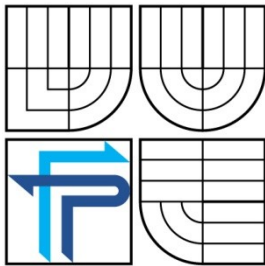


VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA PODNIKATELSKÁ
ÚSTAV MANAGEMENTU

FACULTY OF BUSINESS AND MANAGEMENT
INSTITUTE OF MANAGEMENT

NÁVRH ZMĚN INFORMAČNÍHO SYSTÉMU FIRMY

PROPOSAL FOR CHANGES IN THE COMPANY INFORMATION SYSTEM

DIPLOMOVÁ PRÁCE
MASTER'S THESIS

AUTOR PRÁCE
AUTHOR

Bc. DAVID KECLÍK

VEDOUCÍ PRÁCE
SUPERVISOR

Ing. PETR DYDOWICZ, Ph.D.

BRNO 2011

ZADÁNÍ DIPLOMOVÉ PRÁCE

Keclík David, Bc.

Řízení a ekonomika podniku (6208T097)

Ředitel ústavu Vám v souladu se zákonem č.111/1998 o vysokých školách, Studijním a zkušebním řádem VUT v Brně a Směrnicí děkana pro realizaci bakalářských a magisterských studijních programů zadává diplomovou práci s názvem:

Návrh změn informačního systému firmy

v anglickém jazyce:

Proposal for Changes in the Company Information System

Pokyny pro vypracování:

Úvod
Cíle práce, metody a postupy zpracování
Teoretická východiska práce
Analýza problému
Vlastní návrhy řešení
Závěr
Seznam použité literatury

Seznam odborné literatury:

- BURLESON, K. Oracle Tuning, 5. vydání. North Carolina, USA : Rampant TechPress, 2009. ISBN 0-9744486-2-1.
- INMON, W. Building the Data Warehouse, 4. vydání. Indianapolis, USA : Wiley Publishing, Inc., 2005. ISBN 13-978-0-7645-9944-6.
- JENSEN, CH. Multidimensional Databases and Data Warehousing, Denmark : Morgan&Claypool Publishers, 2010. ISBN 97-816-0-8455-386.
- KIMBALL, R. Mastering Data Warehousing Aggregates, Indianapolis, USA : Wiley Publishing, Inc., 2006. ISBN 13-978-0-471-77709-0.
- ORACLE Corporation. SQL Plan Management in Oracle Database 11g, [online]. 2010 . Dostupné z www: <<http://www.oracle.com/technetwork/database/features/bi-datawarehousing/twp-sql-plan-management-11gr2-133099.pdf>>.

Vedoucí diplomové práce: Ing. Petr Dydowicz, Ph.D.

Termín odevzdání diplomové práce je stanoven časovým plánem akademického roku 2010/2011.

L.S.

PhDr. Martina Rašticová, Ph.D.
Ředitel ústavu

doc. RNDr. Anna Putnová, Ph.D., MBA
Děkan fakulty

V Brně, dne 28.03.2011

Abstrakt

Cílem diplomové práce je návrh změn datového skladu ve společnosti. Předmětem této práce bude nejenom analýza, implementace a uplatnění datového skladu, ale také využití pokročilých optimalizačních metod pro co nejvyšší dosažení přidané hodnoty pro uživatele. První část práce je zaměřena na výběr vhodného konceptu datového skladu, jeho implementaci a možné problémy. Součástí druhé části práce poté porovnáím přínos tohoto systému ve firmě s nezbytnými náklady na jeho vybudování, vývoj a správu.

Abstract

The main aim of this diploma thesis is to propose for changes in the company information system. Document content is focus on analyse, implementation and using data warehouse and usage advance to optimaze methods for achievement added value to users. The first part of this thesis is focused on choice of suitable concept of data warehouse, implementation and possible problems. In the second part, I compare assets of this system with costs for development and administration.

Klíčová slova

Datový sklad, dimenzionální model, multidimenzionální model, faktové tabulky, dimenzionální tabulky, partitioning, indexizace, hint, granularita, Oracle

Key words

Data warehouse, dimensional model, multidimensional model, fact tables, dimensional tables, partitioning, index, hing, granularity, Oracle

Bibliografická citace VŠKP

KECLÍK, D. *Návrh změn informačního systému firmy*. Brno: Vysoké učení technické v Brně, Fakulta podnikatelská, 2011. 91 s. Vedoucí diplomové práce Ing. Petr Dydowicz, Ph.D.

Čestné prohlášení

Prohlašuji, že předložená diplomová práce je původní a zpracoval jsem ji samostatně. Prohlašuji, že citace použitých pramenů je úplná, že jsem ve své práci neporušil autorská práva (ve smyslu Zákona č. 121/2000 Sb., o právu autorském a o právech souvisejících s právem autorským).

V Brně dne 28. května 2011

.....

Poděkování

Na tomto místě bych velmi rád poděkoval všem, kteří mi při psaní této diplomové práce byli nápomocni. Zejména Ing. Petru Dydowiczovi, Ph.D za jeho odborné vedení, konzultace a čas, který věnoval mé diplomové práci.

Dále bych rád poděkoval společnosti HCI, a.s., za poskytnutí materiálů, znalostí a velkého prostoru pro seberealizaci v této firmě. Velké poděkování také patří mé manželce a dvěma synům, kteří mě podporovali po celou dobu délky mého studia.

Obsah

Úvod	11
1. Vymezení problému a cíle práce.....	12
2. Teoretická východiska práce	14
2.1 Vývoj DWH a Business Intelligence	14
2.2 Historie multidimenzionálních databází	16
2.3 Datové sklady	17
2.4 Rozdíly mezi datovým skladem a provozním systémem.....	19
2.5 Technologie datových skladů	19
2.5.1 Principy činnosti DWH.....	20
2.6 Charakteristika datového skladu	22
2.7 Charakteristika datového skladu dle Williama Inmona	22
2.7.1 Granularita	23
2.7.2 Hlavní zásady Inmona	24
2.7.3 ODS	26
2.7.4 Datamart – datové tržiště	27
2.7.5 Data mining pro pokročilé analýzy dat.....	27
2.8. Charakteristika datového skladu dle Ralpa Kimballa.....	28
2.8.1 Provozní systémy	30
2.8.2 Data Staging Area.....	30
2.8.3 Data Presentation Area	31
2.8.4 Data Access Tools	31
2.9. Kimball vs Inmon	31
2.9.1 Jednotlivé přístupy	33
2.9.2 Struktura řešení implementace.....	34
3. Analýza problému a současné situace.....	35
3.1 Představení firmy	35
3.1.1 Základní údaje o firmě	35
3.2. Analýza současného stavu IS ve společnosti.....	38
3.2.1 Jak vznikl DWH / BI systém v HCI.....	39
3.2.2 Častá nepochopení při komunikaci se zákazníky	40
3.3.3 Co je cílem pro zákazníky	41
3.4 DWH/BI systém v HCI: současný stav.....	41
3.5 HOS 8 analýza současného DWH	42
3.5.1 Výsledek analýzy HOS 8.....	43

3.6 Porterův model pěti sil pro oblast IS	44
3.7 SWOT analýza	45
3.8 SMART analýza	47
3.9 Vyhodnocení analýzy	48
4. Vlastní návrhy řešení.....	49
4.1 Možné varianty řešení.....	49
4.1.1 Vytvoření systému na klíč	49
4.1.2 Projekt na zelené louce	51
4.1.3 CPM	52
4.1.4 Rozšíření stávajícího systému.....	54
4.2 Výsledné řešení pro HCI.....	54
4.3 Postup řešení	55
4.3.1 Inepční fáze	55
4.3.2 Elaborační fáze	56
4.3.3 Konstrukční fáze	57
4.3.4 Přechodová fáze	58
4.4 Datový model.....	58
4.4.1 Vzorový dimenzionální model.....	59
4.5. Optimalizační metody	61
4.5.1 Indexování	61
4.5.2 Partitioning tabulek.....	63
4.5.3 Hints.....	64
4.6 Ekonomické zhodnocení projektu	66
4.6.1 TCO	66
4.6.2 Náklady spojené se změnou systému ve společnosti HCI.....	68
4.6.3 Reportovací nástroj	68
4.6.4 Pořizovací náklady.....	70
4.6.5 Náklady na provoz a vývoj systému	70
4.6.6 Výsledná kalkulace nákladů	71
4.7 Přínos nového řešení	72
4.7.1 Kvalita dat.....	72
4.7.2 Přínos uživatelům	73
4.7.3 Podpora primárního systému	73
4.7.4 SOLUS.....	74
4.7.5 Marketingové kampaně.....	75

4.7.6 Telemarketing	75
4.8 Vyhodnocení prospěšnosti systému.....	76
5. Závěr	78
Seznam použité literatury	79
Knihy:	79
Články v časopisu:	80
Elektronické zdroje:	80
Seznam tabulek, obrázků a grafů.....	82
Přílohy	83

Úvod

Za posledních přibližně dvacet let došlo ke změně struktury výrobních faktorů z původních tří na čtyři. K původním faktorům práce, půda, kapitál přibyl čtvrtý – informace.¹ Díky vývoji informačních technologií lze informace nejen sbírat, ale i je uchovávat a následně je využívat jako podpora podnikových rozhodnutí. Informace se staly pro firmy klíčovým nástrojem jak dosáhnout konkurenční výhody ať již formou snižování nákladů nebo dosažením lepšího strategického managementu. Zpracovávání informací a jejich důležitost pro strategické rozhodování se tak staly v průběhu času samostatným úzce specializovaným IT řešením. Již na počátku 90. let tak vznikl pojem Data Warehouse (Datový sklad) jehož autorství je připisováno Williamu Inmonovi, který byl průkopníkem v této oblasti. William Bill Inmon definoval datový sklad jako „podnikově strukturovaný depozitář subjektivě orientovaných, integrovaných, časově proměnlivých, historických dat použitých na získávání informací a podporu rozhodování“. V datovém skladu jsou uložena detailní (atomická) i sumární data.² Srozumitelněji řečeno je datový sklad systém, ve kterém jsou komplexní data uložena ve struktuře, která umožňuje analýzu a efektivní dotazování. Je tedy nejen fyzickým úložištěm dat, ale i metodologií, která zajišťuje zpracování dat tak, aby byla přesná, konzistentní, účelně agregovaná, jednoduše dosažitelná a jednoznačně interpretovatelná. Datové sklady patří k nejkompexnějším systémům, jakými společnosti disponují. Jsou napojeny na klíčové produkční systémy a integrují v sobě sofistikované procesy extrakce, transformace, čištění, uložení, agregace, analýzy a distribuce dat. Na konci tohoto složitého řetězce procesů jsou data skladu využívána pro podporu operativních i strategických činností a rozhodování. Výsledkem řízení těchto dat lze dosáhnout detailního pohledu na procesy nebo jejich částí uvnitř firmy. Zdroje mohou pocházet z různých interních systémů a nejčastěji se pro ukládání a indexaci využívají relační databáze Oracle, Sybase, Informix nebo MS SQL Server.

¹ VOŘÍŠEK J. *Strategické řízení informačního systému a systémová integrace*. 1. Vydání. Praha: Management Press, 2002. 12 s. ISBN 80-85943-40-9

² INMON, W. *Building the Data Warehouse*, 4th ed. Indiana: Wiley Publishing, Inc., 2005. 495 s. ISBN 13-978-0-7645-9944-6

1. Vymezení problému a cíle práce

V této diplomové práci se nadále budu zabývat datovým skladem, který je implementován nad relační databází Oracle, kterou považuji za nejvhodnější variantu pro zpracovávání velkých objemů dat, v řádech desítek a stovek Terabyte. Cílem této práce bude dokázat a vylepšit přidanou hodnotu stávajícího datového skladu ve firmě, ve které pracuji na pozici Business Intelligence and DataWarehouse Specialist. K dosažení tohoto cíle budu využívat nejen nejnovějších optimalizačních metod, které databáze Oracle nabízí, ale také moderní metody při tvorbě datového skladu s důrazným ohledem na architekturu, tedy tvorby dimenzionální vrstvy, dimenzních a faktových tabulek, datamartů atd. dle Ralpa Kimballa. Ralph Kimball spolu s Williamem Inmonem jsou v současné době jedinými lidmi, které lze nazvat guru v oblasti Data Warehouse tedy datových skladů (dále jen již DWH). Každý z nich používá odlišný postup a pohled na implementaci DWH a každé řešení má své pro i proti, jež detailně popíši v kapitole 2.7 a 2.8.

Ruku v ruce s tvorbou DWH se musí nutně počítat i s BI neboli Business Intelligence nástroji. Jedná se hlavně o reportovací nástroje sloužící k výstupu výsledných reportů, grafů, statistik a jiných napočítaných dat, které chce uživatel datového skladu publikovat. Na trhu je široká škála těchto programů od velmi sofistikovaných Business Object či Oracle Business Intelligence až po využití Office balíku např. Excelu. Nákladové kalkulace na pořízení těchto nástrojů se také pohybují v dosti rozdílných cenových stupnicích, a proto je nanejvýš nutné analyzovat, co uživatel skutečně bude vyžadovat od tohoto nástroje, jaké funkce bude nezbytně využívat a např. jaké uživatelské prostředí mu bude nejvíce vyhovovat. Z celkového hlediska lze přínos DWH / BI systému definovat jako poskytování přidané hodnoty na všech úrovních rozhodovacích procesů organizace. Relativní náklady lze docela přesně změřit dle metody TCO. Mezi další analytické metody použité v této práci lze jmenovat HOS 8, která nabízí ucelený pohled na informační systém v podniku, dále Porterův model zaměřený s ohledem na informační systém. Využiji také SWOT analýzy k odhalení nejzávažnějších nedostatků, ale také k objevení silných stránek DWH a využiji metodu SMART k definování cíle navrženého řešení. CPM analýzu budu aplikovat z pohledu zavádění a implementace nového informačního systému. Zájem o datové sklady v posledních letech neustále roste. Více než 60 % firem předpokládá, že datový sklad

bude využívat většina jejich zaměstnanců a zákazníků, téměř 50 % předpokládá jeho zpřístupnění svým obchodním partnerům a dodavatelům a přes 20 % dokonce uvažuje o veřejném zpřístupnění prostřednictvím Internetu.³ Důvodem neustálého růstu trhu týkajícího se datových skladů a BI je jeho nezávislost na hospodářských cyklech, DWH je jediným prostředkem, který umožní nejen nalézt způsob optimalizace nákladů, ale i obchodní příležitosti. Dnes již tedy není otázkou, zda budovat nebo nebudovat datový sklad, ale jak ho zavést co nejrychleji.

³ JEGER, D. Datové sklady a business intelligence. *CIO Business World*. 2011. č.1, s.15

2. Teoretická východiska práce

2.1 Vývoj DWH a Business Intelligence

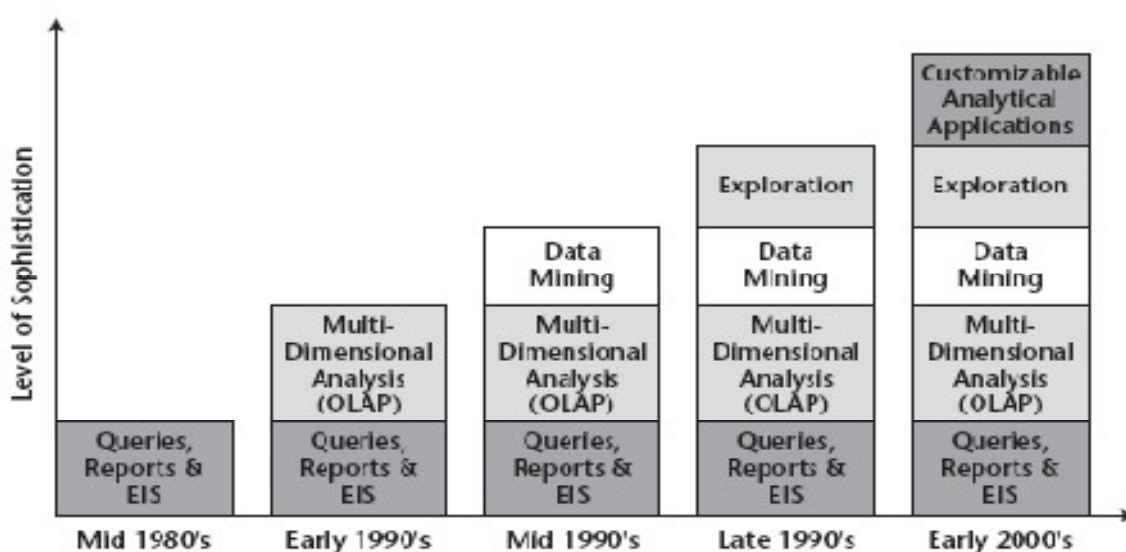
Pod označením business intelligence (BI) si lze představit zejména výkonné analytické, vykazovací, reportovací a jiné nástroje sloužící pro podporu rozhodování. Za pomoci firemních dat lze vytvářet analýzy, studovat jevy nastalé v minulosti, ale také v současnosti nebo predikovat či měnit možné jevy, které nastanou v budoucnosti. Business intelligence lze tak brát zejména jako nástroj, který nám umožní dosažení konkurenční výhody. BI se vylepšují a vyvíjí již více jak dvacet let.

Stejně jak nastává křivka trendu prodeje nového zboží, tak i společnosti přicházely postupně na chuť technologii BI nástrojů. Prvními průkopníky jsou tzv. inovátoři, kteří vyhledávají změnu i za určitého přijímaného rizika, následováni počátečními osvojiteli, kteří jsou součástí počáteční většiny usilující o co nejrychlejší osvojení novinek na trhu bez zbytečného rizika. Dále to je již pozdní většina, nedůvěřující, konzervativní lidé s cílem nakoupit výrobek, v tomto případě BI nástroj až po té, co je již nakupován většinou. V konečném stádiu to jsou tzv. opozdilci, kteří nakupují, až když je výrobek za vrcholem svého cyklu. Z pohledu vývoje BI je to podobné jak s běžným hmotným zbožím. Software má z pohledu tržního produktu své zvláštnosti oproti jiným běžným výrobkům. Představení nové technologie a její uvolnění na trh je většinou doprovázeno těžkostmi, jako je vyšší cena, nedokonalostmi jako např. nepřívětivé uživatelské prostředí a také spoustami aplikačních chyb. V průběhu času jsou chyby jednotlivými záplatami opraveny, zlepšuje se dostupnost produktu a nástroj tak lze používat ku prospěchu, ke kterému byl vyvinut. Pro představu lze použít dobrou komparaci u mobilních telefonů s novými operačními systémy, které jsou uvolňovány na trh s opravdu velkým množstvím chyb a to bez výjimky. A stejně tak jak byly mobilní přístroje v minulosti velké, těžké a drahé a v současnosti jsou již cenově dostupné a výkonné tak si i BI nástroje prošli podobným vývojem. Podobné to bylo i s vývojem datových skladů. Datovým skladům předcházely manažerské informační systémy (EIS), které měly za úkol snadno zpřístupnit výkonným představitelům klíčové informace k dosažení plánovaných cílů. EIS byly založeny na multi dimenzionálním zpracování a uložení dat. Hlavním nedostatkem EIS byla velmi enormní zátěž na systém pro poskytování požadovaných dat cílovým uživatelům, z důvodu nevhodné či ne k tomu

určené architektuře. Získávání informací neboli extrakce, transformace a jejich výsledná agregace (ETL) je sada komplexních aktivit, jejichž hlavním účelem je dosažení nejpřesnějších a integrovaných dat, kterých lze dosáhnout pouze pomocí datového skladu. Postupem času se vyvinulo velké množství nástrojů a technologií, které podporují datové sklady a většina procesů je automatizovaná.⁴

Jeden z nejvýznamnějších vývojových trendů během posledních 10 let byla široce přijatá architektura k tomu, aby podporovala všechny BI technologické požadavky. Tato architektura rozpoznala, že EIS přístup měl několik významnějších kazů, a to EIS datové struktury, které byly velmi často plněny přímo ze zdrojových systémů, což mělo za následek vysoké hardwarové požadavky a vysoké nároky na lidskou kapacitu z důvodu udržování a správy. V České republice se od počátku devadesátých let také začínají objevovat tyto produkty EIS, ale ruku v ruce vznikají řešení založená na datových skladech a datových tržištích (DMA – Data Mart).

Obr. 1: Vývoj BI technologií



Zdroj: Imhoff, Galemno, Geiger, 2004

To umožňuje zpracovávání velkého objemu dat a jeho dolování (DMI Data Mining) prostřednictvím BI nástrojů.

Součástí vývoje je vznik a používání různého názvosloví. Čím dál častěji IT manažeři operují s takovými pojmy jako reporting, data warehousing (DWH), data mining,

⁴ IMHOFF C., GALEMMO N., GEIGER J. *Mastering data warehouse design relational and dimensional technique*. 1st ed. Indiana: Wiley Publishing, Inc. 2003. 5 s. ISBN: 0-471-32421-3

Campaign Management, text mining, data integration, intelligence storage, customer intelligence, systémy na podporu rozhodování, manažerské aplikace, Executive Information Systems, datové sklady, datová tržiště, zkratkami EIS, MIS, HOLAP, ROLAP, MOLAP, DWH, DM, CRM a řadou jiných. Všechny výše uvedené nástroje, metodiky a řešení slouží témuž – poskytnout uživatelům kvalitnější informace, na jejichž základě budou moci efektivněji a lépe rozhodovat.⁵

2.2 Historie multidimenzionálních databází

Multi dimenzionální databáze nemají původ v databázové technologii, ale pocházejí z mnohorozměrného maticového počtu, které byly využívány pro ruční datovou analýzu v šedesátých letech devatenáctého století. Během těchto let dvě společnosti, IRI a Comshare nezávisle na sobě vyvíjely systém, který byl později označen jako multi dimenzionální databázový systém. IRI Express nástroj se stal velmi populární pro oblast marketingové analýzy koncem roku 1970 a počátku roku 1980. Tento nástroj, který se stal na trhu velmi brzy nejužívanějším, získala společnost Oracle. Souběžně společnost Comshare vyvinula System W, který se využíval zejména pro finanční plánování, analýzu a reportování během roku 1980. V roce 1991 byl za specifickým účelem vytvořen multi uživatelský a multi dimensionální databázový server, který vyústil v řešení Essbase systém, později licencovaný IBM pro spojení do DB2. Arbor a Codd pak v roce 1993 vytvořili OLAP. Další významný vývoj přišel v roce 1990 s příchodem datových skladů, které byly typicky založeny na schématu STAR (neboli hvězda) nebo snowflake (sněhová vločka). V roce 1998 poprvé Microsoft vytvořil MS OLAP Server, který byl první multi dimenzionální systém zaměřený pro trh. To vedlo k vývoji dalších multi dimenzionálních systémů, které byly schopny ukládat data o produktech a zboží do relačních databázových systémů pro analýzu dat.⁶

⁵ KHUDHUR P. Business intelligence: Je třeba přemýšlet. *ComputerWorld*. 2007, č.10, s. 28

⁶ JENSEN S., TORBEN P., THOMSEN C. *Muldtidimensional database and data warehousing*. 1st ed. Denmark: Morgan&Claypool publishers, 2010. 19 s. ISBN: 9781608455379

2.3 Datové sklady

Zkratka pro anglický termín Data Warehouse - datový sklad není ani informačním systémem, ani databází, i když jak s databází, tak s informačním systémem určitým způsobem spolupracuje.

Datové sklady lze nazvat úložištěm velkého množství dat, které byly vyprodukovány za použití určitého informačního systému, programu či algoritmu. Jsou navrženy tak, aby poskytovaly uživatelům ucelená data o klientech, produktech, výrobě, prodejkách, financích a mnohé další informace. Jsou datovou základnou pro detailní analýzu dat. S těmito daty je poté nutno nadále nějakým způsobem pracovat – hledat mezi nimi určité vazby, analyzovat je, získávat z nich skryté údaje a ty posléze zpracovávat. Činnost, při které se získávají dosud neznámá a ukrytá data se nazývá Data mining. Jedná se o import dat, jejich transformaci do struktur datového skladu, přípravu dat pro reporting a analýzu a v neposlední řadě i nástroje zpřístupňující data uživatelům. Datový sklad lze v celém procesu dolování dat nazvat jedním z důležitých a ideálních vstupních zdrojů, ze kterých se získávají data.

