

ČESKÁ ZEMĚDĚLSKÁ UNIVERZITA

FAKULTA PROVOZNĚ EKONOMICKÁ
Obor Informatika



DIPLOMOVÁ PRÁCE

Téma: Datový sklad firmy

Vypracoval:

Jaromír Vajgert

Vedoucí diplomové práce:

doc.ing. Vojtěch Merunka, PhD.

PROHLÁŠENÍ

Prohlašuji, že jsem diplomovou práci na téma Datový sklad firmy zpracoval samostatně za použití uvedené literatury a po odborných konzultacích s doc.ing. Vojtěchem Merunkou, PhD.

V Klatovech dne 27. 11. 2011

.....
(podpis autora práce)

PODĚKOVÁNÍ

Děkuji tímto panu doc.ing. Vojtěchu Merunkovi, PhD. za odborné vedení a cenné rady při zpracování diplomové práce.

Souhrn

Práce ukazuje postup budování datového skladu fiktivní obchodní firmy. Krátce popisuje nástroje a postupy, které její autor v minulosti používal při práci na projektech typu Business Intelligence. Mimo jiné vysvětluje varianty architektury datového skladu a některé osvědčené postupy při vytváření dimenzí v OLAP kostkách. Teoretických myšlenek „otců zakladatelů“ data warehousingu Ralpa Kimballa a Williama Inmona se dotýká jen okrajově, neboť jsou v současnosti již široce známé.

V návaznosti na to pak ozřejmí uvedené principy pomocí analýzy a (z důvodu rozsahu jen částečné) realizace datového skladu obchodní firmy. Jako technologická platforma pro realizaci byl zvolen relační databázový stroj Microsoft SQL Server 2008, jehož součástí je i multidimenzionální databázový stroj Microsoft Analysis Services.

Klíčová slova

Agregace, agregační funkce, analytická funkce, bitmapový index, b-tree index, databáze, datový sklad, datové tržiště, Data Staging Area (DSA), denormalizace, dimenze, drill-down, drill-up, drill-through, Extraction, Transformation & Load (ETL), extrakce dat, konsolidace dat, kvalita dat, materializované view, partitioning, perzistence, primární systém, reporting, On-Line Analytical Processing (OLAP), OLAP kostka, pomalu se měnící dimenze, portál, prezentační vrstva, proces, query rewrite, sklad provozních dat (ODS), třetí normální forma.

Abstract

This work describes procedure for building a data warehouse of fictitious business enterprise. Briefly describes the tools and techniques that the author used in the past when working on projects such as Business Intelligence. Among other things, explains the variations of the data warehouse architecture and some best practices when creating dimensions in the OLAP cubes. Author implicates the Ralph Kimball's and William Inmon's theories marginally only, because he consider them to be well-known.

Author explains variants of Data Warehouse architecture and some best practices when creating dimensions in the OLAP cubes.

Subsequently, it becomes clear these principles through analysis and (because of the scale only partial) implementation of data warehouse of business venture. As a technology platform for the implementation was chosen relational database management system Microsoft SQL Server 2008, because its part is a multidimensional database engine Microsoft Analysis Services.

Keywords

Aggregation, aggregation function, analytic function, bitmap index, b-tree index, database, data consolidation, data extraction, data mart, data quality, data warehouse, Data Staging Area (DSA), denormalisation, dimension, drill-down, drill-up, drill-through, Extraction, Transformation & Load (ETL) , materialized view, On-Line Analytical Processing (OLAP), OLAP cube, Operational Data Store (ODS) , partitioning, persistence, presentation layer, primary system, portal, query rewrite, reporting, slowly changing dimension, third normal form.

Obsah

1. Úvod	7
2. Cíl práce a metodika	10
2.4. Terminologie	13
3. Datový sklad jakou součást IS firmy	14
3.1. Architektura datového skladu	14
3.1.1. Základní – minimalistická architektura.....	14
3.1.2. Architektura s DSA	15
3.1.3. Architektura s DSA a závislými datamarty.....	17
3.2. Dimenze	17
3.2.1. Slowly changing dimensions.....	18
3.2.2. Časová osa.....	19
3.2.3. Dimenze „intervalového“ typu.....	19
3.2.4. Nepravidelná struktura dimenze	19
3.3. Kvalita dat	20
3.4. Nástroje pro ETL.....	21
3.5. Diskuse	21
4. Požadavky na datový sklad formou business-process analýzy	24
4.1. Hodnototvorné procesy obchodní firmy	24
4.1.1. Řízení vstupních operací	24
4.1.2. Řízení výstupních operací	25
4.1.3. Servisní služby - Reklamacce prodeje	27
4.1.4. Servisní služby - Reklamacce nákupu	28
4.2. Ukazatelé výkonnosti obchodních procesů a jejich členění.....	29
5. Projekt datového skladu	33
5.1. Reverse engineering zdrojových systémů	33
5.1.1 Class diagram entit zdrojového systému.....	33
5.1.1.1 Zboží	33
5.1.1.2 Nákup	35
5.1.1.3 Prodej	36
5.1.1.4 Banka	37
5.1.2. Popis entit.....	38
5.2. Návrh architektury datového skladu.	61
5.3. Návrh struktury datového skladu	61
5.3.1. Dimenze	61
5.3.2. Míry a jejich seskupení do kostek.....	63
5.3.3 Datový model.....	68
6. Závěr	75

1. Úvod

William Inmon [Inmon, 1992] definoval datový sklad (data warehouse, DW) takto:

A data warehouse is subjekt-oriented, integrated, nonvolatile, and time-variant collection of data in support of management's decisions. The data warehouse contains granular corporate data. Data in the data warehouse is able to be used for many different purposes, including sitting and waiting for future requirements which are unknown today.

(Výše uvedenou definici jsem vysvětloval již ve své bakalářské práci [Vajgert]. Dále uvedená charakteristika je přepracováním a rozšířením odpovídajícího textu z uvedené bakalářské práce.)

Charakterizujme jednotlivé pojmy této definice:

- **Předmětově orientovaný** (*subject-oriented*)
Provozní systémy organizace podporují v typickém případě práci jednoho či více oddělení na jednom či několika fragmentech (mnohdy ani ne ucelených podprocesech) procesů v organizaci. Úkolem datového skladu je poskytnout pohled na procesy v organizaci v celém jejich rozsahu, od začátku do konce.
- **Sjednocený** (*integrated*)
Datový sklad obsahuje a integruje data získaná z různých provozních systémů (dále budou nazývány primárními systémy). V nich mohou být stejné objekty hospodářského světa reprezentovány různými, vzájemně nesouvisejícími datovými strukturami, mohou být identifikovány různými klíči a popsány s různým stupněm detailnosti, přesnosti a spolehlivosti. Databáze provozních systémů mimoto poskytují dobrý obraz o přítomném okamžiku, méně dobrý o minulosti. Úkolem datového skladu je sjednotit jejich popis do jednotné datové struktury, co možno nejpresněji a nejspolehlivěji.
- **Neměnná historická data** (*nonvolatile, and time-variant*)
Datový sklad uchovává sérii snímků databází produkčních systémů. Pozdější snímek nemění data snímku předchozího. Změny provedené v produkčních

systémech v čase jsou uchovány jako posloupnost stavů tak, aby při přijatelném objemu úložného prostoru zachovaly celou trajektorii změn.

Více např. Slowly changing dimensions – viz dále.

- Kolekce dat

Data datového skladu jsou uložena v úložišti, kterým je (relační či objektová) databáze. Ta obsahuje data přenesená z provozních systémů, jakož i data externí (poskytnutá třetími stranami). Její součástí jsou i metadata popisující strukturu datového skladu.

- Potřeby rozhodování

Návrh datového modelu provozních systémů vychází z potřeby průběžně on-line modifikovat uložená data. Proto je vhodné jej pro zajištění integrity jeho dat normalizovat. Návrh datového modelu datového skladu je naproti tomu veden nutností dosáhnout dostatečného výkonu při provádění složitých dotazů; musí rovněž zohlednit potřeby nástrojů pro podporu rozhodování.

- Úroveň podrobnosti (*granularity*)

Pro účely porovnání je nezbytné, aby data uchovávaná v datovém skladu byla převedena na srovnatelné měrné jednotky, vypovídala o srovnatelných časových intervalech a byla zachycována na srovnatelné úrovni podrobnosti.

- Požadavky na datový sklad jsou (analytikem a posléze investorem) formulovány k jistému pevnému okamžiku. Praxe ovšem ukazuje, že již shromážděná, vyčištěná a konsolidovaná data v datovém skladu je možno s výhodou využít i k uspokojení informačních potřeb, které nebyly v okamžiku analýzy, plánování a zadání projektu k realizaci zdaleka zamýšleny.

Je proto osvědčenou praxí uchovávat v datovém skladu nejen požadované ukazatele v požadovaném členění v denormalizovaném tvaru, ale uchovávat kompletní výsledek procesu extrakce, čištění a konsolidace. I data, která nejsou po provedení výpočtu bezprostředně požadovaných ukazatelů zapotřebí, mohou být upotřebena poté, co se požadavky na ukazatele změní. Cenou za jejich uchování je pouze cena diskového prostoru.

Z technického pohledu lze datový sklad charakterizovat jako kolekci databází konsolidovaných historických dat optimalizovanou pro ad-hoc dotazování, reportingové a OLAP aplikace [Humphries, 2001]. Část jeho datového modelu (přinejmenším pak část sloužící pro potřeby ETL procesů – viz dále) je přímo odvozena z datových modelů zdrojových systémů a bývá alespoň ve 3. normální formě.

Každodenní práce uživatele s datovým skladem se ovšem děje pomocí nástrojů pro on-line analýzu dat (OLAP). K tomu slouží další část datového skladu. Ta obsahuje data převedená do denormalizovaného tvaru, přičemž její návrh respektuje požadavky kladené dotazovacími nástroji. Může být uložena nejen v relační databázi (ROLAP), ale i v nerelační (MOLAP) či hybridní (HOLAP) podobě.

V užším významu slova se datovým skladem rozumí pouze taková databáze, která obsahuje data všech relevantních systémů v organizaci, tady celý (až na data typu GIS, fulltext apod.) datový model organizace. Databáze pokrývající pouze část datového modelu organizace se nazývají datová tržiště (datamarts) [Humphries, 2001],

Skladem provozních dat (Operational Data Store, ODS) pak rozumíme databázi, ve které jsou sjednoceně uložena provozní data k podpoře provozního sledování. Na rozdíl od statických a historických dat datového skladu obsahují data nestálá data popisující současný okamžik. [Humphries, 2001]

Bývá uváděno, že datový sklad obsahuje data souhrnná, sklad provozních dat naproti tomu data detailní. Toto tvrzení je platné pouze v některých projektech.

Je zřejmé, že se veličiny uchovávané v datových skladech používají v prvé řadě pro měření obchodních procesů podnikatelských subjektů. Historická povaha uchovávaných dat pak umožňuje analyzovat je i z hlediska delší časové perspektivy, což je služba, kterou primární systémy zhusta neposkytují. I proto se datové sklady staly samozřejmou a nepostradatelnou součástí podnikové informatiky.

2. Cíl práce a metodika

2.1. Cílem práce je

- objasnit roli datového skladu v rozhodovacích procesech obchodní firmy
- podat výklad praktických postupů při budování datových skladů nad relační databází Microsoft SQL Server 2008.
- na základě analýzy obchodních procesů a datového modelu ERP systému firmy navrhnout strukturu datového skladu; tento návrh pak zčásti realizovat.

Vzhledem k šíři problematiky si neklade ambici pokrýt celý její rozsah, důraz klade do oblasti implementace.

2.1. Výstupy práce

Práce postupně podá

- výklad základních prakticky použitelných variant architektury datového skladu,
- analýzu hlavních hodnototvorných procesů obchodní firmy, která bude sestávat z
 - popisu procesu
 - diagramu procesu v notaci BPMN 1.1
- definici požadavků na datový sklad, která bude obsahovat
 - popis sdílených dimenzí
 - popis jednotlivých ukazatelů a jejich členění
- popis a datový model primárních systémů
- výstupy analýzy a návrhu datového skladu
 - návrh architektury datového skladu
 - návrh struktury datového skladu skládající se z
 - návrhu sdílených dimenzí a jejich zdrojových dat v primárních systémech
 - návrhu jednotlivých kostek skládající se z
 - názvu kostky
 - výčtu dimenzí, dle kterých bude členěna

- výčtu ukazatelů a jejich zdrojových dat v primárních systémech
 - datového modelu jednotlivých komponent datového skladu.

Výše uvedený materiál definuje požadavky na ETL procesy probíhající v datovém skladu. Plní tedy navíc funkci jejich zadání. Detailně projektovat ETL procesy v takto malém systému je sice možné, ne však účelné. Vypracování detailního projektu ETL procesů je totiž pracnější, než jejich naprogramování a odladění.

Součástí práce je implementace části navrhovaného systému.

2.3. Metodika

Kapitola 3 - Datový sklad jakou součástí IS firmy je z převážné části literární rešerší. Obsahuje však i obecněji využitelné poznatky, kterých autor práce nabyl při realizaci projektů datových skladů a manažerských informačních systémů.

Kapitola 4 - Požadavky na datový sklad formou business-process analýzy popíše hlavní procesy fiktivní obchodní firmy. Pojem hlavní proces je zde užit ve smyslu Porterova modelu hodnotového řetězce jako proces realizující primární aktivity podniku, tedy aktivity, které se přímo podílejí na dodání výrobku či služby zákazníkovi. Porter člení hlavní procesy na:

- Řízení vstupních operací (*Inbound Logistics*)
- Výroba a provoz (*Operations*)
- Řízení výstupních operací (*Outbound Logistics*)
- Marketing a odbyt (*Marketing and Sales*)
- Servisní služby (*Service*)

Z důvodu rozsahu budou podrobně diskutovány procesy Řízení vstupních operací, Řízení výstupních operací a Servisní služby. V obchodní firmě nemá smysl zabývat se procesem Výroba a provoz. Odbytová část procesu Proces Marketing a odbyt bude diskutována v rámci procesu Řízení výstupních operací, část marketingová pak nebude

diskutována vůbec. Proces Servisní služby se v praxi obchodní firmy rozpadá na procesy Reklamace prodeje a Reklamace nákupu, které budou diskutovány samostatně. Zmíněné hlavní procesy budou charakterizovány slovním popisem a diagramem v notaci BPMN 1.1. Slovní popis bude obsahovat

- Zahájení procesu – událost, která vyvolá spuštění procesu,
- Akce procesu
- Výsledek procesu
- Vlastníka (*Owner*) procesu a vykonávající (*Performers*).

BPMN diagramy budou vytvářeny pomocí nástroje BizAgi Process Modeler.

Následně budou určeny metriky (*Metrics*) procesů - ukazatele, které charakterizují výkon a kvalitu jednotlivých procesů, a jejich členění - dimenze. Jejich výčet stanoví požadavky na datový sklad, který bude analyzován, projektován a částečně budován v kapitole následující.

Procesy budou popisovány nikoli za účelem jejich zlepšení či reengineeringu, ale za účelem určení metrik. Pro tento účel postačí méně podrobná úroveň jejich popisu.

Procesy podpůrné a řídicí zde budou z důvodu rozsahu rovněž opominuty.

Kapitola 5 - Vlastní projekt (analýza, návrh, implementace části systému) vyjde ze seznamu požadavků, které vzniknou v předchozí kapitole. Následně bude proveden

- reverse engineering primárních systémů
- návrh odpovídající architektury datového skladu
- návrh datového modelu datového skladu
 - návrh dimenzí a kostek
 - návrh Data Staging Area (bude-li DSA v navrženém řešení použita)
 - návrh datového skladu

Část datového skladu bude realizována pomocí nástrojů MS SQL Server Management Studio a MS SQL Server Business Intelligence Studio. Veškeré diagramy v jazyce UML2 budou vytvářeny pomocí nástroje StarUML.

2.4. Terminologie

Tento odstavec je rozšířením obdobného odstavce mé bakalářské práce [Vajgert].

