014-10)

##### Čas

- Hodina (číslo 0-23)

##### Segment

- rozdělení zákazníků do segmentů podle výše jejich celkových potenciálních tržeb na telco trhu.(HOME, SOHO, SME, LE, ME)

---

<sup>38</sup> Pominu zcela služby, které nejsou fakturované, ale existují. Například dokud zákazník nezačne na lince volat, nic neplatí.

## NACE

- klasifikace ekonomických činností (např. zemědělství, stavebnictví atd.)

## Produkt

- rozdělení GTS hlasových a kombinovaných služeb podle typu služby (pod stejným jménem produktu jsou prodávány zákazníkům)

## Obchodník

- hierarchicky řazené jméno obchodníka a jeho týmu

## Destinace hovoru<sup>39</sup>

- Hierarchicky členěné hovory podle směrů (národní, mobilní, mezinárodní), podle destinací (např. mobilní síť Albánie) a podle samotné ceníkové položky na faktuře.

### 4.6.3 Fakta

Hodnotící ukazatele, které jsou uživateli požadovány jako fakta, jsou následující:

- Délka hovoru skutečná – délka hovoru ve vteřinách, tak jak je zpracována ústřednou od spojení uživatelů po ukončení (zavěšení) hovoru.
- Délka hovoru účtovaná -(úctovaná podle nastaveného tarifu, tj. např. první minuta vždy celá a následně po vteřinách - > tarifkace 60+1).
- Počet hovorů – počet jednotlivých uskutečněných hovorů ve vybrané agregaci.
- Cena za hovory – vyúčtování hovoru podle nastaveného tarifu, objemové slevy a slevy do vybraného směru.
- Cena za paušál – rozpočtené měsíční poplatky za službu na nejvyšší možný detail.
- Náklady na hovory – (to je ale optional požadavek). Opět v nejnižší možné granularitě.

---

<sup>39</sup> V případě příchozích volání se jedná o originaci hovoru

#### 4.6.4 KPI

Příkladem KPI (klíčového ukazatele) je překročení délky všech hovorů v určitém období. KPI se definují třemi prvky: Cílem (Target), Stavem (Status) a Trendem (Trend). Dále pak sledovanou hodnotou (dále v příkladu jako **SH**).

Cíl je úroveň (hodnota), kterou chceme hlídat. Takže například na jednu službu denně bude cílem limit 1000 minut. Budeme hlídat služby s nezvykle velkým provozem. Zde je možné riziko hlasových podvodů.

Stav pak porovnává aktuální situaci u služby z faktové tabulky s požadovaným cílem. Kolik doposud v daném dni provolala služba (zákazník) minut. Stav používá vizualizačního prvku, velmi často dvou (zelená/červená) nebo třístavového (zelená/oranžová/červená) semaforu nebo tachometru.

Trend je podobný Stavů. Porovnává určitá časová období a sleduje vývoj, který opět graficky vyhodnotí. Typické jsou šipky (nahoru dolů rovně). V tomto případě porovnávám délky hovorů aktuálního dne SH a den předchozího SH<sub>-1</sub>

Prakticky pak bude KPI vypadat:

**Cíl:**

*1000 minut*

**Stav:**

*Když  $\frac{SH}{Cíl} \geq 1$  pak zobraz zelenou*

*Když  $1 > \frac{SH}{Cíl} \geq 0,8$  pak zobraz oranžovou*

*Když  $\frac{SH}{Cíl} < 0,8$  pak zobraz červenou*




**Trend:**

*Když  $SH - SH_{-1} > 1$  pak zobraz šipku nahoru*

*Když  $SH - SH_{-1} < 1$  pak zobraz šipku dolů*

*Když  $SH - SH_{-1} = 1$  pak zobraz šipku rovně*

Skutečné KPI je pak vidět na následujícím obrázku

Display Structure	Value	Goal	Status	Trend
 KPI_delka_hovoru	1733	10000		

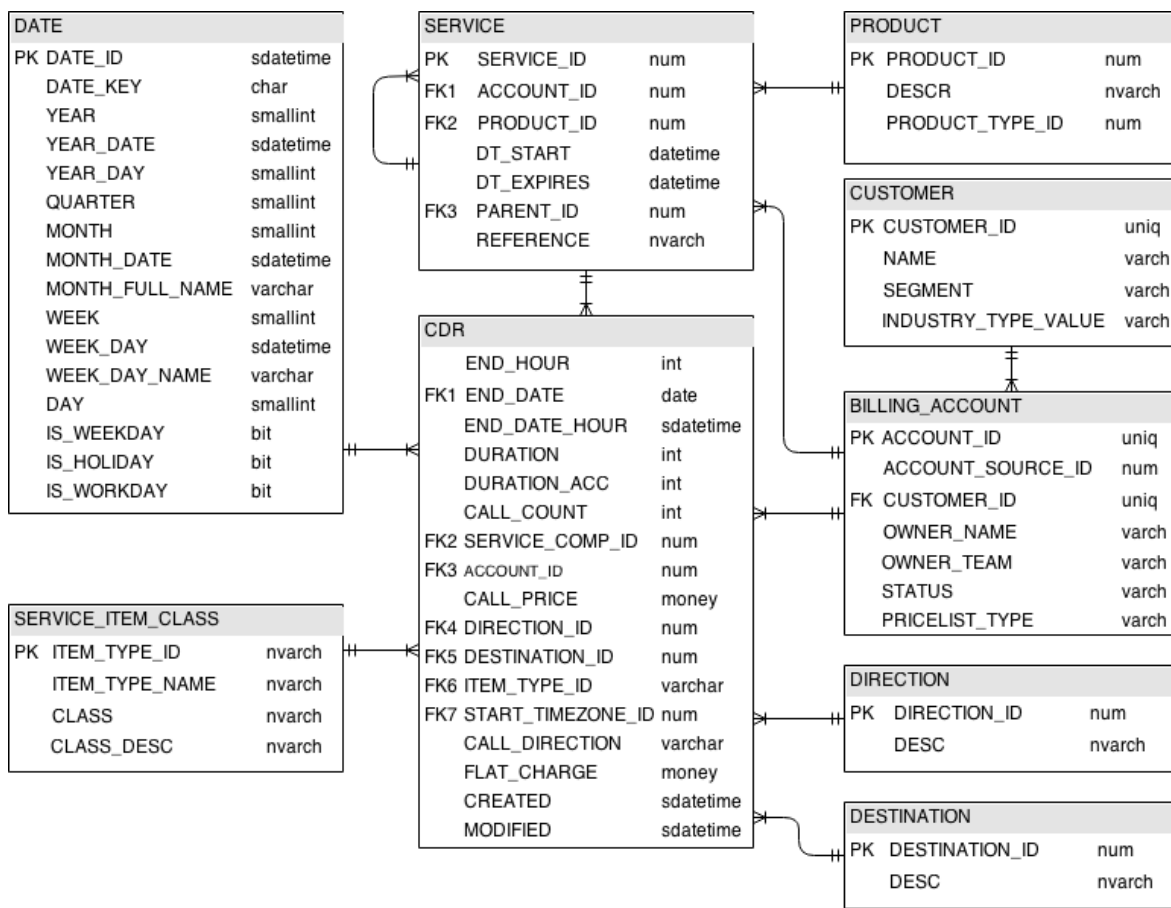
Obrázek 4.5 KPI  
Zdroj: vlastní

#### 4.6.5 Příprava podkladové relační databáze

Pro přehlednost jsem si vytvořil vlastní samostatnou databázi v DWH a do ní vytvořil jednotlivé tabulky, potřebné pro tvorbu dimenzí a fakt. Tabulky jsem se již při tvorbě snažil přiblížit budoucímu schématu kostky. Toto má výhodu ve snadnější orientaci při budoucích změnách (např. přidání atributů nebo tabulek) a dokumentaci. Vazby mezi tabulkami jsou řešeny 1:N a zachyceny pomocí primárních/cizích S\_klíčů (PK/FK).

Na počátku vytváření relačních tabulek jsem pracoval v běžné relační struktuře s vysokou normalizací a vytvořil následující schéma, viz Obrázek 4.6 Relační diagram, jenž pak ukazuje jednotlivé vazby a atributy ve vytvořených tabulkách včetně jejich datových typů. Vzhledem k tomu, že pracuji s MS SQL Server, používám datové typy platné pro tuto variantu SŘBD<sup>40</sup>. Tento model má tabulky dimenzí zcela bez historizace.

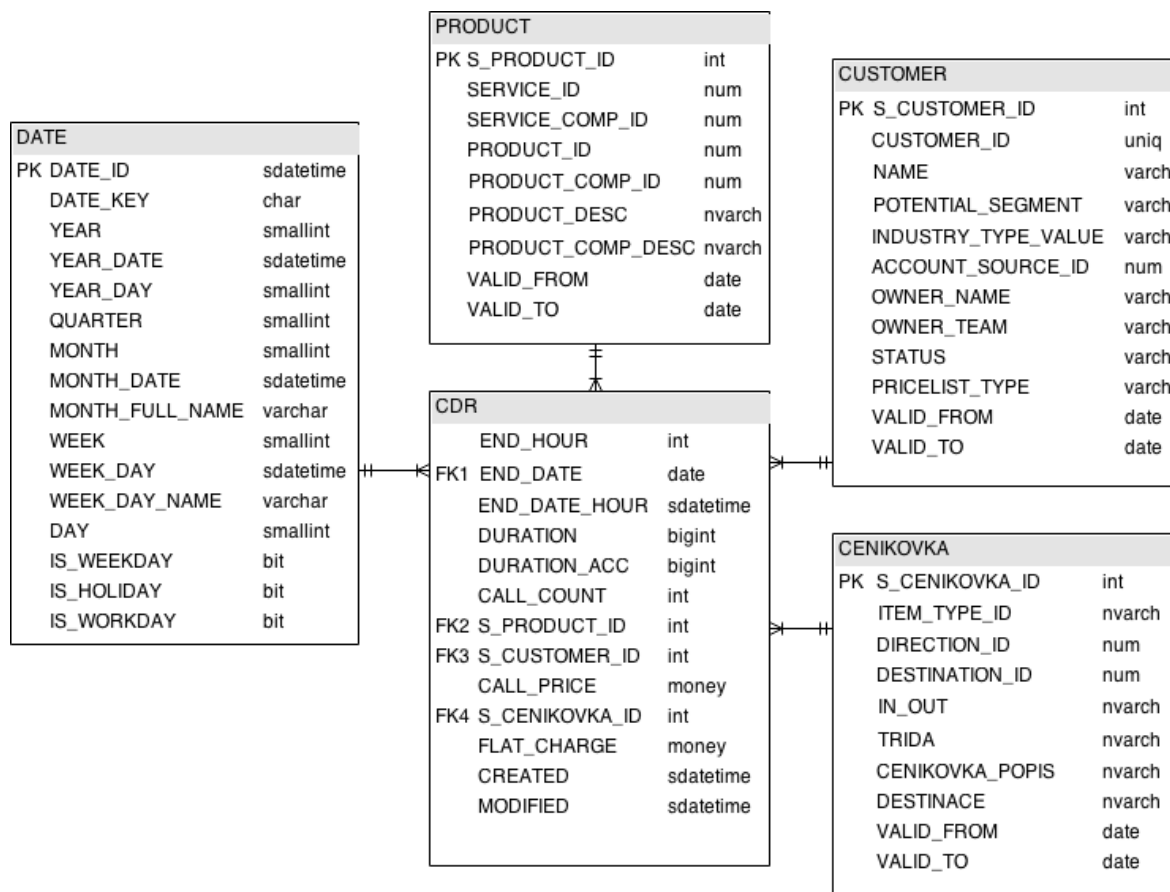
<sup>40</sup> SŘBD MySQL, ORACLE nebo PostgreSQL mohou mít například jiné číselné datové typy.