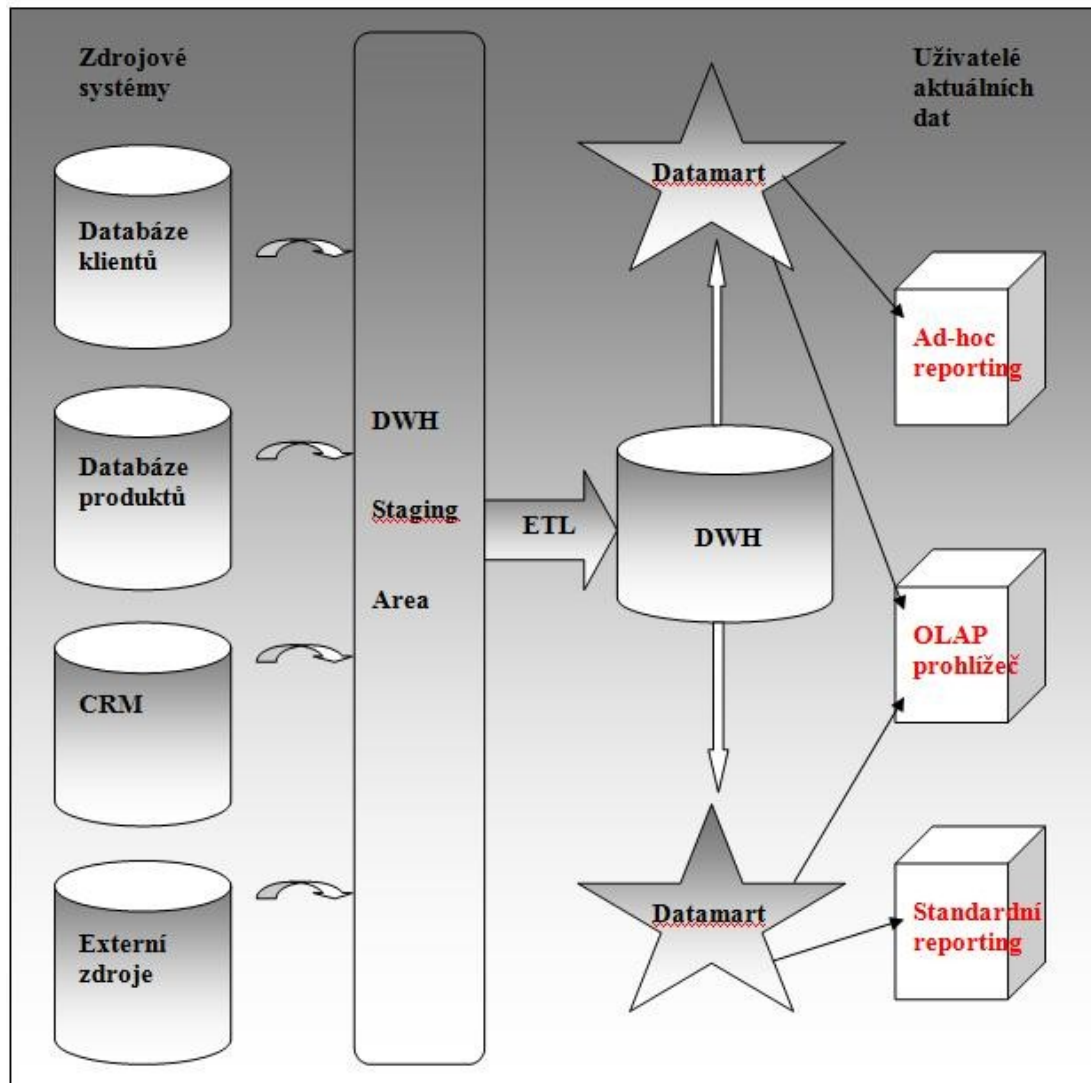
Datové sklady jsou základním stavebním kamenem řešení Business Intelligence (BI). Poskytují jednotlivým komponentám BI potřebná data, připravená v takové podobě, v jaké je tyto nástroje pro své funkce vyžadují. Hlavním impulzem pro vybudování datového skladu bývá skutečnost, že data, která organizace shromažďuje, se uchovávají roztržštěně v databázích mnoha provozních systémů, v různých datových strukturách a v různé kvalitě. Následné získání uceleného pohledu na jakoukoli oblast (např. na klienty) je složité, ne-li přímo nemožné. Posláním datového skladu je tato roztržštěná data integrovat na jednom místě, a to data konsolidovaná (sloučení stejných informací z více zdrojů do jedné, sjednocení číselníkových hodnot, odstranění duplicit atd.) a vyčištěná (např. nahrazení nepřesných/nekvalitních informací informacemi z ověřených zdrojů, doplnění chybějících hodnot, příp. označení informací, které nejsou korektní). Dobře navržený a plně implementovaný datový sklad funguje tak, že pravidelně přebírá a zásobuje se daty z jednoho nebo více provozních informačních systémů, eventuálně jiných zdrojů. Tato data potom ukládá v určité formě přímo do své primární databáze. Nepřebírají se tedy všechna data, ale jen taková, která souvisí s oblastmi, které budou předmětem dalšího zkoumání. V této primární databázi datového skladu jsou data

uložena tak, aby poskytovala určitý obraz vybrané části provozního systému, a uchovávají se zde včetně historie.

V datových skladech se kromě dat aktuálních uchovávají také data historická (dochází k tzv. historizaci dat), což nebývá běžnou praxí u provozních systémů.

Z pohledu hierarchie je datový sklad podřízen produktům, které vznikají v oblasti dolování dat – tyto produkty datový sklad využívají.

Obr. 2: Schéma datového skladu



Zdroj: Vlastní tvorba

Na druhou stranu databáze, ve kterých jsou data původně uložena, jsou datovým skladům nadřazena – datový sklad uskladňuje pouze část z nich.⁷

⁷ZDRAŽILOVA, I. *Datový sklad* [online]. 2011 [cit. 2011 – 01-28] Dostupné z [www: <http://kisk.phil.muni.cz/wiki/Datov%C3%BD_sklad>](http://kisk.phil.muni.cz/wiki/Datov%C3%BD_sklad)

Na obrázku č. 2 jsem vyobrazil základní schéma datového skladu. Jedná se o vyjádření principu fungování DWH a jeho mechanismu činnosti. Zdrojové systémy zde znázorňují provozní systémy, které běží nad podnikovými aplikacemi a kde dochází k přírůstkům nebo modifikacím jednotlivých dat. Tyto datové manipulace jsou zaznamenány a intervalově přenášeny do Data Staging Area (více o DSA píší v kapitole 2.5.1 Principy činnosti DWH). V této oblasti se data konsolidují a čistí. Následně v Extract Transform Loadu (rozvedeno taktéž v kapitole 2.1.1 Principy činnosti DWH) se data agregují a transformují do výsledné podoby pro potřeby analytiků. Data jsou tedy již připravena pro prohlížení, analýzu či reportování.

2.4 Rozdíly mezi datovým skladem a provozním systémem

Provozní systémy se budují s důrazem na rychlé ukládání a aktualizaci dat, nikoli pro jejich následné poskytování uživatelům. Jedná se tedy zejména o transakční způsob ukládání dat a jejich modifikaci. Jakýkoli reporting nad těmito daty bývá velmi pomalý, příp. neúnosně zatěžuje systém a tím snižuje jeho schopnost plnit základní funkce, pro které byl navržen.⁸ Provozní systém není navržen a ani s tím datový model či hardwarové a softwarové prostředky nepočítají, aby bylo možno dosáhnout efektivního reportingu v reálném čase. Oproti tomu datové sklady jsou navrženy a optimalizovány pro co možná nejrychlejší vyhledávání a poskytování odpovědí na dotazy. Je důležité, ale také nezaměňovat datový sklad s on-line systémem. Neboť transformace, čištění dat a ostatní úlohy, které datový sklad plní, vyžaduje určitý strojový čas.

2.5 Technologie datových skladů

Datový sklad byl stvořen zejména k rozhodování podnikového managementu a minimalizaci rizika špatných rozhodnutí. Z pohledu managementu je nanejvýš důležité rychle vyhodnocovat velké množství nesourodých informací z mnoha zdrojů.

Datový sklad představuje ucelené řešení poskytující nejen prostředky pro ukládání dat, ale rovněž sadu nástrojů pro jejich analýzu. Jeho stěžejním úkolem je podpora plánování a řízení firmy, a to nejen na strategické úrovni.

⁸ BRODNÍČEK D. *Úvod do problematiky datových skladů*. Automatizace, 2009, č. 7, s. 428

DWH je postaven na relační databázi. Její vlastnosti a fungování ovlivňují chod celého řešení. Relační databáze hrají v datových skladech klíčovou roli, podstatně důležitější než tomu je například v ERP systémech. OLTP databáze jsou optimalizovány pro zpracování velkého množství malých transakcí a uchovávají pouze aktuální data, kdežto databáze v datových skladech analyzují historická data a jsou neustále rozšiřovány bez jakékoli redukce jejího obsahu. Zvyšuje se důraz na stabilitu výkonu, administraci a škálovatelnost databáze. Značná část funkcionality multi dimenzionálních datových struktur OLAP aplikací i aplikací dolování dat se přesouvá do relační databáze. Další vlastností, kterou by databáze měla mít, je otevřenost neboli schopnost spolupracovat s komponentami jiných výrobců v jakékoli oblasti řešení datových skladů.⁹

2.5.1 Principy činnosti DWH

Pojmem DWH označujeme systém, který plní tyto základní úlohy:

1. Realizuje ETL procesy:
 - a. Extrakce dat z primární datových zdrojů: databází provozních (primárních) systémů, externích databází, textových souborů, XML souborů, provozních logů, průmyslových čidel, atd.
 - b. Transformace extrahovaných dat: výpočty odvozených hodnot, čištění a konsolidace dat, datový audit, nahrazení přirozených klíčů umělými, atd.
 - c. Nahrání (Load) dat do prezentační vrstvy datového skladu.

Vytváření ETL je klíčové pro celkový úspěch celého projektu. Při budování DWH je věnováno 50% času právě ETL.¹⁰

Z primárních systémů je potřeba vyextrahovat potřebná data. Transakční systém obsahuje spoustu dat, která nemusí být nutně zajímavá pro analýzu, jako například adresa bankovního terminálu apod. Extrakční procesy vybírají z datového zdroje pouze ta data, kterých je pro další zpracování třeba. Extrakce také zahrnuje načítání dat z nejrůznějších zdrojů, nejen z databází. Například z textových souborů, XML dokumentů apod.

⁹ HABÁŇ J. *Technologie datových skladů*. [online]. 2004 [cit.2011-01-30] Dostupné z [www: <http://www.cvis.cz/hlavni.php?stranka=novinky/clanek.php&id=60>](http://www.cvis.cz/hlavni.php?stranka=novinky/clanek.php&id=60)

¹⁰ WREMBELL R., KONCILIA CH. *Data warehouse and olap concepts architecture and solutions*. 1st ed. LONDON: IRM Press, 2007. 22 s. ISBN 1-59904-364-5

Během ETL procesů se transformují data z primárních systémů na ukazatele (metriky), uložené v tabulkách faktů, a na atributy popisující kategorie, podle nichž se ukazatele analyzují, uložené v tabulkách dimenzí. Prvky dimenzí vytvářejí hierarchie (například den - týden - rok, prodejna - obec/město- region), které zachycují obchodní hlediska používaná při analýze ukazatelů. Vedle obecných dimenzí, jako je čas, se vytváří mj. dimenze zákazníků, dimenze dodavatelů apod., v nichž jsou kromě analyticky významných hierarchií uloženy i další popisné atributy jako například adresa sídla nebo trvalého pobytu, věk, počet zaměstnanců apod.¹¹ Pro co nejjednodušší vysvětlení lze charakterizovat faktovou tabulku jako entitu, která si uchovává numerické hodnoty o sledované veličině (např. zůstatek účtu) a odkazy na záznamy v dimenzionálních tabulkách. Údaje v dimenzionálních tabulkách představují kontext, ve kterém jsou fakta prezentována a zkoumána (např. typ účtu). Matematicky řečeno, fakta jsou funkcemi dimenzí.

Pro účely extrakce a transformace se využívá tzv. Data staging area (DSA), což je pracovní datová oblast, do které nemají koncoví uživatelé přístup – DSA je přístupná pouze pro ETL procesy.

2. Poskytuje uživatelům přístup k datům prostřednictvím tzv. prezentační vrstvy DWH, ve které jsou data uložena ve vhodné publikační formě pro podporu rozhodování s důrazem na jednotnost, srozumitelnost a přístupnost z pohledu koncových uživatelů.

K realizaci prezentační vrstvy je typicky využívána relační databáze.

DSA lze také realizovat v relační databázi (nejčastější způsob), nebo lze využít služeb souborového systému (textové soubory), nebo proprietární formáty specializovaných ETL nástrojů, popřípadě se lze bez DSA zcela obejít (zpracování dat v paměti). DSA lze označit jako prostor mezi provozními systémy a data presentation area (více o ní v kap. 2.8.2). Uvnitř DSA dochází k ETL procesům a umožňují tak společnosti získávat

¹¹ VAVRUŠKA J. *ETL a kvalita dat.* [online]. 2003 cit [2011-01-31]. Dotupné z [www: <http://www.systemonline.cz/clanky/etl-a-kvalita-dat.htm>](http://www.systemonline.cz/clanky/etl-a-kvalita-dat.htm)

data z různých zdrojů a následně je modifikovat, upravit nebo očistit a přenést do jiného úložiště.

2.6 Charakteristika datového skladu

V předchozích kapitolách jsem se věnoval problematice datových skladů z pohledu historického, technologického a funkčního. V teoretické části této práce, ale považuji za klíčové popsat právě charakteristiku a její popis tohoto systému a to z pohledu dvou nejznámějších osobností v tomto oboru, Williama H. Inmona a Ralpa Kimballa. Právě výběr jednoho z přístupů nebo kombinací obojího je z velké části závislé na výsledku konečného projektu z pohledu konečného uživatele. Oba dva přístupy jsou natolik rozdílné, že považuji za důležité se nyní jednotlivým přístupům věnovat podrobněji a podrobit je jejich vzájemné komparaci.

2.7 Charakteristika datového skladu dle Williama Inmona

William Inmon je označován za otce datových skladů. Je autorem více jak čtyřiceti knížek a 600 článků zabývajících se dolováním dat, datovými sklady, designem návrhů a managementem v oblasti datového zpracování. Jeho publikace byly přeloženy do osmi jazyků a je také vlastníkem dvou softwarových patentů.¹²

Dle Inmona lze charakterizovat datový sklad jako: „předmětně orientovaná, integrovaná, časově proměnná, nevolatilní kolekce dat umožňující rozsáhlé analýzy“.¹³ Jeho koncepce dělí DWH do tří datových vrstev.

- Staging Area což je jak jsem uváděl výše základní vrstva datového skladu, neboli základní vrstva pro dočasné uložení a zpracovávání dat extrahovaných z provozních systémů
- Centrální úložiště dat, tedy normalizovaná relační databáze, ve které se uchovávají konsolidovaná, vyčištěná a historizovaná data. Jedná se o hlavní vrstvu datového skladu.

¹² INMON W. *Bulding the data warehouse*. 4th ed. Indianapolis: Wiley Publishing, 2005. 9 s. ISBN-13: 978-0-7645-9944-6

¹³ LIBOR G., POUR J., TOMAN P. *Podniková informatika*. 1. vydání. Praha: Grada Publishing, 2006. 458 s. ISBN 80-247-1278-4

- Datová tržiště (Data Marts), které obsahují data transformovaná z centrálního úložiště do podoby vhodné pro zamýšlený účel využití, např. přípravu reportů, analýzy, data mining, využití v aplikacích Business Intelligence atd.

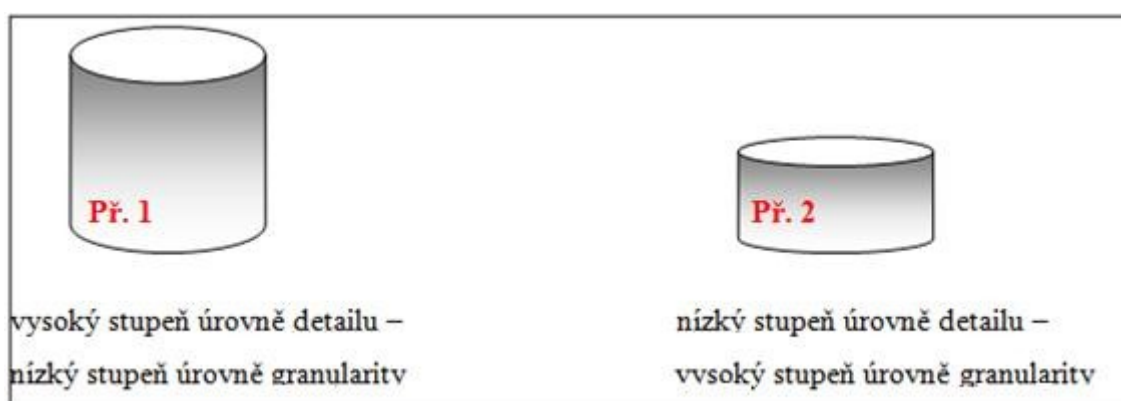
Mezi jednotlivými vrstvami se data postupně přenášejí, a to jednosměrně ze základní vrstvy Staging Area přes centrální úložiště do datamartů. Další důležitou součástí datového skladu jsou tzv. metadata, která v podstatě datový sklad popisují a řídí jeho provoz.¹⁴

Základními myšlenkami Billa Inmona je vytvoření architektury, která by co možná nejvíce minimalizovala redundantní data, a zároveň minimalizovala interface mezi produkčními systémy a datovým skladem. Díky tomu vznikl název Centrální datový sklad, z něhož vychází definice z počátku této kapitoly.

2.7.1 Granularita

Granularita lze označit jako úroveň detailu či souhrn dat v datových skladech. Čím více je detailu, tím méně je granularity a naopak čím méně je detailu, tím více je granularity. Inmon je zastáncem co největšího obsahu datového skladu. DWH by tedy měl být co nejvíce granulární (zrnitý.) Nelze předvídat všechny možné otázky, které budou uživatelé pokládat a hledat na ně odpovědi v prezentační vrstvě DWH. Z tohoto důvodu je vhodné a potřebné, aby datový sklad obsahoval atomická data na nejnižší dostupné úrovni granularity (společně s daty sumarizovanými pro podporu nejčastějších dotazů).

Obr. 3: Granularita – úroveň detailu



Zdroj: vlastní tvorba

¹⁴ BRODNÍČEK D. *Dolování dat z procesních databází*. Automatizace, 2009, č. 7, s. 429

Z obrázku č. 3 u objektu Př. 1 si lze představit úroveň detailu například ve formě každého detailu telefonního hovoru zákazníkovi za měsíc. U Př. 2 lze uvést naopak součet všech telefonních hovorů zákazníkovi za měsíc.

Granulární data v datovém skladu jsou klíčem k mnoha využitím informací, neboť většinou uživatelé data využívají vícero způsoby.

Například, uvnitř společnosti, jsou stejná data využívána pro uspokojení potřeb marketingu, obchodu, a účetnictví. Všechny tři oddělení se dívají na stejná základní data. Marketing může chtít vidět prodeje na měsíčním základu dle zeměpisné oblasti, oddělení obchodu může chtít vidět jen objem prodeje prodejce na týdenním základu, a finance mohou chtít vidět rozpoznatelný příjem na čtvrtletním základu dle sortimentu. Tyto druhy informací jsou si blízké příbuzné a přesto mírně odlišné. Díky datovému skladu se tak organizace může dívat na data tak, jak zrovna potřebuje.

2.7.2 Hlavní zásady Inmona

Mezi hlavní zásady Williama Inmona patří

- Subjektová orientace – datový sklad obsahuje data související s předmětem podnikání, nikoli pouhé zápisy z jednotlivých transakcí v provozním systému. Obecně jsou v provozních systémech data shromažďována okolo aplikací dané společnosti. Jako příklad Inmon uvádí pojišťovací firmu, která má aplikace jako auto, havárie apod. Nejdůležitějším atributem společnosti je, ale zákazník, bonusy, pojistné plnění atd.
- Integrovanost – toto je spojené s problémem mnoha firem, a sice že datový sklad zpracovává data z několika různých systémů a tak je potřeba data nejdříve naformátovat dle předem stanovených pravidel. Např. formát času, odlišné metrické soustavy. Znamená to, že např. všechny délky jsou převedeny na metr atd.
- Nízká proměnlivost – data se do datového skladu nahrávají k určitému okamžiku a nelze je na rozdíl od provozních systémů nikterak modifikovat. Data jsou uložena v databázi jako snímek k určitému času a pokud dojde ke změně na provozním systému, data budou opět jako

snímek přeneseny na DWH a tím automaticky vzniká historizace. Automatická refresh jednotlivých přenosů dat z provozní databáze do DWH je předem nastavena v metadatech.

- Historizace – jedna z nejsilnějších vlastností datového skladu je historizace dat. V provozním systému jsou data stále aktuální a stále měněna na rozdíl od DWH, kde jsou data uložena bez jakýchkoli modifikací a nabízejí tak pohled do historie všech událostí, které nastaly v provozním systému. Jako příklad lze uvést přechod jednotlivých stavů smlouvy od schváleného návrhu, aktivního stavu po stav ukončené smlouvy.

Koncepce, s níž přišel Bill Inmon (obsahující dočasné úložiště, centrální úložiště a datová tržiště) se označuje jako třívrstvá architektura datového skladu. Jedná se o nejjistší řešení, které ovšem vyžaduje vyšší počáteční náklady na analýzu a relativně dlouhou dobu na úplnou realizaci.

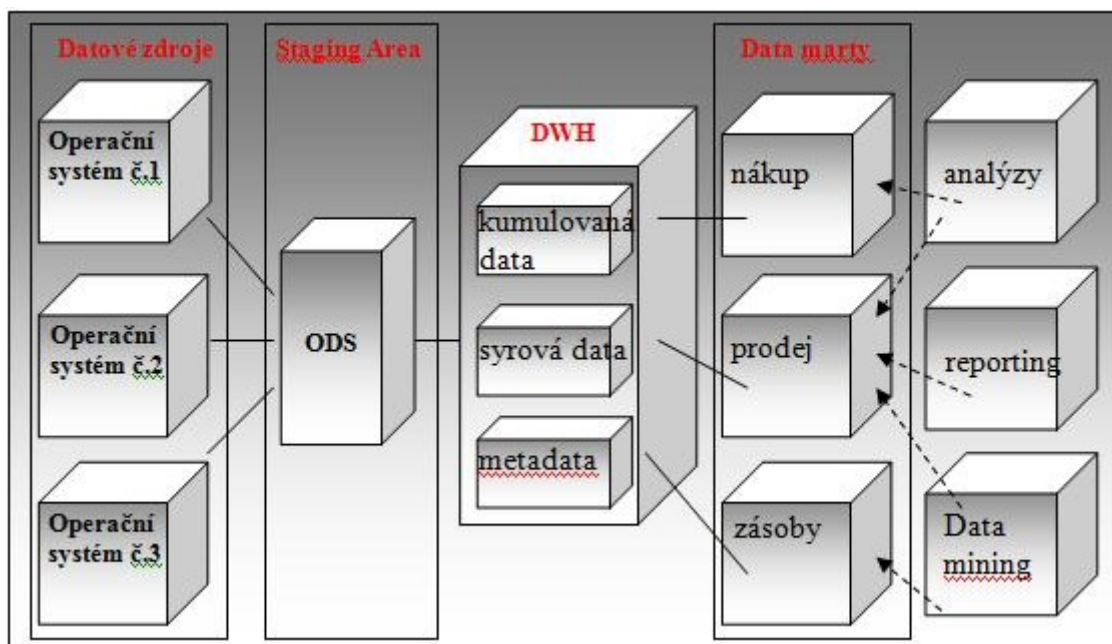
Celopodnikový datový sklad, se stává srdcem podnikové architektury pro podporu rozhodování. Nad tímto datovým skladem jsou pak budovány data marty, které slouží pro podporu rozhodovacích procesů jednotlivých útvarů podniku. Tímto se získá unikátní zdroj celopodnikových detailních dat.