Problémem celého odvětví informačních technologií je nedostatek všeobecně přijímané české odborné terminologie. V textu tedy pro jednoznačnost užívám převážně termínů anglických.

Pro výklad podstatné termíny jsou zavedeny a vysvětleny (či odkázány na citaci) v textu práce. Malý slovník zkratk a termínů, které byly použity okrajově a jejichž vysvětlování v textu by zbytečně odvádělo pozornost čtenáře, následuje.

CASE	computer-aided software engineering, podpora technik a procesů softwarového inženýrství počítačem
constraint	integritní omezení; invariantní podmínka ve schématu relační databáze sloužící ke kontrole integrity dat a pro optimalizaci dotazů
DSA	Data Staging Area; databáze (samostatná instance nebo vyhrazený prostor), do které se nástroji ETL vkládají data extrahovaná z primárních systémů, aby se zde následně provedla jejich transformace, konsolidace a čištění.
join	operace spojení ve smyslu terminologie relačních databází
JavaScript	implementace skriptovacího jazyka ECMAScript vytvořená společností Microsoft
legacy application	aplikace vytvořená na základě dnes již zastaralých technologií, jejíž provoz je však žádoucí pro chod organizace zachovat
metadata	data popisující jiná data
RDBMS	relational database management systém, systém řízení relační báze dat
refresh	operace, která uvádí obsah materializovaného view do souladu s aktuálními daty detailních tabulek
report	sestava či dokument ji obsahující
script	program ve skriptovacím jazyce
skriptovací jazyk	vysokoúrovňový beztypový interpretovaný jazyk
VBScript	skriptovací jazyk společnosti Microsoft, jehož syntax je odvozena od programovacího jazyka Visual Basic
view	pohled ve smyslu terminologie relačních databází
OLAP	On-Line Analytical Processing; uložení kvantitativních dat (veličin, measures) v členění dle vybraných analytických kritérií (dimenzí) do podoby zvané „kostka“ a jejich analýza pomocí dotazovacích nástrojů.
ROLAP	způsob technické realizace OLAP, který používá jako úložiště dat relační databázi

MOLAP	způsob technické realizace OLAP, který používá jako úložiště dat specializovanou multidimenzionální databázi
HOLAP	způsob technické realizace OLAP, který kombinuje postupy a technologie ROLAPu a MOLAPu. <i>Nutno podotknout, že v dokumentaci firmy Microsoft jsou pojmy ROLAP, MOLAP a HOLAP užívány v mírně posunutém smyslu. Tato diplomová práce užívá výše uvedených zkratk v jejich obvyklém významu.</i>

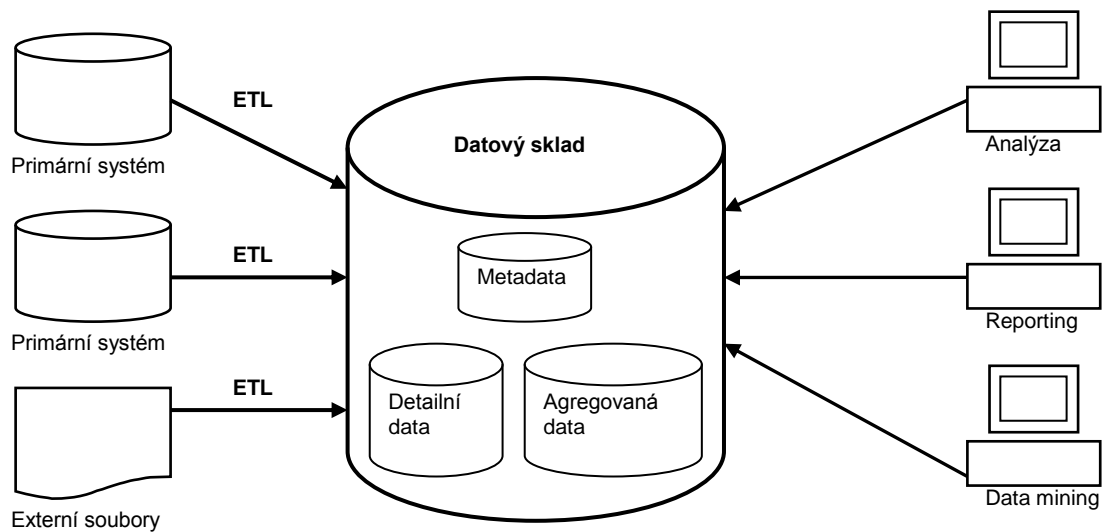
3. Datový sklad jako součást IS firmy

Obdobnou problematiku jsem řešil i ve své bakalářské práci [Vajgert]. Text této kapitoly z ní byl použit a revidován.

3.1. Architektura datového skladu

Tato kapitola je literární rešerší. Zdrojem informací pro tuto kapitolu byl manuál [Oracle, 2].

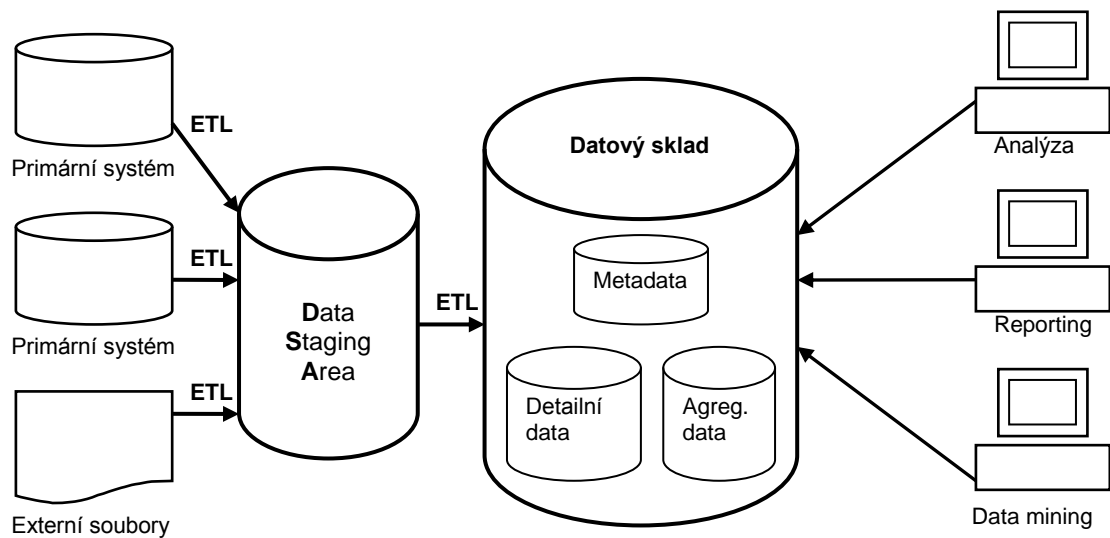
3.1.1. Základní – minimalistická architektura



Obr. 1 – Schéma základní architektury datového skladu

Data z primárních systémů jsou procesy ETL (Extraction, Transformation & Load) importována přímo do datového skladu. ETL procesy provádějí pouze jednoduchou transformaci a konsolidaci, resp. i agregaci dat. Uživatelé přímo přistupují k datům, která byla získána z primárních systémů a k datům z nich odvozeným (větš. agregovaným). Využití agregovaných dat v dotazech umožňuje dosáhnout kratších dob odezvy oproti dotazům směřovaným na detailní tabulky.

3.1.2. Architektura s DSA

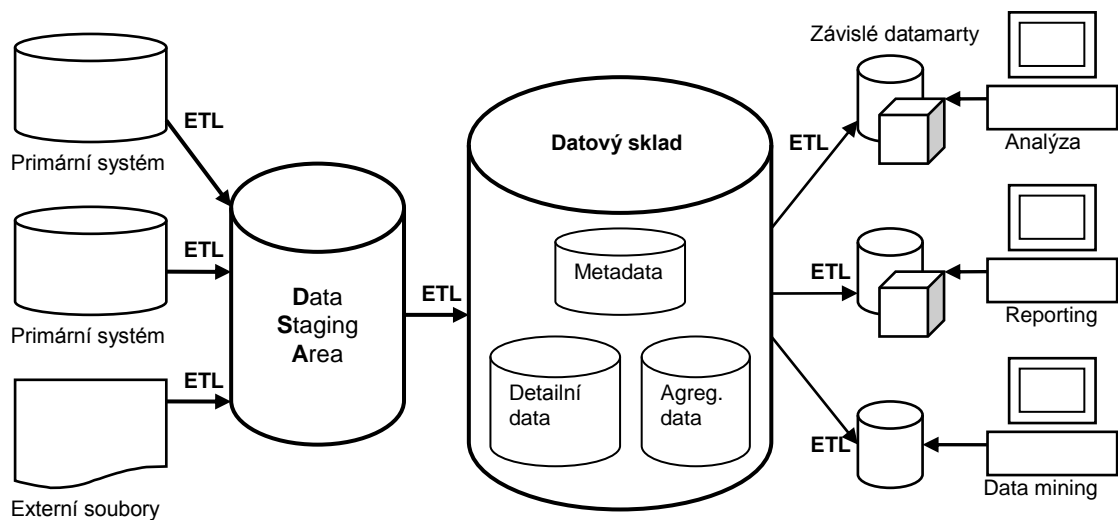


Obr. 2 – Schéma architektury datového skladu s DSA

Tato architektura obsahuje navíc databázi (Data Staging Area, DSA), do které směřují data extrahovaná z primárních systémů. Zde se provádí jejich transformace, konsolidace a čištění. Vzhledem k tomu, že je zde oproti předchozí architektuře k dispozici dodatečný databázový prostor, je možno importovaná data zpracovávat netriviálními a náročnými metodami.

Výsledek je importován do datového skladu. Až v datovém skladu jsou vytvářena agregovaná data. Datový sklad je přímo – nezprostředkovaně k dispozici uživatelům.

3.1.3. Architektura s DSA a závislými datamarty



Obr. 3 – Schéma architektury datového skladu s DSA a závislými datamarty

Z centrálního datového skladu jsou plněny závislé datamarty, ke kterým přistupují uživatelé.

3.2. Dimenze

Dimenze jsou nástrojem ku členění sledovaných veličin (measures) v datovém skladu dle jednotlivých z hlediska uživatelů relevantních hledisek. V primárních systémech jim odpovídají číselníkové entity. Ty jsou ostatně zdrojem dat dimenzních entit v datovém skladu. V některých dále diskutovaných případech je ukázáno, že tato závislost nemusí být zdaleka přímočará.

Podrobný výklad problematiky dimenzí lze nalézt v monografii [Kimball,2002]. Zde jsou diskutovány pouze některé, z hlediska praxe analýzy a projekce zajímavé problémy.

3.2.1. Slowly changing dimensions.

Tato kapitola je převážně literární rešerší, jejímž zdrojem je [Kimball,2002].

Číselníky v primárních systémech nejsou konstantní, ale podléhají pomalým změnám. Datový sklad musí na tyto změny reagovat změnou v dimenzích, a to tak, aby nedošlo ke zkreslení významu dat. Na případné změny musí být struktura dimenzních tabulek předem připravena. Dle situace se volí některá z následujících možností reakce na změnu dimenze :

- Typ 0. Dimenzní tabulky navrženy bez zřetele k případným změnám. Volí se, pokud se očekává, že v dimenzní tabulce budou pouze přibývat řádky.
- Typ 1. Při změně řádku číselníku se přepíše řádek dimenzní tabulky. Starý obsah se neuchovává. Tato strategie je vhodná, pokud se očekává, že změny v číselnících budou pouze opravovat chyby.
- Typ 2. Pro změněný řádek číselníku se založí nový řádek dimenzní tabulky s novou hodnotou dimenzního klíče, starý řádek se ponechá. Jednotlivé verze řádku se číslovají, příp. označují časovým intervalem jejich platnosti od-do.
- Typ 3. Dimenzní tabulka obsahuje sadu sloupců pro původní hodnoty změněného řádku číselníku a datum, do kterého platily. Při změně řádku se dosavadní hodnoty přesunou do sloupců pro původní hodnoty.
- Typ 4. K dimenzní tabulce existuje další tabulka s historickými hodnotami řádků označenými datem vložení. V dimenzní tabulce se nachází aktuální záznam. Před změnou řádku dimenzní tabulky se dosavadní obsah řádku vloží do tabulky historie spolu s datem, do kterého platily.
- Typ 6 – hybridní. Tato metoda kombinuje přístupy metod 1,2 a 3. Pro změněný řádek číselníku se založí nový řádek dimenzní tabulky, ale se stejnou hodnotou dimenzního klíče. Starý řádek se ponechá. Dimenzní klíč zde není primárním klíčem dimenzní tabulky (tím je dvojice dimenzní klíč – datum platnosti od) . Jednotlivé verze řádku se označují časovým intervalem jejich platnosti od-do, příp. i příznakem platného řádku. V tabulce faktů se objevuje dimenzní klíč a ne primární klíč dimenzní tabulky.

Tato metoda je pro svou komplikovanost relativně málo používaná, není ani podporována dotazovacími nástroji.

Je velmi vhodné, aby již návrh primárních systémů počítal se změnou číselníků a jejich jednotlivé platné verze verzoval, jako celek či po řádcích.

3.2.2. Časová osa

Dle postupů, které jsou programátorskou komunitou všeobecně užívány (viz např. [McKnight]), je možno časovou osu v denormalizovaném datovém skladu zobrazit dvěma základními způsoby:

- Jako jedinou časovou dimenzi s hierarchií rok-měsíc-den.
- Jako více dimenzí Rok, Kvartál, Měsíc, Týden, Den, Den v týdnu, resp. pouze některé z nich. Tato alternativa umožňuje na rozdíl od předešlé snadno analyzovat sezónní vlivy na zkoumané ukazatele.

Dimenzní tabulky časových dimenzí se naplňují souvislou posloupností hodnot v požadovaném časovém rozsahu, dimenzní klíče se určují výpočtem. Hodnoty mimo požadovaný časový rozsah a chybějící hodnoty se mapují se na speciální „chybovou“ hodnotu dimenzního klíče.

3.2.3. Dimenze „intervalového“ typu

Často se vyskytne požadavek uživatele definovat dimenzi ne na základě číselníku, ale jako seznam intervalů některého ukazatele z tabulky faktů. Důvod takového požadavku je jednoduchý – pomocí prohlížeče kostek mohou uživatelé snadno získat histogram.

Intervaly od-do se do dimenzní tabulky naplňují buď ručně, či na základě uživatelem definovaného propočtu. Klíč do dimenzní tabulky se stanoví výpočtem na základě hodnoty ukazatele.

3.2.4. Nepravidelná struktura dimenze

V obvyklém případě hierarchická struktura dimenze kopíruje strukturu číselníkových tabulek. Existují však případy, kdy takto sestavená hierarchie (a zejména agregace podle ní prováděné) postrádá pro uživatele vypovídací schopnost. V takovém případě nezbyvá, než zachytit uživatelem požadovanou strukturu dimenze do speciálních tabulek a dimenzi vytvářet na jejich základě.

Dobrym příkladem situace, kdy je takový přístup potřebný, je dimenze Účet (podvojného účetnictví). Dimenze konstruovaná podle účtových tříd a skupin nemá pro uživatele velkou cenu, neboť uživatel požaduje dimenzi respektující strukturu účetních výkazů.

Uživatel může požadovat i možnost sám si definovat prvky dimenze na základě jím definovaných SQL podmínek a následně je dle své vůle seskupovat do hierarchie. Takto definované dimenze jsou v typickém případě využívány pro účely výkaznictví a marketingové segmentace.

3.3. Kvalita dat

Kvalita dat v datovém skladu je určena kvalitou dat v primárních systémech. Z nekvalitních vstupů lze sotva odvodit kvalitní výstupy. V rámci ETL procesů lze jistě aplikovat na jisté typy údajů (jména, adresy, geografické údaje apod.) nástroje pro čištění dat. Ty jistě v drtivé většině případů přispějí ke zkvalitnění dat, zároveň však mohou zanést chyby další. Mohou např. chybně ztotožnit dva zákazníky stejného jména, z nichž každý vykazuje jinou platební morálku. Čištění dat v datovém skladu je tedy třeba chápat – v protikladu k marketingovým tvrzením výrobců nástrojů pro čištění dat - jako nouzové opatření. Za optimální postup nutno považovat čistit data on-line při jejich pořizování.

