Vysoká škola ekonomická v Praze

Fakulta informatiky a statistiky

Security model pro Self service BI tabular cube

BAKALÁŘSKÁ PRÁCE

Studijní program: Aplikovaná informatika

Studijní obor: Aplikovaná informatika

Autor: Nikola Pátková

Vedoucí bakalářské práce: doc. Ing. Jan Pour, CSc.

Praha, květen 2019

Prohlášení

Prohlašuji, že jsem bakalářskou práci Security model pro Self service BI tabular cube vypracovala samostatně za použití v práci uvedených pramenů a literatury.

V Praze dne 3. května 2019 .....

Nikola Pátková

Poděkování

Tímto bych chtěla poděkovat především panu docentu Janu Pourovi, za jeho vedení této bakalářské práce, rychlé reakce, cenné rady a věcné připomínky. Dále bych chtěla poděkovat své rodině a svému příteli za podporu a v neposlední řadě svým kolegům za podporu při výběru tématu a cenným radám při zpracování praktického modelu této práce.

Abstrakt

Cílem této práce bylo vytvoření návrhu a následná implementace funkčního security modelu pro SSAS Tabular datovou kostku. Tabulární datový model byl vytvořen za účelem automatizace manuálního reportu, podpory samostatnosti uživatelů, a tak i zabezpečení tohoto modelu bylo navrženo tak, aby bylo možné ho spravovat odpovědnými uživateli. V rámci tohoto bezpečnostního modelu byla vytvořena stránka v aplikaci sharepoint, která obsahuje seznam uživatelů a jim přidělených práv, který bude spravován pověřenými uživateli. V době vytváření této práce měl přístup k tomuto listu pouze tým vývojářů. Pro takové řešení bylo použito dalších technologií, mezi které patří například Access Control List tabulky nacházející se v relační databázi, kde jsou data ze sharepointu uchovávána, a pro získání a vložení těchto dat do relační databáze je využita technologie SQL Server Integration Services.

V první části práce autorka seznámila čtenáře s technologiemi a s pojmy souvisejícími s tématem této práce, v části druhé bylo analyzováno podobné řešení zabezpečení reportů a následně současný stav tabulárního modelu a jeho zabezpečení. V části následující byl již popsán samotný proces návrhu a implementace security modelu do tabulární datové kostky.

Klíčová slova

ACL, business intelligence, security, sharepoint, SSIS, tabular.

Abstract

The goal of this thesis was to create a design and then implement a functional security model for the SSAS Tabular Data Cube. The Tabular data model was created to automate manual report, support user autonomy, and the security of this model was designed to be managed by responsible users. As a part of this security model, a sharepoint site has been created and contains a list of users and their granted permissions which will be managed by authorized users. At the time this thesis was created, only a team of developers had access to this list. In this solution were used also other technologies such as Access Control List, located in the relational database, where data from the sharepoint is stored, and SQL Server Integration Services to extract and load this data in a relational database.

In the first part of the thesis the reader was introduced to the technologies and concepts related to the topic of this thesis, in the second part a similar solution of report security was analysed and subsequently was described the current status of the tabular model and its security. In the following section, the process of designing and implementing a security model into a tabular data cube was described.

Keywords

ACL, business intelligence, security, sharepoint, SSIS, tabular.

Obsah

Úvod 1
1 Rešerše informačních zdrojů 4
2 Business Intelligence a komponenty 6
2.1 Business Intelligence 6
2.1.1 Vymezení pojmu 6
2.1.2 Shrnutí 7
2.2 Self-Service Business Intelligence 8
2.2.1 Vymezení pojmu 8
2.2.2 Nástroje SSBI 9
2.2.3 Shrnutí 10
2.3 Analytické služby (Analysis Services) 10
2.3.1 Vymezení pojmu 10
2.3.2 Shrnutí 11
2.4 Multidimenzionální datové kostky 11
2.4.1 Multidimenzionální databáze 11
2.4.2 OLAP (Online Analytical Processing) databáze 11
2.4.3 Shrnutí 12
2.5 Tabulární model a jeho engine 12
2.5.1 Představení Tabulárního modelu 12
2.5.2 Engine 13
2.5.3 Shrnutí 15
2.6 Perspektivy 16
2.6.1 Vymezení pojmu 16
2.6.2 Shrnutí 16
2.7 Security handling v tabulárním modelu 16
2.7.1 Databázové Role 16
2.7.2 Shrnutí 17
2.8 Active Directory 17
2.8.1 Vymezení pojmu 17
2.8.2 Shrnutí 18
2.9 Sharepoint 18
2.9.1 Vymezení pojmu 18
2.9.2 Shrnutí 18