Architektura celopodnikového datového skladu s sebou přináší:

- dobře monitorovatelné prostředí
- vytvoření jednoho místa pro zajištění kvality dat
- minimalizaci interface mezi produkčním a DSS prostředím (technologie systémů pro podporu rozhodování (Decision Support System))
- zajištění celopodnikového pohledu na data
- snížení nákladů na HW, SW
- snížení redundantních dat, atd.¹⁵

¹⁵ KUČERA M. *Dva způsoby budování datového skladu*. [online]. 2009 cit [2011-02-02]. Dotupné z [www: <http://www.systemonline.cz/clanky/dva-zpusoby-budovani-datoveho-skladu.htm>](http://www.systemonline.cz/clanky/dva-zpusoby-budovani-datoveho-skladu.htm)

Obr. 4 Integrovaný datový sklad dle Williama Inmona



Zdroj: Vlastní tvorba

V obrázku č. 4 se vyskytují pojmy, které si zaslouží bližší popis.

2.7.3 ODS

Operational Data Store je využívána pro taktické rozhodování, zatímco datový sklad podporuje strategická rozhodnutí. Obě dvě vrstvy mají podobné charakterové rysy, ale v některých hlediscích jsou velmi odlišné. ODS obsahuje aktuální, přesné, integrované data o zákaznících, produktech, inventáři, a tak dále. ODS je přístupná odkudkoli ve společnosti a dá se tak nazvat aplikací specifickou.¹⁶ Nejvýznamnější odlišností od datového skladu je rozdíl četnosti aktualizace dat v rozsahu od několika hodin po téměř dny. Datový sklad nemůže nabídnout natolik čerstvá data, neboť obecně se a jejich agregace do jednotlivých data martů provádí jednou denně a to většinou přes noc, kdy není systém zatížen uživateli. Využití ODS může být např. na call centru, kdy je potřeba u každého zákazníka znát aktuální profil, jeho aktivované či objednané produkty, zařazení do segmentu, poslaných marketingových nabídek apod.

¹⁶ IMHOFF C., GALEMMO N., GEIGER J. *Mastering data warehouse design relational and dimensional technique*. 1st ed. Indiana: Wiley Publishing, Inc., 2003. 14 s. ISBN: 0-471-32421-3

2.7.4 Datamart – datové tržiště

Další komponentu datových skladů tvoří datová tržiště (data marts). Jejich princip je obdobný jako v případě datových skladů. Rozdíl spočívá v tom, že jsou určeny pro pokrytí konkrétní problematiky daného okruhu uživatelů a jejich architektura je již většinou založená na tabulkách faktů (obsahující měřitelné údaje o prodejkách, nákupech, aktivacích...) a tabulkách dimenzí (obsahující číselníky, podle kterých se dá výsledek analyzovat, např. region, produkt, období), což podstatně usnadňuje datovou analýzu. Datová tržiště obsahují rovněž spoustu odvozených a agregovaných atributů často používaných v daném oddělení. Výsledkem jsou pak datová tržiště pro oddělení marketingu, řízení rizik, HR, obchodu, anebo také pro účely data miningu. Při plnění datového skladu je většina jeho komponent nepřístupná. Proto se tento proces spouští v době nejnižšího provozu (o víkendu, v nočních hodinách). Dalším důvodem je rovněž požadavek na minimální zatížení provozních systémů.

2.7.5 Data mining pro pokročilé analýzy dat

Složitější vzorky chování, které nelze prostřednictvím technologie OLAP detekovat, se řeší prostřednictvím technologií data miningu. Data mining se využívá zejména pro predikci (např. odhad budoucího chování klientů, tržeb, pohledávek), segmentaci (shlukování klientů s podobnými vlastnostmi) či pro popis dat. Mezi nejčastěji používané algoritmy predikce patří logistická či lineární regrese, metoda rozhodovacích stromů, neuronové sítě či v poslední době populární support vector machines. Ze shlukovacích algoritmů jsou to pak K-Means a EM clustering.

Výsledky z Data Miningu se zpětně propagují do datových tržišť, centrálního úložiště či přímo do provozních systému (např. CRM), odkud se dále využívají pro cílený marketing (prodej dalších produktů klientům s vysokou pravděpodobností nákupu), predikci odchodu zákazníků, credit scoring (rozhodnutí o schválení nebo zamítnutí žádosti o půjčku, případně před schvalování půjček), detekce pojistných podvodů apod. Data mining vyžaduje především rozsáhlou přípravu dat v podobě analytického datového tržiště a následné nasazení složitějších matematických algoritmů. Tyto zvýšené investice však mohou značně vylepšit nějaký business proces (např. cílení

kampaní) a tím zvýšit zisk celé firmy. Data Mining, OLAP a reporting, patří mezi analytické komponenty BI řešení.¹⁷

2.8. Charakteristika datového skladu dle Ralpa Kimballa

Pokud lze nazývat Williama Inmona otcem datových skladů, Ralpa Kimball může být bezesporu nazýván otcem Business intelligence.

Ralph Kimball je autorem dvouvrstvé architektury konceptu datového skladu. Na rozdíl od Williama Inmona, Kimball prosazuje implementaci datového skladu bez centrálního úložiště, což vede k velmi výrazné úspoře času, lepšímu pochopení datového modelu pro konečné uživatele a často k nižším počátečním nákladům. Jeho metodologie, také známá jako rozměrové modelování, nebo Kimballova metodologie se stala standardní v oblasti podpory rozhodování. Ralph Kimball je autorem mnoha publikovaných článků v renomovaném časopise *Intelligent Enterprise* v oblasti designu datových skladů a je také autorem dvou nejlépe prodávaných knih *Data Warehouse Lifecycle Toolkit* a *The Data Warehouse Toolkit*. Ačkoli Inmonovi ani Kimballovi knihy nikdy nebyly přeloženy do češtiny (což je možná dobře, vzhledem k úzké specializaci a možnému špatnému překladu), lze je zakoupit ve speciálních knihkupectví zaměřených na IT technologie. Ralph Kimball také jako první definoval koncept data martů a popsal využití dimenzionálního modelování včetně „star” a „snowflake” datových struktur.¹⁸ Definice datového skladu dle Ralpa Kimballa zní následovně: „DWH je kopie transakčních dat speciálně strukturovaných pro dotazování a reportování“.¹⁹ Na rozdíl od W. Inmona je definice daleko srozumitelnější a přijatelnější pro běžného člověka. "Data warehouse není nic jiného než sjednocení data martů"²⁰, je další z definic tohoto zakladatele dimenzionálního modelování. Cílem je tedy postupné budování data martů. Komplex těchto data martů poté lze nazývat datovým skladem.

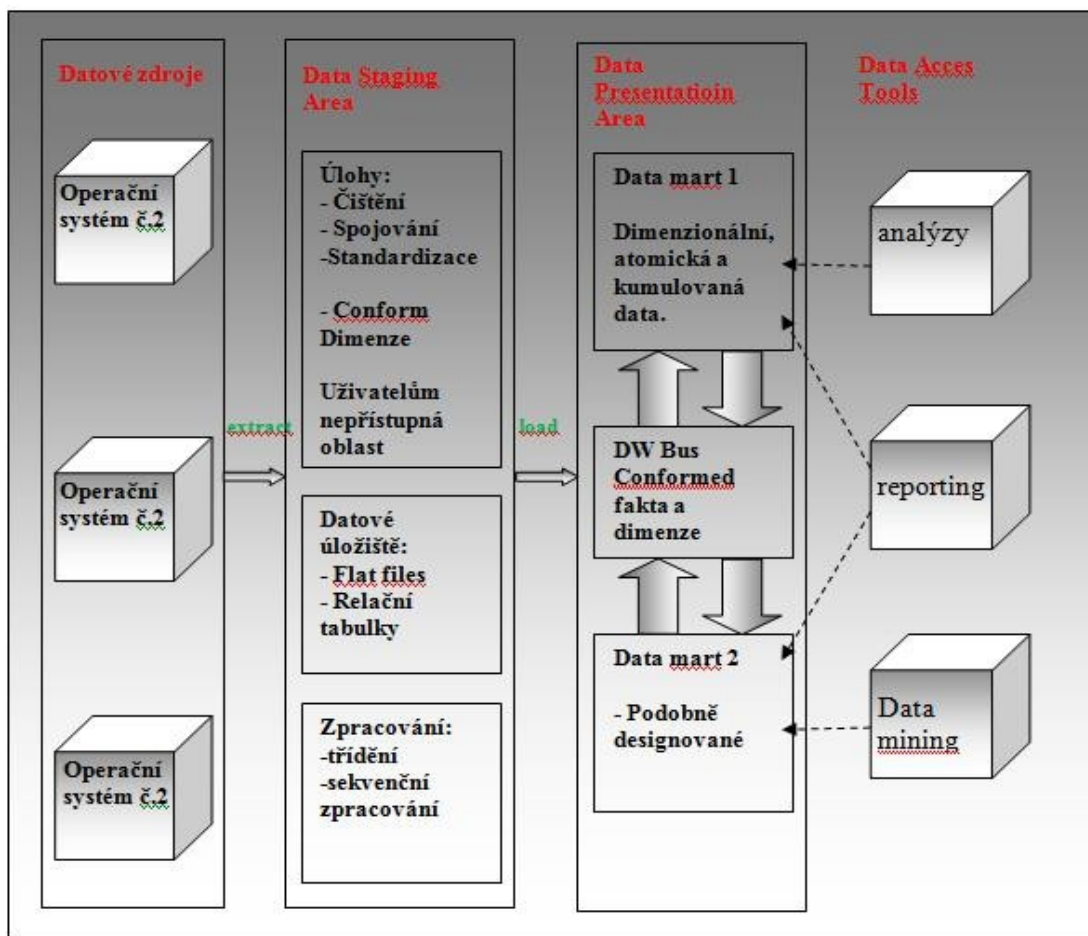
¹⁷ HANUSEK L. Technologie Data Warehousingu a Data Miningu. *Computerworld*, 2007, č. 8, s. 15

¹⁸ V českém jazyce se mluví o schématu hvězdy a vločky

¹⁹ GÁLA L., POUR J., TOMAN P. *Podniková informatika*. 1. vydání. Praha: Grada Publishing, 2006. 103 s. ISBN 80-247-1278-4

²⁰ KIMBALL, R. *The Kimball Group Reader*, 1st ed. Indianapolis: Wiley Publishing, Inc., 2010. 38 s. ISBN: 978-0-470-56310-6

Obr. 5: Integrovaný datový sklad dle Ralpa Kimblla



Zdroj: Kimball, *The Data Warehouse Toolkit*. str. 7. Vlastní úprava

Tento koncept implementace datových skladů se stal brzy velmi populární a to z důvodu, jež jsem poznamenal v odstavci výše:

- rychlá implementace data martu
- srozumitelný datový model
- nižší počáteční investice.

Z obrázku č. 5 je patrné, že Kimballův koncept má 4 odlišné oblasti. Jedná se o:

- Operational source systems (provozní systémy)
- Data Staging Area,
- Data Presentation Area
- Data Acces Tools

2.8.1 Provozní systémy

Provozní systémy zaznamenávají data transakčně. Tato operativní data se transformují do datového skladu typicky pomocí materialized view a jejich fast refresh nebo díky Oracle streaming²¹ z pohledu Oracle technologie. Provozní systémy ukládají data a modifikují obecně velmi rychle, ale z pohledu dávkově objemných operací jsou naprosto nevhodné, od toho tu právě slouží DWH. Velmi často se data do DWH dostávají z mnoha různých aplikací, které, ale nemají vzájemně synchronní data, což datovému skladu velmi ztěžuje situaci.

2.8.2 Data Staging Area

Data staging Area je oblast datového skladu, která plní roli obrazu dat provozních systémů. Zde se aktualizují veškeré databázové tabulky, nad kterými probíhají transakční operace. Zde se také vyskytují tzv. rozdílové tabulky běžně označené koncovkou DIFF, které se díky tzv. triggerům²² plní rozdílovými daty, které v DWH ještě nejsou zpracovány a teprve čekají na svou agregaci do dimenzních či faktových tabulek.

Jedná se čistě o optimalizační řešení, aby se dávkově nezpracovávaly celé objemy naplněných tabulek, ale pouze změnové záznamy. Data Staging Area je tedy pouze prostor mezi operačními systémy a Data Presentation Area, kterou vysvětlím v dalším odstavci. Pro jednodušší vysvětlení si lze představit např. oblast kuchyň a oblastí restaurací. V kuchyni jsou syrové potravinářské výrobky transformované do hotových jídel. V tomto případě jsou syrová data přeměna do forem, která jsou nevhodnější pro uživatelské dotazy. Podobně jak v kuchyni, kam mají přístup pouze zkušení kuchaři, tak i do této oblasti mají přístup pouze zkušení profesionálové, kteří tyto data zpracovávají, neboť v kuchyni by patně zákazníci napáchali spíše více škody nežli užitku a v Data Staging Area by to bylo podobné.

²¹ Zde se pro představu pouze plní funkce databázového systému, jehož technologie umožňuje aktualizovat data z provozního systému do DWH.

²² Opět se jedná o Oracle technologii, kdy tzv. trigger zavěšený nad zdrojovou tabulkou umožní odlišit změnových záznamů do pomocné tabulky sloužící k dalšímu zpracování. Vzhledem k charakteru této práce není potřeba dále tuto funkcionalitu rozvádět.

2.8.3 Data Presentation Area

V Data presentation area jsou data organizovaně uloženy a připravovány pro přímé dotazování uživatelů. Je to oblast, kde lze již generovat sestavy díky dalším analytickým aplikacím. Na rozdíl od předchozí Data staging area, je tato oblast již zpřístupněna uživatelům, v analogii na předchozí odstavec, lze již tedy mluvit o prostředí restaurace, kde zákazníci uspokojují své potřeby. Typicky lze hovořit o této oblasti jako o sérii integrovaných data martů. Ralph Kimball přikládá velký důraz této oblasti. Důležitým aspektem je, že data marty by měly obsahovat taková atomická data, které jsou na nejnižší úrovni detailu. Je to potřeba zejména z důvodu mnoha uživatelských dotazů, která jsou často nepředvídatelná. Díky agregovaným data martům, jsou tyto data v takové formě, která co nejvíce vyhovuje optimalizačním potřebám a tak zpřístupňuje data uživatelům v přijatelném časovém rámci. Nejdůležitějším aspektem je, že tyto data marty se skládají z jednotlivých faktových a dimenzních tabulek, která se nazývají conformed neboli všeobecně přijaté dimenze. Ralph Kimball upozorňuje, že data marty je možné sjednotit pouze za předpokladu tzv. "všeobecně přijatých" dimenzí a fakt. Pokud striktně nedodržíme "všeobecně přijaté" dimenze, pak není možné jednotlivé data marty spojit v celek - data warehouse.²³ Tyto sdílené dimenze a fakta jsou alfou a omegou pro samotný data mart a proto jejich integrace je klíčová z pohledu Ralpha Kimballa pro implementaci datových skladů.

2.8.4 Data Access Tools

Data Access Tools neboli nástroje sloužící pro přístup k datům jsou finální složkou v prostředí datového skladu. Tyto nástroje slouží pro konečné uživatele pro podporu rozhodování. Tyto nástroje jsou fakticky hlavním smyslem celého datového skladu, tedy dolování dat z datového skladu prostřednictvím těchto nástrojů jako Oracle Business Intelligence, Business Object či jen programového balíku Office.

2.9. Kimball vs Inmon

V případě datových skladů je datová architektura prezentační vrstvy zcela klíčovým a určujícím faktorem pro celkový úspěch systému (měřený jeho použitelností, správou,

²³ KIMBALL, R., *The data warehouse toolkit*, 2nd ed. USA: Wiley Computer Publishing 2002. 82 s. ISBN 0-471-20024-7

rozšiřitelností, robustností, výkonností, škálovatelností). Není to jistě nic překvapivého vzhledem k poslání DWH / BI systémů, kterým je publikování dat. V oblasti DWH / BI systémů existují dva hlavní přístupy k řešení datové architektury:

- Corporate Information Factory (CIF) dle Billa Inmona
- Data Warehouse Bus (DWHBus) dle Ralpa Kimballa

Stejně dlouho, jako existují tyto dva přístupy, existuje i nekonečná debata, který, z těchto přístupů je vhodnější. Vychází z představy, že se jedná o dvě soupeřící neslučitelné koncepce. Skutečnost je taková, že se jedná o přístupy komplementární, které mají mnoho společného.

Co mají společného:

- Na začátku musí být zevrubná analýza uživatelských požadavků všech potenciálních uživatelů a analýza všech zdrojových systémů
- Vytváření shora dolů – je nutno začít co nejkompletnějším logickým datovým modelem prezentační vrstvy, vytvořeným na základě analýzy požadavků a analýzy zdrojových systémů
- Inkrementální vytváření prezentační vrstvy i celého DWH / BI systému; rozsáhlý systém nelze vytvořit s přístupem „velký třesk“
- Forma a obsah prezentační vrstvy nejsou diktovány představiteli jednotlivých oddělení, ale vznikají s ohledem na celopodnikové zájmy; prezentační vrstva modeluje data (entity reálného světa a jejich vztahy) tak jak opravdu jsou, nikoli tak, jak je vidí uživatelé
- Prezentační vrstva musí obsahovat atomická data, aby mohla uspokojit všechny (předem neznámé) datové požadavky
- Samostatné nezávislé data nikdy neřeší problém na celopodnikové úrovni, jsou spíše zdrojem problémů a nesrovnalostí (více verzí pravdy, vícenásobné extrakce, neefektivní využití zdrojů)
- Data nikdy nepoužívají dimenzionální model

I když jsem zde napsal výčet toho, co mají spolu tyto dva autoři společného, v mnoha věcech se zásadně liší. Jejich přístupy jsou naprosto odlišné již z podstaty jejich filosofie, designu provedení a implementační strategie.

A v čem se liší:

- Logický model presentační vrstvy
 - o Inmon: Entitně relační návrh – ER model, tabulky 3NF²⁴
 - o Kimball: Procesní návrh – dimenzionální model, faktové tabulky + dimenzionální tabulky 2NF²⁵, 3NF
- Interpretace pojmu datamart (dimenzionální model)
 - o Inmon: problémově orientovaná podmnožina dat presentační vrstvy, která je ve vlastnictví oddělení organizace; je plněn transformací (odvozením z) dat presentační vrstvy; je uložen v dedikovaném úložišti; data mohou být smazána
 - o Kimball: procesně orientovaná pod množina tabulek presentační vrstvy; presentační vrstva je tvořena sjednocením procesně orientovaných data martů.

Oba dva tito autoři spolu velmi důrazně nesouhlasí právě ve volbě použití 2NF (2. normální forma) a 3NF (3. normální forma), kterou blíže vysvětlím v kapitole 2.9.2 a pohledu na data marty. I když se může zdát, že se jedná o dosti složitý technický problém, vysvětlení je zdánlivě jednoduché a pro pochopení a využití v oblasti ekonomické je také nezbytné.

2.9.1 Jednotlivé přístupy

Inmonova filosofie je založena na vybudování velkého centralizované úložiště (CIF) následovně několika satelitními databázemi sloužící k analytickým potřebám pro jednotlivá oddělení, později nazývaná jako data mart. Právě proto je jeho přístup nazýván jako „zhora dolů“.

Kimballova filosofie naopak doporučuje začít budovat okamžitě data marty pro potřeby analýzy jednotlivých oddělení a proto se zase jeho přístup označuje jako „zdola nahoru“. Kimball věří, že různorodé data marty s uloženými informacemi

²⁴ Relační tabulky splňují třetí normální formu (3NF), jestliže splňují 2NF a žádný atribut, který není primárním klíčem, není tranzitivně závislý na žádném klíči.

²⁵ Tabulka splňuje 2NF, právě když splňuje 1NF a navíc každý atribut, který není primárním klíčem je na primárním klíči úplně závislý.

v dimenzionálním modelu jsou rychleji dostupné z hlediska dodání, pro mnoho různě zaměřených oddělení v podniku.

2.9.2 Struktura řešení implementace

Mimo přístupu se tyto dva liší také v myšlenkách struktury při vývoji DWH. Inmon vyznává myšlenku vytvoření relačního modelu ve 3NF a Kimball naopak preferuje multi dimenzionální strukturu (star nebo snowflake schéma). Inmon argumentuje tím, že díky relačnímu modelu je snazší zachovat konzistenci dat v podniku a tím vytvářet dokonalejší data. Kimball naopak argumentuje zcela pochopitelně, že datový model pro technicky neznalé uživatele je příliš nesrozumitelný (právě kvůli 3NF) a volba 2NF umožní okamžitě koncovým uživatelům tvořit analýzy. Bez ohledu na rozdíly využitelnosti dat v této oblasti, jsou oba za jedno, a to v tvorbě multi dimenzionálního modelu neboli oddělení atomických dat od úrovně agregační vrstvy.²⁶

Z výše uvedeného je zřejmé, že R. Kimball preferuje cestu od datových tržišť k datovým skladům, zatímco W. Inmon právě opačnou. Obě pojetí mají své výhody a nevýhody, a záleží vždy na konkrétních podmínkách a potřebách podniku, který datový sklad a tržiště vytváří.

²⁶ ANUPINDI N. *Inmon vs Kimball – An Analysis*. [online]. 2005 [cit. 2011-03-14]. Dostupné z www: <<http://www.nagesh.com/publications/technology/173-inmon-vs-kimball-an-analysis.html>>

3. Analýza problému a současné situace

3.1 Představení firmy

3.1.1 Základní údaje o firmě

Název:	Home Credit International a.s.
Sídlo:	Praha 6, Evropská 2690/17, PSČ 160 41
Sídlo provozu:	Vídeňská 119, Brno 619 00
Právní forma:	Akciová společnost
IČ:	60192666
DIČ:	CZ60192666

Společnost Home Credit International, a. s., je součástí skupiny Home Credit, která je předním poskytovatelem spotřebitelského financování na trzích střední a východní Evropy a střední Asie. Skupina Home Credit náleží do silné mezinárodní finanční skupiny PPF, která se zabývá spotřebitelským financováním a retailovým bankovníctvím, prostřednictvím svého podílu ve společném podniku Generali PPF Holding je účastníkem trhu pojišťovnictví a komplexních služeb v oblasti správy majetku. Skupina PPF rozvíjí své aktivity v České republice, na Slovensku, v Rusku, na Ukrajině, v Kazachstánu, v Bělorusku, ve Vietnamu a v Číně.

Home Credit International, a. s., (dále jen HCI) je také poradenskou a servisní společností orientovanou na členy skupiny Home Credit. Součástí Home Credit International je IT centrum, umístěné v Brně a Ostravě, které je zodpovědné za kompletní IT podporu včetně vývoje a provozu centrálních aplikací kritických pro obchodní aktivity jednotlivých společností skupiny Home Credit. V současné době poskytuje podporu pro společnosti Home Credit v České republice, Slovenské republice, Ruské federaci, Kazachstánu, Bělorusku, Ukrajině, Číně a Vietnamu.

Skupina Home Credit, stále expanduje na řadu světových trhů, přičemž se zaměřuje na tzv. nově se rozvíjející trhy (v angličtině „emerging markets“). Hlavními teritorii Home Creditu jsou střední a východní Evropa, střední a jihovýchodní Asie a nově také po Blízkém východě. Ve skupině Home Credit pracuje 19 300 zaměstnanců a dosud poskytla služby již 25,5 milionu klientů (údaj k 31. prosinci 2010). Home Credit nejen

poskytuje úvěry, ale ve vybraných zemích nabízí i klasické retailové bankovní služby, jako jsou běžné účty a vklady. Například v Rusku je Home Credit jednou z nejziskovějších bank v zemi, v Číně obdržel jako jediný čistě zahraniční subjekt licenci na poskytování spotřebitelského financování a vlastně skoro na každém trhu, kde působí, je jedničkou v oblasti splátkového prodeje.