Součástí výstupu ETL procesů jsou i údaje o chybách, které se vyskytly nebo byly diagnostikovány v průběhu ETL procesu. Z nich lze odvodit ukazatele kvality dat (obvykle četnosti chyb) v členění dle typu, agend, primárních systémů, organizačních útvarů a operátorů) a rovněž je uložit do datového skladu. Lze je pak analyzovat stejnými nástroji, jako ostatní údaje v datovém skladu. Stávají se pak účinným nástrojem při zavádění datového skladu do života organizace a pro jeho další řízení.

Umožňují dlouhodobě sledovat příčiny chyb v datovém skladu a poskytují podklady pro nápravná opatření. Jako takové by měly být pevnou součástí datového skladu.

3.4. Nástroje pro ETL

Úkolem nástrojů pro ETL je poskytnout prostředky pro

- návrh transformací dat v ETL procesu
- návrh a řízení dávkových úloh, ze kterých sestává celý subsystém ETL.

V projektech rozsáhlých datových skladů jsou „nekorunovanými králi“ ETL nástroje Informatica a Oracle Warehouse Builder. Jejich vysoká cena ovšem omezuje jejich použití na prostředí „velkého byznysu“. Pro budování datových skladů malých a středních firem je vhodné hledat nákladově příznivější řešení.

Microsoft SQL Server obsahuje (a to již jako součást základního balíku) mocný a zároveň intuitivní nástroj Integration Services [Microsoft, 1]. Tento nástroj umožňuje graficky navrhovat celý ETL proces, zároveň plánovat jeho vykonávání, řídit je a monitorovat jeho průběh. Jeho význačnou funkcionalitou je řízení ETL procesu, který je rozložen na více serverech.

Microsoft Integration Services je přímým nástupcem nástroje MS Transformation Services. Prostředky tohoto nástroje pro návrh transformací dat obsahovaly v minulosti závažné chyby návrhu (spouštění scriptů v jazyce JScript či VBScript pro každý řádek zdrojové sady dat), které snižovaly použitelnost nástroje na návrh nejelementárnějších transformací dat. Od verze MS SQL Serveru 2005, kde místo scriptovacích jazyků zaujal .NET framework, je možno tento problém pokládat za uspokojivě vyřešený.

3.5. Diskuse

Jedinou výhodou základní architektury je její jednoduchost a z ní vyplývající nenáročnost projekční, realizační a finanční. Nicméně má smysl ji zvolit, pokud realizujeme

- rychlé řešení datového skladu pro malé organizace,
- prototyp (části) zamýšleného datového skladu,

- projekt datového skladu při rozpočtu, který nedovoluje realizovat lepší řešení,
- pilotní projekt, který má vedení organizace přesvědčit o výhodách datového skladu,
- nezávislý datamart.

Datový sklad v základní architektuře je v typickém případě denormalizovaný. Nezávislé datamarty je často účelné řešit pouze pomocí OLAP databáze.

Vyjmenované výhody jsou ovšem převáženy nevýhodami. Mezi ně patří nemožnost provádět náročnější algoritmy transformace a konsolidace dat, jakož i nutnost přepracovat velkou část datového skladu při změně požadavků.

Data Staging Area umožní provádět složité předzpracování dat (transformaci, konsolidaci a čištění) jako přípravu před uložením do datového skladu. Protože jsou tyto operace náročné na spotřebu systémových prostředků, doporučuje se, aby DSA byla umístěna na dedikovaném databázovém serveru. Doba potřebná pro uložení přírůstku do datového skladu, po kterou je funkčnost datového skladu omezena, se tak zkracuje na minimum.

Pokud uživatelé přistupují k datovému skladu přímo, nikoli skrze závislý datamart, je prakticky nezbytné, aby datový sklad obsahoval data v denormalizované podobě. V DSA ovšem vznikají konsolidovaná data ve 3. normální formě, jejichž struktura odpovídá datovému modelu organizace. Ukazuje se, že tato konsolidovaná data představují trvale cenný zdroj informací. Smazat je po provedení přírůstku do denormalizovaného datového skladu znamená utrpět významnou škodu. Naproti tomu se osvědčuje přenést je do datového skladu a trvale je uchovávat. Část datového skladu ve 3. normální formě se významně mění jen tehdy, pokud se významně mění primární systémy. Denormalizovaná část datového skladu naopak podléhá kontinuálním změnám dle stále se vyvíjejících požadavků uživatelů. Změny v denormalizované části se provádějí snadněji, plní-li se denormalizovaná část z relativně stálé části normalizované, nežli v případě, že se pro plnění denormalizované části musí provést kompletní ETL.

Architektura datového skladu s DSA vyhovuje pro drtivou většinu projektů data warehousingu a představuje zde jakousi zlatou střední cestu.

K budování závislých datamartů se přistupuje z následujících důvodů:

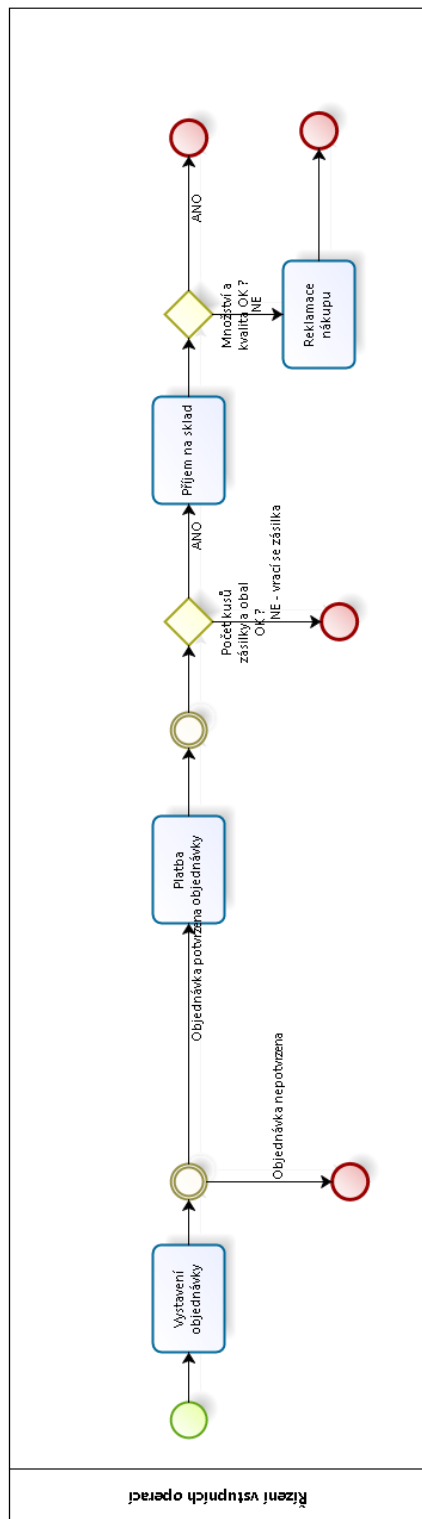
- Závislé datamarty snižují zatížení centrálního datového skladu a zabraňují tomu, aby dotaz jednoho uživatele přespříliš zhoršil odezvu všem ostatním uživatelům datového skladu. To je výhodné zejména v prostředí velkých organizací.
- Závislé datamarty umožňují lépe zabezpečit data. Do datamartů se kopíruje jen ta část datového skladu, k jejímuž využití mají uživatelé jednotlivých datamartů právo.
- Závislé datamarty mohou být budovány nad takovou platformou, která nejlépe vyhovuje potřebám daného útvaru.
- V závislé datamarty se mohou změnit „legacy applications“ původně plněné přímo z primárních systémů. Při plnění z centrálního datového skladu jejich funkčnost zůstane zachována.
- Datový sklad je vybudován v 3. normální formě, uživatelé potřebují mít k dispozici multidimenzionální přístup. Kombinace datového skladu ve 3NF a denormalizovaného závislého datamartu je velmi výhodná, neboť změny v závislém datamartu se provádějí snáze a laciněji nežli změny v datovém skladu.

Ačkoli přímý přístup uživatelů do centrálního datového skladu není vyloučen, má smysl pouze v těch případech, kdy závislý datamart nepokrývá jejich požadavky.

4. Požadavky na datový sklad formou business-process analýzy

4.1. Hodnototvorné procesy obchodní firmy

4.1.1. Řízení vstupních operací



Powered by
bizagi
Modeler

Zahájení:

Požadavek nákupčího na doplnění skladových zásob

NEBO

zahájeno procesem Řízení výstupních operací při nedostatku disponibilních zásob zboží pro uspokojení přijaté objednávky.

Akce:

Vystavení objednávky

(Ne)Potvrzení objednávky.

Přijetí zprávy od dodavatele s potvrzením/nepotvrzením objednávky.

Platba objednávky.

Vystavení příkazu k převodu a jeho odeslání bance.

Přijem zboží na sklad.

Při této operaci se provádí kontrola množství a kvality. Neshoduje-li se počet kusů zásilky nebo je-li porušen obal zásilky, provede se Odmítnutí zásilky. Neshoduje-li se množství či kvalita, zahájí se proces Reklamace nákupu.

Výsledek:

Stav požadovaných skladových zásob požadovaného druhu

zboží se zvýší o požadovaný počet jednotek.

Vlastník procesu: Vedoucí nákupu

Účastníci: Referent nákupu, Finanční účetní, Skladník, Převravec

4.1.2. Řízení výstupních operací

Zahájení:

Příjem objednávky některým z elektronických prodejních kanálů či ručním zadáním Referentem odbytu.

Akce:

Kontrola disponibility zboží

- Kontrola disponibility zboží ve skladu
- Kontrola disponibility zboží u dodavatelů

Rezervace zboží ve skladu

Objednání zboží u dodavatelů (proces Řízení vstupních operací)

Potvrzení objednávky. *Zákazníkovi je odeslána zpráva o potvrzení, částečném potvrzení či nepotvrzení jeho objednávky.*

Platba zboží u dodavatelů (požadují-li dodavatelé platbu předem)

Kontrola došlé platby (byla-li požadována platba předem).

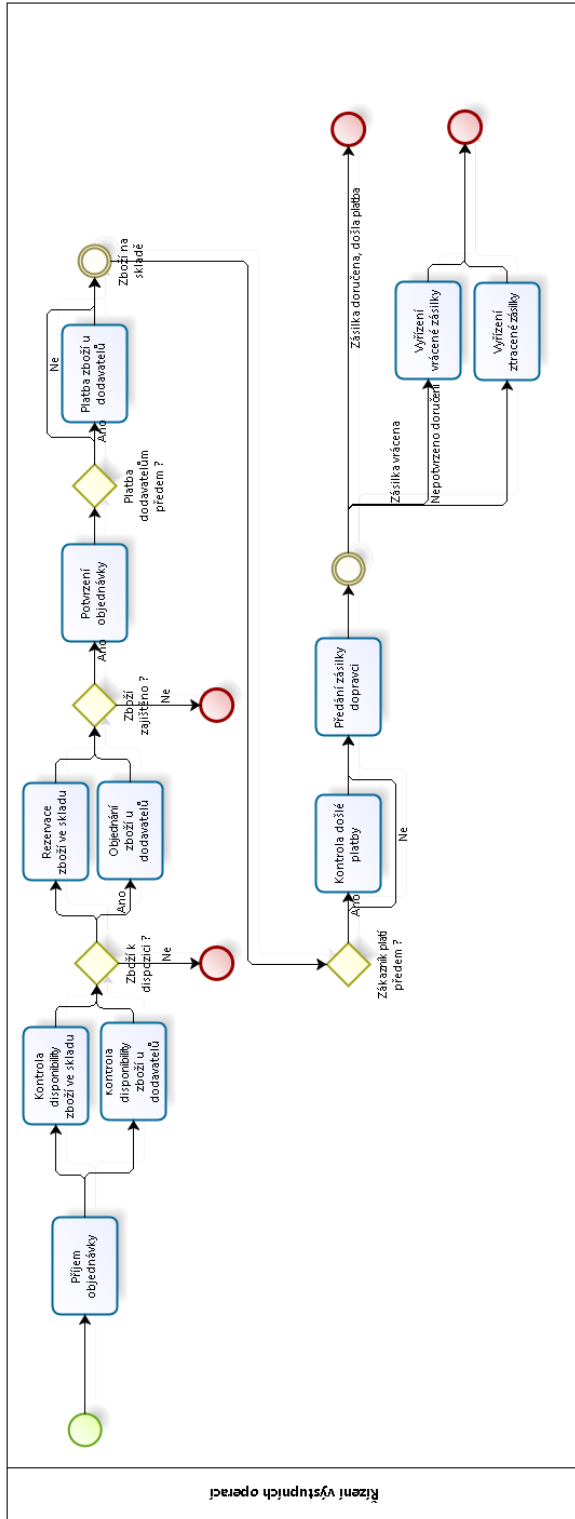
Kompletace zásilky

Předání zásilky dopravci

Kontrola došlé platby (byla-li požadována platba dobírkou). *Pokud nebyla zásilka vrácena dopravcem a částka dobírky nebyla připsána na účet v přiměřeném (definovaném) čase, reklamuje se částka dobírky u dopravce.*

Vyřízení vrácené zásilky. *U poškozené zásilky se zkontroluje obsah za přítomnosti převrvice. Uplatní se nárok na náhradu škody za ztracené či poškozené zboží dle sjednaných smluvních podmínek přepravy. Zásilka se rozbálí, zboží se naskladní.*

Vyřízení ztracené zásilky. *Vůči dodavateli se uplatní nárok na náhradu škody dle sjednaných smluvních podmínek přepravy.*

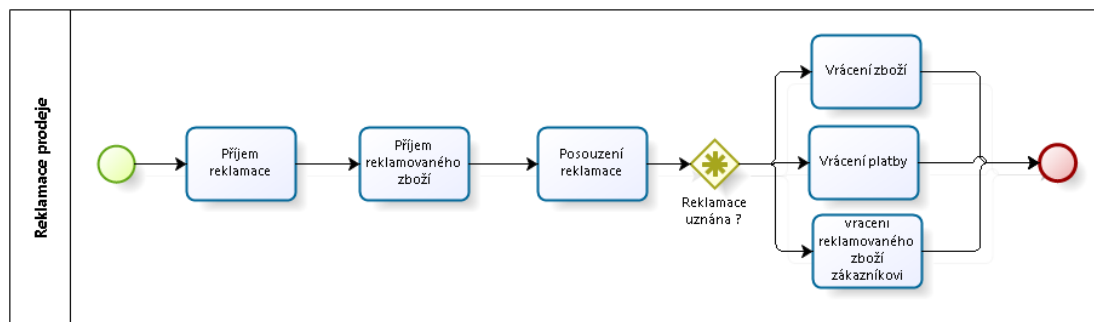


Výsledek: Zákazník obdrží požadované zboží, firma pak inkasuje tržbu za jeho prodej.

Vlastník procesu: Vedoucí odbytu

Účastníci: Referent odbytu, Finanční účetní, Skladník, Přepravce

4.1.3. Servisní služby - Reklamacce prodeje



Powered by
bizagi
Modeler

Zahájení:

Firma obdrží zprávu, ve které zákazník popíše důvod reklamace a návrh na způsob jejího řešení.

Akce:

Přijem reklamace. *Referent odbytu zadá údaje o reklamaci do systému. Komunikací se zákazníkem odstraní případné nejasnosti.*

Přijem reklamovaného zboží

Posouzení reklamace. *Reklamované zboží je posouzeno Technikem zákaznické podpory.*

Možné výsledky posouzení jsou:

po položkách akceptováno / odmítnuto/ částečně akceptováno.