**Obrázek 4.6** Relační diagram vysoká normalizace  
Zdroj: vlastní

Nicméně dimenzionální struktura tabulek již nepožaduje tak vysokou normalizaci a naopak požaduje minimálně základní udržování historie + STAR strukturu, viz Obrázek 3.12. Tabulky dimenzí budou mít vlastní nový unikátní klíč<sup>41</sup>, který je bude spojovat s tabulkou faktů a validační data pro platnost záznamů (Atributy VALID\_FROM a VALID\_TO). Toto se řeší z důvodu ošetření budoucích změn v tabulkách dimenzí tak, aby se původní hodnoty novými záznamy nepřepisovaly. Pokud bych toto neošetřil, tak například pokud by zmizela nějaká země ze zdrojové tabulky v primárním systému, automaticky by zmizela i v mé dimenzi a všechny doposud proběhlé hovory do dané destinace by se mi vymazaly. Proto jsem vytvořil vhodnější strukturu, viz následující obrázek.

<sup>41</sup> Abych oddělil ostatní použité klíče od mnou vytvořeného primárního dimenzionálního klíče, tento budu nazývat S klíč (Source) a v názvech atributů bude vždy začínat písmenem S.



Obrázek 4.7 Relační diagram nízká normalizace

Zdroj: vlastní

Vytvoření tabulek a první vložení dat do tabulek jsem provedl ručně (bez procedury ETL) a jednorázově. Příklady SQL skriptů jsou k dispozici v příloze. Bylo třeba dodržet následující postup vytvoření tabulek, protože Primární klíče vznikaly až s plněním dimenzionálních tabulek:

1. Vytvořit tabulky dimenzí
2. Vytvořit tabulky fakt
3. Naplnit tabulky dimenzí (vytváří se inkrementálně S\_PK)
4. Naplnit tabulky fakt (jako S\_FK zde použít vytvořené S\_PK)

V následující tabulce jsem sumarizoval základní popisy pro aktualizace vytvořených a jednorázově naplněných tabulek. Všechny tabulky, kromě tabulky CDR jsou kopiemi existujících tabulek, CDR pak je již vychází z vytvořených dimenzí a je doplněna o určité modifikace ohledně identifikace času a agregace počtu hovorů, délky hovorů a ceny.

Tabulky, kde je uvedena aktualizace rozdílová budou vždy jednou denně porovnány se zdrojem a rozdíly ošetřeny zneplatněním původního záznamu (atribut VALID\_TO se vyplní aktuálním dnem) v případě změny například názvu a přidáním záznamu nového (VALID\_TO bude null). Nové záznamy, u kterých se nenajde shoda s existujícím, se přírůstek pouze vloží. V případě smazání celého záznamu ve zdrojové tabulce opět v tabulce dimenze záznam zneplatním, ale nikdy zcela nemažu. Tím zajistím částečnou historizaci. Tabulka CDR pak bude řešena přírůstkem. V rámci četnosti update bude pouze tabulka CDR plněna 2x v různých intervalech. Základní agregovaná hodinová data o počtu hovorů,

jejich délce a jejich ceně budou přibývat 2x denně. Jak ale vyřešit přírůstek u faktové tabulky hovorů a navázat přesně na již vložené záznamy? Jak nezpůsobit duplicity v tabulce a zároveň nějaké záznamy nevynechat? Pro faktovou tabulku CDR se mi jeví jako bezkolizní řešení provést aktualizaci smazáním záznamů celého aktuálního dne a dne předchozího a jejich opětovným vložením.

Jednou za měsíc pak budou do již stávajících řádků dodatečně doplněny částky za flat paušál z vystavené faktury za uplynulý měsíc, Tyto částky budou rozpočítané do hodinových agregací pro dané služby podle délky hovorů.

Př.

Zákazník platí 1000 Kč měsíčně jako flat paušál a provolal v říjnu 2014 celkem 250 minut (do směrů národní, INT\_0 a mobilní). V CDR tabulce najdu říjnové volání pro danou službu a v jednotlivých hodinových rádcích, kde je cena 0 (protože má flat tarif, hovor existuje, ale s nulovou cenou, nikoliv délkou!) pak doplním součin 4\*počet volaných minut do připraveného pole FLAT\_CHARGE a atribut aktualizuji. Takto získám přesnější přehled o skutečné zaplacené ceně za hovor.

V jednotlivých tabulkách se již dají odhadnout očekávané měsíční přírůstky dat a podle toho tedy vhodně rozvrhnout databázi. V mém případě se jedná přibližně o 270 MB absolutní přírůstek měsíčně při stávajícím obsahu databáze. Za pět let to bude 16,2 GB dat.

**Tabulka 4.5 Metadata vytvořených tabulek**

Název	Zdroj v DWH	Aktualizace	Typ aktualizace	Historizace	Četnost update	Velikost Aktuální (MB)	Delta MB /měsíc
DATE	Kopie	Ne	-	Ne	-	3	-
CUSTOMER	Kopie	Ano	Rozdíl + přírůstek	Částečná	Denní	1,5	<1
PRODUCT	Kopie	Ano	Rozdíl + přírůstek	Částečná	Denní	2,1	<1
CDR	Dotaz nad DWH + dimenze	Ano	Přírůstek + Update	Ne	2x denně + měsíční	1240	265
CENIKOVKA	Kopie	Ano	Rozdíl + přírůstek	Částečná	Denní	<1	< 1



## 4.7 Implementace do SSAS

Pokud je dobře připravený podklad dimenzí a fakt v relační databázi, samotná základní implementace na analytický server (SSAS) je poměrně jednoduchá. Skládá se ze tří částí. Namodelování (vizualizace kostky), nasazení do SSAS databáze (Deploy) a provedení (Process)[28].

Namodelováním se rozumí vizuální příprava těla kostky. V mém případě jsem pracoval s nástrojem MS Visual studio 2013, které má pro implementaci OLAP na SSAS připraven vlastní projekt. Při modelování jsou potřebné následující kroky:

- **Výběr datového zdroje** – jedná se o výběr relační databáze (případně jiné SSAS databáze), ze které budu čerpat data pro svoji kostku.
- **Vytvoření view** – specifikování relačních tabulek, které použiji pro vytvoření dimenzí a fakt. View může existovat více, ale kostka může být vytvořena právě z jednoho view. Nicméně ve view lze vytvořit virtuální tabulky (v mém případě jsem použil *Select distinct* nad faktovou tabulkou CDR a vytvořil si virtuální tabulku HOUR pouze s 24 hodnotami 0-23 )
- **Namapování dimenzí a fakt** – Z vytvořeného view definuji fakta a dimenze. Zde mohu dimenze přejmenovat, změnit datové typy.
- **Výběr atributů dimenzí** – U jednotlivých dimenzí si vybírám, které atributy bude uživatelů v kostce moci použít, které budou viditelné.
- **Vytvoření hierarchií** v dimenzích – typicky pro dimenzi času (Rok -Měsíc-Den), kterou jsem použil i v mé kostce nebo pro Lokaci (Země – Kraj - Město ) atd.
- Kontrola nebo **vytvoření vazem** mezi fakty a dimenzemi – vytvořím logické vazby logických vazeb a případně porovnávám s relačním modelem.
- **Vytvoření počítaných polí** (Measures) – umožňuje vytvořit vlastní atribut, který je dopočítaný z jiných atributů. Například v případě chvil i vytvořit počítané pole Marže (v %) použiji pro výpočet faktové atributy  $(Cena - Náklady) / Cena$ . Toto pole si potom může uživatel zobrazit jako další faktové pole. Vzorce jsou tvořeny jazykem MDX, který se používá pro práci s OLAP.
- **Vytvoření KPI** – popsáno v podkapitole 4.6.4

- **Použití dalších BI počítaných polí** – Visual Studio 2013 nabízí již předpřipravené počítané pole. V mém případě jsem využil prvky tzv. Time Intelligence<sup>42</sup>. Ty mi snadno umožnily přidat například následující výpočty, pracující s fakty a časem.
  - YTD (Year to date) – sečte všechny fakta za kalendářní rok do vybraného dne/měsíce. (například vyberu 15. Listopad 2014 a YTD výpočet mi vrátí součet faktových hodnot od 1. 1. 2014 do 15. 11. 2014)[29].
  - MTD (Month to date) – stejně jako YTD, pouze období je platné pouze pro aktuální měsíc.
  - Day over Day Growth – porovnání mezi vybraným a předchozím dnem buď v relativní (%) nebo absolutní výši
  - Month over Month Growth – porovná aktuální a předešlý měsíc s absolutním nebo relativním výsledkem

Nasazení (Deploy) vlastně znamená, že vytvořenému modelu postavíme odpovídající databázi v SSAS v patřičné struktuře a s potřebnými metadaty. Zároveň v této fázi nasazení se automaticky kontrolují vytvořené vazby mezi dimenzemi, přístupová práva, hierarchie dimenzí. Zde také určují, jak budou data uložena v SSAS databázi. Volím metodu MOLAP, viz kapitola 3.8.4.

Pokud v budoucnu změníme nějaké atributy kostky, např. přidáme dimenzi nebo fakta, je nutné znovu provést proces nasazení tak, aby se veškeré změny validovaly a převedly do analytické databáze. Kostku jsem tedy v tomto kroku nasadil. Visual tools projekt upozorní na nedostatky (například právě nevhodně vytvořené hierarchie) nebo dokonce vůbec nedovolí model vytvořit pro fatální nedostatky.