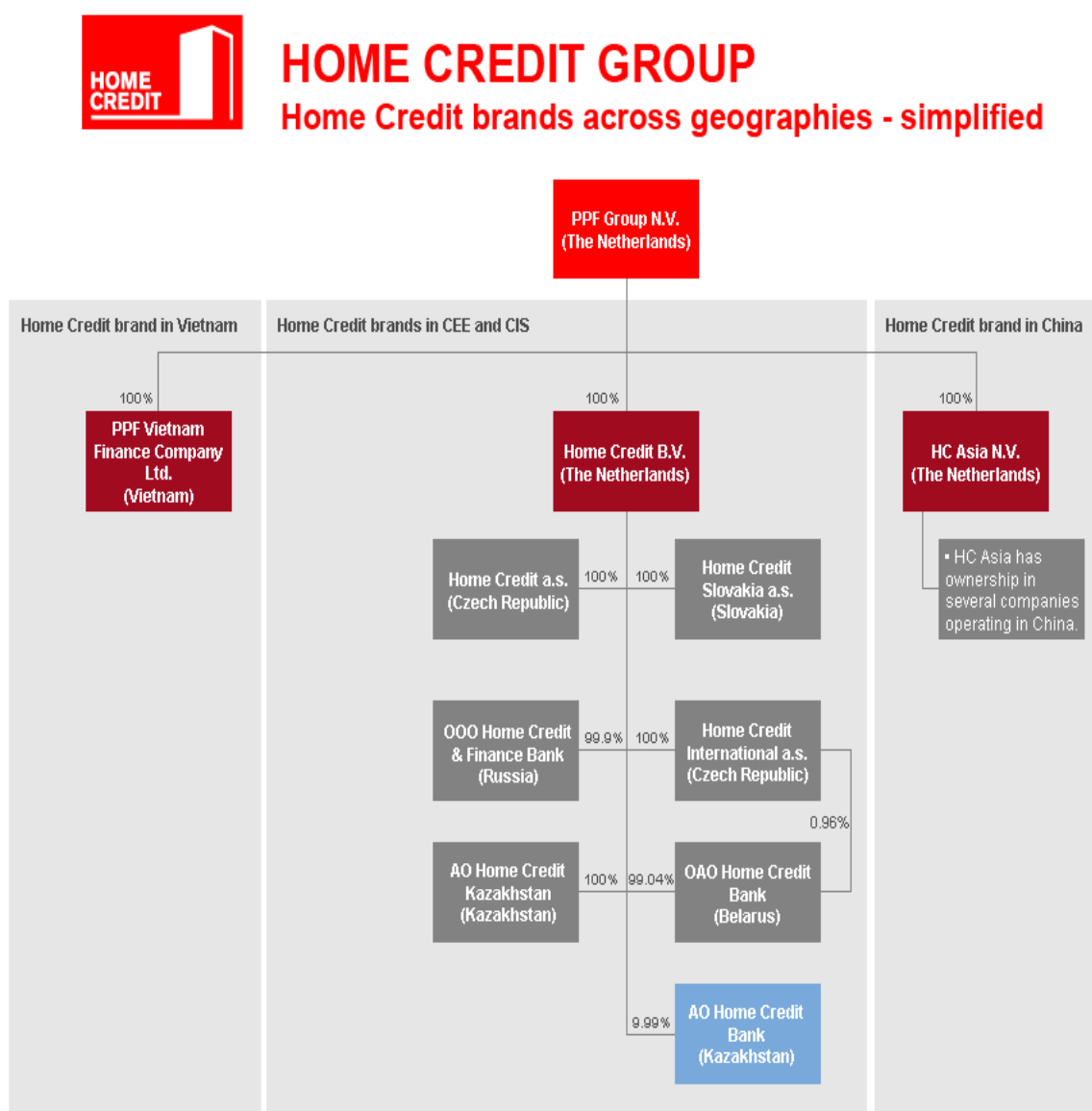
Obr. 6: Působnost Home Creditu



Skupinu Home Credit plně vlastní PPF Group N.V. („PPF“), která je jednou z největších investičních a finančních skupin ve střední a východní Evropě. Jejím zakladatelem a více než devadesátiprocentním akcionářem je Petr Kellner. PPF vlastní aktiva ve výši zhruba 12 miliard EUR (údaj k 30. 6. 2010) a zahrnuje různorodé aktivity od bankovníctví a pojišťovnictví přes nemovitosti, oblast energetiky a těžbu nerostů až po největší ruský obchodní řetězec se spotřební elektronikou. Působnost PPF sahá ze střední a východní Evropy přes Rusko až do Asie.

Tolik o obecném představení společnosti, ve které pracuji jako specialista pro „business intelligence and data warehouse systém (dále jen BI&DW) o kterém bych se rád na následujících stránkách podrobněji rozepsal.

Obr. 7: Home Credit Struktura.



Zdroj: <http://www.homecredit.net/pub/en/about-Us/structure.html>

Obrázek č.7 zobrazuje nejdůležitější společnosti Home Credit Group z pohledu finanční oblasti ve spotřebitelském financování.

Home Credit B.V.vlastní také společnost Home Credit International a.s., která zajišťuje veškerou IT podporu a řešení v klíčových oblastech všem zahraničním společnostem Home Credit.

3.2. Analýza současného stavu IS ve společnosti

Jak jsem psal již výše, HCI a.s. se zabývá podporou, provozem, vývojem a testováním centrálních aplikací kritických pro obchodní aktivity jednotlivých společností skupiny Home Credit. Nedílnou součástí těchto aplikací je i datový sklad, který vyvíjíme, implementujeme a spravujeme v České republice, na Slovensku, v Rusku, na Ukrajině, v Kazachstánu, v Bělorusku, ve Vietnamu a v Číně. Vzhledem k obrovským objemům dat, které zpracováváme a významu datového skladu pro naše zákazníky, jehož dlouhodobější nefunkčnost má výrazný dopad na „business“ si nemůžeme dovolit dělat chyby. Nicméně vývoj datových skladů v této společnosti, který začal v roce 1999, nemá příliš mnoho společného s moderními koncepty a architekturou, kterou jsem blíže vysvětlil v teoretické části. S postupným nárůstem objemů dat je správa a výkon datového skladu v HCI stále více a více náročnější a to jak na čas, tak na strojový výkon čehož důsledkem jsou delší prodlevy mezi aktuálními změnami v provozním systému a jejich promítnutí do DWH tabulek resp. uživatelům. Ne zcela optimální iniciální návrh datových struktur, postupný nárůst dat, zvyšující se počty uživatelů a jejich dotazů do datového skladu, implementace nových a nových funkčních vylepšení a mnohé jiné faktory způsobují, že postupně roste provozní režie datového skladu. To se nutně nemusí projevit jen rostoucími náklady na provoz, ale také například postupným snižováním počtu požadavků, které jsou realizovány, prodlužováním doby realizace požadavků na změny, poklesem kvality výstupů datového skladu, nekonzistentními či nesprávnými informacemi v reportech a podobně.

Problém samozřejmě začíná již při návrhu a následné implementaci skladu. Špatně zvolený postup má za následek vytvoření jakéhosi hybridu, který nemá dlouhodobějšího využití.

3.2.1 Jak vznikal DWH / BI systém v HCI

Navzdory své rozšiřitelnosti, otevřenosti, robustnosti a srozumitelnosti je v českém prostředí Kimballův přístup používán méně často než přístup Inmonův.

Přístup dle Inmona je často (a mnohdy nevědomky) použit v prostředí, které má některé z následujících rysů:

1. Datový sklad nebyl budován na základě věcného zadání od koncových uživatelů – analýza nevychází z uživatelských požadavků (anglicky *business requirement-driven analysis*) – ale na základě existující datové struktury primárních zdrojů – analýza vychází z datového modelu primárních zdrojů (anglicky *data-driven analysis*).
2. Realizace DWH / BI systému byla svěřena odborníkům, kteří mají zkušenosti s realizací provozních systémů. Své znalosti se přirozeně snaží uplatnit i při vývoji DWH / BI systému – výsledkem je normalizovaný datový model prezentační vrstvy DWH.
3. Řešení bylo potřeba dodat co nejrychleji s maximálním využitím stávajících znalostí a minimalizací času stráveného studiem jiných koncepcí a diskuzemi s uživateli.
4. Jedná se o první generaci DWH / BI systému budovaného v prostředí, kde byla tato problematika relativně nová a kde neexistovala znalost přístupu dle Kimballa.

Všechny výše zmíněné rysy jsou bohužel typické také obecně pro české prostředí (ale nejenom pro ně). Nejčastěji jsme se setkávali s následujícím scénářem:

1. Z vrcholového managementu přišlo zadání na vybudování datového skladu. Jsou alokovány potřebné prostředky a stanoveny termíny.
2. Úkol byl buď svěřen přímo oddělení IT (pokud se řeší interně), nebo častěji externí firmě za podpory interního oddělení IT. Postavení expertů ze zadavatelské organizace je často okrajové. Zejména externí firma často přichází s:
 - a. přehnanou vírou ve vlastní znalosti a schopnosti, která je založena spíše na zvládnutí určitých nástrojů a technologií než na komplexním pochopení problematiky BI,

- b. připraveným generickým schématem (které však bylo ověřeno u mnoha významných zákazníků v daném oboru po celém světě),
 - c. otázkou „Tak kde máte ta data, kterými to schéma naplníme?“.
3. Při řešení externím dodavatelem nedošlo k žádnému předání znalostí mezi externími konzultanty a zaměstnanci organizace tak, aby mohl následný rozvoj probíhat interně. Jednak to nebylo definováno jako jeden z cílů, jednak bohužel často není co předávat.
 4. Projekt byl často prohlášen za ukončený a úspěšný v okamžiku, kdy je vytvořena a pravidelně plněna prezentační vrstva datového skladu se schématem 3NFMod. Bohužel 3NFMod je nevhodný pro koncové uživatele – a to i tehdy, kdy je navržen a plněn daty na základě analýzy jejich požadavků a požadovaná data tam určitě někde jsou.
 5. Uživatelé často museli ještě čekat na dodání a nasazení nástrojů pro přístup k datům, protože na tento aspekt se při definici cílů projektu často nemyslí (datový sklad je přece hlavně o těch datech, s tím ostatním si nějak poradíme potom).

3.2.2 Častá nepochopení při komunikaci se zákazníky

Často dochází k nedorozumění, co je podstatou DWH / BI systému (čím takový systém je). Uvádím výčet věcí, za které je často mylně považován:

1. Produkt: DWH / BI systém se nedá koupit. Žádný produkt nemůže obsáhnout všechny oblasti – ETL procesy, prezentační vrstvu, aplikační OLAP vrstvu, reportovací služby, nástroje pro přístup k datům, uživatelský portál, správu metadat, bezpečnost, atd. (Dodavatelé se však snaží nabízet platformu pro DWH / BI systémy složenou z různých produktů.)
2. Projekt: Zavádění DWH / BI systému zahrnuje mnoho projektů. Plánování a analýzu je nutno provádět na celopodnikové globální úrovni, avšak realizace musí být inkrementální, kdy typicky jeden projekt pokrývá jeden datamart (obchodní proces). DWH / BI systém je spíše než projekt nikdy nekončící iterativní proces, kdy objem práce v jednotlivých iteracích (projektech) často narůstá (vývojový tým často zajišťuje jak rozvoj, tak i údržbu a provoz).

3. Datový model: Samotný datový model nic neřeší. Bez ETL procesů, přístupu k datům, správě metadat a dalších vyřešených oblastí je i ten nejlepší datový model k ničemu.
4. Kopie provozního systému: Typická chyba je domnívat se, že zkopírováním provozních dat do samostatného reportovacího systému vzniká DWH / BI systém. Přenesení dat bez jejich restrukturalizace nepřináší nic podstatného z pohledu uživatele.

3.3.3 Co je cílem pro zákazníky

Cíl při vývoji datového skladu musí být dán mírou úspěchu využívání a spokojenosti jeho uživatelů a přidanou hodnotou, kterou jim přináší v jejich každodenní obchodní praxi.²⁷

Úspěšný DWH / BI systém musí mít následující vlastnosti:

- Slouží jako základna pro kvalitní rozhodování
- Umožňuje snadný přístup k informacím organizace: rychlý přístup ke srozumitelným a intuitivním datům
- Presentuje informace konzistentně a věrohodně
- Je adaptivní a schopný pružně a rychle reagovat na změny
- Je bezpečným místem, které chrání citlivá informační aktiva
- Je přijat uživateli

3.4 DWH/BI systém v HCI: současný stav

Současný DWH / BI systém byl v době svého vzniku určitě vhodným a dostačujícím řešením a jistě vznikl dle nejlepšího vědomí a svědomí jeho tvůrců a zodpovědných vedoucích pracovníků, s využitím všech tehdy dostupných znalostí a technologií. Jeho hlavním a zcela legitimním cílem byla podpora provozního reportingu. V průběhu času se jeho součástí staly i specifické analytické úlohy pro podporu provozních činností.

²⁷ Se vzestupem výkonu HW vybavení lze pozorovat, že DWH / BI systémy kromě podpory pro strategické rozhodování začínají přebírat i úlohu reportovacích systémů a jsou schopny realizovat provozní reporting v reálném čase.

Současný stav ve společnostech Home credit (narůstající objemy dat, narůstající počty uživatelů, další provozované instance DWH / BI systému) a vývoj v oboru (nové poznatky a technologické platformy) vybízejí k vyhodnocení stavu současného řešení a k úvahám o jeho rozvoji nebo dokonce opuštění a přechodu na jiný model.

Aktuální rozvojový cíl – nahrazení Business Objects v roli reportovací platformy – je bezpochyby zcela legitimní a věcně správný, ale bohužel nedostatečný. Pro další úspěšný rozvoj systému je nutno zabývat se nejprve společným jmenovatelem všech současných problémů a komplikací, kterým je neoptimální datová architektura.

3.5 HOS 8 analýza současného DWH

HOS 8 slouží k posouzení úrovně informačních systémů v podniku. Metoda je založena na hodnocení osmi oblastí a to: hardware, software, orgware, peopleware, dataware, customers, suppliers a management IS.²⁸

Pro každou oblast byl vytvořen dotazník, který je přiložen v příloze s odpověďmi, na jejichž základech byla sestavena tabulka č. 1.

Tab. 1: Hodnocení jednotlivých oblastí metodou HOS 8

Oblast metody HOS 8	Hodnocení	Slovní hodnocení
Hardware	4	vysoká úroveň oblasti
Software	2	nízká úroveň oblasti
Orgware	3	střední úroveň oblasti
Peopleware	2	nízká úroveň oblasti
Dataware	4	vysoká úroveň oblasti
Customers	3	střední úroveň oblasti
Suppliers	4	vysoká úroveň oblasti
Management IS	4	vysoká úroveň oblasti

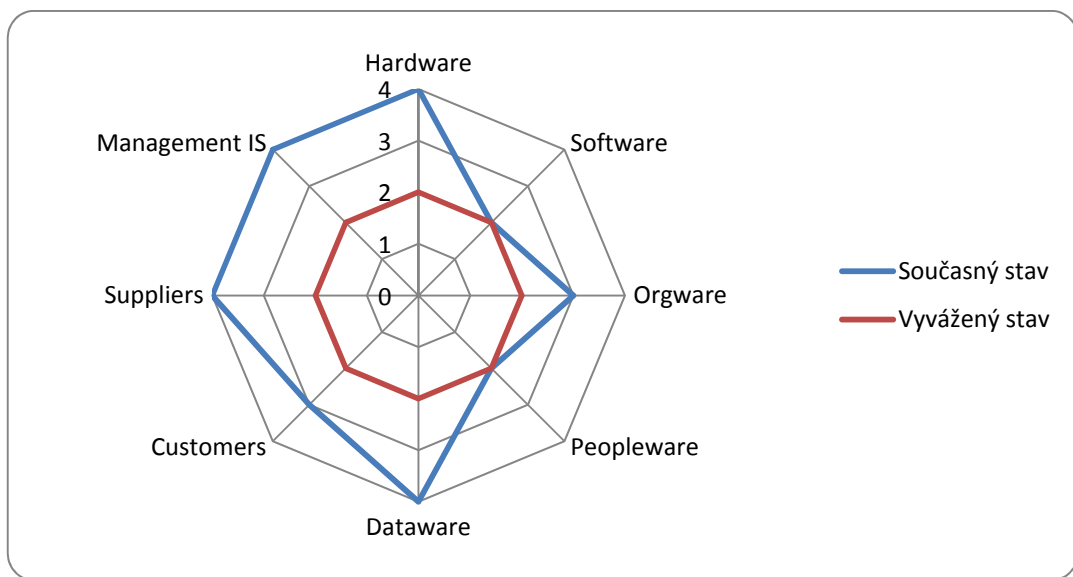
²⁸ KOCH M., DOVRTĚL J. *Management informačních systémů*. 1. vydání. Brno: Akademické nakladatelství CERM, 60 s.

3.5.1 Výsledek analýzy HOS 8

Souhrnný stav dosahuje firma na úrovni nejnižší hodnocené oblasti. V tomto případě se tedy jedná o číslo 2, což značí nízkou úroveň stavu informačního systému. Taktéž není systém dle této metody vyvážený neboť odchýlené hodnoty od nejnižší napočítané hodnoty je více nežli 3 a také rozptyl mezi např. Suppliers a Peopleware je také 2, když peopleware je označen dvojku a suppliers čtyřkou. Datový sklad ve firmě HCI je tak nevyvážený a potvrzuje tak i můj názor.

Z grafu č. 2 je patrné, že mezi nejslabší články tohoto skladu je software a peopleware. Software oblast zahrnuje zkoumání programového vybavení, jeho funkcí, snadnosti používání a ovládání. Peopleware oblast naopak zahrnuje zkoumání uživatelů informačních systémů ve vztahu rozvoji jejich schopností, k jejich podpoře při užívání informačních systémů a vnímání jejich důležitosti.²⁹

Graf 1: Grafické hodnocení jednotlivých oblastí metodou HOS 8



Zdroj: vlastní tvorba

V oblasti software je nekritičtější faktorem rychlost a aktuálnost dostupných dat, na jejichž základě analytické oddělení vytváří nejrůznější ekonomické reporty a tak celý smysl datového skladu, tedy systém pro podporu strategických a operativních

²⁹ Metoda HOS 8 si neklade za cíl hodnotit odborné kvality uživatelů či míru jejich schopností.

rozhodnutí, které má nabízet je do značné míry omezen. Za připomínku také stojí nedostatečný komentář k jednotlivým tabulkám a sloupcům, čímž se velmi ztíží práce uživatelům, kteří si musí neustále zjišťovat význam jednotlivých datových struktur. V oblasti peopleware se lze setkat s velmi častým jevem nezastupitelnosti lidí v jednotlivých modulech skladu. A tak se velmi často stává, že jednotlivé požadavky uživatelů musí čekat, nežli se příslušný autor vrátí z dovolené či se uzdraví. Z velmi omezených kapacitních důvodů z pohledu lidských zdrojů také není čas k „vylepšování“ systému, ale naopak se nejkritičtější chyby nesystémově „zaplátují“.

Obecně datový sklad není z pohledu businessu kritický klíčový a jednodenní výpadek neohrozí chod společnosti, ale je třeba si uvědomit, že jakýkoli výpadek je potřeba poté ve skladu „dohnat“ a získat tak aktuální data. Zpoždění má tedy i pak vliv na délku doby dohrání dat a jejich aktualizaci. Čím větší odstávka tak bude, tím i déle bude trvat, nežli systém bude akceschopný z pohledu uživatelů. Z toho důvodu by mělo být cílem společnosti dostat datový sklad na úroveň „střední“ a tak by měly všechny oblasti být ohodnoceny minimálně stupněm číslo 3.

3.6 Porterův model pěti sil pro oblast IS

Porterův model pěti sil, asi nejznámější, nejpropracovanější a nejčastěji používaný model analýzy odvětví a konkurenčního prostředí, považovaný za převážně statický model, umožňuje identifikovat hlavní síly působící uvnitř konkurenčního prostředí. Souhrnné působení těchto pěti sil určuje intenzitu konkurence v odvětví a zároveň spolurozhoduje o úspěšnosti podniku v daném odvětví. Současně odráží i podstatnou skutečnost, že konkurence v odvětví daleko přesahuje hranice daného odvětví. Dodavatelé kapitálu, technologie, surovin, jako i zákazníci, substituty, potenciální konkurenti, ti všichni jsou nebo mohou být z hlediska vývoje odvětví významní.³⁰

Pro oblast IS je Porterův model rozšířen pro další pochopení strategického významu informačního systému.

Tento model zahrnuje:

- Analýza současné konkurence na trhu v oblasti IS a její řešení v případě zjištění možnosti dosažení konkurenční výhody.

³⁰ SEDLÁČKOVÁ H. Trendy v chápání zdrojů podniku při tvorbě strategie podniku. *Acta Oeconomica Pragensia*, 2007, č. 2, s. 107

- Analýza vyjednávací síly zákazníků a zjištění zda tuto sílu má možnost IS ovlivňovat
- Analýza vyjednávací síly dodavatelů a zjištění zda tuto vyjednávací sílu má možnost IS také ovlivňovat
- Analýza hrozeb vstupu nových konkurentů, a zda má IS možnosti vytvářet nové bariéry vstupu
- Analýza hrozeb substitučních produktů a schopnost IS produkovat nové výrobky³¹

3.7 SWOT analýza

Analýza silných a slabých stránek, příležitostí a hrozeb podniku (SWOT - Strengths, Weaknesses, Opportunities, Threats), tvoří základ strategické analýzy. Teprve po zpracování odhadu vnitřní a vnější situace firmy lze uvažovat o výběrech realizovatelné strategie. Tyto výběry by mohly vycházet z analýzy. Výběry mohou být jen hrubě roztříděny, jako například rozšíření výrobního programu, růst, nebo maximalizace krátkodobého peněžního toku. Způsob, jakým podniková jednotka soupeří na trhu, nemůže být generalizován, protože je zcela závislý na předchozí analýze. Dobrá strategie by měly stavět na zdrojích síly a využívat příležitosti. Logika tohoto přístupu naznačuje, že každá firma bude čelit rozdílnému souboru příležitostí a hrozeb a každá bude mít rozdílné zdroje síly. Strategie, které z toho vyplynou, budou u každé firmy specifické. V nezkušených rukou bohužel vede SWOT analýza k vytvoření dlouhého seznamu problémů a čím je tento seznam delší, tím nejasnější je obraz vznikající strategie.³²

SWOT analýzu je možné implementovat i na informační systém, v tomto případě na datový sklad.

Analýza silných stránek

- Silný nástroj pro podporu rozhodování
- V dlouhodobém časovém hledisku levný a efektivní nástroj
- Rychlé a účelné grafy a reporty
- Přehled o hospodářském a ekonomickém průběhu firmy

³¹ GÁLA, L. *Podniková informatika*. Praha: Grada Publishing, a.s., 2006. ISBN 80-247-1278-4.

³² MALLYA, T. *Strategické řízení*, 1.vyd. Praha: Grada Publishing, a.s., 2006. 31 s. ISBN 80-247-1911-8

- Moderní nástroj
- Přizpůsobitelnost

Analýza slabých stránek

- Špatný návrh datového modelu
- Nelze koupit jako „krabicové řešení“
- Špatná komunikace se zákazníky
- Nákladná počáteční školení potřebná k provozu
- Vzrůstající požadavky zákazníka vedoucí ke zvyšování finančních nákladů na HW

Analýza příležitostí

- Naplnění potřeb vyššího a nejvyššího managementu
- Výrazně lepší hospodaření a rozhodování firmy na základě vysoce sofistikovaných reportů
- Blízká doba návratnosti investice
- Vytváření nepřeborného množství kampaní, nabídek apod.