Referent odbytu spolu s Technikem zákaznické podpory posoudí navrhovaný způsob řešení reklamace.

Vrácení zboží zákazníkovi / vrácení platby zákazníkovi

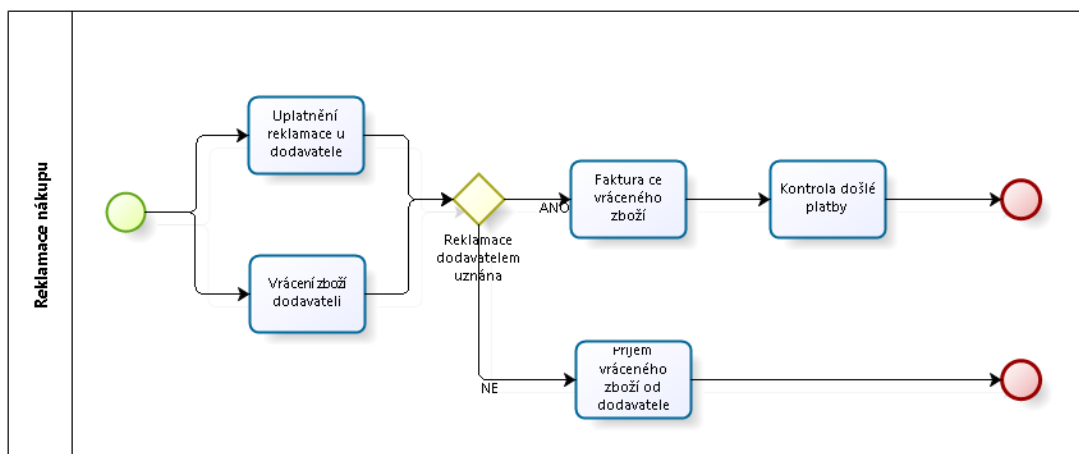
Vrácení zboží z uznané reklamace dodavateli (procesem Dodavatelská reklamace)

Výsledek: Je odstraněn důvod reklamace vrácením bezvadného zboží zákazníkovi či vrácena platba za zboží. Reklamované zboží je vráceno dodavateli.

Vlastník procesu: Vedoucí odbytu

Provádějí: Referent odbytu, Finanční účetní, Skladník, Převážce, Technik zákaznické podpory.

4.1.4. Servisní služby - Reklamace nákupu



Powered by
bizagi
Modeler

Zahájení:

Proces Reklamace nákupu je vyvolán

- neshodou množství či jakosti při naskladnění zboží,
- neshodou jakosti při interní kontrole,
- neshodou jakosti na základě reklamace zákazníka.

Akce:

Uplatnění reklamace u dodavatele. *Dodavatel je informován o druhu a množství reklamovaného zboží vč. dodávky, ve které bylo zboží dodáno, a vytýkané vadě. Zároveň je navržen způsob řešení.*

Vrácení zboží dodavateli

Faktura na úhradu škody

Kontrola došlé platby

Výsledek: Reklamované zboží je vráceno dodavateli. Je inkasována platba za vrácené zboží.

Vlastník procesu: Vedoucí nákupu

Provádějí: Referent nákupu, Finanční účetní, Skladník , Dopravce, Dodavatel

4.2. Ukazatelé výkonnosti obchodních procesů a jejich členění

Ukazatelé:

Počet přijatých objednávek, počet položek přijatých objednávek

Množství v základních měrných jednotkách (měrná jednotka je neaditivní dimenze)

Hodnota v měně (Měna je neaditivní dimenze)

Hodnota v Kč

Daň z přidané hodnoty

Členěny dle dimenzí:

Měrná jednotka (jen pro množství, neaditivní)

Měna (neaditivní)

Zboží

Zákazník, Dealerská kategorie (možno srovnání dealerských kategorií v čase)

Sazba DPH (jen pro DPH)

Stav objednávky

Prodejní akce (při objednávce za akční ceny)

Rok, Měsíc, Den v měsíci, Den v týdnu, Pracovní/nepracovní den

Prodejní kanál, Prodejce

Ukazatelé:

Počet dodávek, Počet položek v dodávce

Hodnota dodávky

Logistické náklady

Objem (dm³)

Hmotnost (kg)

Čas vyřízení (od objednávky po příjem)

Členěny dle dimenzí:

Dodavatel

Dopravce

Položka logistických nákladů

Ukazatelé:

Počet přijatých faktur, počet položek přijatých faktur

Počet přijatých dodacích listů, počet položek přijatých dodacích listů

Množství dodané v základních měrných jednotkách

Množství fakturované v základních měrných jednotkách

Hodnota v dodacím listu v měně

Hodnota v dodacím listu v Kč

Hodnota na faktuře v měně

Hodnota na faktuře v Kč

Daň z přidané hodnoty

Vedlejší náklady pořízení v měně

Vedlejší náklady pořízení v Kč

Členěny dle dimenzí:

Měrná jednotka (jen pro množství, neaditivní)

Měna

Zboží

Zákazník, Dealerská kategorie (možno srovnání dealerských kategorií v čase)

Sazba DPH (jen pro DPH)

Stav faktury

Rok, Měsíc, Den v měsíci, Den v týdnu, Pracovní/nepracovní den
Prodejní kanál, Prodejce

Ukazatelé:

Počet vydaných reklamací, počet položek vydaných reklamací
Množství v základních měrných jednotkách (měrná jednotka neaditivní dimenze)
Hodnota v měně
Hodnota v Kč
Daň z přidané hodnoty
Dobropisováno v měně
Dobropisováno v Kč
Dobropisovaná DPH
Náklady na reklamace

Členěny dle dimenzí:

Nákladová položka
Měrná jednotka (jen pro množství, neaditivní)
Měna
Zboží
Zákazník, Dealerská kategorie (možno srovnání dealerských kategorií v čase)
Sazba DPH (jen pro DPH)
Stav reklamace
Rok, Měsíc, Den v měsíci, Den v týdnu, Pracovní/nepracovní den
Nákupní kanál, Nákupčí
Dopravce, Druh dopravy

Ukazatelé:

Počet vydaných faktur, počet položek vydaných faktur
Množství v základních měrných jednotkách (měrná jednotka neaditivní dimenze)
Hodnota v měně

Hodnota v Kč

Daň z přidané hodnoty

Členěny dle dimenzí:

Měrná jednotka (jen pro množství, neaditivní)

Měna

Zboží

Dodavatel

Sazba DPH (jen pro DPH)

Stav faktury

Rok, Měsíc, Den v měsíci, Den v týdnu, Pracovní/nepracovní den

Ukazatelé:

Počet přijatých reklamací, počet položek přijatých reklamací

Množství v základních měrných jednotkách (měrná jednotka neaditivní dimenze)

Hodnota v měně

Hodnota v Kč

Daň z přidané hodnoty

Dobropisováno v měně

Dobropisováno v Kč

Dobropisovaná DPH

Náklady na reklamace

Členěny dle dimenzí:

Nákladová položka

Měrná jednotka (jen pro množství, neaditivní, možno použít pro rozštěpení kostky do více kostek)

Měna

Zboží

Zákazník, Dealerská kategorie (možno srovnání dealerských kategorií v čase)

Sazba DPH (jen pro DPH)

Stav reklamace

Prodejní akce (při objednávce za akční ceny)

Rok, Měsíc, Den v měsíci, Den v týdnu, Pracovní/nepracovní den

Prodejní kanál, Prodejce

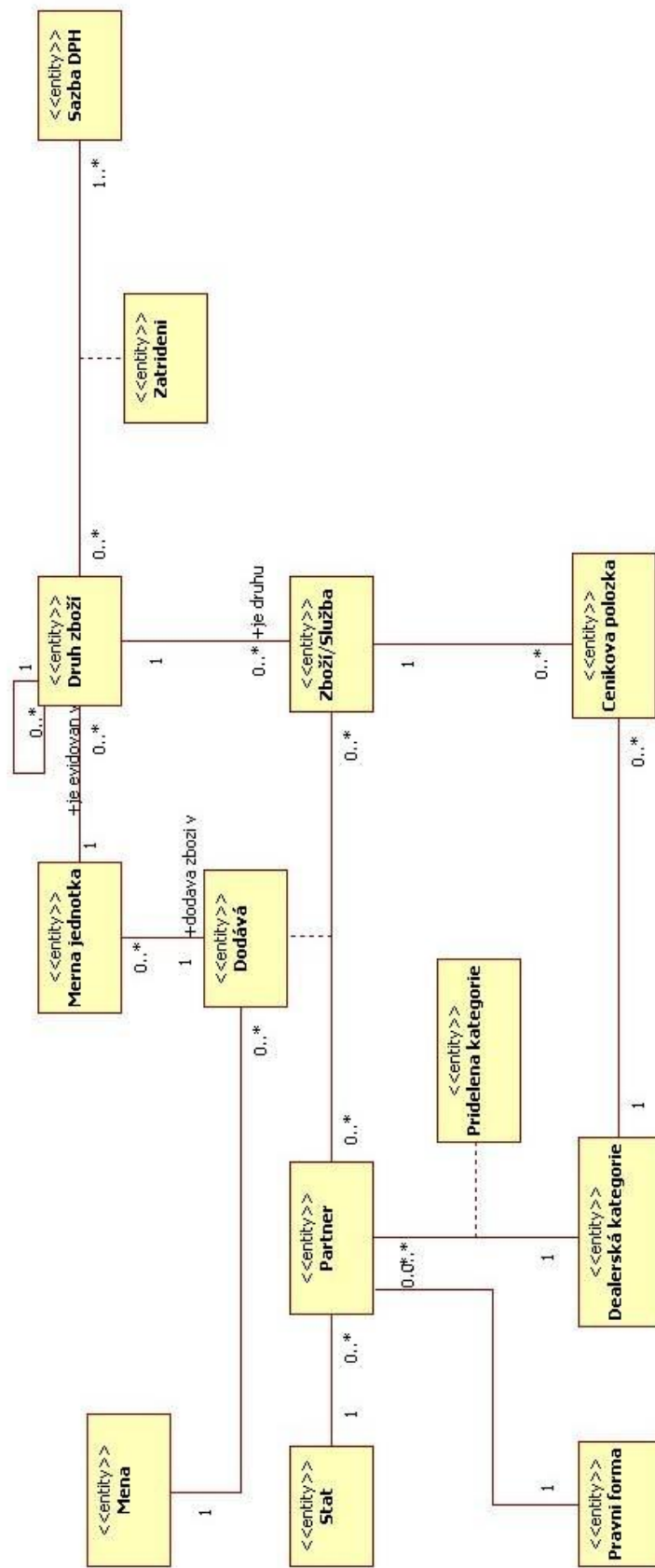
Dopravce, Druh dopravy

5. Projekt datového skladu

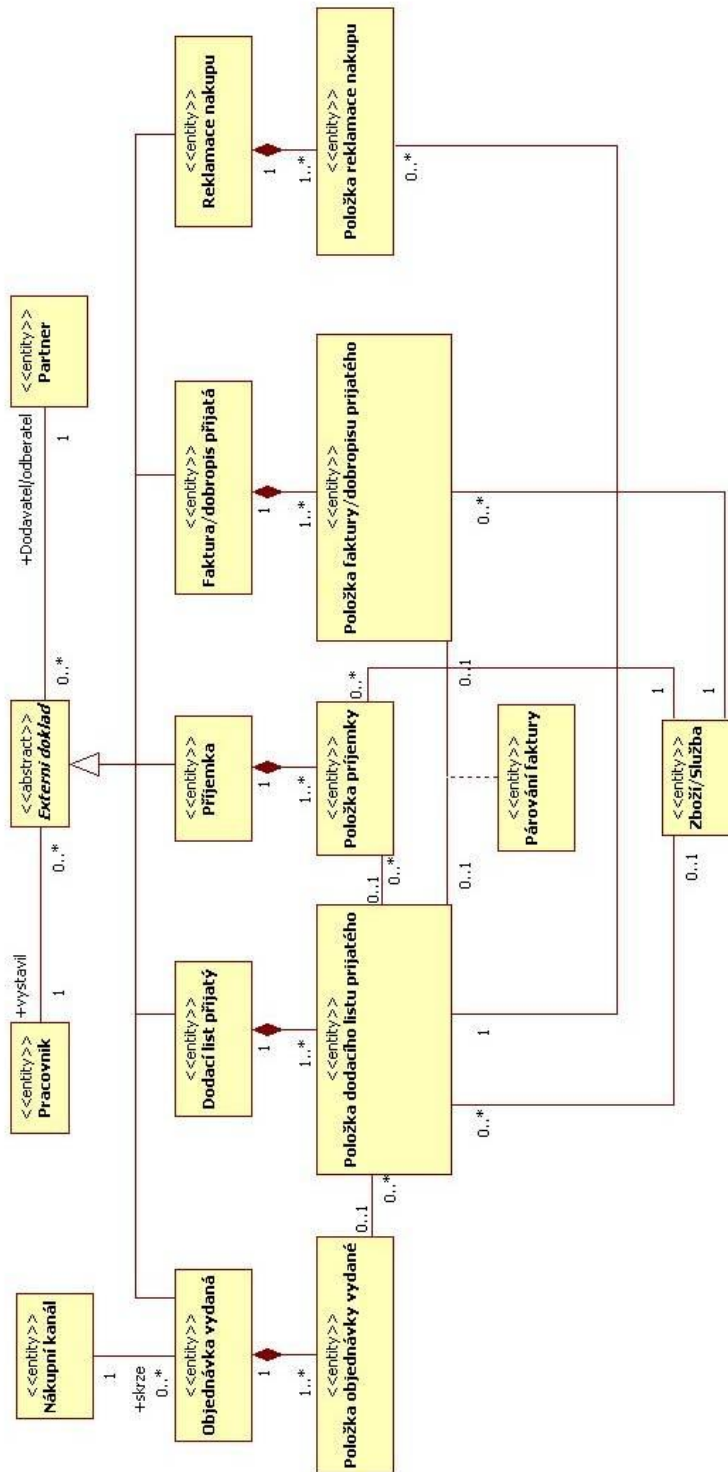
5.1. Reverse engineering zdrojových systémů

5.1.1 Class diagram entit zdrojového systému

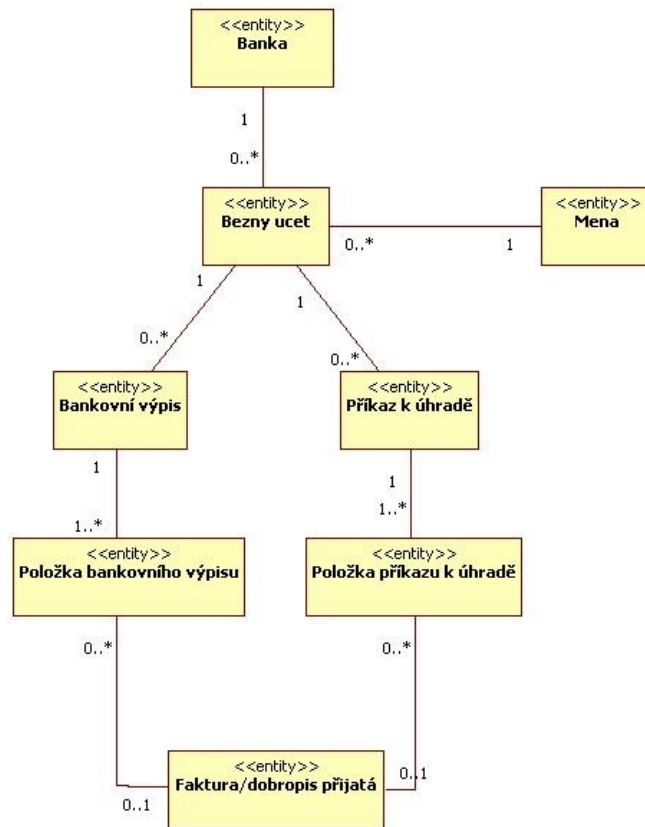
5.1.1.1 Zboží



5.1.1.2 Nákup



5.1.1.4 Banka



5.1.2. Popis entit

Entita: **Měrná jednotka**

Tabulka: **MERNA_JEDNOTKA**

Číselník měrných jednotek pro nákup, prodej a evidenci zboží.