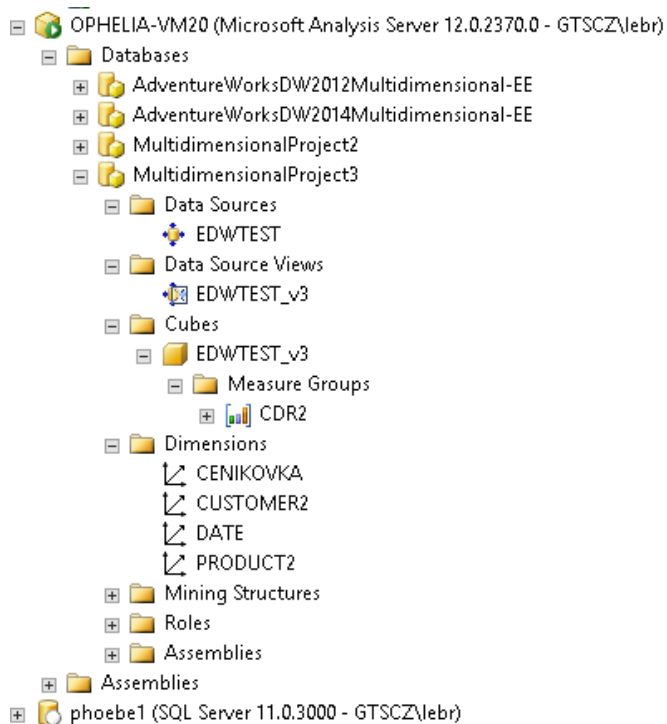
Třetím krokem je Provedení (Process) což už v podstatě je pouze naplnění vytvořené databáze daty a update minoritních změn v modelu. Pokud je tedy třeba aktualizovat data, stačí spustit právě Provedení (ať už přímo pomocí MS Management Studio v SSAS nebo v MS Visual Studio) a data jsou aktualizována. Kostka je během aktualizace uležena v paměti (cache), takže nadále pro uživatele funguje, dokud aktualizace nedoběhne. Teprve po dokončení aktualizace se aplikují změny a jsou zobrazena nová data a cache je

---

<sup>42</sup> Dalšími typy BI Measures pak jsou např. Account a Dimension Intelligence.

uvolněna. Provedení je možné i navrhnout do paralelních úkolů v případě více-jádrových procesorů. V mém případě jsem nechal proces proběhnout již bez změn.

Následně jsem zkontroloval vytvoření databáze v SSAS pomocí MS Management Studio



**Obrázek 4.8** Nasazená databáze v SSAS v prostředí MS Management Studio

Zdroj: vlastní

## 4.8 Návrh ETL

Pro ETL řešení této implementace mohu použít dvě řešení a to buď přímo v SSIS nebo pomocí uložených procedur v SQL serveru.

### 4.8.1 SSIS řešení

Pomocí SSIS namodelovat celý proces ve Visual Studiu. Jednotlivé tzv. krabičky v ETL si pak po spuštění budou připojovat k zdrojovým tabulkám a provádět definované transakce a následné nahrání do cílových tabulek. Velkou výhodou tohoto řešení je, že používá tzv. datový stream a nezatěžuje tolik transaction log (paměť) databázového serveru. Podstatou je kontinuální (proudové) provádění celého procesu ve velmi malých dávkách. Je tedy vhodné pro práci nad velkými objemy dat, kdy vím, že tabulky budou v řádech stovek tisíc řádků. Přípravené ETL řešení se pak spouští pomocí SQL Agentů, ve kterém se vytvoří tzv. Job, kde si nastavím v jaké časy se má vybrané ETL řešení spouštět. Celý Job proces lze ještě rozfázovat do jednotlivých kroků, např. pokud spouštíte více ETL řešení nebo kombinaci ETL a nějakých procedur.

The image shows a screenshot of the 'New Job Schedule' dialog box in SQL Server Enterprise Manager. The dialog is titled 'New Job Schedule' and contains the following fields:

- Name: spusteni ETL
- Schedule type: Recurring
- Enabled:
- One-time occurrence:
  - Date: 2/ 2/2015
  - Time: 10:56:10 AM
- Frequency:
  - Occurs: Daily
  - Recurs every: 1 day(s)
- Daily frequency:
  - Occurs once at: 12:00:00 AM
  - Occurs every: 12 hour(s) (selected)
  - Starting at: 12:00:00 AM
  - Ending at: 11:59:59 PM
- Duration:
  - Start date: 2/ 2/2015
  - End date: 2/ 2/2015
  - No end date:
- Summary:
  - Description: Occurs every day every 12 hour(s) between 12:00:00 AM and 11:59:59 PM. Schedule will be used starting on 2/2/2015.

Buttons: OK, Cancel, Help

**Obrázek 4.9 SQL Agent Job**  
Zdroj: vlastní

V jednotlivých Jobech lze nastavit i chování při neprovedení nějakého kroku, např. zaslání emailu.

Výhodou tohoto řešení je grafické zobrazení celého procesu a možnost tvorby řešení bez znalosti SQL.

#### **4.8.2 Řešení s uloženými procedurami**

Všechny potřebné úkony budou přidány do samostatných uložených procedur na SQL Serveru, kde jsou umístěny cílové relační tabulky<sup>43</sup>. Jednotlivé procedury pak budou spouštět samotné SQL dotazy, které extrahují potřebná data z potřebných zdrojů, transformují do potřebné podoby a vloží do cílových tabulek. V rámci procedur pak lze opět nastavit chování při chybě.

Uložené procedury využívají zcela transaction log a proto je vhodnější je využít pro menší tabulky, menší objemy dat.

Výhodou tohoto řešení je obrovská svoboda v možnosti SQL zápisu procedury. Nevýhodou pak nutná znalost MS SQL syntaxe a jazyka.

V případě tvorby ETL pro tuto práci jsem využil tento způsob.

#### **4.8.3 Popis ETL procesu**

Samotný ETL proces lze z pohledu času rozdělit na tři, na sobě nezávislé, části, viz 4.6.5.

V procesu jsem ošetřil chyby tak, že systém vygeneruje email s popisem chyby (resp. s názvem neprovedeného kroku) a zašle na určené adresy. Proces pak dál nepokračuje a bude třeba ruční zásah.

Pokud proběhnou všechny procedury (kroky) v pořádku, jako poslední krok je Provedení datové kostky, která aktualizuje svá data v SSAS databázi.

Konkrétní časy spouštění lze měnit podle požadavků uživatelů, tak aby měli informace dříve, např. denní cyklus zkrátit na půldenní nebo update faktové tabulky provádět po hodině. V základním nastavení se ale budou ETL spouštět:

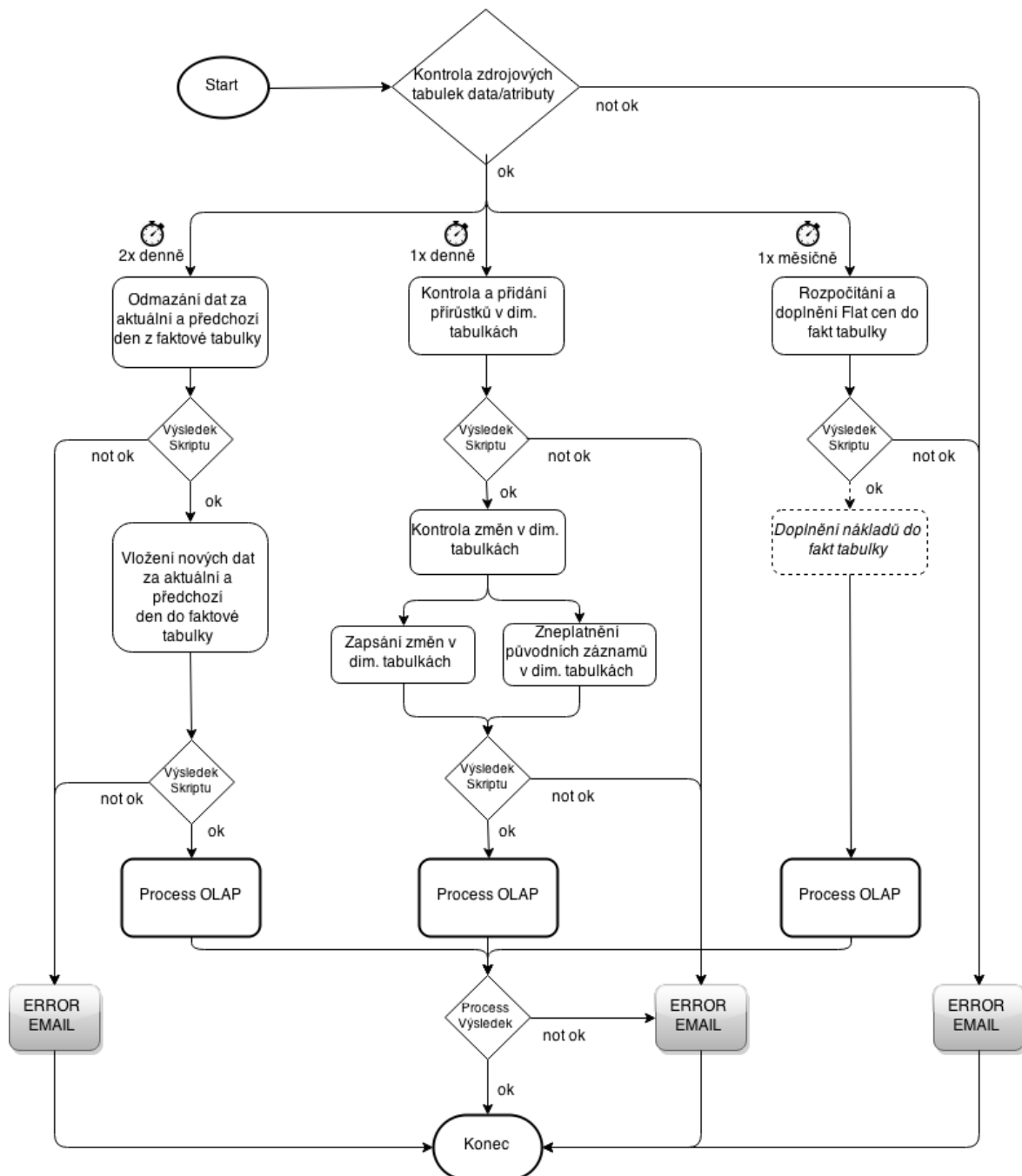
- ETL faktové tabulky 2x denně ve 00:00, resp. 12:00
- ETL dimenzí denní 03:00

---

<sup>43</sup> Zcela vhodné řešení je ale umístit uložené procedury na SQL Server určený pouze pro ně.

- ETL faktové tabulky měsíční vždy 15. den<sup>44</sup> v měsíci v 02:00

Následující obrázek charakterizuje všechny tři ETL procesy.



Obrázek 4.10 ETL proces jako vývojový diagram

Zdroj: vlastní

<sup>44</sup> V té době již bude účetně uzavřený předchozí měsíc a budeme znát potřebná data

## 4.9 Autorizace a autentizace

Nastavení uživatelských rolí v rámci BI ošetřím na dvou místech:

- V samotné SSAS databázi u přístupu k datům
- Ve firemním webovém sdíleném úložišti SharePoint

V mém případě použiji pouze dvě role: Administrátor a User:

- Administrátor je role bez omezení s možností modifikace, čtení i mazání dat i OLAP kostky. U této role se očekává vysoká znalost logiky OLAP a databází jak relačních, tak SSAS.
- User bude role pouze s možností čtení (prohlížení) dat. Právě tuto roli nastavím testovacím uživatelům a to i na SharePoint i v SSAS databázi.