Analýza hrozeb

- Neprofesionální provoz systému
- Špatná optimalizace systému vedoucí ke zvyšování finančních nákladů na HW
- Při špatném navržení mnohamiliónový a špatně sloužící systém

Tab. 2: tabulka SWOT analýzy (nejdůležitější fakta)

Silné stránky	Slabé stránky
<ul style="list-style-type: none"> • Silný nástroj pro podporu rozhodování • Přizpůsobitelnost a sofistikovanost • Moderní a rozvíjející nástroj 	<ul style="list-style-type: none"> • Dlouhý vývoj • Počáteční náklady • Nelze pořídit hotové řešení
Příležitosti	Hrozby
<ul style="list-style-type: none"> • Přidaná hodnota pro management • Zvýšení přehledu o hospodaření • Blízká doba návratnosti investice 	<ul style="list-style-type: none"> • Špatná správa • Špatná optimalizace • Špatný návrh datového modelu

3.8 SMART analýza

SMART je souhrn pravidel, která pomáhají především v rámci projektového managementu (Project Management) efektivně definovat rámec či cíl projektu a navrhovaného řešení. Pravidla SMART lze ale uplatnit i v jiných oblastech, kde je třeba nějakým efektivním způsobem definovat cíle. V okamžiku, kdy v počáteční fázi projektového managementu se definuje obsah a cíle projektu, je nutno zamyslet se nad kvalitativními stránkami tohoto vymezení. Řešení v rámci projektového managementu by mělo být definováno podle pravidel známých souhrnně jako SMART.

Specific - Navrhované řešení nebo příležitost by měly být přesně popsány.

Measurable - Navrhované řešení by mělo být měřitelné.

Action Oriented - Řešení musí odpovídat potřebám svého příjemce.

Realistic - Řešení musí být realistické.

Timed – Přesně definovaný časový rámec pro uvedení řešení v praxi.³³

Specific

V oblasti specific je nutné si přesně definovat, co je potřeba a co od projektu očekávat. Přehnaná očekávání a nároky mohou v konečném důsledku uživatele velmi zklamat a demotivovat ho k využívání finálního produktu, v tomto případě tedy DWH. Je tedy nutné přesně popsat, jak navrhované řešení bude fungovat a co je možné od něho očekávat a co už ne.

Measurable

Měřit výkonnost či úspěch takového projektu číselně je značně obtížné. Osvojení si DWH je běh na dlouhou trať a výsledky se dostávají průběžně, ale dlouhodobě. Hlavním cílem je přidaná hodnota tohoto produktu a lze ji v dlouhodobém měřítku porovnávat např. s výsledkem hospodaření podniku.

Action Oriented

Aby řešení odpovídalo potřebám svého příjemce, je potřeba se vrátit k pravidlu „Specific“ a z něho vycházet. Je velmi důležitá analýza a zpětná vazba zákazníka v čase celého průběhu vývoje.

³³ SPIDY, P. *Definice cíle SMART* [online]. 2011 [cit. 2011 - 03 - 21] Dostupné z: < <http://www.finance-management.cz/080vypisPojmu.php?IdPojPass=39&X=Definice+cile+SMART+Project+Management> >

Realistic

O tom, že řešení musí být realistické, by nemělo být pochyb již od samého počátku analýzy a vývoje. Podhodnocené ceny např. HW mohou značně zkomplikovat celý proces na jeho konci a tak výrazně ztížit dokončení produktu. Přehnané požadavky a nereálné termíny by měl rozpoznat projektový manažer již na samém počátku.

Timed

Toto pravidlo je snad nejsložitější z pohledu dodržení dodání jednotlivých částí i konečného produktu. Není snadné termíny dodržet ať již z důvodu nenadálých technologických, logistických či kapacitních důvodů a to, ani když nejsou termíny značně přiškrčené. Vývoj velkého projektu ať už je jakýkoli sebou nese rizika zpoždění dodávky, na kterých se většinou podílí obě strany, tedy zákazník i dodavatel.

3.9 Vyhodnocení analýzy

Z výše uvedeného vyplývá, že současné řešení je zastaralé a je nezbytně nutné vyvinout nové řešení resp. aktualizovat stávající řešení. Současný systém byl nepochybně navrhnout a naprogramován s těmi nejlepšími úmysly, a dle nejlepšího vědomí avšak bez pevné koncepce s výhledem na budoucnost. S konstantním nárůstem objemu dat a počtu požadavků od zákazníků se současný systém stále častěji obtížněji vyrovnává. Na základě provedených analýz byl zjištěn a definován problém zejména s aktuálností dat v datovém skladu, který má a v budoucnu by měl stále delší odezvu. V tomto případě systém nepracuje správně a ztrácí tak na své největší přednosti neboli přidanou hodnotu pro uživatele. Mezi další problémy lze uvést nedostatečný popis systému, čímž se datový model stává pro zákazníka či analytika značně nepřehledným a velmi tím ztěžuje práci těmto koncovým uživatelům. Nejvýraznějším problémem, je nekoncepční datový model, který je náročný na správu a zároveň chaotický jak pro dodavatele tak odběratele. V případě vývoje nového systému musí být dána jasná pravidla pro vývoj tohoto modelu, musí být také nastaveny procesy pro jednotný postup a samozřejmě nesmí chybět podrobná dokumentace, která v současnosti chybí, čímž by se zamezilo aktuální nezastupitelnosti. Základem pro vyřešení těchto problémů je také navýšit kapacitu lidských zdrojů, neboť nedostatečný počet DWH vývojářů vede pouze k odsunování současných problémů nebo jejich dočasné řešení, což je opět značně nekoncepční.

4. Vlastní návrhy řešení

4.1 Možné varianty řešení

Jak jsem psal již v podkapitole 3.4, byl datový sklad v HCI vybudován určitě vhodným a dostačujícím řešením. Nicméně s narůstajícími objemy dat, uživatelů a nově známými technologiemi je současné řešení již nedostatečné a nereflektuje tak požadavky uživatelů. Ať je již DWH koncipován různými metodami a postupy, jeho výsledek je vždy dán přidanou hodnotou pro uživatele, kteří se tak stávají hlavními hodnotícími prvky na konci celého projektu.

Tvorbu datového skladu lze svěřit odborníkům (outsourcing) a takto si ho pronajmout. Lze si ho i objednat a nechat si ho vytvořit na klíč, ale nelze si ho v žádném případě koupit jako krabicové řešení. Poslední variantou a podle mě tou nejproduktivnější, nejlevnější a zároveň nejefektivnější je využít vlastní kapacity a vytvořit datový sklad tzv. na zelené louce.

4.1.1 Vytvoření systému na klíč

Tato varianta je z krátkodobého i střednědobého hlediska finančně zajímavější než li si ho objednat na klíč. Má to, ale velké ALE. Za hlavní negativum bych označil sdílení dat outsourcingovou firmou. Přece jenom podnikové účetnictví, výsledky hospodaření a obecně celé know how firmy má k dispozici externí firma. Samozřejmě, že by tento outsourcing musel být právně ošetřený ve smlouvách tak, aby nemohlo docházet ke zneužívání dat, ale už jen pocit, že se kdokoli může dívat na data, která jsou určena jen podnikovým zaměstnancům či nejvyššímu managementu, je přinejmenším odstrašující. V takovém případě by zřejmě nejkritičtější moduly např. účetnictví či mzdy nebyly zahrnuty v koncepci datového skladu. Dalším velice silným argumentem pro nepřijetí tohoto řešení by byla také značně velká úměra času na vybudování tak složitého systému jakým datový sklad je. Externí dodavatel by se musel velmi detailně seznámit s datovým modelem provozního systému, musel by také být stále v kontaktu se zákazníkem a neustále konzultovat své kroky tak, aby se minimalizovala přehnaná očekávání či nevznikaly nepotřebné modely. Posledním a asi nejsilnějším argumentem by byla cena, která by mohla dosáhnout a to z vlastní zkušenosti mohu potvrdit, stovek miliónů korun. Nicméně pro společnosti bez vlastního IT oddělení nebo malé či středně

malé firmy, které by rády datový sklad využívaly, nemají většinou na výběr a volbu této varianty minimálně zvaží, pokud hned nepřijmou. Přece jen zaměstnávat programátory na správu několika GB skladu je zbytečný luxus a také vybudování tohoto DWH by nedosahovalo tak závratných cen jak je uvedeno níže.

Pro příklad firmy HCI a.s. pro vybudování cca 10TB datového skladu může sloužit následující tabulka.

Tab. 3: Přibližná cenová kalkulace na vybudování 10 TB DWH

Datový sklad				
	Počet	Denní sazba	Dní	Celkem
Externí pracovníci	10	200 000	365	73 000 000
	Počet	Cena školení na osobu za 2 dny	Dní	
Školení interních zaměstnanců	10	40 000	10	2 000 000
	Počet na 1 CPU	Cena		
Oracle Standard Edition Licence	10	150 000		1 500 000
OWB Licence	10	100 000		1 000 000
Oracle Business Intelligence licence	10	100 000		1 000 000
Hardware – disková pole, paměť, procesory ..				10 000 000
Celkem				88 500 000

Jak je vidět, tak nechat si udělat DWH na klíč je velmi finančně náročný projekt. Specialistů na vybudování tohoto systému není v České republice stovky, ale spíše desítky a dle toho také vypadají hodinové sazby, které se nezdá kdy pohybuji méně než 2500 korun na hodinu. Ovšem po uskutečnění se náklady výrazně sníží, pouze již na mzdy interních IT zaměstnanců, kteří tento projekt již jen spravují. Mezi firmy, které nabízejí toto řešení, patří například Oracle, Ambica, Asecco, Gordic a další. Ceny, které

v tabulce č. 3 uvádím, jsou orientační, ale nijak výrazně se neliší od skutečných. Nejnákladnější položkou jsou externí pracovníci, které si firma najme na analýzu, vývoj a implementaci datového skladu. Školení interních zaměstnanců je na nejnižší možné úrovni, pravděpodobnější budou vyšší náklady potřebné k zaškolení oracle nástrojů a BI nástrojů. Školení uživatelů by poté mohlo z úsporných potřeb provádět již samotné IT oddělení. Oracle Standard Edition je již samotná databáze, nad kterou datový sklad bude implementován. Licenční politika společnosti Oracle umožňuje kombinovat mnoho produktů za různé ceny a tak by obchodní zástupce mohl vyjednat určité zajímavé ceny. OWB neboli Oracle Warehouse Builder je nástroj sloužící k budování DWH, tedy tvorby mapování neboli ETL jež jsou popsány v podkapitole 2.5.1. Nakonec licence potřebné pro konečné uživatele k vytváření dashboardů a sestav, tedy k nástroji Oracle Business Intelligence.

4.1.2 Projekt na zelené louce

Projekt na zelené louce je nejžádanější z pohledu vývojářů, ale zejména analytiků tedy uživatelů. Zároveň je to i nejrizikovější projekt z pohledu investora. Ve své krátké praxi jsem se již setkal s několika projekty, které měli ambice začít na zelené louce, ve kterých byla velká očekávání, a bylo do nich investováno značné množství peněz, ale nakonec po delším čase skončili tzv. v šuplíku či koši nebo byly za ještě značnějších investic optimalizovány, opravovány a předělávány třetí stranou. V horším případě vznikly všelijaké nepoužitelné hybridy. Očekávání a představy jsou velmi pestré a chuť do této práce a motivace je velmi vysoká a však nakonec z důvodu neodbornosti nebo nepředvídatelných nákladů, které již společnost nechce vynakládat, není projekt dokončen. Pokud IT firma nedisponuje skutečně fundovanými odborníky na danou oblast, vzniká tu obrovské riziko prodlužování termínu dokončení a s tím spojené minimálně navyšování mzdových nákladů. Ve většině projektů na zelené louce v oblasti IT se jedná o předělání informačního systému z jedné programovací platformy do druhé vzhledem k technologickým pokrokům a možnostem. Jde např. o přepis jednotlivých modulů z visual basic do Javy, nebo Javy do .NET apod. V případě datového skladu v HCI se jedná o vytvoření naprosto nového datového modelu na bázi faktových a dimenzionálních tabulek využívající stávající technologii databáze Oracle. Samozřejmě za přispění vhodné kombinace obou koncepcí Ralpa Kimballa a Billa Inmona.

4.1.3 CPM

Pro první dvě zvolené varianty je vhodné využít tzv. CPM. Critical Path Method - metoda kritické cesty nebo také metoda síťové analýzy (časový rozbor dílčích činností projektu), kdy se pro každou činnost nejprve určuje nejdříve možný začátek a následně zpětně nejpozději možný konec. Stavby projektu ležící na kritické cestě mají shodný nejdříve možný začátek a nejpozději možný konec, kritické činnosti leží na cestě mezi těmito uzly.³⁴

Tab. 4 :vstupní údaje

Číslo hrany	Před. Činnost	Návrh zákazníka	Délka trvání ve hodinách
A	-	Analýza požadavku	16
B	A	Odhad pracnosti	8,5
C	A	Finanční odhad	24
D	A	Vyjednávání o ceně	24
E	A	Odhad HW a SW vybavení	4
F	B, C, D	Návrh a podpis smlouvy	8,5
G	F	Rozdělení pracovních činností	2
H	G	Návrh datového modelu	40
CH	G	Návrhu designu aplikace	35
I	H	Programování datového modelu	80
J	CH	Realizace interface	100
K	J	Programování funkčních částí modulu	120
L	K, I	Interní testování 1. části projektu z funkčního hlediska	32
M	L	Začátek 2. Části projektu	2
N	M	Naplnění databáze reálnými daty	4,5
O	M	Prizpůsobení aplikace legislativním podmínkám	6
P	N, O	Ladění aplikace	24
Q	P	Finální interní testy	40
R	P	Testování aplikace u zákazníka	60
S	R	Tvorba dokumentace	40
T	R, Q	Předání aplikace zákazníkovi	40
U	T	Údržba	4

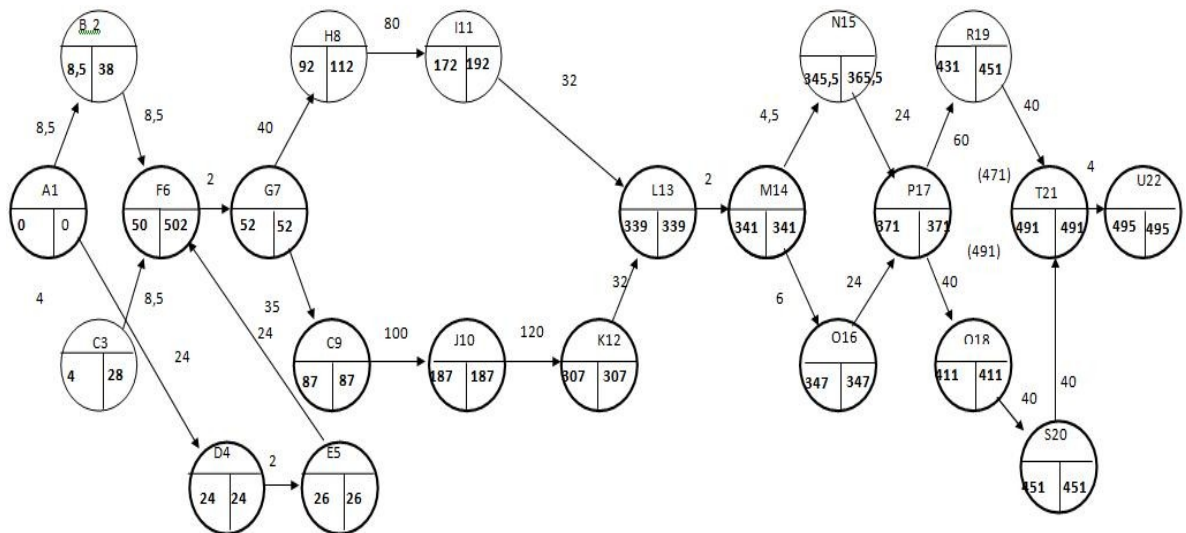
U vývoje nového systému se metoda CPM přímo nabízí. Použiji metodu kritické cesty, pomocí níž je tato složitá činnost rozložena na několik dílčích činností, mezi nimiž existuje časová návaznost a podmíněnost.

³⁴BUSINESS, C. CPM, [online 2010] poslední revize: neuvědno. [cit 10.12.2010] Dostupné z: <<http://business.center.cz/business/pojmy/p987-CPM.aspx>>

Cílem projektu je řešení složitých časových posloupností s cílem dosáhnout maximálního zkrácení celkového průběžného času, který je potřebný pro vývoj nové aplikace.

Následuje schéma na obrázku č. 8, vyplývající ze vstupních dat v tabulce č. 4.

Obr 8: schéma CPM



Tab 5: Výsledky výpočtu charakteristik

činnost		trvání	nejdříve možný		nejpozději přístupný		rezervy		
i	j	t_{ij}	začátek ZM	konec KM	začátek ZP	konec KP	RC	RV	RN
1	2	8,5	0	8,5	29,5	38	29,5	0	0
1	3	4	0	4	24	28	24	4,5	0
1	4	24	0	24	0	24	0	0	0
2	6	8,5	8,5	17	28,5	37	20	33	3,5
3	6	8,5	4	12,5	24	32,5	20	47,5	13,5
4	5	2	24	26	24	26	0	0	0
5	6	24	26	50	26	50	0	0	0
6	7	2	50	52	50	52	0	0	13
7	9	35	52	87	52	87	0	0	0
7	8	40	52	92	72	112	20	0	0
9	10	100	87	187	87	187	0	0	0
8	11	80	92	172	112	192	20	0	0
10	12	120	187	307	187	307	0	0	0
11	13	32	172	202	192	222	20	135	0
12	13	32	307	339	307	339	0	0	0
13	14	2	339	341	339	341	0	0	0
14	15	4,5	341	345,5	361	365,5	20	0	0
14	16	6	341	347	341	347	0	0	0
15	17	24	345,5	369,5	365,5	389,5	20	0	0
16	17	24	347	371	347	371	0	0	0
17	19	60	371	431	391	451	20	0	0
17	18	40	371	411	371	411	0	0	0
18	20	40	411	451	411	451	0	0	0
19	21	40	431	471	451	491	20	20	0
20	21	40	451	491	451	491	0	0	0
21	22	4	491	495	491	495	0	0	0

Výsledek CPM

Podle hodnot ze sloupce RC(Celková rezerva) lze vyčíst kritickou cestu, která vede po činnostech, které mají RC = 0, tedy činnosti: 1->4; 4->5; 5->6;6->7; 7->9; 9->10; 10->12;12->13;13->14;14->16;16->17;17->18;18->20;20->21;21->22. Celková délka kritické cesty je 495 hodin. Samotný vývoj nového softwaru je čistě hierarchicky předem znám a naplánován. Největší vliv na zpoždění je tak spíše obchodního charakteru, kdy se licituje o ceně celého projektu a mění se jeho požadavky.³⁵

4.1.4 Rozšíření stávajícího systému

Z pohledu společnosti HCI a.s., která plní funkci IT centra pro ostatní společnosti skupiny PPF Group., a je zde tedy soustředěna velmi silná skupina odborníků z pohledu IT oboru, je zbytečné si najímat třetí stranu pro potřeby programování či testů. I když i zde k tomu zřídka kdy dochází u menších projektů z kapacitních i časových důvodů zejména pro potřeby testování. Nicméně pro oblast skladů má společnost vyhrazených 15 lidí což není jistě plnohodnotný počet lidí pro tak velký projekt (je nutné spravovat nadále i původní řešení po nezbytně dlouhou dobu), ale s velkou mírou snahy lze dosáhnout velmi úspěšného projektu. Abychom se vyvarovali předešlých možných komplikací, které jsem zmiňoval v případě plně outsorcového řešení, tedy možnosti neodborného a nesystémového přístupu bych zvolil minimálně dva externí specialisty - konzultanty, kteří by měli roli dohlížející a školící. Vzhledem k tomu, že nový model koncepce datových skladů se v blízké budoucnosti bude dotýkat všech zemí, kde home credit operuje, je tato investice u vývoje prvního datového skladu na největší instanci tedy Ruska, zvláště na místě.

4.2 Výsledné řešení pro HCI

Z výše uvedeného je pro společnost Home Credit International nejlepší řešení kombinace projektu na zelené louce spolu s koordinací expertů ze společnosti Oracle pro podporu vytvoření datového modelu, konceptu a obecně přístupu při vývoji nového datového skladu. Z pohledu vývoje na zelené louce jde především tedy o úplně nový datový model, který je základem pro každý projekt postavený nad už jakoukoli

³⁵ RAIS, K, *Základy optimalizace a rozhodování*. 1. vyd. Brno 2004. 134 s. ISBN 80-214-23280

databází. Externí pracovníci jsou nutné zejména v počátku, tedy při prvním inkrementu nebo při vytvoření DWH v jedné zemi. V ostatních zemích již nebudou tato outsorcová řešení potřeba, neboť datový model z 90% zachováme a také již nabudeme potřebné znalosti.

4.3 Postup řešení

Současné řešení je rozsáhlé a přechod do cílového stavu nemůže být skokový. Bude nutno budovat nové řešení inkrementálně, paralelně se současným řešením, a stávající funkčnost nahrazovat novými komponentami.

Následující fázový plán se snaží popsat jednotlivé klíčové kroky, kterými bude nutno projít. Fáze jsou sekvenční, určité aktivity v rámci fází mohou probíhat souběžně (závislosti nejsou znázorněny).

Podstatné jsou především činnosti popsané v inceptní a elaborační fázi. Konstrukční a přechodová fáze neobsahují nic překvapivého, pouze standardní aktivity při vývoji softwarového díla.

4.3.1 Inceptní fáze

Interní aktivity (prováděné v HCI)

- Interní osvěta a vzdělávání
- Analýza současného stavu a vytvoření dokumentace architektury současného řešení: datově-procesní model, zmapování datových a servisních rozhraní, vymezení modulů provozní podpory a jejich rozhraní
- Konceptní návrh cílové architektury a testování technologických platform pro její realizaci
 - Prezentační vrstva: relační DB Oracle nebo relační DB MS SQL Server 2005
 - Aplikační vrstva: OLAP databáze MS Analysis Services (Unified Dimensional Model – UDM) nebo univerzum BO
 - ETL: MS Integration Services nebo proprietární řešení (uložené procedury, Perl), případně kombinace obojího
 - Klientský přístup k datům: MS platforma (Reporting Services, Excel, ...) nebo BO

Srovnání DWH / BI řešení Oracle/BO a MS SQL Server 2005 a argumentace, proč právě MS SQL Server 2005 by měl být testován a zvažován, přesahují rámec této práce.

- Vytvoření prototypových řešení prezentační, aplikační a klientské vrstvy s maximálním využitím současného řešení (prototyp bude nadstavba nad současnou prezentační vrstvou L1).
- Příprava vývojového a provozního prostředí
- Analýza rizik
- Volba cílového zákazníka pro implementaci pilotního řešení

Externí aktivity (vyžadují součinnost zákazníka)

- Osvěta uživatelů (analytiků, manažerů), zdůraznění klíčové role uživatelů a jejich požadavků při návrhu BPDMod³⁶.
- Získání podpory uživatelů.
- Analytici mají vytvořeny spousty BO reportů a na MS SQL Serveru velké množství tabulek a uložených procedur. Nebude snadné přesvědčit je o zásadních změnách.
- Obsazení role *business sponzora (patrona)* – vysoce postaveného výkonného manažera na straně zákazníka, který bude aktivním členem, prvkem a zastáncem zamýšlených změn.

4.3.2 Elaboráční fáze

Interní aktivity

- Konečná volba technologií
- Detailní návrh cílové architektury, identifikace komponent a jejich rozhraní
- Mapování cílové architektury na současnou (např. BPDMod na 3NFMod³⁷)

³⁶ Dimenzionální model, kde sledovanými veličinami (fakty) jsou údaje o procesech organizace – fakta jsou kvantitativní údaje popisující sledované procesy (získávané měřením procesů). Sousloví vyjadřuje skutečnost, že návrhu BPDMod musí předcházet analýza uživatelských požadavků, která identifikuje klíčové procesy organizace: kvantitativní veličiny, které reprezentují sledované procesy, a jejich odpovídající dimenzionální – většinou kvalitativní (kategorické, popisné) – veličiny.