Sloupec	Typ	Význam
ID_MJ	int	interní identifikátor – primární klíč
ID_MJ_ZAKLADNI	int	id základní měrné jednotky; pro zákl. mj NULL
ZKRATKA	nvarchar(20)	zkratka měrné jednotky
NASOBITEL	decimal(12,6)	násobitel převodu na základní měrnou jednotku
VIDITELNE	char(1)	příznak viditelnosti; A=viditelný řádek

Entita: **Sazba DPH**

Tabulka: **SAZBA_DPH**

Tabulka sazeb daně z přidané hodnoty.

Sloupec	Typ	Význam
ID_DPH	int	interní identifikátor – primární klíč
ZKRATKA	nvarchar(20)	označení (základní/snížená/osvobozeno)
SAZBA	decimal(4,1)	sazba DPH v procentech
PLATNOST_OD	date	datum počátku platnosti sazby
PLATNOST_DO	date	datum konce platnosti sazby

Entita: **Druh zboží**

Tabulka: **DRUH_ZBOZI**

Číselník druhů zboží ve smyslu sortimentního zařazení.

Sloupec	Typ	Význam
ID_DRUH	int	interní identifikátor – primární klíč
ID_DRUH_NADRIZENY	int	id nadřizovaného druhu zboží
ID_MJ	int	id měrné jednotky pro evidenci zboží
NAZEV	nvarchar(80)	zkratka měrné jednotky
VIDITELNE	char(1)	příznak viditelnosti; A=viditelný řádek

Entita: **Sazba zboží**

Tabulka: **SAZBA_ZBOZI**

Zatřídění jednotlivých druhů zboží (sortimentních skupin) do jednotlivých sazeb DPH.

Sloupec	Typ	Význam
ID_DPH	int	id sazby DPH
ID_DRUH	int	id druhu zboží
PLATNOST_OD	date	datum počátku platnosti sazby pro daný druh zboží
PLATNOST_DO	date	datum konce platnosti sazby pro daný druh zboží

Entita: **Zboží**

Tabulka: **ZBOZI**

Zboží a jeho trvalé charakteristiky

Sloupec	Typ	Význam
ID_ZBOZI	int	interní identifikátor – primární klíč
ID_DRUH	int	id druhu
NAZEV	nvarchar(80)	název zboží
POPIS_KRATKY	nvarchar(2000)	popis zboží určený do přehledů
POPIS_DLOUHY	text	popis zboží určený pro detailní zobrazení
VIDITELNE	char(1)	příznak viditelnosti; A=viditelný řádek

Entita: **Partner**

Tabulka: **PARTNER**

Obchodní partneři – dodavatelé a zákazníci.

Sloupec	Typ	Význam
ID_PARTNER	int	interní identifikátor – primární klíč
ID_FORMA	int	id právní formy obchodního partnera
ID_STAT	int	id státu
NAZEV_SPOL	int	název firmy
JMENO	nvarchar(20)	jméno fyzické osoby
PRIJMENI	decimal(12,6)	příjmení fyzické osoby
TITUL_PRED	nvarchar(20)	titul fyzické osoby před jménem

TITUL_ZA	nvarchar(20)	titul fyzické osoby za jménem
ULICE	nvarchar(80)	adresa - ulice
CISLO_DOMU	varchar(10)	adresa- číslo domu
OBECE	nvarchar(30)	adresa - obec
PSC	nvarchar(20)	adresa – poštovní směrovací číslo

Entita: **Právní forma**

Tabulka: **PRAVNI_FORMA**

Číselník právních forem obchodních partnerů.

Sloupec	Typ	Význam
ID_FORMA	int	interní identifikátor – primární klíč
FP	char(1)	fyzická osoba=F, právnická osoba=P
ZKRATKA	nvarchar(20)	zkratka právní formy

Entita: **Stát**

Tabulka: **STAT**

Číselník států obchodních partnerů

Sloupec	Typ	Význam
ID_STAT	int	interní identifikátor – primární klíč
NAZEV	nvarchar(80)	název státu
ZKRATKA	nvarchar(20)	zkratka státu
OMEZENI	text	omezení obchodu s daným státem
VIDITELNE	char(1)	příznak viditelnosti; A=viditelný řádek

Entita: **Dodává**

Tabulka: **PARTNER_DODAVA_ZBOZI**

Který partner a za jakých podmínek dodává jaké zboží (asociační třída).

Sloupec	Typ	Význam
ID_PARTNER	int	id obchodního partnera
ID_ZBOZI	int	id zboží
ID_MJ	int	id měrné jednotky
KOD_CENIKU_	varchar(32)	kód zboží v ceníku dodavatele

PARTNERA		
MNOZSTVI_MIN	decimal(12,6)	minimální dodané množství
MNOZSTVI_MAX	decimal(12,6)	maximální dodané množství
CENA_MJ	decimal(12,2)	cena/mj
ID_MENA	int	id měny, ve které je stanovena cena
DODACI_LHUTA	time	obvyklý čas od objednávky a zaplacení k dodávce

Entita: **Měna**

Tabulka: **MENA**

Číselník měn

Sloupec	Typ	Význam
ID_MENA	int	interní identifikátor – primární klíč
ZKRATKA	nvarchar(20)	zkratka měrné jednotky
VIDITELNE	char(1)	příznak viditelnosti; A=viditelný řádek

Entita: **Dealerská kategorie**

Tabulka: **DEALERSKA_KATEGORIE**

Číselník dealerských kategorií

Sloupec	Typ	Význam
ID_DEAL_KATEGORIE	int	interní identifikátor – primární klíč
NAZEV	nvarchar(30)	název dealerské kategorie
SLEVA	decimal(4,1)	obvyklá (výchozí) sleva v %
VIDITELNE	char(1)	příznak viditelnosti; A=viditelný řádek

Entita: **Ceníková položka**

Tabulka: **CENIKOVA_POLOZKA**

Ceník zboží pro danou dealerskou kategorii v daném čase.

Sloupec	Typ	Význam
ID_CEN_POLOZKA	int	id ceníkové položky
ID_DEAL_KATEGORIE	int	id dealerské kategorie

ID_ZBOZI	int	id zboží
CENA_MJ	decimal(8,2)	cena za jednotku daného zboží pro danou dealerskou kategorii
PLATNOST_OD	date	datum počátku platnosti ceny
PLATNOST_DO	date	datum konce platnosti ceny

Entita: **Přidělená kategorie**

Tabulka: **PRIDELENA_KATEGORIE**

Dealerská kategorie přidělená partnerovi-odběrateli v daném čase.

Sloupec	Typ	Význam
ID_DEAL_KATEGORIE	int	id dealerské kategorie
ID_PARTNER	int	id obchodního partnera
PLATNOST_OD	date	datum počátku platnosti dealerské kategorie
PLATNOST_DO	date	datum konce platnosti dealerské kategorie

Entita: **Banka**

Tabulka: **BANKA**

Číselník bank

Sloupec	Typ	Význam
ID_BANKA	int	interní identifikátor – primární klíč
KOD_BANKY	char(4)	kód banky
ZKRATKA	nvarchar(20)	zkratka banky
NAZEV	nvarchar(80)	název banky
VIDITELNE	char(1)	příznak viditelnosti; A=viditelný řádek

Entita: **Bankovní účet**

Tabulka: **BANKOVNI_UCET**

Číselník bankovních účtů firmy

Sloupec	Typ	Význam
ID_UCET	int	interní identifikátor – primární klíč
ID_BANKA	int	id banky

ID_MENA	int	id měny, ve které je veden účet
ZKRATKA	nvarchar(30)	zkratka účtu
PREDCISLI	vvarchar(6)	předčíslí bankovního účtu
CISLO_UCTU	vvarchar(10)	číslo bankovního účtu
IBAN	vvarchar(24)	mezinárodní číslo účtu
VIDITELNE	char(1)	příznak viditelnosti; A=viditelný řádek

Entita: **Bankovní výpis**

Tabulka: **BANKOVNI_VYPIS**

Bankovní výpisy přebírané on-line či na nosičích dat.

Sloupec	Typ	Význam
ID_VYPIS	int	interní identifikátor – primární klíč
ID_UCET	int	id účtu, ke kterému náleží výpis
CISLO_VYPISU	char(4)	číslo bankovního výpisu
DATUM_VYPISU	date	datum bankovního výpisu
PERIODA	nvarchar(15)	perioda výpisu
NAZEV_UCTU	nvarchar(80)	název účtu (dle banky)
STAV_POCATEK	decimal(12,2)	stav na počátku období
STAV_KONEC	decimal(12,2)	stav na konci období
POHYB_CREDIT	decimal(12,2)	suma kreditních pohybů
POHYB_DEBET	decimal(12,2)	suma debetních pohybů

Entita: **Příkaz k úhradě**

Tabulka: **PRIKAZ_K_UHRADE**

Příkazy k úhradě

Sloupec	Typ	Význam
ID_PRIKAZ	int	interní identifikátor – primární klíč
ID_UCET	int	id účtu, ke kterému náleží příkaz k úhradě
CISLO_PRIKAZU	char(4)	číslo příkazu k úhradě
DATUM_PRIJETI	date	datum odeslání příkazu k úhradě bance
DATUM_SPLATNOSTI	date	datum splatnosti příkazu k úhradě

Entita: **Položka bankovního výpisu**

Tabulka:

BANKOVNI_VYPIS_POLOZKA

Položky bankovního výpisu

Sloupec	Typ	Význam
ID_VYPIS	int	id bankovního výpisu
CISLO_RADKU	tinyint	pořadové číslo řádku
PRIKAZCE_KOD_BANKY	varchar(4)	kód banky příkazce
PRIKAZCE_PREDCISLI	varchar(6)	předčísli bankovního účtu příkazce
PRIKAZCE_CISLO_UCTU	varchar(10)	číslo bankovního účtu příkazce
PRIKAZCE_IBAN	varchar(24)	mezinárodní číslo účtu příkazce
PRIKAZCE_NAZEV	nvarchar(80)	název příkazce
DATUM_TRANSAKCE	date	datum transakce
CASTKA	decimal(12,)	částka v měně, ve které je veden účet
CASTKA_JINA_MEN	decimal(12,)	částka v měně převodního příkazu
ID_MENA	int	měna převodního příkazu
KURS	decimal(12,6)	kurs použitý při připsání na účet
KONST_SYMBOL	varchar(4)	konstantní symbol
VAR_SYMBOL	bigint	variabilní symbol
SPEC_SYMBOL	bigint	specifický symbol

Entita: **Položka příkazu k úhradě**

Tabulka:

PRIKAZ_K_UHRADE_POLOZKA

Položky příkazů k úhradě

Sloupec	Typ	Význam
ID_PRIKAZ	int	id příkazu k úhradě
CISLO_RADKU	tinyint	pořadové číslo řádku
PRIJEMCE_KOD_BANKY	varchar(4)	kód banky příjemce

PRIJEMCE_PREDCISLI	varchar(6)	předčísí bankovního účtu příjemce
PRIJEMCE_CISLO_UCTU	varchar(10)	číslo bankovního účtu příjemce
PRIJEMCE_IBAN	varchar(24)	mezinárodní číslo účtu příjemce
DATUM_TRANSAKCE	date	datum transakce
CASTKA	decimal(12,)	částka
ID_MENA	int	měna převodního příkazu
KONST_SYMBOL	varchar(4)	konstantní symbol
VAR_SYMBOL	bigint	variabilní symbol
SPEC_SYMBOL	bigint	specifický symbol

Entita: **Nákupní kanál**

Tabulka: **NAKUPNI_KANAL**

Číselník nákupních kanálů

Sloupec	Typ	Význam
ID_NKANAL	int	id nákupního kanálu
NAZEV	varchar(40)	název nákupního kanálu
VIDITELNE	char(1)	příznak viditelnosti; A=viditelný řádek

Entita: **Prodejní kanál**

Tabulka: **PRODEJNI_KANAL**

Číselník prodejních kanálů

Sloupec	Typ	Význam
ID_PKANAL	int	id prodejního kanálu
NAZEV	varchar(40)	název prodejního kanálu
VIDITELNE	char(1)	příznak viditelnosti; A=viditelný řádek

Entita: **Pracovník**

Tabulka: **PRACOVNIK**

tav dodacího listu vydaného

Pracovník oprávněný pracovat se systémem.

Sloupec	Typ	Význam
ID_PRACOVNIK	int	id pracovníka
LOGIN	varchar(30)	přihlašovací jméno
DOMENA	varchar(30)	doména Windows 2008 Serveru, ve které je platný login
JMENO	varchar(30)	jméno pracovníka
PRIJMENI	varchar(30)	příjmení pracovníka

Abstraktní třída: **Externí doklad**

Předek všech dokladů obchodního systému

Sloupec	Typ	Význam
ID_DOKLAD	int	id dokladu
ZNAK_DOKLADU	varchar(3)	rozlišení číselné řady, ve které jsou číslovány doklady
CISLO_DOKLADU	int	číslo dokladu v číselné řadě určené pro dané účetní období znakem dokladu
UCETNI_ROK	varchar(5)	účetní období; různá účetní období v jednom roce se rozeznávají dodatečným 5. znakem
ID_PARTNER	int	dodavatel/odběratel
STAV_DOKLADU	char(1)	stav dokladu

Entita: **Změna dokladu**

Tabulka: **DOKLAD_ZMENA**

Chronologický záznam o změnách dokladu.

Sloupec	Typ	Význam
ID_DOKLAD	int	id dokladu
CAS_ZMENY	datetime	čas, kdy došlo ke změně
ID_PRACOVNIK	int	pracovník, který provedl změnu
STAV_DOKLADU	char(1)	stav dokladu po uskutečněné změně
ZMENY	varchar(2000)	automaticky generovaný text shrnující uskutečněné změny

Entita: **Číselná řada**

Tabulka: **CISELNA_RADA**

Předek všech dokladů obchodního systému

Sloupec	Typ	Význam
ZNAK_DOKLADU	varchar(3)	rozlišení číselné řady, ve které jsou číslovány doklady
UCETNI_ROK	varchar(5)	účetní období; různá účetní období v jednom roce se rozeznávají dodatečným 5. znakem
POSLEDNÍ_CISLO_DOKLADU	int	naposledy použité číslo dokladu v číselné řadě určené pro dané účetní období znakem dokladu

Entita: **Objednávka vydaná**

Tabulka: **OBJEDNAVKA_VYDANA**

Objednávka adresovaná dodavatelům.

Sloupec STAV nabývá následujících hodnot:

- K – koncept (rozpracovaná)
- V - vystavená
- S – stornovaná
- P – potvrzená
- D – odmítnutá
- C – na cestě

- V - vyřízená

Potomek abstraktní třídy Externí doklad; kromě sloupců definovaných pro tuto abstraktní třídu obsahuje ještě následující:

Sloupec	Typ	Význam
ID_NKANAL	int	id nákupního kanálu
DATUM_VYDANI	date	datum vydání objednávky
ID_DOPRAVA	int	požadovaný způsob dopravy
TERMIN_DODANI	datetime	požadovaný termín dodání

Entita: **Položka objednávky vydané**

Tabulka:

OBJEDNAVKA_VYDANA_POLOZKA

Položky vydané objednávky.

Sloupec	Typ	Význam
ID_DOKLAD	int	id dokladu
PORADI	tinyint	pořadové číslo řádku
ID_ZBOZI	int	objednané zboží
KOD_CENIKU_PARTNERA	varchar(32)	kód zboží v ceníku dodavatele
CENA_MJ	decimal(12,2)	cena/mj
ID_MENA	int	id měny, ve které je stanovena cena
ID_MJ	int	měrná jednotka objednávky
MNOZSTVI	decimal(12,4)	objednané množství

Entita: **Dodací list přijatý**

Tabulka: **DODACI_LIST_PRIJATY**

Dodací list v podobě, v jaké byl firmou přijat.