## 4.10 Školení uživatelů, testy a předání BI řešení do produkce

V rámci Demo fáze ještě zůstávají data statická a ETL procesy nejsou aktivní. Například naplním kostku (dimenze + fakta) daty za poslední 3 měsíce, což by mělo být dostatečné pro testy uživatelů. Dále provedu základní školení klíčových uživatelů, kdy vysvětlím principy fungování kostky, zajistím propojení kostky do vybraných testovacích nástrojů. V našem případě se bude jednat o MS Excel 2010 a MS SharePoint 2014. Uživatelé provedou v následujících dnech a týdnech akceptační testy, během kterých ověří, zda byly splněny jejich business požadavky a zpětnou vazbu mi předají. Na základě zpětných vazeb zjistím, zda je třeba nějakých změn v logice a provedení OLAP kostky. Tyto změny případně analyzuji a implementuji do stávajícího řešení. Další variantou zjištění na základě zpětných vazeb bude, že uživatelé nerozumí logice kostky. Zde pomůže opakování školení a lepší podpůrné školící materiály.

Pokud jsou již akceptační testy (případně opakované testy po implementaci změn) schváleny, lze nástroj nasadit do produkční fáze a aktivovat ETL procesy.

V rámci předání je třeba mít zajištěnou dokumentaci všech potřebných komponent, a to jak SQL skriptů, tak popisů jednotlivých tabulek, zodpovědností a rolí.

## 4.11 Zhodnocení výsledků implementace a další doporučení

Celkově technologicky implementace proběhla celkem hladce. To se povedlo zejména díky:

- Výborně zpracovaným primárním zdrojům, resp. datovému skladu, ze kterého jsem již mohl data sbírat – ETL bylo tedy jednodušší.
- Podpoře BI oddělení při nasazování a spolupráci klíčových uživatelů při testech a anketě.

Hodnocení implementace s klíčovými uživateli jsem procházel pouze ústně. Vzhledem k tomu, že jsem implementaci prováděl v jednom uceleném kroku s rovnou viditelným výsledkem, hodnocení již bylo pouze celkové. Hlavní pozitivní přínos nasazení kostky vnímají v tom, že uživatelé:

- Ušetří čas při získání informací, týkajících se hlasového provozu s rozpadem až na jednotlivé služby v libovolných nadefinovaných pohledech. Informace jim kostka poskytne okamžitě.
- Nemusí zakládat pro každý potřebný pohled (dimenze) na fakta nový požadavek na IT, ale sami si informace rychle získají.

Výhody optimalizací i nové OLAP kostky jsou velmi obtížně měřitelné v penězích nebo v čase. Pokud ale bude mít např. produktový manažer správné informace v potřebném detailu včas, může se rozhodnout lépe než při nejistotě. Takové rozhodnutí může ušetřit firmě velikosti GTS Czech desítky tisíc korun, například při stanovování cen nových hlasových tarifů.

Díky tomu, že můžeme využít stávající infrastrukturu a pouze jí rozšířit, dochází k úspoře oproti vystavění BI řešení „na zelené louce“ v řádech desítek (až stovek) tisíc Kč.

Je také možné ušetřit na nákladech za externí pracovníky z oddělení BI, kteří jednotlivé reporty budují a nasazují. Datová kostka toto zcela eliminuje. Při ceně externího pracovníka 3500-5000 Kč/den, kdy stavba reportu trvá průměrně 1-2 pracovní dny (spolu s testy) a nových reportů připravují měsíčně několik.

Dalším doporučením je jednoznačně datovou kostku rozšířit o další business oblasti (např. CRM, provisioning, troubleshooting atd.) Kostka se dá velmi snadno a rychle doplnit o další dimenze i fakta.



## 5 Závěr

Cílem této práce byla rešerše literárních zdrojů o BI aplikacích a nástrojích a na jejím základě postavená případová studie s návrhem řešení (optimalizací) v BI oblasti vybrané společnosti.

V teoretické části práce jsem vysvětlil pojem BI, jeho podstatu, komponenty, využití. Detailněji jsem se věnoval popisu datového skladu – jedné ze základních komponent BI. Vysvětlil jsem pojem metadata a proč jsou v BI (ale i v jiných oblastech IT) tak důležitá.

Dále jsem charakterizoval pojem reporting a jeho nástroje. Zmínil jsem použití analytických aplikací a BPM komponent BI pro taktické a operativní rozhodování a strategii. Hlavní obsah teoretické části je věnován vysvětlení OLAP (datové kostce) a charakteristice celého procesu BI implementace s detailním popisem jeho fází. V teoretických částech jsem se věnoval detailněji oblastem BI, u kterých jsem očekával pozdější využití v praktické části práce. A naopak, např. velmi zajímavou oblast dolování dat (Data mining) jsem zcela opominul. K té bych se rád později vrátil a věnoval se jí samostatně.

V praktické části práce jsem aplikoval metodiku nestrukturovaných pohovorů a ankety mezi vybranými klíčovými uživateli pro analýzu stávajícího stavu ve společnosti GTS Czech. Další analýzu jsem provedl nad technologickou platformou, DWH a ETL. Třetí oblast analýzy je zaměřena nejen na používané, ale i dostupné a nevyužívané reportovací nástroje. Moje pracovní hypotéza v úvodu práce, že uživatelé nebudou rozumět BI nástrojům (datové kostce) a většinu z nich ani umět používat, se potvrdila. Kromě běžných reportů v tabulkách MS Excel nejsou využívány další možnosti a nástroje BI, ale uživatele se nebrání novým znalostem a BI možnosti je zajímají.

Hlavním výstupem této práce je vytvoření a nasazení datové kostky s fakty hlasového provozu podle specifikace uživatelů v rámci optimalizace stávajícího stavu.

Má klíčová doporučení (návrhy na zlepšení) u zjištěných neoptimálních skutečností jsou následující:

- přechod na near-to realtime stav vybraných reportů
- revize metadat v DWH, dokumentace ETL procesů, revize všech spouštěných uložených procedur (triggerů) a jejich konsolidace

- školení uživatelů na BI nástroje, workshopy, zajistit jim demo výukové prostředí a materiály
- upgrade a rekonfigurace SQL Serveru a LAN s důrazem na komfort uživatelů BI

V budoucnu budu pokračovat se zkoumáním BI možností a jejich prezentací (a implementací) uživatelům nejenom na platformě Microsoft. Například právě v oblasti dolování dat.

## Seznam literatury a použitých zdrojů

- [1] GÁLA, Libor, Jan POUR a Zuzana ŠEDIVÁ. Podniková informatika. 2., přeprac. a aktualiz. vyd. Praha: Grada, 2009, 496 s. Expert (Grada). ISBN 978-80-247-2615-1.
- [2] MICHALEWICZ, Zbigniew. *Adaptive business intelligence*. New York: Springer, c2007, xiii, 246 p. ISBN 978-354-0329-282.
- [3] TURBAN, Efraim. *Business intelligence: a managerial approach*. 2nd ed. Boston: Prentice Hall, c2011, xx, 292 p. ISBN 01-361-0066-X.
- [4] POUR, Jan, Miloš MARYŠKA a Ota NOVOTNÝ. *Business intelligence v podnikové praxi*. 1. vyd. Praha: Professional Publishing, 2012, 276 s. ISBN 978-80-7431-065-2.
- [5] NOVOTNÝ, Ota. *Business intelligence: jak využít bohatství ve vašich datech*. 1. vyd. Praha: Grada, 2005, 254 s. ISBN 80-247-1094-3.
- [6] INMON, William H. *Building the data warehouse*. 4th ed. Indianapolis, Ind.: Wiley, c2005, xxviii, 543 p. ISBN 07-645-9944-5.
- [7] *Understanding metadata*. Bethesda, MD: NISO, 2004. ISBN 18-801-2462-9.
- [8] LEBR, Radek. *Reportovací nástroje MSSQL Server 2012 ve společnosti GTS Czech*. Praha, 2013. Bakalářská práce. Vysoká škola manažerské informatiky, ekonomiky a práva.
- [9] SABHERWAL, Rajiv a Irma BECERRA-FERNANDEZ. *Business intelligence: practices, technologies, and management*. Hoboken, NJ: Wiley, c2011, vii, 295 p. ISBN 04-704-6170-5.
- [10] HUBBARD, Douglas W. *How to measure anything: finding the value of "intangibles" in business*. 2nd ed. Hoboken: John Wiley, c2010, xv, 304 s. ISBN 978-0-470-53939-2.
- [11] ECKERSON, Wayne. *Performance dashboards: measuring, monitoring, and managing your business*. Hoboken: Wiley, c2006, xviii, 301 s. ISBN 978-0-471-72417-9.
- [12] EVANGELU, Jaroslava Ester a Jiří NEUBAUER. *Testy pro personální práci: jak je správně vytvářet a používat*. 1. vyd. Praha: Grada, 2014, 142 s. Manažer. ISBN 978-802-4750-569.

## Webové stránky:

- [13] BITTLE. Bittle - Online dashboards and reporting software [online]. 2012 [cit. 2013-02-12]. Dostupné z: <http://www.bittle-solutions.com/en>
- [14] Reporting a analýzy. MIBCON A.S. Mibcon [online]. 2010 [cit. 2013-03-14]. Dostupné z: <http://www.mibcon.cz/cs/co-delame/epm/reporting-a-analyzy>
- [15] Eclipse BIRT Project. ECLIPSE FOUNDATION. Eclipse - The Eclipse Foundation open source community website [online]. 2012 [cit. 2013-02-10]. Dostupné z: <http://www.eclipse.org/birt/phoenix/>
- [16] JReport. JINFONET SOFTWARE. Java Reporting Software | Jinfonet Software [online]. 2012 [cit. 2013-02-10]. Dostupné z: <http://www.jinfonet.com/product/jreport-designer>
- [17] Reporting a analýzy. MIBCON A.S. Mibcon [online]. 2010 [cit. 2013-03-14]. Dostupné z: <http://www.mibcon.cz/cs/co-delame/epm/reporting-a-analyzy>
- [18] Reporting Services SSRS. MICROSOFT CORPORATION. Microsoft Developer Network [online]. 2012 [cit. 2013-01-27]. Dostupné z: <http://msdn.microsoft.com/en-us/library/ms159106.aspx>
- [19] Report Builder. MICROSOFT CORPORATION. MSDN – the Microsoft Developer Network [online]. 2012 [cit. 2013-02-10]. Dostupné z: <http://msdn.microsoft.com/en-us/library/ee633670.aspx>
- [20] BOEHM, B. A spiral model of software development and enhancement. *ACM SIGSOFT Software Engineering Notes*. 1986, vol. 11, issue 4, s. 22-42. DOI: <http://dx.doi.org/10.1145/12944.12948>.
- [21] CHANDLER, Neil, Nigel RAYNER, John VAN DECKER a James HOLINCHECK. *Magic Quadrant for Corporate Performance Management Suites*. Gartner RAS Core Research Note, 2009. Gartner RAS Core Research Note. Dostupné z: <https://www.gartner.com/doc/959212/magic-quadrant-corporate-performance-management>
- [22] Types of OLAP Systems. *Olap.com* [online]. 2013 [cit. 2014-09-07]. Dostupné z: <http://olap.com/types-of-olap-systems/>
- [23] GEISHECKER, Lee a Nigel RAYNER. Corporate Performance Management: BI Collides With ERP. *Gartner.com* [online]. 2001. vyd. 2001 [cit. 2014-09-07]. Dostupné z: <https://www.gartner.com/doc/350794/corporate-performance-management-bi-collides>