- Vytvoření plánu přechodu ze současného stavu do cílového
- Zavedení vývojového a provozního prostředí
- Vytvoření iteračního plánu pro konstrukční fázi

Externí aktivity

- Analýza uživatelských požadavků
 - Vstupy: pohovory s uživateli, stávající řešení (dokumentace současné a cílové architektury, reporty, OLAP kostky), prototypová řešení
 - Výstup: BPDMod, dokumentace cílové architektury, požadavky na frekvenci obnovy dat, prioritizace požadavků, rozsah řešení pro konstrukční fázi
- Definice a odsouhlasení akceptačních kritérií
- Stanovení hranice mezi HCI a jeho zákazníky:
 - Identifikace datového rozhraní, které HCI poskytuje svým zákazníkům. Možnými kandidáty jsou prezentační vrstva, dimenzionální univerzum BO, OLAP databáze MS Analysis Services (UDM). Nyní je tímto rozhraním univerzum BO.
 - Definice rolí, zodpovědností a pravidel spolupráce.
 - Určení provozních komponent, jejichž užívání a správu je účelné převést k zákazníkům (např. poskytnout analytikům datové rozhraní pro marketingové výběry).

4.3.3 Konstrukční fáze

Interní aktivity

- Vývoj řešení
- Testování (vyhodnocování kvality)
- Tvorba dokumentace (vývojové, testovací, uživatelské)

³⁷ Normalizované schéma je díky svým vlastnostem vhodné pro využití ve víceuživatelských víceúlohových provozních systémech pracujících v reálném čase, jejichž úkolem je realizovat každodenní provozní aktivity organizace a v daném okamžiku uspokojovat souběžné požadavky mnoha uživatelů, které zahrnují čtení / modifikaci / vytváření záznamů o individuálních objektech.

- Optimalizace procesů
- Sdílení know-how
- Sledování plánu, řízení zdrojů, řízení rizik

Externí aktivity

- Komunikace se zákazníkem: změny vyplývající z vývoje, změnové požadavky přicházející od zákazníka

4.3.4 Přechodová fáze

Interní aktivity

- Testování
- Změny na základě zpětné vazby od zákazníka

Externí aktivity

- Školení uživatelů
- Akceptační testování uživateli
- Získání zpětné vazby
- Akceptace řešení
- Poskytnutí konečného řešení uživatelům
- Provozní podpora uživatelů

4.4 Datový model

V teoretické části této práce jsem rozebíral datový model z pohledu vývoje DWH. Tento model se skládá z dimenzionálních a faktových tabulek, jedná se o tzv. BPDMod, kde jsou sledované veličiny zaznamenané ve faktových tabulkách. Jsou to údaje o procesech v organizaci a jedná se také o data, která vznikla těmito procesy v organizaci. Může se jednat např. o založení smlouvy, vložení vymáhací metody, schválení úvěrů či odprodej smluv třetí straně neboli sekuritizaci. Faktové tabulky tedy zachycují sledované ukazatele na rozdíl od dimenzí, které určují úhel pohledu jako čas, produkt či zákazník. Dimenze obsahují klíčová data, kvalitativní, které jsou např. kategorické či popisné. Dimenze určují hierarchii (1:N). Dimenzionální modelování klade velký důraz

na srozumitelnost pro uživatele a je optimalizován pro vyhledávání dat a složité analýzy. Mezi jeho výhody patří lehká rozšiřitelnost bez dopadu na aplikace a snadné přidání nových dimenzí a fakt se stejnou granularitou. K dimenzím jsou faktové tabulky spojeny pomocí cizích klíčů.

Jak takový dimenzionální model může vypadat, jsem nastínil na obrázku č. 9. Záměrně jsem nevytvářel jednoduchý, vzorový model, který se předkládá v odborných člancích či publikacích, neboť jsou velmi zjednodušené a ve skutečnosti se neobjevují. Vzorový příklad bývá většinou zobrazován ve formě jedné faktové tabulky např. prodeje, kolem níž jsou dimenzionální tabulky obchodu, data, zákazníka a popřípadě produktu. Ve skutečnosti jsou multi dimenzionální modely daleko sofistikovanější a složitější. Model obsahuje stovky, ale spíše tisíce tabulek a číselníků a milióny vazeb mezi nimi. Model, který jsem vytvořil, obsahuje sedm dimenzionálních tabulek včetně dvou číselníků a dvou faktových tabulek. Tento model není konečný, ale dále se rozvětjuje do veliké hierarchie dalších a dalších faktových a dimenzionálních tabulek, nicméně pro vzorový příklad je dostačující.

4.4.1 Vzorový dimenzionální model

Uprostřed modelu se nachází faktová tabulka nazvaná jako FT Contract AD (accumulation daily) neboli smlouva. Nad touto tabulkou se nachází primární klíč a následné cizí klíče spojující ji s jednotlivými dimenzemi.

FT_CONTRACT_AD : je jak jsem psal výše faktová tabulka obsahující data vztahující se ke smlouvě. Tabulka obsahuje všechny schválené smlouvy s údaji o datu schválení, statusu, výše anuity či počtu dnů v prodlení.

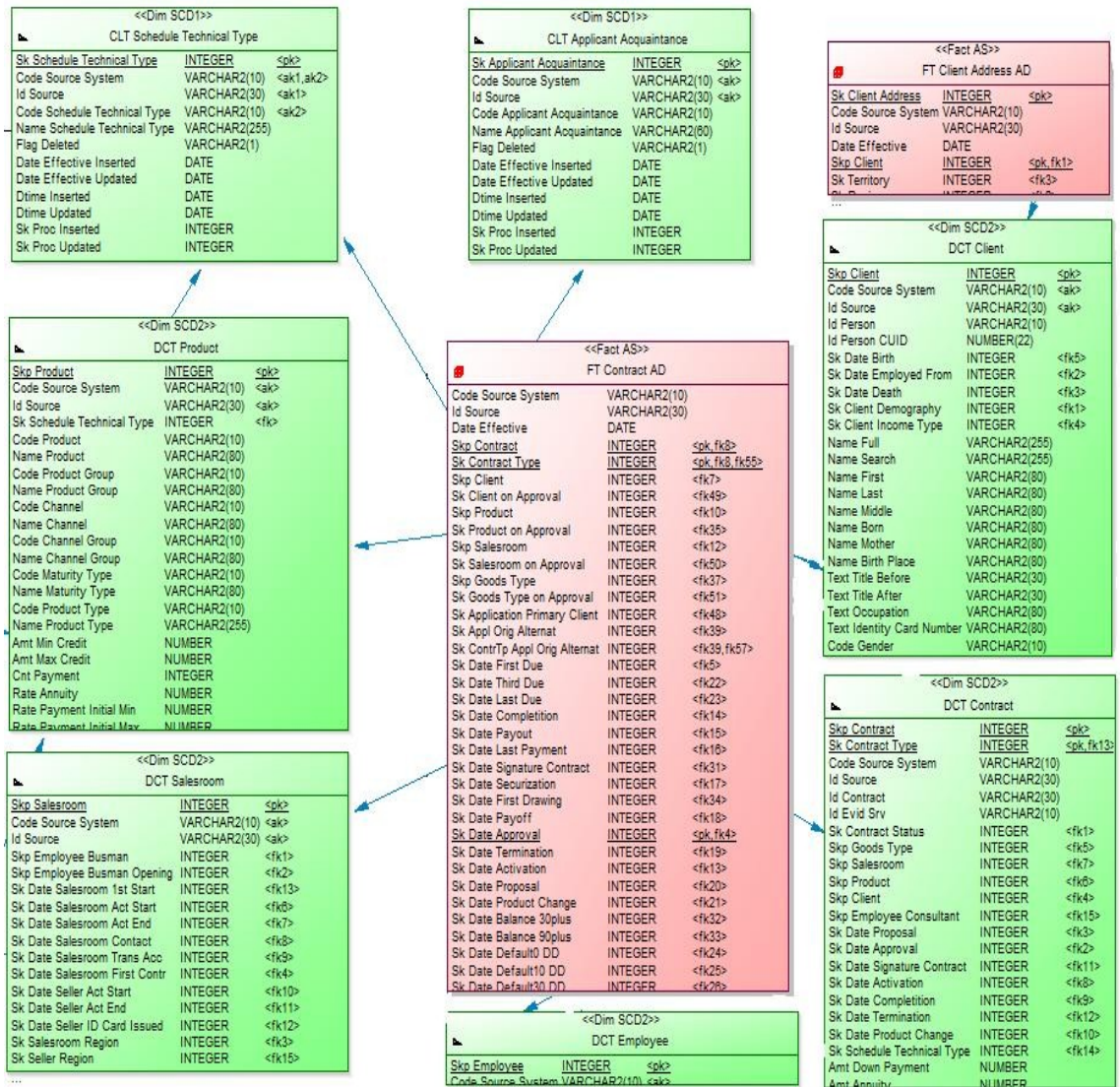
CLT Schedule Technical Type: je číselník obsahující typy předpisů a splátkového procesu využívaný pro produkty a smlouvy.

DCT Product: je označení pro dimenzionální tabulku produktů. Tato tabulka obsahuje informace o produktu, jeho typu, skupině a detailu.

DCT Salesroom: je dimenze obchodního partnera, kde je půjčka uzavřena. Obsahuje hierarchie regionu, síť kategorie, subjektu apod.

CLT Application Acquaintance: tento číselník je využíván pro schvalovací proces neboli označení smlouvy jako nové, zamítnuté, schválené apod.

Obr. 9: datový model vztahující se ke smlouvě



Zdroj: vlastní tvorba

FT Client Address: takto faktová tabulka obsahuje adresu všech klientů napříč všemi adresnými typy.

DCT Client: dimenzní tabulka sloužící k uchování všech klientů v přímé vazbě na smlouvu.

DCT Contract: poslední dimenzionální tabulka označená jako DCT Contract obsahuje úplně všechny schválené smlouvy.

Tento výsek datového modelu se může zdát na první pohled jako složitý a nepřehledný, ale opak je pravdou. Běžně se lze setkat s takovými modely, které vyžadují delší čas k jejich pochopení a proniknutí do logiky. Z části k tomu pomáhá vhodně vytvořená dokumentace, která ale není vždy k dispozici z důvodu ať už kapacitních či časových nebo jen z důvodu lenosti a nezáživnosti u tohoto druhu práce.

4.5. Optimalizační metody

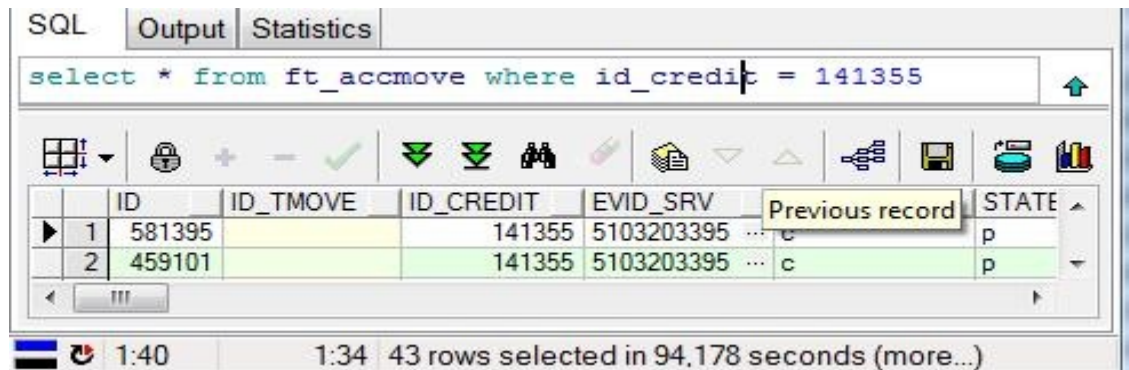
V úvodu této práce jsem se zmínil o využití optimalizačních metod vyšší úrovně, které umožňuje databáze Oracle. Nestává se často, že se s nimi běžný uživatel setká, ale v nadnárodních společnostech či ve společnostech, které mají databáze o velikosti několika Terabyte, jsou tyto metody nezbytné pro perfektní běh tak, aby sloužili co nejlépe a nejefektivněji uživatelům. Není přímo nezbytné, aby uživatelé museli znát dotazovací jazyk SQL³⁸, ale v mnoha případech se vyžaduje a představuje tak na pracovním trhu velkou výhodu. Představit tento jazyk v celé jeho šířce by přesahovalo rozsah této práce a tak jen velmi stručně popíši, co tento jazyk obnáší. Je to zejména pět základních příkazů a to: SELECT, INSERT, UPDATE, DELETE a ALTER přičemž pro uživatele je nejdůležitější si osvojit příkaz SELECT, který využívá uživatel k výběru dat, se kterými chce pracovat. Tyto příkazy jsou jednoduché a dají se naučit v průběhu několika dní maximálně týdnů. Hlubší znalost samozřejmě vyžaduje delší časový interval k získání zkušeností a využití maximální síly SQL. Pro práci business analytika je často velmi nezbytné tyto znalosti mít, stejně jak je nezbytné pro DWH a jeho vývojáře znát možnosti optimalizací a využívat je právě z důvodu obrovského množství dat.

4.5.1 Indexování

Indexy jsou základním optimalizačním nástrojem a jsou nezbytné pro zvyšování výkonnosti. Index se tvoří nad jedním či více sloupci tabulky. Index je vhodné vytvořit v případě, kdy se často potřebuje méně než 15 % řádků z velké tabulky. Uvedené procento samozřejmě závisí na mnoha faktorech. Dalším důvodem pro vytvoření indexu

³⁸ Zkratka SQL značí Structured Query Language. Jazyk v sobě zahrnuje nástroje pro tvorbu databází (tabulek) a dále nástroje na manipulaci s daty (vkládání dat, aktualizace, mazání a vyhledávání informací).

je urychlení spojování tabulek. Naopak je zbytečné vytvářet indexy pro malé tabulky. Jakmile ale dotaz začne být pomalý, tabulka přestává být malá. Příkladem může být faktová tabulka FT_ACCMOVE účetních pohybů, která obsahuje cca 24 milionů záznamů, což lze pokládat za průměrně velkou tabulku v HCI.



Jednoduchý dotaz do této tabulky vyhledávající jednu smlouvu trvá přes minutu a půl což je neúměrně dlouho. Pokud bychom měli napsat složitější dotaz s připojenými dalšími tabulky pro výpis dalších atributů ke smlouvě, jako jsou např. vlastněné karty, bankovní účet, adresa a jiné mohl by dotaz trvat i několik desítek minut. V prováděcím plánu je vidět, že Oracle čte tabulku úplně celou viz. TABLE ACCESS FULL. Znamená to, že databáze musí projít všech 24 milionů záznamů, aby nám vrátila 43 vytoužených řádků.

Description	Object owner	Object name	Cost	Cardinality	Bytes
SELECT STATEMENT, GOAL = ALL_R...			991...	3279	2259...
TABLE ACCESS FULL	SA	FT_ACCMOVE	991...	3279	2259...

Po vytvoření indexu nad odkazovaným sloupcem ID_CREDIT tedy identifikátorem smlouvy je výsledek značně svižnější. Z původních 94 vteřin se dotaz zkrátil na půl vteřiny, což je značný rozdíl a rozhodně komfortnější pro uživatele.

SQL		Output	Statistics			
<pre>select * from ft_accmove where id_credit = 141355</pre>						
	ID	ID_TMOVE	ID_CREDIT	EVID_SRV	TYPE_MOVE	S
▶	1	581395	141355	5103203395 ...	c	p
	2	459101	141355	5103203395 ...	c	p
1:40 47 rows selected in 0.468 seconds (more...)						

Dle prováděcího plánu je vidět, že databáze již neprochází všechny záznamy v tabulce. Index je jakýmsi rejstříkem, který umožní rychlé vyhledání požadovaných dat bez nutnosti prohledat celou databázi.

Optimizer goal					
Description	Object owner	Object name	Cost	Cardinality	Bytes
SELECT STATEMENT, GOAL = ALL_R...			74	84	57876
TABLE ACCESS BY INDEX ROWID	SA	FT_ACCMOVE	74	84	57876
INDEX RANGE SCAN	SA	ID_CREDIT_ACC_IDX	3	84	

Samozřejmě, že to není zadarmo, index zabírá určité místo na disku, a proto není vhodné zakládat indexy nad všemi sloupci, ale pouze nad těmi nejselektivnějšími či nad těmi, které se nejčastěji objevují v podmínkách dotazů.

Indexů existuje celá řada od normálních, jaký jsem použil v tomto případě, jedná se o:

- cluster indexy: přístup k datům je ještě rychlejší, ovšem náklady za to jsou vyšší
- unique indexy: znamená, že hodnota ve sloupci je jedinečná. Nevýhodou je pomalejší zápis do tabulky, kdy databáze právě tuto jedinečnost musí hlídat
- složené indexy: slouží k indexování dvou sloupců zároveň. Je to velmi vhodné, pokud jsou tyto sloupce použity v podmínkách dotazu.
- bitmapové indexy: je vhodný, pokud počet různých hodnot ve sloupci je menší než 1% počtu řádků v tabulce

4.5.2 Partitioning tabulek

Další z nejvíce využívajících optimalizačních metod vedoucí ke zvýšení výkonu patří partitioning využívající se zejména u velkoobjemových tabulek. Slouží k fyzickému rozdělení rozsáhlých datových tabulek do menších částí. To se děje za pomoci

logického členění dat v tabulce např. dle data či číselné sekvence. Takto se tabulka, která, může mít několik giga prostoru, se rozdělí na x tabulek s menší velikostí. Dotaz do této tabulky bude směřovat pouze do té části, kterou uživatel bude právě selektovat. Příkladem může být tabulka smluv, která obsahuje datum schválení smlouvy. Pokud uživatel bude chtít odfiltrovat všechny smlouvy schválené po určitém datu a tato tabulka bude napartišnovaná právě dle hodnoty tohoto sloupce, tak databáze nebude prohledávat celou obrovskou tabulku smluv, ale právě jen tu část, která je ve vyhledávacím intervalu. Jednoduše si lze představit partitioning po měsíci a poté bude po roce používání tabulka smluv rozdělena do dvanácti samostatných částí i když se bude zobrazovat jako jedna jediná. Vzniká tak tabulka s dynamicky rostoucím počtem partitions, tzn., že každý měsíc přibude jedna nová.

Partitioning je jedna z nejsilnějších optimalizačních metod a proto je také hojně využívána v oblasti DWH.

4.5.3 Hints

Žádný software není dokonalý a s databází to není jinak. Nežádá kdy se Oracle rozhodne pro naprosto nevhodný prováděcí plán a dotazy uživatelů se tak mohou neplánovaně velmi prodloužit a v horším případě naplánované úlohy přes noc ani nedoběhnou. Proto se v pravidelném intervalu provádí tzv. statistiky nad kritickými tabulkami. Tyto statistiky sice trvají často velmi dlouhou dobu a jsou proto plánovány na dobu, kdy není systém zatížen, např. o víkendu, ale databáze si tak aktualizuje počet řádků v tabulce a jiné atributy sloužící pro co nejvhodnější zvolení prováděcího plánu při dotazování nad těmito tabulkami. Někdy, ale bohužel ani toto nestačí a tak si musí zkušený vývojář nebo uživatel vypomoci s tzv. hinty. Optimalizační hinty jsou vlastně doporučení Oracle SQL optimalizátoru, které by byly za normálních okolností použity, ale prováděcí plán je nepoužije.

Hinty dokumentované v Oraclu lze rozdělit do čtyř skupin:

- optimalizační: slouží pro co nejrychlejší výsledek dotazu např. `FIRST_ROWS` pro co nejrychlejší dodání prvních řádků
- přístupové: prováděcí plán někdy místo, aby využil založených indexů, zvolí jinou (často špatnou) variantu a dotaz je tak velmi pomalý. Pomocí hintu `FULL`, `INDEX`, `NO_INDEX` a další, poradíme databázi jak by měla dotaz zpracovávat.

- pořadí spojení: tyto hinty jsou zvláště důležité a hojně využívané, jedná se např. o LEADING či ORDERED kdy říkáme jakou tabulku by měl při čtení prováděcí plán upřednostnit a dále se jedná o způsob slučování tabulek, to jsou např. USE_HASH, USE_NL nebo USE_MERGE
- paralelní: určující paralelní zpracování PARALLEL či PARALLEL_INDEX

Pokud budu vycházet z předchozího příkladu, kdy jsem nad faktovou tabulkou založil index pro rychlejší zpracování dotazu, může se stát, že plán bude tento index ignorovat a dotaz bude stále trvat minutu a půl. V tomto případě se využije hint INDEX, a vnutím tak plánu tento index použít.

Příklad využití tohoto hintu je následující:

```
,select /* + INDEX (ft_accmove ID_CREDIT_ACC_IDX)*/ * from ft_accmove where id_credit = 141355'
```

Přičemž ID_CREDIT_ACC_IDX je název našeho založeného indexu.

Všechny tyto hinty slouží k nasměrování jít po správně cestě, neboť plán využívá jen statistiky a ty mohou být buď zastaralé, nebo nedostatečně podrobné. Navíc není možné sesbírat všechny možné statistiky, protože by to vyžadovalo příliš mnoho výkonu.

Hinty mají i své stinné stránky. Pokud tyto hinty vývojář či uživatel zanesou do svých SQL dotazů trvale, mohou v budoucnu nadělat spíše více škody nežli užitku. Tabulka může např. v čase růst a tato neznalost programátora, který si myslí, že bude vždy malá, v konečném důsledku po optimalizaci způsobí snížení výkonu. Další negativní důsledkem této optimalizační metody je odsouvání či spíše odmítnutí skutečné nápravy jak zvýšit výkon databáze. Pro příklad si lze představit špatně provedenou dekompozici, která způsobí značné snížení výkonu. Uživatel nebo programátor raději využije hint pro zvýšení výkonu neboť je to rychlejší a méně pracné nežli provést dekompozici pořádně a správně, což rozhodně není správná volba.

Nicméně všechny optimalizační metody, které jsem uvedl, jsou naprosto nezbytné při vývoji nového DWH a nesmějí být opomíjeny.

4.6 Ekonomické zhodnocení projektu

Spočítat, jaké náklady si vyžádá projekt DWH, kolik se dá ušetřit a jak dlouho bude trvat návratnost tohoto systému je často velmi diskutabilní. Jednou ze základních vlastností u těchto projektů je právě velmi nízká úroveň dosažení těchto informací neboli „Cost Adjustment“, tedy vyčíslení přímých vazeb mezi náklady a přínosy systému DWH.

DWH systém slouží zejména jako podpora pro strategické a operativní rozhodování vrcholového managementu. Jinými slovy vždy záleží na schopnostech a umu koncových uživatelů využít a vydolovat potřebná data a využít je nejlepším možným způsobem pro dosažení cílů firmy. DWH systém jako takový slouží především, jako přidaná hodnota těmto koncovým uživatelům mezi kterými jsou zejména analytici a střední management.

Při určování výnosu z tohoto systému lze vycházet jen z přibližných odhadů, získaných ze zkušeností z obdobných projektů. Tyto informace, lze však jen těžko brát za všeobecné, vzhledem k sofistikovanému řešení a navíc se většinou jedná o řešení konkurenčních subjektů a získat tak tato data je velmi malá. Určit relativní náklady lze určit docela přesně. Jedná se zejména o režijní náklady, které doprovází tento systém po celou dobu jeho životnosti. Tyto náklady lze změřit pomocí tzv. TCO neboli „Total Cost of Ownership“

4.6.1 TCO

TCO stanovuje, kolik pořízení a provozování systému bude stát v průběhu let, kdy se bude využívat.