Sloupec STAV nabývá následujících hodnot:

- K – koncept (pořízený)
- K – akceptovaný
- D – odmítnutý/vrácený

Potomek abstraktní třídy Externí doklad; kromě sloupců definovaných pro tuto

abstraktní třídu obsahuje ještě následující:

Sloupec	Typ	Význam
DATUM_VYDANI	date	datum vydání dod. listu
CISLO_OBJEDNAVKY	varchar(32)	číslo objednávky tak, jak je uvedeno na dodacím listu; slouží k párování s vydanými objednávkami.
ID_DOPRAVA	int	způsob dopravy
TERMIN_DODANI	datetime	termín dodání
MIME	varchar(32)	MIME typ uloženého dodacího listu
DOKUMENT	varbinary(max)	dodací list v podobě, jakou byl přijat (scan dokumentu, elektronický dokument – XML, pdf s metadaty apod.)

Entita: **Položka dodacího listu přijatého** Tabulka:

DODACI_LIST_PRIJATY_POLOZKA

Položka přijatého dodacího listu

Sloupec	Typ	Význam
ID_DOKLAD	int	id dokladu
PORADI	tinyint	pořadové číslo řádku
ID_ZBOZI	int	dodané zboží
KOD_CENIKU_PARTNERA	varchar(32)	kód zboží v ceníku dodavatele
CENA_MJ	decimal(12,2)	cena/mj
CENA_CELKEM	decimal(12,2)	celková cena
ID_MENA	int	id měny, ve které je stanovena cena
ID_MJ	int	měrná jednotka
MNOZSTVI	decimal(12,4)	dodané množství

Entita: **Párování dodacího listu s
objednávkou**

Tabulka: **DODLIST_PRIJATY_OBJEDNAVKA**

Každý dodací list musí být odkontrolován vůči objednávce; neobjednané zboží nesmí být přijato. Jakákoliv neshoda (s výjimkou případu, že dodavatel dodal menší než objednané množství) vede k reklamaci.

Sloupec	Typ	Význam
ID_DOKLAD_OBJEDNAVKA	int	id objednávky
PORADI_OBJEDNAVKA	tinyint	pořadové číslo řádku objednávky
ID_DOKLAD_DODLIST	int	id dodacího listu
PORADI_DODLIST	tinyint	pořadové číslo řádku dodacího listu
MNOZSTVI	decimal(12,4)	dodané množství

Entita: **Příjemka**

Tabulka: **PRIJEMKA**

Sloupec STAV nabývá následujících hodnot:

- K – koncept (rozpracovaná)
- V - vystavená
- C – oceněná
- U – zaúčtovaná
- S – stornovaná

Potomek abstraktní třídy Externí doklad; kromě sloupců definovaných pro tuto abstraktní třídu obsahuje ještě následující:

Sloupec	Typ	Význam
ID_STRED_MD	int	id střediska (=skladu)
ID_STRED_D	int	id střediska, ze kterého bylo zboží naskladněno (interní pohyby)
POHYB	char(2)	druh pohybu; definuje provádění kontroly při editaci příjemky a

		kontaci při exportu do účetnictví
--	--	-----------------------------------

Entita: **Položka příjemky**

Tabulka: **PRIJEMKA_POLOZKA**

Potomek abstraktní třídy Externí doklad; kromě sloupců definovaných pro tuto abstraktní třídu obsahuje ještě následující:

Sloupec	Typ	Význam
ID_DOKLAD	int	id dokladu
PORADI	tinyint	pořadové číslo řádku
ID_ZBOZI	int	objednané zboží
ID_DOKLAD_DODLIST	int	id dodacího listu
PORADI_DODLIST	tinyint	pořadové číslo řádku dodacího listu
CENA_MJ	decimal(12,2)	cena/mj v CZK
CENA_CELKEM	decimal(12,2)	celková cena v CZK
MNOZSTVI	decimal(12,4)	dodané množství ve skladové měrné jednotce určené zbožím

Entita: **Faktura/dobropis přijatá**

Tabulka: **FAKTURA_PRIJATA**

Sloupec STAV nabývá následujících hodnot:

- K – koncept (rozpracovaná)
- A - akceptovaná
- V - vrácená
- U – zaúčtovaná
- S – stornovaná

Potomek abstraktní třídy Externí doklad; kromě sloupců definovaných pro tuto abstraktní třídu obsahuje ještě následující:

Sloupec	Typ	Význam
DATUM_VYDANI	date	datum vydání faktury
DATUM_SPLATNOSTI	date	datum splatnosti faktury

CASTKA	decimal(12,2)	částka faktury
KOD_BANKY	char(4)	kód banky
MENA	varchar(3)	id měny, ve které je veden účet
ZKRATKA	nvarchar(30)	zkratka účtu
PREDCISLI	varchar(6)	předčísli bankovního účtu
CISLO_UCTU	varchar(10)	číslo bankovního účtu
IBAN	varchar(24)	mezinárodní číslo účtu
KONST_SYMBOL	varchar(4)	konstantní symbol
VAR_SYMBOL	bigint	variabilní symbol
SPEC_SYMBOL	bigint	specifický symbol
MIME	varchar(32)	MIME typ uložené faktury
DOKUMENT	varbinary(max)	faktura v podobě, jakou byla přijata (scan dokumentu, elektronický dokument – XML, pdf s metadaty apod.)

Entita: **Položka přijaté**

Tabulka:

faktury/dobropisu

FAKTURA_PRIJATA_POLOZKA

Potomek abstraktní třídy Externí doklad; kromě sloupců definovaných pro tuto abstraktní třídu obsahuje ještě následující:

Sloupec	Typ	Význam
ID_DOKLAD	int	id dokladu
PORADI	tinyint	pořadové číslo řádku
ID_ZBOZI	int	objednané zboží
KOD_CENIKU_PARTNERA	varchar(32)	kód zboží v ceníku dodavatele
CENA_MJ	decimal(12,2)	cena/mj
CENA_CELKEM	decimal(12,2)	celková cena
ID_MENA	int	id měny, ve které je stanovena cena

ID_MJ	int	měrná jednotka
MNOZSTVI	decimal(12,4)	dodané množství
ID_DPH	int	id sazby DPH
DPH	decimal(12,4)	částka DPH

Entita: **Reklamace nákupu**

Tabulka: **REKLAMACE_NAKUPU**

Sloupec STAV nabývá následujících hodnot:

- K – koncept (rozpracovaná)
- A - akceptovaná
- N - neakceptovaná
- V – vyřízená
- S – stornovaná

Potomek abstraktní třídy Externí doklad; kromě sloupců definovaných pro tuto abstraktní třídu obsahuje ještě následující:

Sloupec	Typ	Význam
DATUM_UPLATNĚNÍ	date	datum uplatnění reklamace u dodavatele
DATUM_VYŘÍZENÍ	date	datum vyřízení reklamace dodavatelem
POZNAMKA	text	poznámky k vyřízení reklamace

Entita: **Položka reklamace nákupu**

Tabulka:

REKLAMACE_NAKUPU_POLOZKA

Potomek abstraktní třídy Externí doklad; kromě sloupců definovaných pro tuto abstraktní třídu obsahuje ještě následující:

Sloupec	Typ	Význam
ID_DOKLAD	int	id dokladu
PORADI	tinyint	pořadové číslo řádku
ID_DOKLAD_DODLIST	int	id přijatého dodacího listu

PORADI_DODLIST	tinyint	pořadové číslo řádku dodacího listu
ID_NESHODA	int	id druhu neshody
POPIS_NESHODY	text	popis neshody

Entita: **Párování faktury přijaté**

Tabulka:

PAROVANI_FAKTURY_PRIJATE

Potomek abstraktní třídy Externí doklad; kromě sloupců definovaných pro tuto abstraktní třídu obsahuje ještě následující:

Sloupec	Typ	Význam
ID_DOKLAD_FAKTURA	int	id přijaté faktury
PORADI_FAKTURA	tinyint	pořadové číslo řádku přijaté faktury
ID_DOKLAD_DODLIST	int	id přijatého dodacího listu
PORADI_DODLIST	tinyint	pořadové číslo řádku dodacího listu
ID_SLOZKA_CENY	int	identifikátor složky ceny
CASTKA	decimal(12,2)	částka složky ceny

Entita: **Způsob dopravy** Tabulka: **ZPUSOB_DOPRAVY**

Potomek abstraktní třídy Externí doklad; kromě sloupců definovaných pro tuto abstraktní třídu obsahuje ještě následující:

Sloupec	Typ	Význam
ID_DOPRAVA	int	identifikátor způsobu dopravy
NAZEV	varchar(30)	název způsobu dopravy
K_DODAVATELI	char(1)	A – použit ve vydaných objednávkách
K_ZAKAZNIKOV	char(1)	A – použit v přijatých objednávkách
DOPRAVNE	decimal(6,2)	dopravné (přijaté objednávky)
BALNE	decimal(6,2)	balné (přijaté objednávky)

Entita: **Složka ceny**

Tabulka: **SLOZKA_CENY**

Číselník složek pořizovací ceny zboží.

Sloupec	Typ	Význam
ID_SLOZKA_CENY	int	identifikátor složky ceny
NAZEV	varchar(30)	název složky pořizovací ceny

Entita: **Druh neshody**

Tabulka: **DRUH_NESHODY**

Číselník druhů neshod v reklamacích.

Sloupec	Typ	Význam
ID_NESHODA	int	identifikátor druhu neshody
NAZEV	varchar(30)	název druhu neshody

Entita: **Objednávka**

Tabulka: **OBJEDNAVKA_PRIJATA**

přijatá

Objednávka adresovaná dodavatelům.

Sloupec STAV nabývá následujících hodnot:

- K – koncept (rozpracovaná)
- T - přijatá
- S – stornovaná
- P – potvrzená
- D – odmítnutá
- C – na cestě
- V - vyřízená

Potomek abstraktní třídy Externí doklad; kromě sloupců definovaných pro tuto abstraktní třídu obsahuje ještě následující:

Sloupec	Typ	Význam
ID_PKANAL	int	id prodejního kanálu
DATUM_VYDANI	date	datum přijetí objednávky
ID_DOPRAVA	int	požadovaný způsob dopravy
TERMIN_DODANI	datetime	požadovaný termín dodání

Entita: **Položka**

Tabulka: **OBJEDNAVKA_PRIJATA_POLOZKA**

objednávky přijaté

Položky vydané objednávkou.

Sloupec	Typ	Význam
ID_DOKLAD	int	id dokladu
PORADI	tinyint	pořadové číslo řádku
ID_CEN_POLOZKA	int	objednané zboží (ceníková položka)
MNOZSTVI	decimal(12,4)	objednané množství

Entita: **Dodací list vydaný**

Tabulka: **DODACI_LIST_VYDANY**

Sloupec STAV nabývá následujících hodnot:

- K – koncept (pořízený)
- V – vystavený
- S – stornovaný

Potomek abstraktní třídy Externí doklad; kromě sloupců definovaných pro tuto abstraktní třídu obsahuje ještě následující:

Sloupec	Typ	Význam
DATUM_VYDANI	date	datum vydání dod. listu
ID_DOPRAVA	int	způsob dopravy
TERMIN_DODANI	datetime	termín dodání

Entita: **Položka dodacího listu**

Tabulka:

vydaného

DODACI_LIST_VYDANY_POLOZKA

Položka přijatého dodacího listu

Sloupec	Typ	Význam
ID_DOKLAD	int	id dokladu
PORADI	tinyint	pořadové číslo řádku

ID_DOKLAD_VYDEJKA	int	id výdejky
PORADI_VYDEJKA	tinyint	pořadové číslo položky výdejky
ID_CEN_POLOZKA	int	ceníková položka
CENA_MJ	decimal(12,2)	prodejní cena/mj
CENA_CELKEM	decimal(12,2)	celková prodejní cena
ID_MENA	int	id měny, ve které je stanovena cena
ID_MJ	int	měrná jednotka
MNOZSTVI	decimal(12,4)	dodané množství

Entita: **Výdejka**

Tabulka: **VYDEJKA**

Sloupec STAV nabývá následujících hodnot:

- K – koncept (rozpracovaná)
- V - vystavená
- C – oceněná
- U – zaúčtovaná
- S – stornovaná

Potomek abstraktní třídy Externí doklad; kromě sloupců definovaných pro tuto abstraktní třídu obsahuje ještě následující:

Sloupec	Typ	Význam
ID_STRED_MD	int	id střediska, na které bylo zboží vydáno (interní pohyby)
ID_STRED_D	int	id střediska(=skladu))
POHYB	char(2)	druh pohybu; definuje prováděné kontroly při editaci příjemky a kontaci při exportu do účetnictví

Entita: **Položka výdejky**

Tabulka: **VYDEJKA_POLOZKA**

Potomek abstraktní třídy Externí doklad; kromě sloupců definovaných pro tuto

abstraktní třídu obsahuje ještě následující:

Sloupec	Typ	Význam
ID_DOKLAD	int	id dokladu
PORADI	tinyint	pořadové číslo řádku
ID_DOKLAD_PRIJEMKA	int	id příjemky, kterou bylo zboží přijato
PORADI_PRIJEMKA	tinyint	pořadové číslo příjemky, kterou bylo zboží přijato
ID_DOKLAD_OBJ_VYDANA	int	id přijaté objednávky, která je plněna touto položkou výdejky
PORADI_OBJ_VYDANA	tinyint	pořadové číslo položky přijaté objednávky, která je plněna touto položkou výdejky
CENA_MJ	decimal(12,2)	cena/mj v CZK
CENA_CELKEM	decimal(12,2)	celková cena v CZK
MNOZSTVI	decimal(12,4)	dodané množství ve skladové měrné jednotce určené zbožím

Entita: **Faktura/dobropis vydaná**

Tabulka: **FAKTURA_VYDANA**

Sloupec STAV nabývá následujících hodnot:

- K – koncept (rozpracovaná)
- A - akceptovaná
- V - vrácená
- U – zaúčtovaná
- S – stornovaná

Potomek abstraktní třídy Externí doklad; kromě sloupců definovaných pro tuto abstraktní třídu obsahuje ještě následující:

Sloupec	Typ	Význam
---------	-----	--------

DATUM_VYDANI	date	datum vydání faktury
DATUM_SPLATNOSTI	date	datum splatnosti faktury
CASTKA	decimal(12,2)	částka faktury
ID_UCET	int	bankovní účet
KONST_SYMBOL	varchar(4)	konstantní symbol
VAR_SYMBOL	bigint	variabilní symbol
SPEC_SYMBOL	bigint	specifický symbol

Entita: **Položka vydané**

Tabulka:

faktury/dobropisu

FAKTURA_VYDANA_POLOZKA

Potomek abstraktní třídy Externí doklad; kromě sloupců definovaných pro tuto abstraktní třídu obsahuje ještě následující:

Sloupec	Typ	Význam
ID_DOKLAD	int	id dokladu
PORADI	tinyint	pořadové číslo řádku
ID_DOKLAD_DODLIST_VYDANY	int	id vydaného dodacího listu
PORADI_DODLIST_VYDANY	tinyint	pořadové číslo řádku vydaného dodacího listu
CENA_MJ	decimal(12,2)	cena/mj
CENA_CELKEM	decimal(12,2)	celková cena
ID_MENA	int	id měny, ve které je stanovena cena
ID_MJ	int	měrná jednotka
MNOZSTVI	decimal(12,4)	dodané množství
ID_DPH	int	id sazby DPH
DPH	decimal(12,4)	částka DPH

Entita: **Reklamacce prodeje**

Tabulka: **REKLAMACE_PRODEJE**

Sloupec STAV nabývá následujících hodnot:

- K – koncept (rozpracovaná)
- A - akceptovaná
- N - neakceptovaná
- V – vyřízená
- S – stornovaná

Potomek abstraktní třídy Externí doklad; kromě sloupců definovaných pro tuto abstraktní třídu obsahuje ještě následující:

Sloupec	Typ	Význam
DATUM_UPLATNĚNÍ	date	datum uplatnění reklamace u dodavatele
DATUM_VYŘÍZENÍ	date	datum vyřízení reklamace dodavatelem
POZNAMKA	text	poznámky k vyřízení reklamace

Entita: **Položka reklamace prodeje**

Tabulka:

REKLAMACE_PRODEJE_POLOZKA

Potomek abstraktní třídy Externí doklad; kromě sloupců definovaných pro tuto abstraktní třídu obsahuje ještě následující:

Sloupec	Typ	Význam
ID_DOKLAD	int	id dokladu
PORADI	tinyint	pořadové číslo řádku
ID_DOKLAD_DODLIST_VYDANY	int	id vydaného dodacího listu
PORADI_DODLIST_VYDANY	tinyint	pořadové číslo řádku vydaného dodacího listu
ID_NESHODA	int	id druhu neshody
POPIS_NESHODY	text	popis neshody

5.2. Návrh architektury datového skladu.

Vzhledem k rozsahu práce je sotva možné provést projekci datového skladu v některé z rozsáhlejších variant. Byla proto zvolena základní varianta datového skladu.