- [24] BusinessFinance: Best Practices for Finance Executives. In: COLBERT, John. *Captain Jack and the BPM Market: Performance Management in Turbulent Times* [online]. 2009 [cit. 2014-09-07]. Dostupné z: <http://businessfinancemag.com/business-performance-management/captain-jack-and-bpm-market-performance-management-turbulent-times>
- [25] NOVELL. *Executive Dashboards: Elements of Success* [online]. Novell White Paper, 2009, Duben 2009 [cit. 2014-09-07]. Dostupné z: [http://www.novell.com/docrep/2009/03/Executive%20Dashboards Elements of Success White Paper en.pdf](http://www.novell.com/docrep/2009/03/Executive%20Dashboards%20Elements%20of%20Success%20White%20Paper%20en.pdf)
- [26] Microsoft Agrees to Acquire ProClarity, Enhancing Business Intelligence Offering. In: *Microsoft* [online]. 2006 [cit. 2014-12-09]. Dostupné z: <http://news.microsoft.com/2006/04/03/microsoft-agrees-to-acquire-proclarity-enhancing-business-intelligence-offering/>
- [27] Power View (SSRS). MICROSOFT CORPORATION. Microsoft Technet [online]. 2012 [cit. 2013-02-01]. Dostupné z: <http://technet.microsoft.com/cs-cz/library/hh213579.aspx>
- [28] Business Intelligence: Building Your First Cube. MICROSOFT. *TechNet Magazine* [online]. 2009 [cit. 2015-02-01]. Dostupné z: <https://technet.microsoft.com/en-us/magazine/ee677579.aspx>
- [29] YTD (MDX). MICROSOFT. *Microsoft Developer Network* [online]. 2012 [cit. 2015-02-01]. Dostupné z: <https://msdn.microsoft.com/en-us/library/ms146039.aspx>

## Seznam obrázků

- Obrázek 3.1 Adaptivní BI
- Obrázek 3.2 Uživatelská architektura BI
- Obrázek 3.3 BI schema
- Obrázek 3.4 ODS funkce
- Obrázek 3.5 Přírůstkové řešení
- Obrázek 3.6 ETL proces
- Obrázek 3.7 Microsoft Reporting a OLAP nástroje dle přístupu a pravidelnosti
- Obrázek 3.8 Dělení BI spotřebitelů
- Obrázek 3.9 Bittle thin client
- Obrázek 3.10 Vystavení reportu pomocí Report Builderu
- Obrázek 3.11 BARC průzkum spokojenosti
- Obrázek 3.13 STAR schema
- Obrázek 3.14 SNOWFLAKE schema
- Obrázek 3.15 OLAP kostka
- Obrázek 3.16 Dashboard Operativní
- Obrázek 3.17 Životní cyklus SW
- Obrázek 3.18 BI implementace
- Obrázek 4.1 Primární zdroje GTS Czech v DWH
- Obrázek 4.2 Využití DWH v průběhu dne
- Obrázek 4.3 Porovnání přístupu k vytváření reportů
- Obrázek 4.4 Příklad uživatelského reportu v prostředí SharePoint (web)
- Obrázek 4.5 KPI
- Obrázek 4.6 Relační diagram vysoká normalizace
- Obrázek 4.7 Relační diagram nízká normalizace
- Obrázek 4.8 Nasazená databáze v SSAS v prostředí MS Management Studio
- Obrázek 4.9 SQL Agent Job
- Obrázek 4.10 ETL proces jako vývojový diagram

## Seznam tabulek

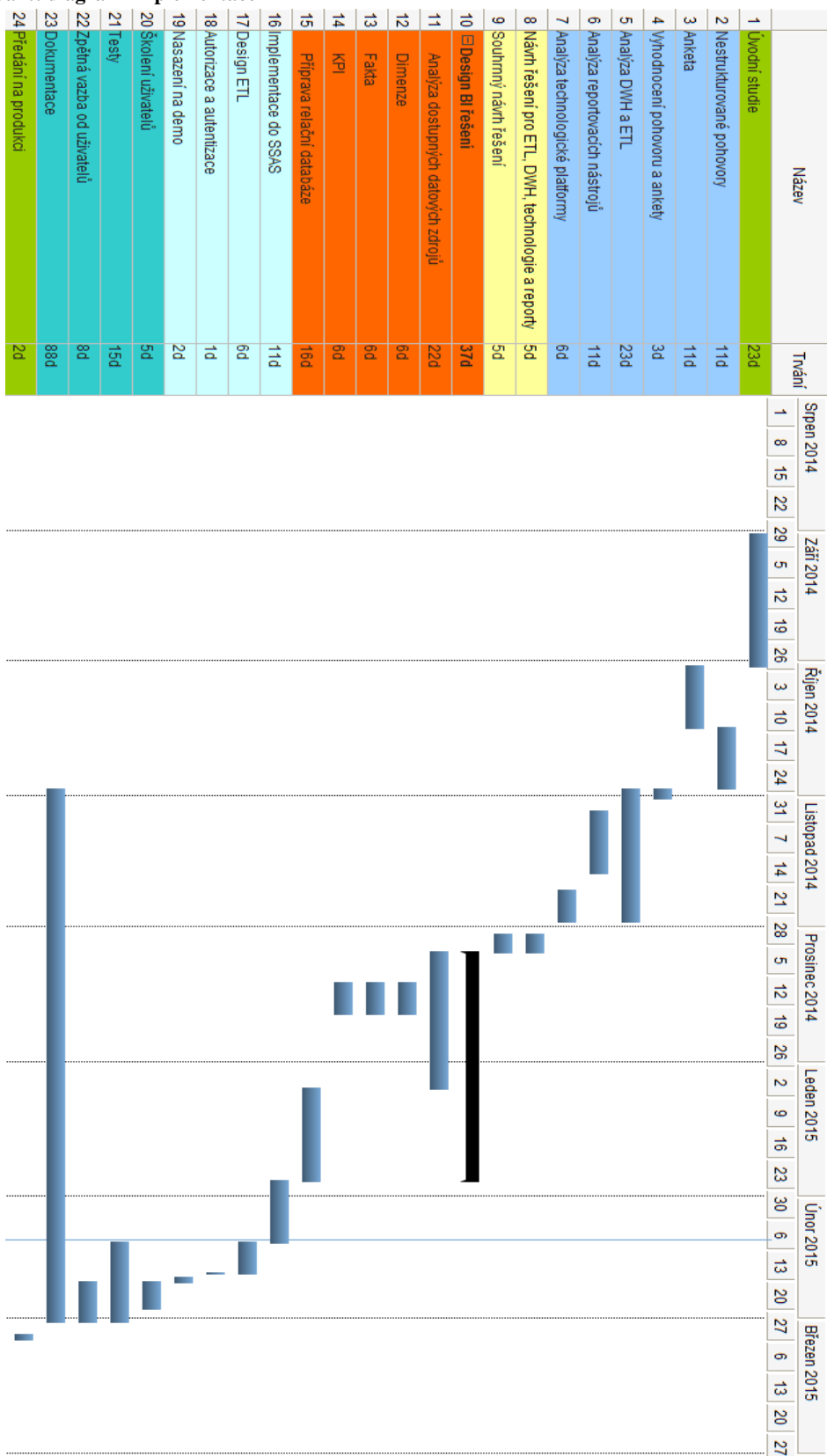
- Tabulka 3.1 Srovnání hierarchických přístupů
- Tabulka 3.2 Porovnání BI nástrojů
- Tabulka 3.3 Porovnání SNOWFLAKE a STAR modelu
- Tabulka 4.1 Seskupené a zjednodušené odpovědi uživatelů
- Tabulka 4.2 Porovnání web GUI SharePoint BI nástrojů
- Tabulka 4.3 Náklady na optimalizaci BI
- Tabulka 4.4 Uživatelské požadavky na změny BI
- Tabulka 4.5 Metadata vytvořených tabulek

## **Seznam příloh**

- Příloha 1 Gantt diagram Implementace BI
- Příloha 2 Výstup ankety respondentů 1-3
- Příloha 3 Výstup ankety respondentů 4-6
- Příloha 4 Výstup ankety respondentů 7-9
- Příloha 5 Výstup ankety respondentů 10-12
- Příloha 6 Výstup ankety respondentů 13-15
- Příloha 7 Výstup ankety respondentů 16-17
- Příloha 8 Seznam primárních zdrojů v GTS Czech
- Příloha 9 Přehled BI řešení
- Příloha 10 Příklad použitých vytvářecích a update SQL skriptů

# Přílohy

Příloha 1 Gantt diagram Implementace BI





**Příloha 2 Výstup ankety respondentů 1-3**

Číslo respondenta	1	2	3
<b>Divize</b>	Finanční	Produkt	Marketing
<b>Pozice</b>	Zaměstnanec	Zaměstnanec	Manažer
<b>Otázky:</b>			
<b>Zákazník?</b>	Interní	Oba	Oba
<b>1</b>	Excel, Access, Sharepoint, MS SQL Management studio	Excel, Access, SharePoint	Access, Reporty na Sharepointu, Externí dB a Excel
<b>2</b>	Ano	Ano	Ano
<b>3a</b>	-		
<b>3b</b>		1)aktuální provoz mobilních služeb na číslo - critical ; 2)aktuální provoz fixních hlas služeb na číslo - optional; 3)napárování whsl nákladů na retailový provoz (jde o INT)-critical; 4)napárování nákladů k mobilním službám - optional; 5)report fakturací provozu+ dat na mobilních službách není na shp-critical	Optional
<b>4</b>	Hlasový provoz -critical; Tržby - optional	GM report (2 měsíce zpětně kvůli nákladům) - optional	Máme 2 CRM systémy, které nejsou zcela On-line - optional
<b>5</b>	U některých nadefinovaných reportů je možné se dostat o úroveň níže, u některých by bylo vhodné mít možnost udělat agregaci u určitých položek nebo naopak větší detail (vazba na OLAP).	Reportů hlasu je dostatek s různým detailem a lze si vybrat – nevýhoda, při změně filtru nutno upravit všechny.	Ano je
<b>6a</b>	Ano, data se rozbálí. Přepočty pro určitou oblast grafu by byly vhodné.	Ne, ale rád bych (Optional)	Občas v Excelu
<b>6b</b>	Nejsou		Ano
<b>7a</b>	Některé jsou pomalé	Reporty hlasového provozu trvají dlouho – velký detail.	Access by mohl mít rychlejší odezvu.
<b>7b</b>	Critical – čím rychlejší výstup, tím samozřejmě lepší.	Optional	Optional
<b>8a</b>	Nevyužívám, generuji si individuální dotazy, ale možnost měnit pohledy OLAP přes např. grafický interface by bylo příjemné.	minimální	Malé
<b>8b</b>	Např. U GM reportů by byla vhodná možnost kombinace pohledů zákaznický vs. produktový atp.	Critical	Ano - optional
<b>9a</b>	Záleží na konkrétním reportu a správci reportu.	Půl napůl	Spíše ano
<b>9b</b>	7-14 dní	14dni-měsíc	1-3 dny - sami vytváříme
<b>9c</b>	Ano. Občas se k řešení požadavků dostanou až po dlouhé době.	Optional	Reporty si s 95% vytváříme sami - optional