Komponenty TCO

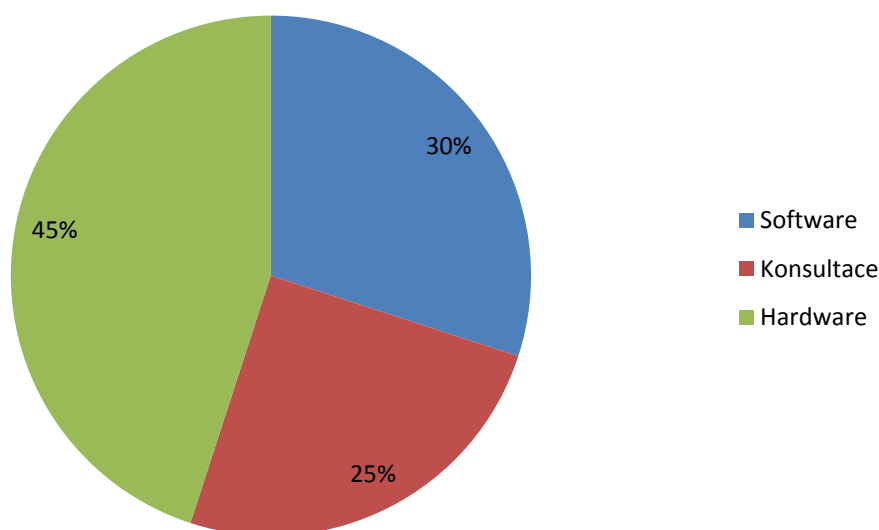
V zásadě lze náklady na vlastnictví systému rozdělit na:

- pořizovací náklady
- průběžné náklady na obsluhu systému
- průběžné náklady na úpravy, přizpůsobování a funkční rozšiřování systému

- náklady na rozšiřování (velikosti) systému

Při odhadu celkových nákladů lze vycházet z podkladů společnosti Meta Group a zveřejněných výsledků oficiálních benchmarkových testů nejznámějších technologií datových skladů. Z analýzy společnosti Meta Group (2009) vyplývá, že u dosavadních projektů v oblasti datových skladů vypadá typické rozložení nákladů následovně:

Graf 2: Odhad celkových nákladů DWH dle TCO



Zdroj:

http://www.readme.cz/web/read_me.nsf/e370b809ef7dabab4125686a0046ff2c/c3f1bcd53a09ef77412568850055e4c5?OpenDocumen, Vlastní tvorba

Ta stejná studie udává, že náklady na provoz typického datového skladu za pět let představují trojnásobek původního rozpočtu. Důvodem je zejména to, že nárůst velikosti datového skladu představuje reálně v průměru kolem 40% ročně, což je 4 x více, než většina projektů optimisticky předpokládá.³⁹

³⁹ KYJONKA, V. *TCO* [online]. 2011 [cit. 2011-04-26] Dostupné z www: <http://www.readme.cz/web/read_me.nsf/e370b809ef7dabab4125686a0046ff2c/c3f1bcd53a09ef77412568850055e4c5?OpenDocumen>

4.6.2 Náklady spojené se změnou systému ve společnosti HCI

Náklady spojené se změnou systému nebudou tak velké jak při implementaci nového řešení. V zásadě se jedná o nákup dodatečného hardware, neboť stávající řešení musí paralelně pracovat s původní, starší verzí což, ale zvyšuje mzdové náklady z důvodu správy stávajícího DWH a zároveň vývoje nového DWH. Mimo to budou implementovány také nové BI nástroje pro práci uživatelů a z toho vyplývající náklady na školení jak programátorů, tak samotných uživatelů k osvojení si znalostí a funkcí těchto nástrojů. Mezi nezanedbatelné náklady patří také konzultace s odborníky, kteří mají bohaté zkušenosti s implementací obdobných systémů. Dle TCO je také nezbytné kalkulovat s náklady budoucími jako je správa a nákup nových diskových polí a procesorů dalších nezbytných hardwarových prostředků, upgrade současných BI nástrojů a také spravování datového skladu.

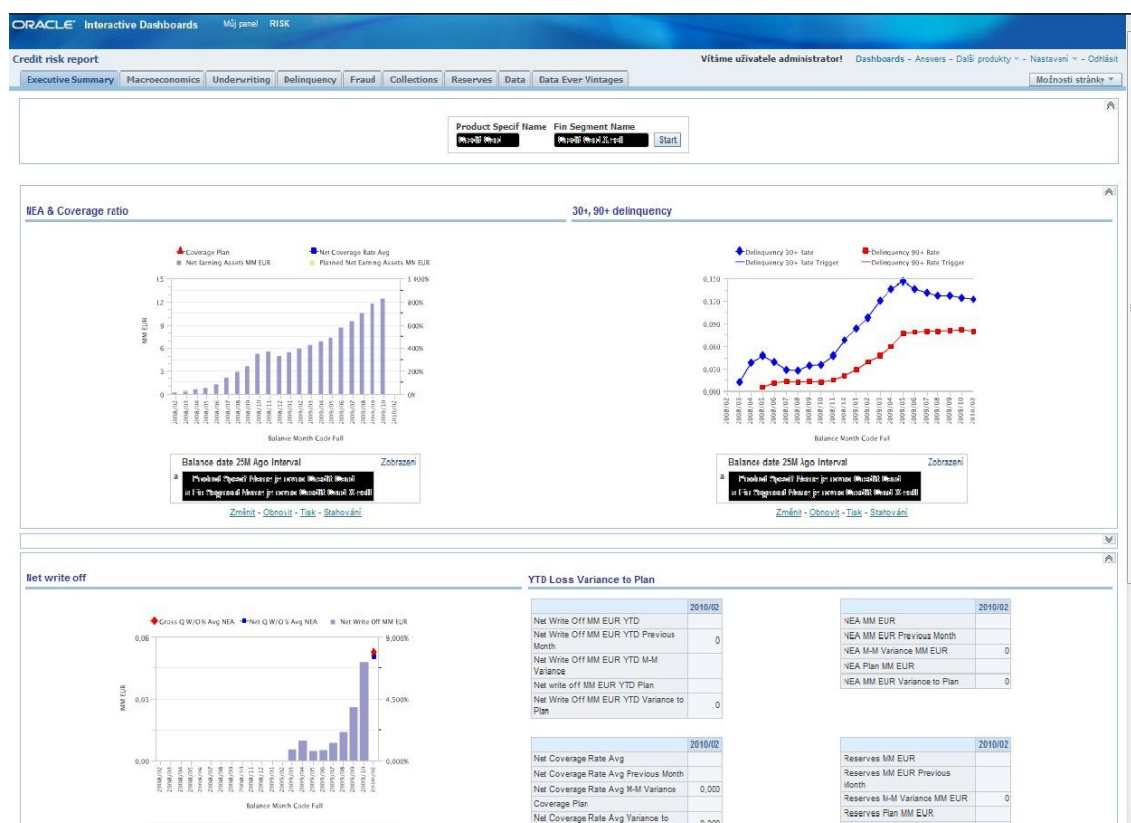
4.6.3 Reportovací nástroj

Velkou výhodou při užívání databáze Oracle je i využití i nástrojů z této firmy zabezpečující dokonalou kompatibilitu. Většinu nástrojů a software společnost Oracle buď naprogramovala ve vlastní režii, ale spíše je odkoupila jako hotová řešení a zároveň takto eliminovala konkurenci. Z pohledu nákladů, implementace a užívání není tato informace pro uživatele či investora rozhodujícím faktorem. Hlavním rozhodujícím faktorem je cena a přidaná hodnota pro uživatele. I když nákup nového BI řešení není nezbytným krokem, protože při návrhu změny aktuálního systému již reportovací nástroje, se kterými uživatelé operují, existují (jsou to Business Objects), přece jen je vhodné i toto řešení aktualizovat a přejít na novější a modernější řešení.

Společnost Oracle uvedla na trh tzv. nástroj Oracle Business Intelligence (dále jen již OBI) což je vysoce škálovatelná sada pro zákazníky, sloužící pro získání lepšího přehledu o chodu firmy. Prostředí poskytuje přístup a analýzu dat nacházejících se v relačních, OLAP a XML zdrojích dat. OBI poskytuje zodpovězení všech problematických podnikových otázek. To sice může znít jako laciný reklamní slogan, ale šikovný uživatel se správnými zdrojovými daty by měl skutečně umět odpovědět na jakoukoli otázku z pohledu businessu a přetvořit ji do vzhledného reportu, grafu či dashboardu. OBI umožňují zkoumání informací zobrazující se v datových grafech,

kontingenčních tabulkách a sestavách. Uživatel dále může tyto data ukládat, třídít a sdílet jejich obsah. Požadavky vytvořené v OBI lze ukládat do různých formátů včetně formátu „html“ pro možnosti vystavení reportu na internetu nebo vytvářet tzv. Dashboardy. Dashboard je tzv. interaktivní řídicí panel, který může obsahovat grafy, tabulky, obrázky, mapy atd. a to vše na jednom místě. Příklad takového dashboardu je na obrázku č. 10.

Obr. 10: Příklad Dashboardu



Zdroj: Vlastní tvorba

Tento nástroj je interaktivní a velmi intuitivní pro uživatele. Z toho vyplývá velmi přátelské uživatelské prostředí umožňující i uspořeni nákladů za školení. Mířím tím k možnosti vyškolit pouze určitý segment uživatelů, kteří poté zaškolí ostatní uživatele. Rovněž lze čerpat z rozsáhlé dokumentace a propracovaného helpu.

Náklady na pořízení jedné licence tohoto produktu se pohybuje v ceně okolo 100 000 Kč. Náklady na zaškolení jednoho pracovníka poté 40 000 Kč. Náklady na deset licencí

OBI a vyškolení dvou pracovníků poté vycházejí na 1 080 000 Kč. Tento nástroj byl vybrán na základě nejlepšího poměru cena x výkon a zejména také na doporučení uživatelů, kteří s ním mají dlouhodobě dobré zkušenosti.

4.6.4 Pořizovací náklady

Z důvodu rozšiřování stávajícího řešení není naštěstí potřeba zakupovat další licence na databázi Oracle neboť společnost vlastní multilicenci, což je z pohledu licencové politiky pro HCI velmi výhodné. Rovněž není potřeba kupovat licence na vývojový software nazvaný Oracle Warehouse Builder na tvorbu jednotlivých mapování neboť tyto licence jsou od verze 10gR2 součástí každé edice Oracle databáze (tedy Standard Edition One, Standard Edition a Enterprise Edition), tzn. Oracle Warehouse Builder je zdarma (dříve byl součástí balíku Oracle Developer Suite v ceně \$5.000,-). Nevýhodou může být, že tento produkt již není podporován a navíc obsahuje větší množství chyb, a tak je nutné pro práci s ním delší zkušenost a také velkou dávku trpělivosti. Čemu se, ale nelze vyhnout jsou dodatečné náklady na nákup potřebného hardware a to již je nezanedbatelná položka v rozpočtu. Z tohoto pohledu bude nutné nakoupit procesory, disková pole, paměti a další nezbytné prostředky pro běh systému. V kostce se dají tyto investice odhadnout na minimálně 10 000 000 Kč.

4.6.5 Náklady na provoz a vývoj systému

Největší část rozpočtu vyhrazenému pro tento projekt budou mzdové, konzultační a také školící náklady. Vzhledem k tomu, že systém bude spravován a vyvíjen přímo společností HCI a.s. je nutné najmout na trvalý zaměstnanecký poměr minimálně pět nových vývojářů, kteří se přidají k již šestičlennému DWH týmu. Tato investice bude ovšem z dlouhodobého pohledu výhodná, neboť nových projektů v podobě datových skladů bude nadále plánovaně přibývat a tak neexistuje efektivnější a levnější řešení. Je tak nespornou úsporou mít kolektiv lidí znalých problematiky datových skladů z původního řešení, neboť outsorcovat tento vývoj lidmi z externích firem by stálo desítky miliónů. Rovněž najmutí odborníků a konzultantů pro pilotní projekt a první inkrement je sice velmi drahá záležitost, ale z dlouhodobého pohledu výhodná a nezbytná. Poslední položkou budou potřebná školení na vývojový software Oracle

Warehouse Builder a Oracle Business Intelligence. Toto ovšem budou jednorázové neopakující se investice a navíc do vlastních zaměstnanců.

Konzultační náklady se budou skládat z najmutí minimálně 4 externích lidí na dobu šesti měsíců s náklady 20 000 za člověka na den. To ve výsledku činí náklady ve výši 9 600 000 Kč. Dále je potřeba na trhu práce získat minimálně pět nových vývojářů s průměrně hrubým platem 40 000 Kč měsíčně což pro zaměstnavatele znamená měsíční náklady ve výši 53 600 Kč. V konečném součtu jen pro projekt, který se plánuje dokončit, vynaložit 321 600 Kč. Posledními náklady spojenými s lidskými zdroji budou školení na OWB a OBI. Tato školení jsou minimálně dvoudenní a jejich cena za člověka na den se pohybují okolo 40 000 Kč. Minimální počet školitelů je pět. Po součtu nákladů za tyto služby se cena bude pohybovat okolo 400 000 Kč.

4.6.6 Výsledná kalkulace nákladů

V předchozích bodech jsem vyčíslil nezbytné náklady potřebné k úspěšnému dokončení projektu. Některé z těchto nákladů jsou pouze jednorázové, a při zavádění dalšího datového skladu, do nové země již nebudou potřeba. Jedná se zejména o školení na OWB či OBI a externí konzultanty, přičemž cena za tyto služby je velmi vysoká. Některé náklady, ale naopak budou neustále stoupat. V tomto případě mám na mysli zvyšující se nároky na hardwarové prostředky spolu se zvyšujícím se objemu dat v systému a také tlak na vyšší mzdové nároky v budoucnosti. V tabulce č. 6 je uvedena kalkulace pro vývoj nového řešení datového skladu.

Tab 6:kalkulace nákladů

Náklady na vybudování DWH	
Licence Oracle Business Intelligence	1 000 000,-
Hardware	10 000 000,-
Konzultace	9 600 000,-
Mzdové náklady	321 600,-
Školení	400 000,-
Celkem	21 321 600,-

4.7 Přínos nového řešení

Každá společnost se snaží investovat své prostředky do sofistikovaných BI řešení a očekávají, že díky komplexní analýze dat z provozních systému získá dříve data, která dříve nebylo možné vydolovat a umožní jim tak lépe poznat potřeby zákazníků a tím zvyšovat svůj tržní podíl a dosahovat lepších výsledků podniku. Klíčem tedy je zajištění kvality dat a následně informací, zejména z pohledu konsolidace, úplnosti a korektnosti, a také aktuálnosti a konzistence.

Studie IDC, která hodnotila 62 implementací datových skladů, prokázala celkovou návratnost investic na projekt datového skladu ve výši 401 %, s průměrnou dvou- až tříletou návratností. Při stejném výzkumu data martů byla zaznamenána ROI 533 %. Ostatní typické ukazatele přínosů včetně míry úspory nákladů a měřitelných ukazatelů efektivnosti se rovněž výrazně zlepšily.⁴⁰ Následná studie je důkazem toho, že datový sklad přínosný určitě je, záleží ovšem na mnoha faktorech vedoucí k úspěšnému řízení tohoto systému.

4.7.1 Kvalita dat

Jedním z klíčových faktorů, vedoucí k úspěšnému systému, který je pro společnost přínosem, jsou kvalitní data. Není třeba rozebírat odbornými detaily pojem kvality informací. Zjednodušeně se dá říct, že kvalita informace = soulad s realitou. Příkladem může být zákazník vlastníci spotřebitelský úvěr a debetní kartu. Všechny tyto informace by měly být vidět u záznamu našeho zákazníka. Navíc by měly být správně vyplněny i další údaje jakými jsou adresa (i více), telefonní kontakty a další. Hlavním problémem je nedostatek výše uvedených údajů a jejich roztroušenost po celém systému, několikrát duplikovaných a v různých verzích. Úlohou datového skladu je tyto informace synchronizovat a správně prezentovat uživateli. Jistě se již každému vlastníku účtu stalo, že mu byla nabídnuta např. debetní karta prostřednictvím telefonického hovoru, poté nabídka přišla poštou a konečně nabídka nechybí ani v měsíčním poštovním vyúčtování, přestože první dvě nabídky byly odmítnuty. V konečném důsledku je

⁴⁰ KELLEY, CH. *The IDC data warehousing ROI study: An analysis* [online]. 2011 [cit. 2011-05-05]. Dostupné z [www : <http://searchsqlserver.techtarget.com/tip/The-IDC-data-warehousing-ROI-study-An-analysis>](http://searchsqlserver.techtarget.com/tip/The-IDC-data-warehousing-ROI-study-An-analysis)

minimálně značně poškozena důvěryhodnost společnosti. Tyto aktivity jsou vyústěním práce jednotlivých oddělení, která pracují separátně a využívají různé zdroje dat. Právě vhodně navržené BI prostředí tyto nedostatky odstraní. Automaticky se očekává, že uživatelé nového řešení budou pečlivě analyzovat data v systémech a výsledky analýz pak pomohou k dalšímu růstu obratu, prodeji nových produktů a oslovení více zákazníků.

4.7.2 Přínos uživatelům

Datový sklad je natolik úspěšný, nakolik jsou schopní jeho uživatelé spolu s kooperací středního a vyššího managementu ho využívat. DWH lze asociovat s populární dětskou stavebnicí Merkur, pomocí níž dokázal Oto Wichterle zkonstruovat přístroj na odstředivé lití kontaktních čoček v roce 1963. Stejně tak jak tato populární dětská stavebnice umožňuje nepřeborné možnosti realizací, tak i datový sklad nabízí vytvářet jakékoli analýzy, jež jsou omezené pouze uživatelským umem. Hlavním přínosem tohoto systému je přidaná hodnota koncovým uživatelům, kteří tak jeho prostřednictvím mohou naplňovat cíle firmy. Tento přínos nelze exaktně číselně měřit neboť sestavy, které analytici vytvářejí, jsou většinou informativního charakteru pro operativní a strategické rozhodování středního a vyššího managementu. Může se jednat o denní reporty informující o počtu schválených smluv, druhů produktu typu spotřebitelský účet, revolvingový účet, car loan, počtu dlužníků, počet dní po splatnosti apod. Tyto a obdobné reporty nepřinášejí okamžitě finanční úsporu či zisk, ale na jejich základě se buduje strategie firmy v krátkodobém i dlouhodobém období. Záleží tak šikovnosti, nápadu a efektivnosti zadavatele reportu, jeho využití a samozřejmě na kvalitě dat.

4.7.3 Podpora primárního systému

Datový sklad neslouží jen pro analytiku jakožto koncové uživatele pro dolování dat a tvorbu reportů, ale také z velké části ulehčuje zátěži provoznímu systému. Primární systém musí mít v každém ohledu co možná nejmenší dobu odezvy pro co nejpohodlnější práci uživatelů. To může být mnohdy problém, zvláště když systém musí např. na základě vstupních parametrů provádět tzv. ‚scoring‘ neboli ohodnotit klienta a nabídnout mu adekvátní výši půjčky. V tomto případě se jedná o velmi složitý algoritmus, který musí vyhodnocovat mnoho podmínek a provádět tyto výpočty na

úrovni primárního systému by jej velmi zatížilo a tím i zpomalilo. Datový sklad tak před počítává veškerá data potřebná pro tuto funkcionalitu, naplní příslušné tabulky v primárním systému a ty již jen pouze zobrazí výsledek. Takových podpor je velké množství a vypomáhá tak snížit uspořít strojový čas pro aktivity výhradně sloužící v primárním systému.

4.7.4 SOLUS

V předešlých odstavcích jsem popisoval přínos datového skladu z pohledu uživatelů či podpory primárního systému. Tento přínos není exaktně změřitelný a slouží zejména pro potřebu zkvalitnění a pohodlnosti práce samotných uživatelů a tím zajišťující i lepší výsledky společnosti. Nyní bych se rád zaměřil na pár úloh, které datový sklad plní, a ze kterého lze přesně vyčíslit úspory pro firmu. V tomto případě se jedná o SOLUS.

SOLUS je zájmové sdružení právnických osob, jehož cílem je v rámci tzv. odpovědného úvěrování přispívat k prevenci předlužování klientů, k prevenci růstu počtu dlužníků v prodlení, přispívat ke zvyšování vymahatelnosti stávajících dluhů po splatnosti a snižovat potenciální finanční ztráty věřitelů. Sdružení SOLUS bylo zaregistrováno a fakticky zahájilo činnost v červnu 1999. SOLUS vytváří dva negativní registry klientských informací, které shromažďují informace o klientech, kteří se dostali do problémů se splácením svých závazků u některého z členů sdružení SOLUS. Jde o Registr FO, do kterého jsou zařazovány fyzické osoby (spotřebitelé), a Registr IČ, do kterého jsou zařazovány fyzické osoby podnikatelé a právnické osoby. Negativní registry klientských informací SOLUS fungují na základě on-line dotazů a jednoznačně interpretovatelných odpovědí, které jsou poskytovány v čase kratším než jedna vteřina.

Aktivity sdružení SOLUS mají hmatatelné přínosy nejen pro členské společnosti, ale také pro spotřebitele. Klienti, kteří splácejí bezproblémově své závazky u členů sdružení a nemají v registru klientských informací SOLUS záznam, mají často snazší přístup ke službám členů sdružení.⁴¹

Cílem datového skladu je složitý algoritmus, který pro všechny smlouvy v systému vypočítává počet splátek, počet dlužných splátek, výši případného dluhu a mnoho

⁴¹ STAŠKA, A. *Zájmové sdružení právnických osob* [online]. 2011 [cit. 2011-06-05] Dostupné z <<https://www.solus.cz/hlavni-stranka/>>

dalšího. Tyto výpočty by z výkonnostních důvodů nikdy nemohly být prováděny v provozním systému. Vstupních dat je obrovské množství a počet všech atributů, na jejichž základě se rozhoduje o tvrdosti příznaku lze počítat na desítky. Z toho pohledu plní datový sklad velmi důležitou roli jak pro ostatní partnery ve sdružení SOLUS tak i pro společnost samotnou, neboť při poskytování finančních služeb je jedno z nejdůležitějších faktorů pro schválení, právě dotaz do SOLUSu. V případě, že klient figuruje v tomto sdružení z důvodu neplnění podmínek, resp. vzniklého dluhu, má již jen malou možnost získat další produkt. Další úlohu, kterou plní DWH v oblasti SOLUSu je tvorba a nápočet detailních reportů. V případě, že klient poprvé vstoupí do této databáze dlužníků, ocitne se v reportu, který je poté podkladem pro zaslání informujícího dopisu. Ve velkém množství případů má tento dopis silný psychologický efekt a klient dlužné splátky uhradí. Jedná se v průměru o několik tisíc lidí měsíčně a úspora z toho plynoucí je v průměru 300 000 Kč, což ročně ušetří společnosti na vymáhání minimálně 3 600 000 Kč a to vše pouze díky složitým nápočtům v DWH, který by primární systém nebyl schopný napočítat.

4.7.5 Marketingové kampaně

V DWH se vytváří velké množství podnikových kampaní zaměřených na cíleně vybrané konkrétní skupiny zákazníků s cílem nabídnout produkt optimálním způsobem. Nezbytným předpokladem úspěšného využití je kvalitní co možná nejrozsáhlejší databáze firemních údajů, které slouží k přesnému výběru konkrétních adres ze zvolených cílových skupin, což datový sklad splňuje. Jedná se o kampaně s nabídkami např. o překlopení stříbrné karty na zlatou, nabídky karet s určitým rámcem, nabídky osobních úvěrů, zvýšení rámce na kartě apod. Tyto kampaně se setkávají s velkým úspěchem a generují společnosti příjem v průměru půl miliónu měsíčně. V ročním horizontu se tak tento způsob oslovování klientů vyplatí ve finančním hodnocení v minimální výši šesti miliónů.