Datový model primárního systému v dostatečné míře zachycuje historii změn dimenzí a plně zachycuje historii jednotlivých dokladů. Datový model datového skladu může být proto denormalizován v celém svém rozsahu; úkoly, které obvykle plní DSA, ošetřuje již primární systém.

5.3. Návrh struktury datového skladu

5.3.1. Dimenze

Název a popis dimenze	Typ dimenze	Zdroj dimenze (entita primárního systému)
Dodavatel	1	Partner (<i>pouze ti, kteří dodávají nějaké Zboží, jehož Druh zboží nejvyšší úrovně je Zboží; zmíněný druh zboží nejvyšší úrovně je parametrem ETL subsystému</i>)
Dopravce, Druh dopravy	1	Partner (<i>pouze ti, kteří dodávají nějaké Zboží, jehož Druh zboží je Doprava; zmíněný druh zboží je parametrem ETL subsystému</i>)
Měna	0 (neaditivní)	Měna
Měrná jednotka (neaditivní)	0 (neaditivní)	Měrná jednotka
Nákladová položka	1	Druh zboží (<i>nadřazený druh zboží je parametrem ETL subsystému</i>)
Nákupní kanál	1	Nákupní kanál

Nákupčí	1	Pracovník (<i>pouze ti, kteří vystavili nějakou Objednávku vydanou</i>)
Prodejní akce	0	Dealerská kategorie (<i>zdrojový systém přímo nepodporuje prodejní akce; jsou simulovány pomocí dealerských kategorií; dealerské kategorie vybírány pomocí podmínky na název</i>)
Prodejní kanál	1	Prodejní kanál
Prodejce	0	Pracovník (<i>pouze ti, kteří zadali nějakou Objednávku přijatou</i>)
Čas <i>Rok, Měsíc, Den v měsíci, Den v týdnu, Pracovní/nepracovní den</i>	0	Naplněno kalendářem
Sazba DPH	0	Sazba DPH
Stav objednávky vydané	0	Neměnný číselník (<i>hodnoty viz entita Objednávka vydaná</i>)
Stav dodacího listu přijatého	0	Neměnný číselník (<i>hodnoty viz entita Dodací list přijatý</i>)
Stav příjemky	0	Neměnný číselník (<i>hodnoty viz entita Příjemka</i>)
Stav faktury/dobropisu přijatého	0	Neměnný číselník (<i>hodnoty viz entita Faktura/dobropis přijatý</i>)
Stav reklamace nákupu	0	Neměnný číselník (<i>hodnoty viz entita Reklamace nákupu</i>)
Stav objednávky přijaté	0	Neměnný číselník (<i>hodnoty viz entita Objednávka přijatá</i>)
Stav dodacího listu vydaného	0	Neměnný číselník (<i>hodnoty viz entita Dodací list vydaný</i>)

Stav výdejky	0	Neměnný číselník (<i>hodnoty viz entita Výdejka</i>)
Stav faktury/dobropisu vydaného	0	Neměnný číselník (<i>hodnoty viz entita Faktura/dobropis vydaný</i>)
Stav reklamace prodeje	0	Neměnný číselník (<i>hodnoty viz entita Reklamace prodeje</i>)
Druh zboží, Zboží	2	Druh zboží (<i>potomci Druhu zboží nejvyšší úrovně označujícím zZboží; zmíněný druh zboží nejvyšší úrovně je parametrem ETL subsystému</i>) Zboží (<i>pouze potomci výše uvedených druhů</i>)
Zákazník, Dealerská kategorie	3	Partner, Dealerská kategorie (<i>pouze ti Partneři, kteří mají či měli přidělenou nějakou Dealerskou kategorii</i>)

5.3.2. Míry a jejich seskupení do kostek

Kostka: Objednávky přijaté

Popisuje základní kvantitativní údaje o přijatých objednávkách

Zdroj: Objednávka přijatá, Položka objednávky přijaté, Zboží

Míry:

Počet přijatých objednávek, počet položek přijatých objednávek

Množství v základních měrných jednotkách (měrná jednotka neaditivní dimenze)

Hodnota v měně

Hodnota v Kč

Daň z přidané hodnoty

Dimenze:

Měrná jednotka

Měna

Zboží

Zákazník, Dealerská kategorie

Sazba DPH

Stav objednávky

Prodejní akce (při objednávce za akční ceny)

Rok, Měsíc, Den v měsíci, Den v týdnu, Pracovní/nepracovní den

Prodejní kanál, Prodejce

Kostka: Dodané zboží

Zdroj: Dodací listy vydané, Faktury vydané

Míry:

Počet vydaných faktur, počet položek vydaných faktur

Počet vydaných dodacích listů, počet položek vydaných dodacích listů

Množství dodané v základních měrných jednotkách

Množství fakturované v základních měrných jednotkách

Hodnota v dodacím listu v měně

Hodnota v dodacím listu v Kč

Hodnota na faktuře v měně

Hodnota na faktuře v Kč

Daň z přidané hodnoty

Vedlejší náklady pořízení v měně

Vedlejší náklady pořízení v Kč

Čas vyřízení (od přijetí objednávky po odevzdání zboží dopravci)

Pozn. Pro vyčíslení ukazatelů Objem (dm³) a Hmotnost (kg) neobsahuje primární systém potřebné údaje.

Dimenze:

Měrná jednotka

Měna

Zboží

Zákazník, Dealerská kategorie

Sazba DPH

Stav faktury vydané

Stav dodacího listu vydaného

Prodejní akce (při objednávce za akční ceny)

Rok, Měsíc, Den v měsíci, Den v týdnu, Pracovní/nepracovní den

Prodejní kanál, Prodejce

Kostka: Reklamace přijaté

Zdroj: Reklamace přijaté

Míry:

Počet přijatých reklamací, počet položek přijatých reklamací

Množství v základních měrných jednotkách (měrná jednotka neaditivní dimenze)

Hodnota v měně

Hodnota v Kč

Daň z přidané hodnoty

Dobropisováno v měně

Dobropisováno v Kč

Dobropisovaná DPH

Náklady na reklamace

Dimenze:

Nákladová položka

Měrná jednotka (jen pro množství, neaditivní, možno použít pro rozštěpení kostky do více kostek)

Měna

Zboží

Zákazník, Dealerská kategorie (možno srovnání dealerských kategorií v čase)

Sazba DPH

Stav reklamace

Prodejní akce (při objednávce za akční ceny)

Rok, Měsíc, Den v měsíci, Den v týdnu, Pracovní/nepracovní den

Prodejní kanál, Prodejce

Dopravce, Druh dopravy

Kostka: Objednávky vydané

Popisuje základní kvantitativní údaje o vydaných objednávkách

Zdroj: Objednávka vydaná, Položka objednávky vydané, Zboží

Míry:

Počet vydaných objednávek, počet položek vydaných objednávek

Množství v základních měrných jednotkách (měrná jednotka neaditivní dimenze)

Hodnota v měně

Hodnota v Kč

Daň z přidané hodnoty

Dimenze:

Měrná jednotka

Měna

Zboží

Dodavatel

Sazba DPH

Stav objednávky vydané

Prodejní akce (při objednávce za akční ceny)

Rok, Měsíc, Den v měsíci, Den v týdnu, Pracovní/nepracovní den

Nákupní kanál, Nákupčí

Kostka: Nakoupené zboží

Zdroj: Dodací listy přijaté, Faktury přijaté

Míry:

Počet přijatých faktur, počet položek přijatých faktur

Počet přijatých dodacích listů, počet položek přijatých dodacích listů

Množství dodané v základních měrných jednotkách

Množství fakturované v základních měrných jednotkách

Hodnota v dodacím listu v měně

Hodnota v dodacím listu v Kč

Hodnota na faktuře v měně

Hodnota na faktuře v Kč

Daň z přidané hodnoty

Vedlejší náklady pořízení v měně

Vedlejší náklady pořízení v Kč

Dimenze:

Měrná jednotka

Měna

Zboží

Dodavatel

Sazba DPH

Stav faktury přijaté

Stav dodacího listu přijatého

Prodejní akce (při objednávce za akční ceny)

Rok, Měsíc, Den v měsíci, Den v týdnu, Pracovní/nepracovní den

Nákupní kanál, Nákupčí

Kostka: Reklamace vydané

Zdroj: Reklamace vydané

Míry:

Počet vydaných reklamací, počet položek vydaných reklamací

Množství v základních měrných jednotkách (měrná jednotka neaditivní dimenze)

Hodnota v měně

Hodnota v Kč

Daň z přidané hodnoty

Dobropisováno v měně

Dobropisováno v Kč

Dobropisovaná DPH

Náklady na reklamace

Dimenze:

Nákladová položka

Měrná jednotka

Měna

Zboží

Zákazník, Dealerská kategorie (možno srovnání dealerských kategorií v čase)

Sazba DPH

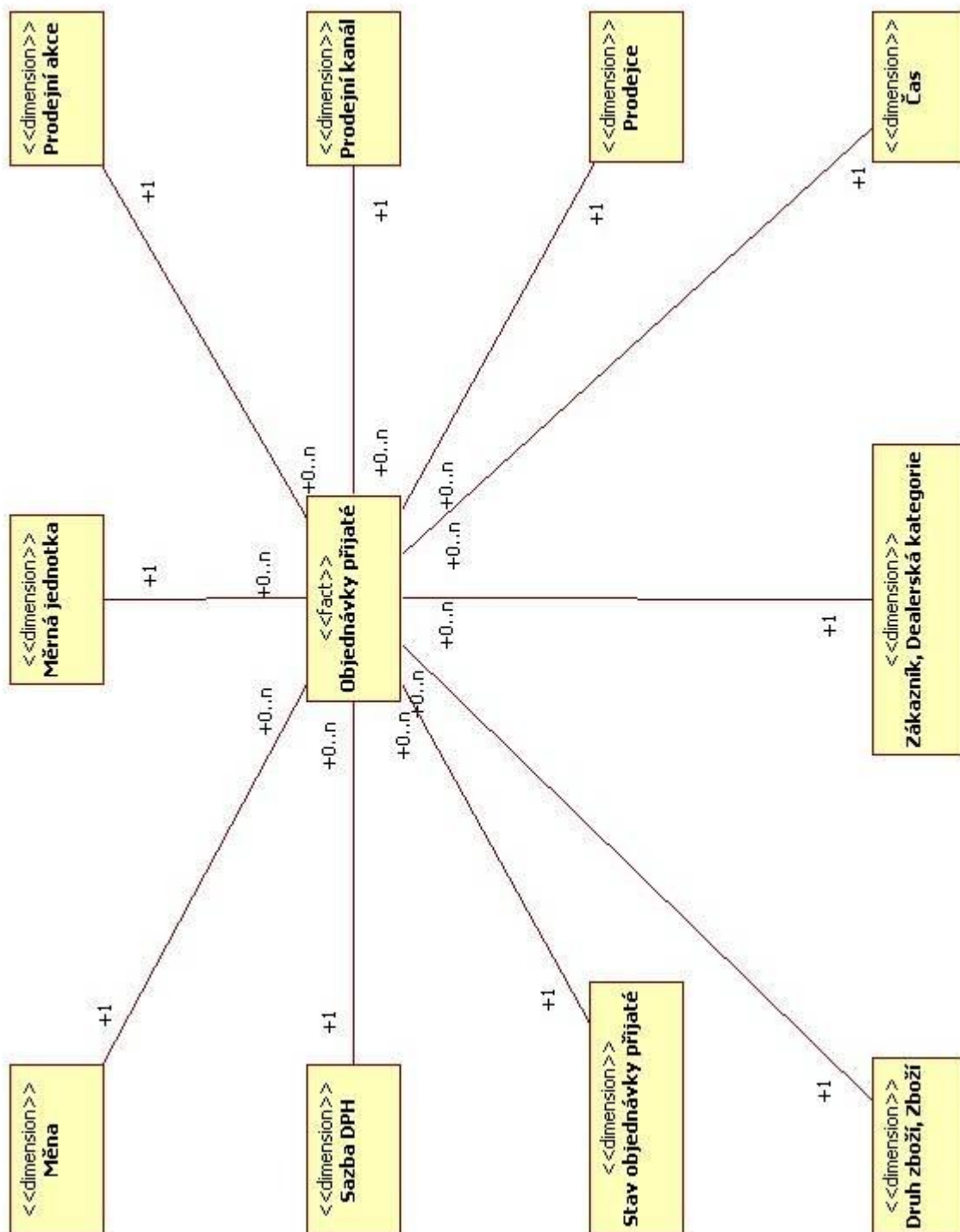
Stav reklamace vydané

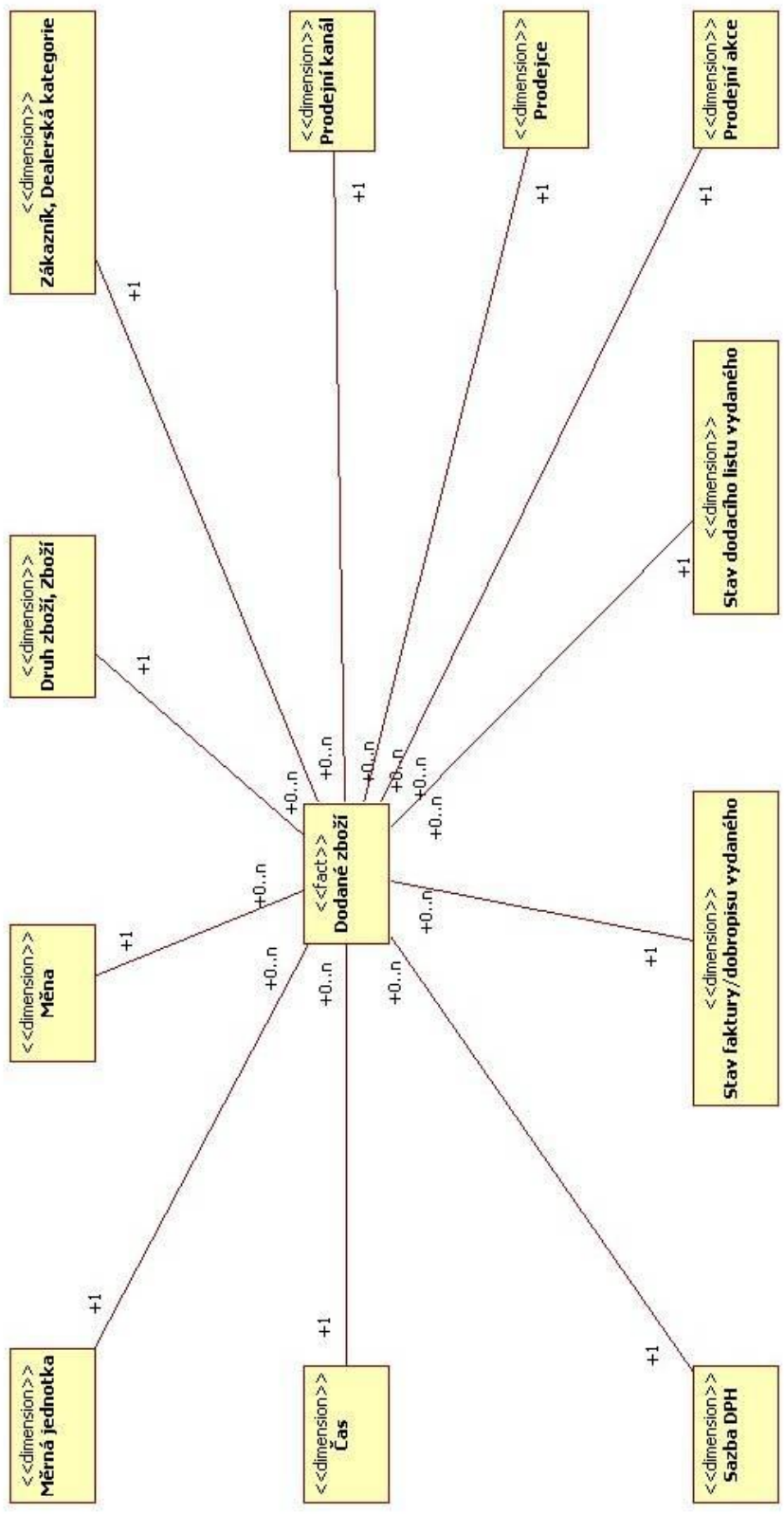
Rok, Měsíc, Den v měsíci, Den v týdnu, Pracovní/nepracovní den