**Příloha 3 Výstup ankety respondentů 4-6**

Číslo respondenta	4	5	6
<b>Divize</b>	Customer Care	Wholesale Woice	Finanční
<b>Pozice</b>	Ředitel	Ředitel	Manažer
<b>Otázky:</b>			
<i>Zákazník?</i>	Externí	Oba	Interní
<b>1</b>	Reporty na Sharepointu	Excel, Sharedpoint, Proclarity	Excel, Reporty na Sharepointu, Essbase, Reporty ze SAPu
<b>2</b>	Ne	Ano	Ano
<b>3a</b>	-	-	nic mi nechybí, jsem úplně spokojený
<b>3b</b>	1)stav (%) plnění plánu intaku optional; 2)report na výpočet provizí optional	-	-
<b>4</b>	dostatečné	Informace o provozu v Proclarity – nyní cca. 2,5 hodiny zpoždění, potřebujeme max. 15 minut-critical	GM Report – rychlejší měsíční aktualizace -critical
<b>5</b>	naprosto dostačující		Detail reportu je OK, případně agregace si dělám sám v Excelu
<b>6a</b>	nevyužívám	ano	ano
<b>6b</b>	-	ano	ano
<b>7a</b>	až na výjimky (např. report churnu) je rychlost dostačující, a to i při práci přes vpn	Většinou OK	Je to OK.
<b>7b</b>	Nice to have	Nice to have	Nice to have
<b>8a</b>	co to je?:-)	žádné	drivejší - Proclarity a sledování WS hlasového provozu
<b>8b</b>	neznám, nemůžu posoudit	-	momentálně ne
<b>9a</b>	nemám požadavky na změnu	ano	Celkovou spokojenost bych označoval za 3. Úpravy a změny trvají dlouho. Požadavku mám jen minimum.
<b>9b</b>	N/A	7-14 dní	V řádu týdnů. Mám minimum požadavku.
<b>9c</b>	N/A	Nice to have	Nice to have

**Příloha 4 Výstup ankety respondentů 7-9**

Číslo respondenta	7	8	9
<b>Divize</b>	Wholesale Voice int	Customer Solutions	Finance
<b>Pozice</b>	Ředitel	Ředitel	Manažer
<b>Otázky:</b>			
<i>Zákazník?</i>	Oba	Oba	Oba
<b>1</b>	Excel, Reporty na Sharepointu, ProClarity	Excel, Access, Reporty na Sharepointu, MS SQL	Excel, Sharepoint, Salesforce
<b>2</b>	ano	ano	ano
<b>3a</b>	nic	Pokročilý selfcare BI nástroj Typu Power Pivot nebo rovnou MS SQL 2014 + BI Power	
<b>3b</b>	-	-	1)implementování ratingu zákazníku do CRM systému - Nice to have; 2) dodatečná automaticke procesu vymáhání pohledávek - Optional; 3)automaticky report opravných položek - Critical
<b>4</b>	1)CDR zústředěn, mediace - optional; 2)Datové kostky AWSB -optional	Problém není aktualizace, ta by šla , ale čistota data a nároky na znalost Business Logiky v rámci DW tabulek, pokud si clovek chce vytvořit report sam.	1)U reportu kde mam možnost editovat data, update az v noci - critical; 2)Saldo prichozich plateb jenom 1x denne, urcite lepsi online - critical
<b>5</b>	detail stávajících reportů je dostatečný	Pro standarni reporty detail OK, pro individualni je to slozitejsi.	Zaleží na uživateli, Excel nebo SharePoint report má možnost si sám nadefinovat ve spolupráci s IT. Když uživatel neví co chce, tak ani IT mu neporadí. Nevím o možnosti sám vytvořit a administrovat report
<b>6a</b>	ano	No tady jsem spíše na začátku toho umět tyto nástroje vytvořit, používám je ale rád.	ne
<b>6b</b>	ano	ano	-
<b>7a</b>	Rychlost je akceptovatelná	Rychlost pokud nedojde o enormní objemy dat což jsou výjimky je OK	Když je report jenom lokální, je odezva je dostatečná.
<b>7b</b>	Nice to have	Optional	v současné době, „nice to have“. U nového reportu určíte critical
<b>8a</b>	OLAP kostky obsahují data pro naše reporty z Proclarity	střední	nemam
<b>8b</b>	nové kostky nepotřebujeme	nice to have	-
<b>9a</b>	ano	OK, máme dobré lidi co je tvoří individuálně..	Moc, ne. Zalezi na priorite IT
<b>9b</b>	Reporty tvoříme sami online v Pr	1-3 dny	Více než jeden měsíc
<b>9c</b>	N/A	nice to have	critical

**Příloha 5 Výstup ankety respondentů 10-12**

Číslo respondenta	10	11	12
<b>Divize</b>	Obchod	Produkt	Technical
<b>Pozice</b>	Manažer	Manažer	Manažer
<b>Otázky:</b>			
<b>Zákazník?</b>	Externí	Oba	Externí
<b>1</b>	Excel, Reporty na Sharepointu, OC	Excel, Sharepoint	Excel, reporty na sharepointu
<b>2</b>	ano	ano	ano
<b>3a</b>	-	-	nevím jestli chybi, ale možná špatně hledám
<b>3b</b>	Okamžitý náhled na výši provize u jednotlivého OP - nice to have	1)Reporty uzavřených casů z VM a BCT - optional; 2)Reporty ala Product Report Package za všechny země - optional; 3)Reporty počtu služeb mimo CZ - optional	1)pocet sluzeb per produkt - optional; 2)chybi možnost reportovat na zaklade jednotlivych zapisu v OCTT, podle me data nejsou v zadne kostce - critical
<b>4</b>	Vše OK	-	Ne, aktualizace každý týden/den mi staci
<b>5</b>	V podstatě dostatečný; V rámci sales budgetů mít možnost rozkliknout detail z jakých OP/zákazníků se údaje o intaku/churnu/vývoje reventýkají - optional	Moc ne případně požadavky řeším individuálně	Nejcasteji delam report nad TT, moji vyhodou je, ze si mohu udelat i report sam, neboť mam zpristupneny DWH. U opakujících se reportu nasledne zadavam HD pro jejich zřízení
<b>6a</b>	ne	ne	ne
<b>6b</b>	-	-	-
<b>7a</b>	Dobré	Stredni, mohlo by to byt lepsi ale jde to	Krome nektery pondelnic ran, je ve Ok
<b>7b</b>	-	nice to have	-
<b>8a</b>	žádné	žádné	dobře
<b>8b</b>	nevím	nice to have	jiz pomohlo
<b>9a</b>	ano	Ano, zejména s Kamilou	Vetsinou potrebují report hned, potreba vychazi z aktualni situace
<b>9b</b>	Nežádám o vlastní reporty, tedy neumím odpovědět	4-7 dní	14dní - měsíc
<b>9c</b>	-	nice to have	nekdy critical, nekdy optional, zalezi na typu reportu

**Příloha 6 Výstup ankety respondentů 13-15**

Číslo respondenta	13	14	15
<b>Divize</b>	Technická	Obchod	Produkt
<b>Pozice</b>	Teamleader	Manažer	Zaměstnanec
<b>Otázky:</b>			
<b>Zákazník?</b>	Oba	Externí	Oba
<b>1</b>	Excel, Reporty na Sharepointu	Reporty na Sharepointu	Excel, Reporty na Sharepointu
<b>2</b>	Ano	Ne	Ano
<b>3a</b>	nic mi nechybí	-	Nic mě nenapadá
<b>3b</b>	-	1) stav (%) plnění plánu intaku - Optional; 2) report na výpočet provizí - Optional	-
<b>4</b>	Občas trva na sharepointu dlouho dotahovani dat (napr. z OC:TT) - Optional	dostatečné	1) GM report - chybí aktuální data - Critical; 2) KD - uvítala bych taky rychlejší zpracování - Optional
<b>5</b>	Pro moje potřeby je stavající detail dostatečný	naprosto dostačující	Detail je pro me OK
<b>6a</b>	ne	nevyžívám	ne
<b>6b</b>	-	-	-
<b>7a</b>	Rychlost je dostatečna.	až na výjimky (např. report churnu) je rychlost dostačující, a to i při práci přes vpn	1) Obvykle pomalejší . Excely bývají rozsáhlé, takže trvá, než se načtou. 2) Sharepoint generuje některé (ne všechny) reporty velmi pomalu.
<b>7b</b>	-	Nice to have	Optional
<b>8a</b>	Vyborne zkušenosti, jsem spokojeny	co to je?:-)	žádné
<b>8b</b>	-	neznám, nemůžu posoudit	nevím, nemám zkušenosti
<b>9a</b>	Občas bych uvítal pružnější reakce na opravy/vytváření nových, ale nastěti moje požadavky nejsou většinou urgentní.	nemám požadavky na změnu	Mohlo by to být trochu lepší, ale celkově nemám problém.
<b>9b</b>	Více než jeden měsíc	N/A	Pokud je myšlen nový report, tak 14dní-měsíc
<b>9c</b>	Optional	N/A	Nice to have