4.7.6 Telemarketing

Telemarketing představuje službu komunikace se zákazníky prostřednictvím telefonu. Dnes se to již stává součástí života každého moderního člověka. Lidé ušetří čas a tak využívají výhod telefonu či internetu a společnosti naopak mohou využít zrychlené

komunikace se zákazníkem. Pro oblast DWH to znamená zejména selektovat vybrané okruhy zákazníků a ty dále předávat našim operátorům. Význam účasti DWH v tomto případě má stejný přínos jak při výběru klientů u marketingových kampaní. Datový sklad, ale může zajít ještě dál a vybírat pouze ty klienty, kteří jsou z dlouhodobého hlediska nesolventní a na základě výběru je mohou telefonistky upozorňovat na blížící se termín splatnosti jejich pohledávky, resp. pravidelné splátky. V případě uzavření smlouvy musí klient vyplnit své soukromé telefonní číslo, číslo do zaměstnání apod. Datový sklad poté vybírá klienty, kteří se dostali do předem daných dnů po splatnosti a systém vygeneruje jejich seznam s výši nesplacené částky a ostatními údaji. Na tento seznam poté odejdou upozorňující sms. Stejně tak jak v případě odeslání informujícího dopisu zákazníkovi o hrozícím vstupu do SOLUSu, který má silný psychologický efekt, tak i tato sms má velmi silný účinek. Z mých databázových výpočtů dochází díky této metodě k uhrazení pohledávky po zaslání sms k více jak 50% úspěšnosti což společnosti uspoří náklady za vymáhání měsíčně minimálně 100 000 Kč a ročně poté 1 200 000Kč. Datový sklad sice nemá přímou úměru na těchto zaplacených pohledávkách, neboť mohlo dojít v mnoha případech pouze o např. opomenutí, ale zásluhu DWH nelze odepřít.

4.8 Vyhodnocení prospěšnosti systému

V předchozích kapitolách jsem popsal některé úlohy datového skladu, jež jsou přínosem pro společnost. Jedná se pouze jen o část úloh, které musí DWH provádět, ale i to stačí na předpokládanou návratnost prostředků vložených do systému v rozmezí dvou až tří let. Nicméně z mého pohledu má datový sklad největší přínos právě v úlohách, které nelze exaktně měřit, a sice analýzu firemních dat, na jejichž základě management může provádět operativní a strategické plánování.

Hlavním přínosem datového skladu je přidaná hodnota pro uživatele a jak jej uživatel využije, závisí už pouze na něm samotném.

V tabulce č. 6 jsem nastínil náklady potřebné pro vybudování datového skladu. Nyní tyto náklady lze porovnat s níže uvedenými přínosy a úsporami za rok. Náklady byly vyčísleny na 21 321 600 Kč.

Tab č. 7: Porovnání nákladů a přínosu DWH

Přínosy a úspory DWH	
SOLUS	3 600 000
Marketingové kampaně	6 000 000
Telemarketing	1 200 000
Celkem	10 800 000,-

Návratnost je v tomto případě do dvou let, čímž se potvrzuje studie IDC, která předpokládá návratnost s průměrnou dvou – až tříletou dobou. V dalších letech bude potřeba do systému nadále investovat peněžní prostředky a to zejména v oblasti hardware neboť zvyšující se objem dat bude čím dál více zatěžovat paměť, strojový čas a v neposlední řadě bude docházet místo na diskových polích. V tomto čase by už měl být, ale každý uživatel včetně managementu přesvědčen o přínosu tohoto systému a jeho potenciálu. Návratnost investic do výkonu systému bude daleko dříve a umožní rychlé výpočty a dolování dat s přijímatelnou dobou odezvy pro co nejefektivnější práci uživatelů. Mezi další náklady lze uvést upgrady databáze a BI nástrojů na nové verze, ale výše těchto nákladů by byla velmi spekulativní, neboť jsou tyto výdaje závislé na uzavřených licenčních podmínkách a tato licenční politika je v mnoha případech pro společnost velmi výhodná, pokud je dobře vyjednána. Spolu s rostoucími náklady by měly jít ruku v ruce i rostoucí výnosy tohoto systému např. ve formě budování nových modulů jako může být např. sekuritizace, neboli odprodej kvalitních smluv třetí straně či rozšíření podpory primárního systému tak, aby byl systém čistě transakční a veškeré datové manipulace se prováděly na úrovni datového skladu.

5. Závěr

Datové sklady mají schopnost pojmout stovky gigabytů podnikových dat z mnoha systémů a tyto pak přeměňovat ze surových na užitečné informace, které se využívají při operativních a strategických rozhodnutí podniků s cílem zvyšování kvality a produktivity, snížení nákladů nebo nalezení slabých míst. Analytici mohou prostřednictvím BI nástrojů denně kontrolovat hospodaření společnosti a okamžitě přijímat opatření. Většina řídicích vedoucích pracovníků a manažerů je již přesvědčeno o přínosu datového skladu, ale již nevědí jak takový datový sklad do firmy implementovat a jak začít. Na tuto otázku jsem se snažil odpovědět v této práci, kdy jsem popsal nejčastěji používané koncepce budování DWH dle Ralpa Kimballa a Williama Inmona. Oba přístupy jsou odlišné a velmi osobité. Řešením je velmi často kombinace obojího dle požadavků firmy. Univerzální softwarové řešení, které by se pouze nakonfigurovalo, neexistuje a nikdy existovat nebude vzhledem k vysoké míře sofistikovanosti tohoto systému. Jedná se vždy o unikátní řešení šité na konkrétního zákazníka, respektujícího používaný informační systém, požadavky na dostupnost informací aj. Náklady na vybudování tak rozsáhlé datové základny jsou relativně vysoké a i přesto, že návratnost je v průměru dvou- až tříletá, je implementace tak unikátního řešení dostupná zejména velkým firmám. Hlavním přínosem datového skladu je pak v největší míře přidaná hodnota pro uživatele. Tato přidaná hodnota není exaktně měřitelná a závisí vždy na schopnostech uživatele, nakolik dokáže prostřednictvím BI nástrojů, využít data agregovaná v DWH ve prospěch firmy. Neméně důležitým přínosem pro firmu je také podpora primárního systému ve formě před počítávání dat či datových manipulací. Tato podpora výrazně odlehčí informačnímu systému, který je tak více uživatelský příjemnější z pohledu strojového času. Mezi další přínos či úsporu, které jsem v této práci uváděl, mohou patřit jednotlivé moduly sloužící k výběru marketingových kampaní, napočítávání dat pro telemarketing či identifikaci dlužníků splňující podmínky pro zařazení do celostátní databáze dlužníků. To že se datový sklad firmě z pohledu návratnosti investic vyplatí, je si myslím neoddiskutovatelné a již mnohokrát prokázané, záleží ovšem vždy na tom, jak s ním naloží samotní uživatelé a nakolik jej společně s BI řešením dovedou využít ke splnění vytyčených cílů firmy.

Seznam použité literatury

Knihy:

- 1) GÁLA L., POUR J., TOMAN P. *Podniková informatika*. 1. vydání. Praha: Grada Publishing, 2006. 103 s. ISBN 80-247-1278-4
- 2) INMON, W. *Building the Data Warehouse*, 4th ed. Indiana: Wiley Publishing, Inc., 2005. 495 s. ISBN 13-978-0-7645-9944-6
- 3) IMHOFF C., GALEMMO N., GEIGER J. *Mastering data warehouse design relational and dimensional technique*. 1st ed. Indiana: Wiley Publishing, Inc. 2003. 5 s. ISBN: 0-471-32421-3
- 4) JENSEN S., TORBEN P., THOMSEN C. *Muldtidimensional database and data warehousing* . 1st ed. Denmark: Morgan&Claypool publishers, 2010. 19 s. ISBN: 9781608455379
- 5) KIMBALL, R *The Kimball Group Reader*, 1st ed. Indianapolis: Wiley Publishing, Inc., 2010. 38 s. ISBN: 978-0-470-56310-6
- 6) KIMBALL, R., *The data warehouse toolkit*, 2nd ed. USA: Wiley Compter Publishing 2002. 82 s. ISBN 0-471-20024-7
- 7) KOCH M., DOVRTĚL J. *Management informačních systémů*. 1. vydání. Brno: Akademické nakladatelství CERM, 60 s.
- 8) MALLYA, T. *Strategické řízení*, 1.vyd. Praha: Grada Publishing, a.s., 2006. 31 s. ISBN 80-247-1911-8
- 9) RAIS, K, *Základy optimalizace a rozhodování*. 1. vyd. Brno 2004. 134 s. ISBN 80-214-23280

10) VOŘÍŠEK J. *Strategické řízení informačního systému a systémová integrace*. 1. Vydání. Praha: Management Press, 2002. 12 s. ISBN 80-85943-40-9

11) WREMBELL R., KONCILIA CH. *Data warehouse and olap concepts architecture and solutions*. 1st ed. LONDON: IRM Press, 2007. 22 s. ISBN 1-59904-364-5

Články v časopisu:

12) BRODNÍČEK D. Úvod do problematiky datových skladů. *Automatizace*, 2009, č. 7, s. 428

13) HANUSEK L. Technologie Data Warehousingu a Data Miningu. *Computerworld*, 2007, č. 8, s. 15

14) JEGER, D. Datové sklady a business intelligence. *CIO Business World*. 2011. č.1, s.15

15) KHUDHUR P. Business intelligence: Je třeba přemýšlet. *ComputerWorld*. 2007, č.10, s. 28

16) SEDLÁČKOVÁ H. Trendy v chápání zdrojů podniku při tvorbě strategie podniku. *Acta Oeconomica Pragensia*, 2007, č. 2, s. 107

Elektronické zdroje:

17) ANUPINDI N. *Inmon vs Kimball – An Analysis*. [online]. 2005 [cit. 2011-03-14]. Dostupné z www:< <http://www.nagesh.com/publications/technology/173-inmon-vs-kimball-an-analysis.html>>

18) HABÁŇ J. *Technologie datových skladů*. [online]. 2004 [cit.2011-01-30] Dostupné z www: <<http://www.cvis.cz/hlavni.php?stranka=novinky/clanek.php&id=60>>

- 19) BUSINESS, C. *CPM*, [online 2010] poslední revize: neuvedno. [cit 10.12.2010]
Dostupné z: [www <http://business.center.cz/business/pojmy/p987-CPM.aspx>](http://business.center.cz/business/pojmy/p987-CPM.aspx)
- 20) KELLEY, CH. *The IDC data warehousing ROI study:An analysis* [online]. 2011
[cit. 2011-05-05] Dostupné z [www : http://searchsqlserver.techtarget.com/tip/The-IDC-data-warehousing-ROI-study-An-analysis](http://searchsqlserver.techtarget.com/tip/The-IDC-data-warehousing-ROI-study-An-analysis)
- 21) KUČERA M. *Dva způsoby budování datového skladu*. [online]. 2009 cit [2011-02-02].
Dostupné z [www: <http://www.systemonline.cz/clanky/dva-zpusoby-budovani-datoveho-skladu.htm>](http://www.systemonline.cz/clanky/dva-zpusoby-budovani-datoveho-skladu.htm)
- 22) KYJONKA, V. *TCO* [online]. 2011 [cit. 2011-04-26] Dostupné z [www: <http://www.readme.cz/web/read_me.nsf/e370b809ef7dabab4125686a0046ff2c/c3f1bcd53a09ef77412568850055e4c5?OpenDocument>](http://www.readme.cz/web/read_me.nsf/e370b809ef7dabab4125686a0046ff2c/c3f1bcd53a09ef77412568850055e4c5?OpenDocument)
- 23) SPIDY, P. *Definice cíle SMART* [online]. 2011 [cit. 2011 - 03 - 21]
Dostupné z [www: <http://www.finance-management.cz/080vypisPojmu.php?IdPojPass=39&X=Definice+cile+SMART+Project+Management>](http://www.finance-management.cz/080vypisPojmu.php?IdPojPass=39&X=Definice+cile+SMART+Project+Management)
- 24) STAŠKA, A. *Zájmové sdružení právnických osob* [online]. 2011 [cit. 2011-06-05]
Dostupné z [<https://www.solus.cz/hlavni-stranka/>](https://www.solus.cz/hlavni-stranka/)
- 25) VAVRUŠKA J. *ETL a kvalita dat*. [online]. 2003 cit [2011-01-31]. Dostupné
z [www:<http://www.systemonline.cz/clanky/etl-a-kvalita-dat.htm>](http://www.systemonline.cz/clanky/etl-a-kvalita-dat.htm)
- 26) ZDRAZILOVA, I. *Datový sklad* [online]. 2011 [cit. 2011 – 01-28] Dostupné
z [www: <http://kisk.phil.muni.cz/wiki/Datov%C3%BD_sklad>](http://kisk.phil.muni.cz/wiki/Datov%C3%BD_sklad)

Seznam tabulek, obrázků a grafů

Tabulka č. 1: Hodnocení jednotlivých oblastí metodou HOS 8

Tabulka č. 2: tabulka SWOT analýzy (nejdůležitější fakta)

Tabulka č. 3: Přibližná cenová kalkulace na vybudování 10 TB DWH

Tabulka č. 4: Vstupní údaje do CPM

Tabulka č. 5: Výsledky výpočtu charakteristik

Tabulka č. 6: Kalkulace nákladů

Obrázek č. 1: Vývoj BI technologií (*Zdroj: Imhoff, Galemno, Geiger, 2004*)

Obrázek č. 2: Schéma datového skladu (vlastní)

Obrázek č. 3: Granularita – úroveň detailu (vlastní)

Obrázek č. 4: Integrovaný datový sklad dle Williama Inmona (vlastní)

Obrázek č.5: Integrovaný datový sklad dle Ralpa Kimblla (*Zdroj: Kimball, The Data Warehouse Toolkit. str. 7. Vlastní úprava*)

Obrázek č. 6: Působnost Home Creditu

Obrázek č.7: Home Credit Struktura.(*Zdroj:<http://www.homecredit.net/pub/en/about-Us/structure.html>*)

Obrázek č 8: schéma CPM

Obrázek č. 9: datový model vztahující se ke smlouvě

Obrázek č. 10: Příklad Dashboardu

Graf 1: Grafické hodnocení jednotlivých oblastí metodou HOS 8

Graf 2: Odhad celkových nákladů DWH dle TCO

Přílohy

Dotazník:

Oblast Hardware:

1) Je možné současné HW vybavení označit za moderní a sledující současné trendy?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

2) Přispívá HW pozitivně k rychlosti a použitelnosti informačního systému?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

3) Nákup nového HW je posuzován s ohledem na ergonomii pro jeho uživatele?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

4) Dá se připojení k počítačovým sítím označit za spolehlivé, dostatečně rychlé a vyhovující?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

5) Jsou klíčové prvky HW dostatečně fyzicky chráněny před krádeží, požárem a povodní?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

6) Je nové HW vybavení pořizováno po zvážení jeho kompatibility s existujícím HW vybavením a softwarem, který na něm bude provozován?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

7) Současné HW neumožňuje účinnou výměnu dat s odběrateli či dodavateli?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

8) Je rychle dostupné záložní vybavení v případě výpadku klíčových HW prvků systému?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

9) Souhlasíte s výrokem, že současné HW vybavení bude do dvou let těžkou použitelné?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

10) Jsou poruchy HW vybavení na denním pořádku?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

36-5-2=29/8=4

Oblast Software

1) Poskytuje zkoumaný software všechny funkce nezbytné pro práci uživatelů?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

2) Je grafické členění plochy pro zadávání, editaci vstupních údajů přehledné a přispívá tak ke snadnosti práce se systémem?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

3) Jsou chybová, varovná hlášení či jiné nestandardní oznámení srozumitelná a poskytují na požádání i bližší vysvětlení vzniklé situace?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

4) Rychlost zpracování úkolů jako tisky, dotazy, vyhledávání se jeví jako dostatečně rychlé?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>

5) Platí, že koncoví uživatelé nesmějí poskytovat podněty pro případné úpravy SW, nové nastavení nebo pořízení nových verzí software?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

6) Je nápověda k softwaru srozumitelná a přehledná?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

7) Má zkoumaný informační systém jednotné ovládání obrazovek, menu, sestav a nápovědy?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

8) Jsou při pořízení nových verzí SW využívány jejich nové vlastnosti?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

9) Je pravda, že snadnost používání softwaru koncovými uživateli nehraje roli při jeho pořízení nebo vývoji?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

10) Existují pravidelné nebo nahodilé kontroly sloužící ke zjištění abnormalit ve využívání systému, jeho nesprávného užívání či zneužívání?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>

24-4-1=19/8=2

Oblast Orgware

1) Existují postupy či směrnice pro zotavení IS z nestandardních a havarijních situací a jsou tyto dokumenty dostatečně známé uživatelům?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

2) Existují doporučené pracovní postupy a procedury běžného provozu pro koncové uživatele a jsou udržovány v aktuálním stavu?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

3) Existují pravidla pro bezpečnost IS a obsahují i ustanovení pro nakládání s dokumenty či přílohami e-mailů získaných z Internetu?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

4) Je pravda, že management příliš nedozírá na dodržování pravidel bezpečnosti a provozu IS?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

5) Má každý pracovník jasně určeno, s jakými úlohami smí pracovat a kdy?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

6) Provádějí jakékoliv rozsáhlejší instalace, změny nastavení, připojení nové techniky pověřené osoby, nikoliv uživatelé?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

7) Jsou ošetřeny odchody zaměstnanců a ukončení platnosti jejich přístupových práv?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

8) Existují pravidla nebo politika bezpečnosti IS a jsou tyto pravidelně aktualizovány?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

9) Umožňuje informační systém efektivní výměnu informací mezi uživateli IS v podniku?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

10) Platí, že pravidla pro provoz a bezpečnost IS jsou nejasná a nelogická?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

30-4-1=25/8=3

Oblast Peopleware

1) Je každý pracovník zaškolen na úlohy, které má s informačním systémem provádět?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

2) Jsou dostupná školení nových pracovníků o používaných informačních systémech, pravidel provozu a bezpečnosti IS?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

3) Je pravda, že pro stávající zaměstnance není třeba školit na nové funkce IS a že školení není dostupné?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

4) Existuje zastupitelnost koncových uživatelů, kteří jsou klíčoví pro chod systému a jeho klíčové výstupy?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

5) Je dokumentace běžných postupů práce s IS jednoduše dosažitelná pro koncové uživatele?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

6) Je si management vědom vlivu firemní kultury na způsob práce koncových uživatelů s informačním systémem?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

7) Jsou dostupná místa uvnitř firmy nebo u externího dodavatele, kam se mohou uživatelé obracet se žádostí o pomoc či konzultaci ohledně IS? (tato místa jsou označována dále jako informační centra)

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

8) Řeší informační centra z předchozího bodu podněty uživatelů obvykle v dostatečné míře a včas?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>

9) Je pravda, že informační centra především „hasí“ palčivé problémy a nemají důvod se snažit o dlouhodobé zlepšení chodu IS?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

10) Podporuje vedení firmy učení koncových uživatelů a jejich školení za účelem zvýšení efektivity fungování IS?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

$$23-4-1=18/8=2$$

Oblast Dataware

1) Mají pracovníci jasně vymezenou odpovědnost za data, která spravují? Tedy, platí zásada, že určitá data smí měnit jen určitý pracovník?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

2) Mají pracovníci určeno, kdy musí jaká data zavést do informačního systému a kdy je musí aktualizovat?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

3) Platí, že uživatelům chybí z informačního systému data pro jejich rozhodování?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

4) Získávají koncoví uživatelé nadbytečná nebo nepřesná data?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

5) Musí pracovníci správy IS pravidelně provádět zálohování dat a dozírá management na dodržování pravidel zálohování?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

6) Uznává management důležitý význam koncových uživatelů pro integritu a správnost zpracování dat?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

7) Existují podrobné plány pro obnovu klíčových dat v informačním systému?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

8) Jsou média se zálohami dostatečně katalogizována a chráněna před zneužitím, krádeží či živelnou pohromou?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

9) Je bezpečnost dat zvažována a řízena i pro hrozby z Internetu nebo jiných počítačových sítí?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

10) Mají pracovníci určeno, s jakými daty smí pracovat a s jakým oprávněním? Platí tedy zásada, že nikdo nesmí získat přístup k datům, která nepotřebuje pro svou práci?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

40-5-3=32/8=4

Oblast Customers

1) Jsou jasně stanoveny základní cíle zkoumaného informačního systému směrem k jeho zákazníkům?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

2) Existují metriky cílů uvedených v předchozím bodu a jsou dostatečně vyhodnocovány?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

3) Je pravidelně zkoumáno, jaké přínosy od informačního systému jeho zákazníci očekávají?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

4) Je pravda, že názory zákazníků IS na zlepšení, změnu či úpravu informačního systému nejsou pro podnik důležité?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

5) Jsou data o zákaznících IS, jejich požadavcích, operacích, atd. ukládány v informačním systému centrálně (tj. nejsou ukládány vícekrát nebo jinak nekonzistentně).

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

6) Přispívá současné hardwarové a softwarové vybavení k dostatečně rychlým odezvám na požadavky zákazníků IS?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

7) Je forma výstupů z informačních systémů volena tak, aby umožňovala jejich snadné využití zákazníkem IS?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

8) Ošetřují pravidla provozu nakládání s citlivými či obchodně cennými daty o zákaznících IS?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

9) Je řízena integrace zkoumaného informačního systému firmy spolu s dalšími IS podniku, které poskytují výstupy pro dané zákazníky?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

10) Mohou zákazníci získávat ze zkoumané IS výstupy pomocí různých komunikačních kanálů, které si zvolí?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

32-4-2=26/8=3

Oblast Suppliers

1) Jsou jasně stanoveny základní požadavky kladené na dodavatele, které jsou nezbytné pro plnění definovaných cílů zkoumaného informačního systému?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

2) Existují metriky hodnocení výše zmíněných požadavků a jsou dostatečně vyhodnocovány.

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

3) Je forma vstupů do zkoumaného IS od dodavatelů volena tak, aby umožňovala jejich snadné převzetí a využití zkoumaným IS?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

4) Jsou v pravidlech provozu definovány kontroly informací od dodavatelů?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

5) Jsou požadavky na dodavatele ve vztahu ke vstupům do zkoumanému IS formulovány tak, aby byla jasně určená požadovaná podrobnost předávaných informací?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

6) Jsou požadavky na dodavatele ve vztahu ke vstupům do zkoumanému IS formulovány také s jasným určením požadované včasnosti jejich dodávání?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

7) Zvažuje firma možnost účelného přizpůsobení či nastavení zkoumaného IS dle návrhů dodavatelů za účelem efektivnější výměny informací?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

8) Je forma výstupů ze zkoumaného IS pro dodavatele řízena s ohledem na efektivní komunikaci s dodavateli?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

9) Je pravda, že výstupy z IS pro dodavatele nejsou řízeny s ohledem na včasnost jejich předání?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

10) Přispívá zkoumaný informační systém ke snadnosti a efektivnosti komunikace s dodavateli?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

36-4-2=30/8=4

Oblast Management IS

1) Trvají manažeři na dodržování pravidel stanovených pro informační systém?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

2) Provádí řízení rozvoje a provozu informačního systému osoba, která této oblasti rozumí?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

3) Je rozvoj IS formulován také ve střednědobé či dlouhodobé perspektivě formou informační strategie vzhledem k cílům firmy?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

4) Je v plánech rozvoje informačních systémů zahrnut případný růst firmy a rozvoj jejich informačních potřeb?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

5) Platí, že plány rozvoje IS neexistují nebo v nich nejsou stanoveny možnosti kontroly jejich plnění?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

6) Je při plánech rozvoje informačního systému, pořizování IS provedeno obhájení dané investice z ekonomického hlediska?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

7) Považuje management informačních systémů koncové uživatele za faktor s vysokou důležitostí pro úspěšný chod informačních systémů?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

8) Usiluje management IS soustavně o zlepšení efektivnosti chodu zkoumaného informačního systému?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

9) Vnímá obecný management informační systém firmy nejen jako výdaje, ale také jako potenciál případného růstu firmy?

Ano	Spíše ano	Částečně	Spíše ne	Ne
-----	-----------	----------	----------	----

<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
--------------------------	-------------------------------------	--------------------------	--------------------------	--------------------------

10) Podporuje obecný management firmy rozvoj informačních systémů, který je odůvodněný přispěním IS k dosažení podnikových cílů?

Ano	Spíše ano	Částečně	Spíše ne	Ne
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

37-5-2=30/8=4