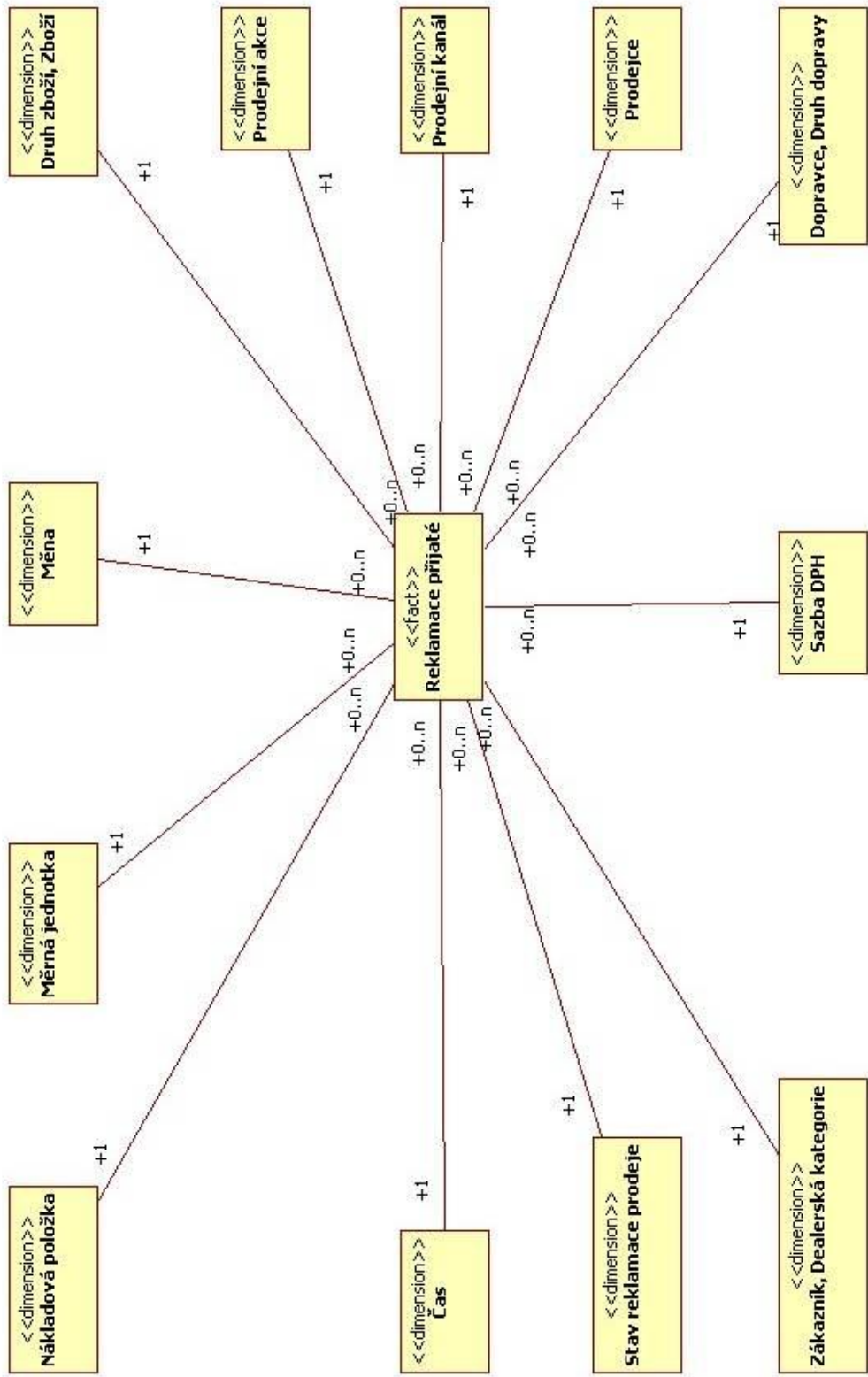
Nákupní kanál, Nákupčí

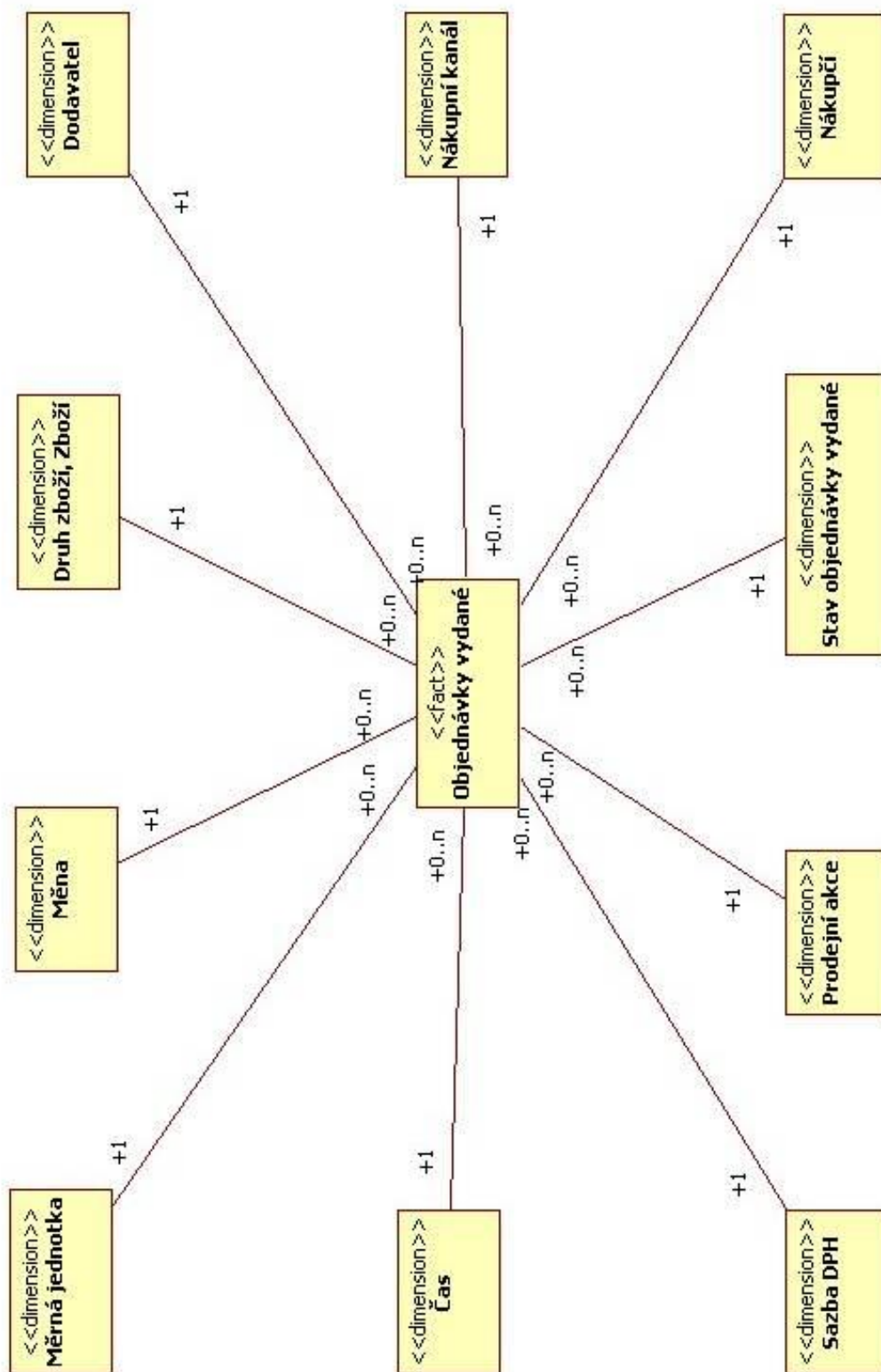
Dopravce, Druh dopravy

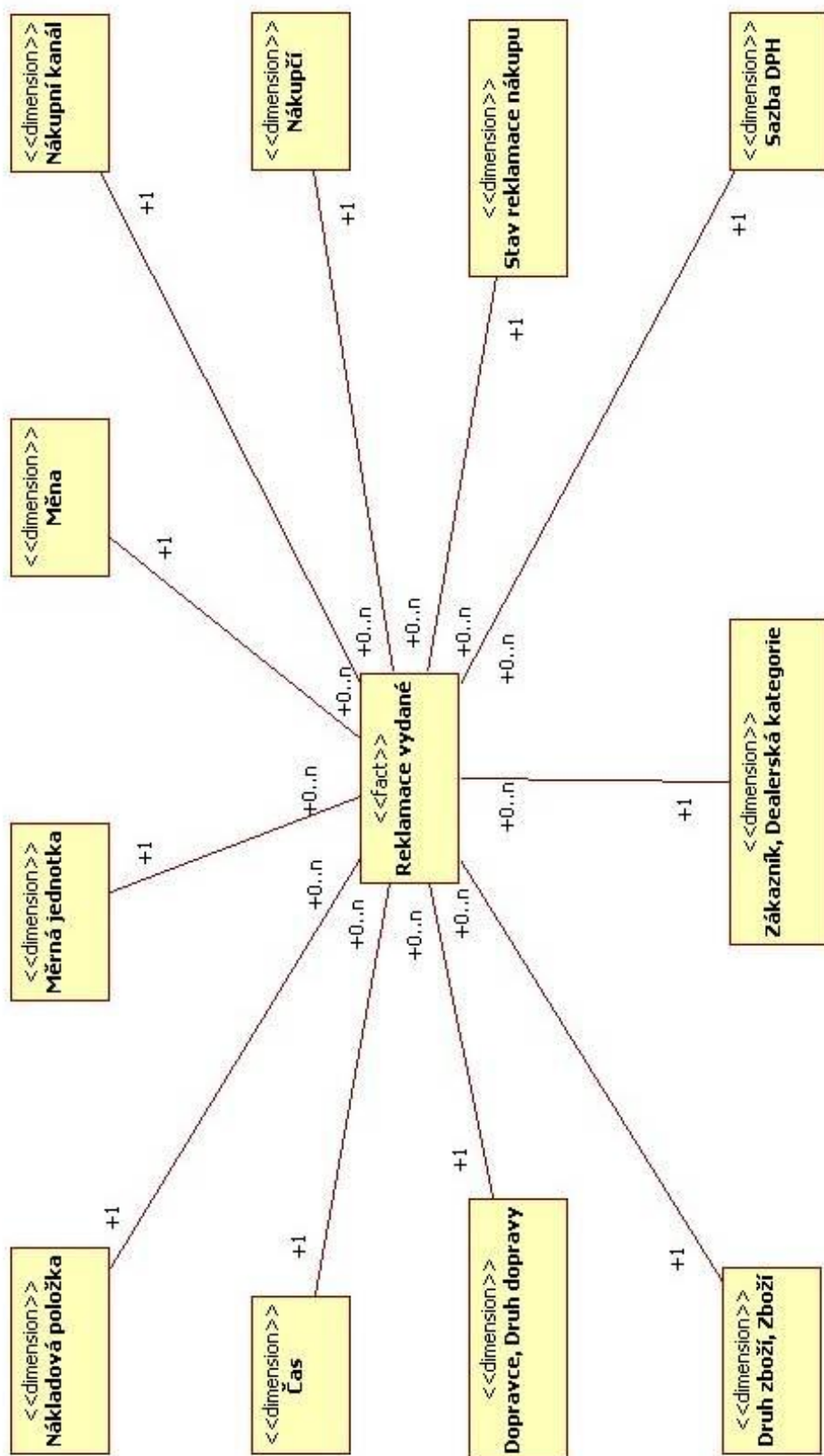
5.3.3 Datový model











6. Závěr

Práce po krátkém vysvětlení variant architektury datového skladu uvádí některé osvědčené postupy při vytváření dimenzí v OLAP kostkách. Poté ukazuje postup budování datového skladu na příkladu informačního systému fiktivní obchodní firmy.

V rámci budování datového skladu byly provedeny následující analytické a projekční části:

- analýza hodnototvorných procesů obchodní firmy
- definice požadavků na datový sklad jako seznam ukazatelů výkonnosti analyzovaných procesů a jejich potřebného členění
- reverse engineering zdrojových systémů
- výběr architektury datového skladu
- návrh struktury datového skladu vč. definování korespondence mezi dimenzemi a mírami datového skladu a zdrojovými systémy.

Dokumentace podrobného technického návrhu datového skladu nebyla z důvodu rozsahu do této práce pojata. Jako důkaz, že byl skutečně podrobný technický návrh proveden, poslouží funkční vzorek částečně realizovaného datového skladu, který bude předveden v rámci obhajoby práce.

Jako technologická platforma pro realizaci byl zvolen relační databázový stroj Microsoft SQL Server 2008, jehož součástí je i multidimenzionální databázový stroj Microsoft Analysis Services.

Seznam literatury

- [Humphries, 2001] Mark Humphries a kol., Data warehousing – návrh a implementace, Computer Press 2001, ISBN 80-7226-560-1
- [Inmon, 1992] William Inmon, Building the Data Warehouse, QED Technical Publishing Group 1992, ISBN 0-89435-404-3
- [Kimball,2002] Ralph Kimball, Margy Ross, The Data Warehouse Toolkit – The Complete Guide to Dimensional Modelling, Second Edition, John Wiley and Sons, Inc, 2002, ISBN 0-471-20024-7
- [Kimball,2004] Ralph Kimball, Joe Caserta, The Data Warehouse ETL Toolkit – Practical Techniques for Extracting, Cleaning, Conforming and Delivering Data, Wiley Publishing, Inc, 2004, ISBN 0-7645-6757-8
- [McKnight] William McKnight, SQL and the data warehouse: Tip on date dimensions, dostupné z: < <http://searchcrm.techtarget.com/> >
- [Oracle, 1] Oracle® Database Concepts 10g Release 2 (10.2), dostupné z : <http://www.oracle.com>
- [Oracle, 2] Oracle® Database Data Warehousing Guide 10g Release 2 (10.2), dostupné z : <http://www.oracle.com>
- [Microsoft, 1] Microsoft SQL Server 2008 Books Online, dostupné jako součást instalace produktu Microsoft SQL Server 2008
- [Porter] Konkurenční vyhoda - jak vytvořit a udržet si nadprůměrný výkon. Praha : Victoria Publishing, 1993, ISBN: 80-85605-12-0
- [Vajgert] Datové sklady. Klatovy, 2007. Bakalářská práce. Česká zemědělská univerzita v Praze, Provozně ekonomická fakulta, Katedra informačního inženýrství