**Příloha 7 Výstup ankety respondentů 16-17**

<b>Číslo respondenta</b>	16	17
<b>Divize</b>	Produkt	Finance
<b>Pozice</b>	Zaměstnanec	Zaměstnanec
<b>Otázky:</b>		
<b>Zákazník?</b>	Oba	Interní
<b>1</b>	Excel, Access, Reporty na Sharepointu	Excel, Access, Sharepoint, MS SQL Management studio, ProClarity
<b>2</b>	Ano	Ano
<b>3a</b>	v současnosti ne	žádné, na své pozici mám přístup ke všem potřebným datům. Dokážu si udělat jakýkoli ad-hoc report svépomocí
<b>3b</b>	-	-
<b>4</b>	-	Vhodné by byla častější aktualizace dat z Durianu (faktury/cdr) - Nice to have
<b>5</b>	Dostatečné	Většinu reportů si vytvářím sám přímo nad daty v DWH, tudíž jsem na nejvyšším možném detailu
<b>6a</b>	Ne	Ano. Tuto možnost nabízejí Proclarity pro WS hlasový provoz. Na své pozici ji ale příliš nevyžívám
<b>6b</b>	-	Ano (bylo by ale vhodné mít podobné možnosti i v reportech týkajících se jiných oblastí)
<b>7a</b>	ok	Náročnější reporty na Sharepointu jsou poměrně pomalé a bylo by vhodné jejich rychlost zrychlit – important (velké reporty s mnoha řádky - např. reporty nad databází okruhů)
<b>7b</b>	Nice to have	Optional
<b>8a</b>	žádné	Zkušenosti s kostkami nemám, používám pouze statické reporty. Kostka mi možná práci ulehčila, ale bohužel jí nemám k dispozici a sám si ji nedokážu vytvořit.
<b>8b</b>	Optional	Ano, pomohla
<b>9a</b>	Ne	Nelze určit (jak kdy). U náročnějších reportů, které si nevytvářím sám, ale jsou vytvářené oddělením IT, trvá vytvoření často poměrně dlouho
<b>9b</b>	4-7 dní	14dní-měsíc
<b>9c</b>	Critical/Optional	Critical

**Příloha 8 Seznam primárních zdrojů v GTS Czech**

Schema	Usage	Description	DataUsage (GB)
AMD	Frozen source	exContactel Billing system	0,243591
ANOA	Source system	GTSCZ Integration platform	0,006325
BASEN	Source system	Monitoring služeb	3,474144
BASH	Frozen source	exGTS Billing system	0,098885
BMODEL	Source system	Výpočetní model	1,328957
BROADSOFT	Source system	VoIP platforma	2,059708
CBCT	Source system	Common Business Calculation Model for CZ	0
COW	Source system	Project and Workflow for Service Provisionig (Customer Order Workflow)	1,068283
CSR	Source system	Customer Service Repository	1,36821
DBLM	Source system	Roster system	9,92
DC	Source system	Active Direcrory Domain Controlers	0,366257
DENVER	Source system	GTSCZ Integration platform	0,443695
DUR	Source system	GTSCZ Billing system	291,3095
IDENT	Source system	CZ technology inventory	6,822105
ISP	Source system	IPS platform	0,416893
JIRA	Source system	GTSCZ Helpdesk system	2,17688
KRAB	Source system	Approval workflow	6,418297
KS	Source system	Attendance system	0,00235
NRM	Source system	Network inventory	0,423409
OCP	Source system	GTSCZ Opportunity/Provisioning system	24,26205
OSS	Source system	Technology platform for DSLAM	0,400566
SAP	Source system	GTSCZ ERP system	0,391228
SFDC	Source system	GTSCZ CRM system	49,32484

SIMA	Source system	Project Tool	0,125092
SLW	Source system	GTSCZ CRM system	8,19854
SPS	Source system	Sharepoint	33,29908
UMS	Source system	Umbrela management system	1,0485

**Příloha 9 Přehled BI řešení**

Název BI řešení	Licence	Společnost	www stránky
1010data	komerční	1010 data	<a href="http://www.1010data.com">www.1010data.com</a>
BIRT	Open source	eclipse	<a href="http://eclipse.org/birt/">http://eclipse.org/birt/</a>
Actuate BIRT Analytics	komerční	OpenText	<a href="http://birtanalytics.actuate.com/">http://birtanalytics.actuate.com/</a>
Alteryx solutions	komerční	Alteryx	<a href="http://www.alteryx.com/">http://www.alteryx.com/</a>
arcplan Enterprise	komerční	arcplan	<a href="http://www.arcplan.com/">http://www.arcplan.com/</a>
Centerprise	komerční	Astera	<a href="http://www.astera.com/">http://www.astera.com/</a>
Birst	komerční	Birst	<a href="http://www.birst.com/">http://www.birst.com/</a>
Artus	komerční	Bitam	<a href="http://www.bitam.com/i">http://www.bitam.com/i</a>
BOARD	komerční	BOARD International S.A.	<a href="http://www.board.com/">http://www.board.com/</a>
Anzo Enterprise	komerční	Cambridge Semantics	<a href="http://www.cambridgesemantics.com/">http://www.cambridgesemantics.com/</a>
Clear Analytics	komerční	Clear Analytics Inc.	<a href="http://www.clearanalyticsbi.com/">http://www.clearanalyticsbi.com/</a>
ClickBase	komerční	ClickBase	<a href="http://www.clickbase.com/">http://www.clickbase.com/</a>
Cloudera Enterprise	komerční	Cloudera	<a href="http://www.cloudera.com/">http://www.cloudera.com/</a>
Cover-All BI	komerční	Cover-All	<a href="http://www.cover-all.com/">http://www.cover-all.com/</a>
DataHero Premium	komerční	DataHero	<a href="https://datahero.com/">https://datahero.com/</a>
DataRPM	komerční	DataRPM	<a href="http://datarpm.com/">http://datarpm.com/</a>
Datawatch	komerční	Datawatch	<a href="http://www.datawatch.com/">http://www.datawatch.com/</a>
Decisyon 360	komerční	Decisyon, Inc.	<a href="http://www.decisyon.com/">http://www.decisyon.com/</a>
The Diver Solution	komerční	Dimensional Insight	<a href="http://www.dimins.com/">http://www.dimins.com/</a>
Domo	komerční	Domo	<a href="http://www.domo.com/">http://www.domo.com/</a>
Operational Business Intelligence	komerční	EXASOL	<a href="http://www.exasol.com/">http://www.exasol.com/</a>
Enterprise Analytics	komerční	GoodData	<a href="http://www.gooddata.com/">http://www.gooddata.com/</a>



Halo BI	komerční	Halo BI	<a href="http://halobi.com/">http://halobi.com/</a>
High IQ Business Intelligence	komerční	High IQ Business Intelligence	<a href="http://highiqbi.com/">http://highiqbi.com/</a>
Cognos	komerční	IBM	<a href="http://www.ibm.com">www.ibm.com</a>
InetSoft	komerční	InetSoft	<a href="https://www.inetsoft.com/">https://www.inetsoft.com/</a>
Infor 10x	komerční	Infor	<a href="http://www.infor.com/">http://www.infor.com/</a>
WebFOCUS	komerční	Information Builders	<a href="http://www.informationbuilders.com/">http://www.informationbuilders.com/</a>
Kofax business intelligence			
and analytics	komerční	KOFAX	<a href="http://www.kofax.com/">http://www.kofax.com/</a>
Logi Info	komerční	Logi Analytics	<a href="http://www.logianalytics.com/">http://www.logianalytics.com/</a>
Looker	komerční	Looker	<a href="http://www.looker.com/">http://www.looker.com/</a>
Magnitude Software	komerční	Magnitude Software	<a href="http://magnitudesoftware.com/">http://magnitudesoftware.com/</a>
Sharepoint, BI tools	komerční	Microsoft	<a href="http://www.microsoft.com">www.microsoft.com</a>
MicroStrategy Analytics	komerční	MicroStrategy	<a href="http://www.microstrategy.com/">http://www.microstrategy.com/</a>
Numerify	komerční	Numerify	<a href="http://numerify.com/">http://numerify.com/</a>
Business Intelligence			
Foundation Suite	komerční	Oracle	<a href="http://www.oracle.com">www.oracle.com</a>
Panorama Necto 15	komerční	Panorama	<a href="http://www.panorama.com/">http://www.panorama.com/</a>
Pentaho Business Analytics	komerční	Pentaho	<a href="http://www.pentaho.com/">http://www.pentaho.com/</a>
Performance Canvas	komerční	Performance Canvas	<a href="http://www.performancecanvas.com/">http://www.performancecanvas.com/</a>
Prognoz Platform	komerční	Prognoz	<a href="http://www.prognoz.com/">http://www.prognoz.com/</a>
Qlik Sense	komerční	Qlik	<a href="http://www.qlik.com/">http://www.qlik.com/</a>
Rapid Insight Analytics	komerční	Rapid Insight	<a href="http://www.rapidinsightinc.com/">http://www.rapidinsightinc.com/</a>
Roambi Analytics	komerční	Roambi	<a href="http://roambi.com/">http://roambi.com/</a>
Salient Collaboration			
Intelligence Suite	komerční	Salient Management Company	<a href="http://www.salient.com/">http://www.salient.com/</a>
BusinessObject BI	komerční	SAP SE	<a href="http://www.sap.com/">http://www.sap.com/</a>
SAS Visual Analytics	komerční	SAS	<a href="http://www.sas.com/">http://www.sas.com/</a>
Sisense	komerční	Sisense	<a href="http://www.sisense.com/">http://www.sisense.com/</a>

Tableau BI	komerční	Tableau Software	http://www.tableau.com/
TARGIT Decision Suite	komerční	TARGIT	http://www.targit.com/
Teradata BI platform	komerční	Teradata	http://www.teradata.com/
ThoughtSpot	komerční	ThoughtSpot	http://www.thoughtspot.com/
TIBCO Spotfire	komerční	Tibco	http://www.tibco.com/

## Příloha 10 Příklad použitých vytvářecích a update SQL skriptů

### Vytvoření a první naplnění

#### dimenze CENIKOVKA

```

USE [EDWTEST]
GO
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
CREATE TABLE [dbo].[CENIKOVKA](
  S_CENIKOVKA_ID INT IDENTITY(1,1),
  ITEM_TYPE_ID [nvarchar](24) NULL,
  [DIRECTION_ID] [numeric](8, 0) NULL,
  [DESTINATION_ID] [numeric](8, 0) NULL,
  [IN_OUT] [nvarchar](4) NULL,
  [TRIDA] [nvarchar](64) NULL,
  [CENIKOVKA_POPIIS] [nvarchar](64) NULL,
  [DESTINACE] [nvarchar](64) NULL,
  [VALID_FROM] date null,
  [VALID_TO] date null,
  CONSTRAINT [PK_CENIKOVKA] PRIMARY KEY CLUSTERED
(
  [S_CENIKOVKA_ID] ASC
)WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY = OFF,
ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON) ON [PRIMARY]
) ON [PRIMARY]
GO

--naplnění první OUT
SELECT distinct
  a.[ITEM_TYPE_ID],
  [DIRECTION_ID]
  ,[DESTINATION_ID],'OUT' as IN_OUT
  into #data02
FROM [DSA].[DUR].[MDEVICE_CDR_OUT] a
where END_DATE_TIME >= '2014-10-01';

insert into [dbo].[CENIKOVKA]
select
a.ITEM_TYPE_ID,a.DIRECTION_ID,a.DESTINATION_ID,a.IN_OUT,d.DESCR,c.DESCR,e.DESCR,GETDATE(),null
from #data02 a
join [DSA].[DUR].[PLIST_ITEM_TYPE] b on a.ITEM_TYPE_ID = b.ID
join [DSA].[DUR].[PLIST_ITEM_TYPE_DESCR] c on a.ITEM_TYPE_ID = c.ITEM_TYPE_ID
and c.LANGUAGE_ID = 'CZECH'
join [DSA].[DUR].[PLIST_ITEM_CLASS_DESCR] d on b.CLASS_ID = d.ITEM_CLASS_ID
and d.LANGUAGE_ID = 'CZECH'
join [DSA].[DUR].[TARIFF_BM_DESTINATION_DESCR] e on a.DESTINATION_ID = e.BM_DESTINATION_ID
and e.LANGUAGE_ID = 'CZECH'
;

```

```

--naplnění první IN
SELECT distinct
  a.[ITEM_TYPE_ID],
  [DIRECTION_ID]
  ,[DESTINATION_ID],'IN' as IN_OUT
  into #data01
FROM [DSA].[DUR].[MDEVICE_CDR_IN] a
where END_DATE_TIME >= '2014-10-01';

insert into [dbo].[CENIKOVKA]
select
a.ITEM_TYPE_ID,a.DIRECTION_ID,a.DESTINATION_ID,a.IN_OUT,d.DESCR,c.DESCR,e.DESCR,GETDATE(),null
from #data01 a
join [DSA].[DUR].[PLIST_ITEM_TYPE] b on a.ITEM_TYPE_ID = b.ID
join [DSA].[DUR].[PLIST_ITEM_TYPE_DESCR] c on a.ITEM_TYPE_ID = c.ITEM_TYPE_ID
and c.LANGUAGE_ID = 'CZECH'
join [DSA].[DUR].[PLIST_ITEM_CLASS_DESCR] d on b.CLASS_ID = d.ITEM_CLASS_ID
and d.LANGUAGE_ID = 'CZECH'
join [DSA].[DUR].[TARIFF_BM_DESTINATION_DESCR] e on a.DESTINATION_ID = e.BM_DESTINATION_ID
and e.LANGUAGE_ID = 'CZECH'
;

select * from [dbo].[CENIKOVKA]
--drop table #data01;
--drop table #data02;

```

## Update tabulky

### dim CENIKOVKA

/\*update tabulky nový záznam s novým ID + DIR\_ID + DEST\_ID\*/

```

insert into [EDWTEST].dbo.[CENIKOVKA]
Select a.ITEM_TYPE_ID,a.DIRECTION_ID,a.DESTINATION_ID,'OUT' as
IN_OUT,d.DESCR,c.DESCR,e.DESCR,GETDATE(),null from (
SELECT distinct a.ITEM_TYPE_ID,a.DIRECTION_ID,a.DESTINATION_ID--, 'OUT' as
IN_OUT,d.DESCR,c.DESCR,e.DESCR,GETDATE(),null
from [DSA].[DUR].[MDEVICE_CDR_OUT] a
where END_DATE_TIME >= DATEADD(DD,-1,GETDATE())
except
select ITEM_TYPE_ID,DIRECTION_ID,DESTINATION_ID--,
,IN_OUT,TRIDA,CENIKOVKA_POPOPIS,DESTINACE,GETDATE(),null
from [EDWTEST].dbo.[CENIKOVKA]
where VALID_TO is null
)a
join [DSA].[DUR].[PLIST_ITEM_TYPE] b on a.ITEM_TYPE_ID = b.ID
join [DSA].[DUR].[PLIST_ITEM_TYPE_DESCR] c on a.ITEM_TYPE_ID = c.ITEM_TYPE_ID
and c.LANGUAGE_ID = 'CZECH'
join [DSA].[DUR].[PLIST_ITEM_CLASS_DESCR] d on b.CLASS_ID = d.ITEM_CLASS_ID
and d.LANGUAGE_ID = 'CZECH'
join [DSA].[DUR].[TARIFF_BM_DESTINATION_DESCR] e on a.DESTINATION_ID = e.BM_DESTINATION_ID
and e.LANGUAGE_ID = 'CZECH'

```

---a to samé pro IN

```

Select a.ITEM_TYPE_ID,a.DIRECTION_ID,a.DESTINATION_ID,'IN' as
IN_OUT,d.DESCR,c.DESCR,e.DESCR,GETDATE(),null from (
SELECT distinct a.ITEM_TYPE_ID,a.DIRECTION_ID,a.DESTINATION_ID--, 'OUT' as
IN_OUT,d.DESCR,c.DESCR,e.DESCR,GETDATE(),null
from [DSA].[DUR].[MDEVICE_CDR_IN] a
where END_DATE_TIME >= DATEADD(DD,-1,GETDATE())
except
select ITEM_TYPE_ID,DIRECTION_ID,DESTINATION_ID--,
,IN_OUT,TRIDA,CENIKOVKA_POPOPIS,DESTINACE,GETDATE(),null

```

```

from [EDWTEST].dbo.[CENIKOVKA]
where VALID_TO is null
)a
join [DSA].[DUR].[PLIST_ITEM_TYPE] b on a.ITEM_TYPE_ID = b.ID
join [DSA].[DUR].[PLIST_ITEM_TYPE_DESCR] c on a.ITEM_TYPE_ID = c.ITEM_TYPE_ID
and c.LANGUAGE_ID = 'CZECH'
join [DSA].[DUR].[PLIST_ITEM_CLASS_DESCR] d on b.CLASS_ID = d.ITEM_CLASS_ID
and d.LANGUAGE_ID = 'CZECH'
join [DSA].[DUR].[TARIFF_BM_DESTINATION_DESCR] e on a.DESTINATION_ID = e.BM_DESTINATION_ID
and e.LANGUAGE_ID = 'CZECH'

/*update tabulky - pouze přepis jména, kdy ID zůstane stejné, vložit nový se stejným Destination klíčem a novým
jménem + musím původní záznam zneplatnit -> Sklíč bude ale nový! */
--OUT směr
SELECT distinct a.ITEM_TYPE_ID,a.DIRECTION_ID,a.DESTINATION_ID,'OUT' as IN_OUT,d.DESCR as
TRIDA,c.DESCR as CENIKOVKA_POPIS,e.DESCR as DESTINACE ,GETDATE() as VALID_FROM,convert (date
,null) as VALID_TO
into #data03
from [DSA].[DUR].[MDEVICE_CDR_OUT] a
join [DSA].[DUR].[PLIST_ITEM_TYPE] b on a.ITEM_TYPE_ID = b.ID
join [DSA].[DUR].[PLIST_ITEM_TYPE_DESCR] c on a.ITEM_TYPE_ID = c.ITEM_TYPE_ID
and c.LANGUAGE_ID = 'CZECH'
join [DSA].[DUR].[PLIST_ITEM_CLASS_DESCR] d on b.CLASS_ID = d.ITEM_CLASS_ID
and d.LANGUAGE_ID = 'CZECH'
join [DSA].[DUR].[TARIFF_BM_DESTINATION_DESCR] e on a.DESTINATION_ID = e.BM_DESTINATION_ID
and e.LANGUAGE_ID = 'CZECH'

where END_DATE_TIME >= DATEADD(DD,-1,GETDATE())
except
select
ITEM_TYPE_ID,DIRECTION_ID,DESTINATION_ID,IN_OUT,TRIDA,CENIKOVKA_POPIS,DESTINACE,GETDATE()
,null
from [EDWTEST].dbo.[CENIKOVKA]
where VALID_TO is null
--vložení záznamu z tmp tabulky
insert into [EDWTEST].dbo.[CENIKOVKA]
Select * from #data03

--a nyní zneplatnění původního záznamu, použiji temporary tabulku
update [EDWTEST].dbo.[CENIKOVKA]
set VALID_TO = GETDATE()
where VALID_TO is null
and IN_OUT = 'OUT'
and S_CENIKOVKA_ID in (select S_CENIKOVKA_ID from

[EDWTEST].dbo.[CENIKOVKA] a
join #data03 b
on a.ITEM_TYPE_ID = b.ITEM_TYPE_ID
and a.DESTINATION_ID = b.DESTINATION_ID --
and a.DIRECTION_ID = b.DIRECTION_ID
where
a.VALID_TO is null
and ( b.CENIKOVKA_POPIS <> a.CENIKOVKA_POPIS or b.TRIDA <> a.TRIDA or b.CENIKOVKA_POPIS
<> a.CENIKOVKA_POPIS or b.DESTINACE <> a.DESTINACE ))
drop table #data03;

--to samé pro IN směr opět s použitím tmp tabulky

SELECT distinct a.ITEM_TYPE_ID,a.DIRECTION_ID,a.DESTINATION_ID,'IN' as IN_OUT,d.DESCR as
TRIDA,c.DESCR as CENIKOVKA_POPIS,e.DESCR as DESTINACE ,GETDATE() as VALID_FROM,convert (date
,null) as VALID_TO
into #data04
from [DSA].[DUR].[MDEVICE_CDR_IN] a

```

```

    join [DSA],[DUR],[PLIST_ITEM_TYPE] b on a.ITEM_TYPE_ID = b.ID
join [DSA],[DUR],[PLIST_ITEM_TYPE_DESCR] c on a.ITEM_TYPE_ID = c.ITEM_TYPE_ID
and c.LANGUAGE_ID = 'CZECH'
join [DSA],[DUR],[PLIST_ITEM_CLASS_DESCR] d on b.CLASS_ID = d.ITEM_CLASS_ID
and d.LANGUAGE_ID = 'CZECH'
    join [DSA],[DUR],[TARIFF_BM_DESTINATION_DESCR] e on a.DESTINATION_ID = e.BM_DESTINATION_ID
and e.LANGUAGE_ID = 'CZECH'

    where END_DATE_TIME >= DATEADD(DD,-1,GETDATE())
    except
    select
ITEM_TYPE_ID,DIRECTION_ID,DESTINATION_ID,IN_OUT,TRIDA,CENIKOVKA_POPIS,DESTINACE,GETDATE()
,null
    from [EDWTEST].dbo.[CENIKOVKA]
    where VALID_TO is null

--vložení záznamu
insert into [EDWTEST].dbo.[CENIKOVKA]
Select * from #data04

-- zneplatnění původního záznamu
update [EDWTEST].dbo.[CENIKOVKA]
set VALID_TO = GETDATE()
where VALID_TO is null
and IN_OUT = 'IN'
and S_CENIKOVKA_ID in (select S_CENIKOVKA_ID from
[EDWTEST].dbo.[CENIKOVKA] a
join #data04 b
    on a.ITEM_TYPE_ID = b.ITEM_TYPE_ID
    and a.DESTINATION_ID = b.DESTINATION_ID --
    and a.DIRECTION_ID = b.DIRECTION_ID
    where
    a.VALID_TO is null
    and ( b.CENIKOVKA_POPIS <> a.CENIKOVKA_POPIS or b.TRIDA <> a.TRIDA or b.CENIKOVKA_POPIS
<> a.CENIKOVKA_POPIS or b.DESTINACE <> a.DESTINACE ))
drop table #data04;

```