2.10 ACL tabulky 19
2.10.1 Vymezení pojmu 19
2.10.2 Shrnutí 19
2.11 Integration services 19
2.11.1 Vymezení pojmu 20
2.11.2 Shrnutí 20
3 Návrh a implementace security modelu 21
3.1 Analýza podobného řešení 21
3.1.1 Popis 21
3.1.2 Shrnutí 22
3.2 Analýza stávajícího řešení 22
3.2.1 Popis 22
3.2.2 Shrnutí 23
3.3 Workflow návrhu 23
3.3.1 Vytvoření sharepoint subsite 24
3.3.2 Návrh sharepointové stránky 24
3.3.3 Vytvoření schématu databází 29
3.3.4 Vytvoření potřebných tabulek 31
3.3.5 ACL SSIS 34
3.3.6 ACL Stored Procedure 45
3.3.7 Import ACL tabulek do tabulárního modelu 47
3.3.8 Vhodné použití a nastavení ACL tabulek v tabulárním modelu 49
3.3.9 Testování 51
Závěr 52
Terminologický slovník 54
Použitá literatura 57

Úvod

Úvod do problematiky a důvod výběru tématu

Tato bakalářská práce se zabývá problematikou bezpečnosti v oblasti Business Intelligence, konkrétně Self Service Business Intelligence. Bezpečnost dat je pro podniky zásadní a proto je nutné omezit přístup k datům pro jednotlivé uživatele přistupující k datům například pomocí reportů, které využívají data z datových kostek či z databází.

Autorka je součástí týmu Business Intelligence developerů v logistické společnosti, kde získává své první zkušenosti v tomto oboru. Během svého působení v tomto týmu se setkala s několika různými úkoly týkající se nejen podpory reportingu, ale také vývoje. Tým má na starosti podporu reportingu a vytváří různé reporty pro různé části podniku a často je nutné vytvořit report, který budou moci používat i uživatelé z konkurenčních částí podniku. To ovšem vyžaduje restrikce.

Autorka se společně s kolegy aktuálně podílí na migraci dat ze starého serveru na nový, což otevírá nové možnosti pro kreativitu a zlepšení procesů nad migrovanými daty. Této příležitosti tým developerů využil k tomu, aby stávající řešení reportingu posunul na úroveň Self Service Business Intelligence. Byla vytvořena datová kostka typu Tabular, o které bude zmínka v praktické části. Před tím, než bude kostka nasazena na produkční server, je potřeba ji důkladně zabezpečit a zamezit tak uživatelům přístup k datům, která pro ně nejsou určena či jsou to data konkurenční části podniku. A touto problematikou se bude zabývat tato práce.

CÍLE PRÁCE

Hlavním cílem této práce je návrh a implementace Security modelu využívající ACL tabulky pro datovou kostku typu Tabular.

Dílčími cíli, které povedou k dosažení hlavního cíle práce, jsou následující:

- Vytvoření stránky na sharepointu
- Návrh a implementace vhodného designu sharepointové stránky
- Vytvoření schématu v databázi pro tento projekt
- Vytvoření ACL tabulek
- Vytvoření ETL datové pumpy získávající data ze sharepointové stránky
- Propojení vytvořených ACL tabulek s Tabulární datovou kostkou z pohledu zabezpečení dat
- Testování vytvořeného Security modelu

Výsledkem práce bude přiblížit se k nejvhodnějšímu zabezpečení tabulární kostky, které bude splňovat podmínky omezení přístupu k datům na základě uživatelského přihlašovacího jména a jemu přidělené podnikové části. Toto řešení však bude podporovat samostatnost managementu v udělování přístupu k datům.

ZPŮSOB DOSAŽENÍ CÍLE

K úspěšnému dosažení cílů definovaných v předchozí podkapitole autorce pomohla především spolupráce na výše uvedené migraci dat, znalost podnikových systémů a procesů, avšak největším přínosem pro ni byla spolupráce s kolegy se seniorními znalostmi. Pro výběr řešení zabezpečení datové kostky ji bylo inspirací stávající zabezpečení několika reportů.

Dále bylo třeba spolupráce některého z kolegů k testování funkčnosti bezpečnostního modelu. Autorka se musela také seznámit s tvorbou a návrhem stránek v aplikaci sharepoint.

Pro správné zpracování dat ze sharepointu musela autorka nejprve analyzovat stávající řešení zabezpečení několika reportů, která pro vyhodnocení přístupu k datům pro jednotlivé uživatele využívají právě data uvedená na sharepointu.

K dosažení cíle bylo nutné seznámit se s vytvořeným tabulárním modelem a vytvoření vhodného návrhu relačních tabulek, které budou splňovat 3. normální formu a lze je importovat do vytvořeného tabulárního modelu.

PŘEDPOKLADY A OMEZENÍ PRÁCE

Předpokladem pro úspěšné dosažení cílů bakalářské práce je zájem společnosti DHL omezit přístup k datům z důvodu konkurence různých částí podniku. K tomu je vhodná komunikace s koncovými uživateli a vytvoření návrhu bezpečnostního modelu, který bude schválen jak koncovými uživateli, tak vývojáři.

Dalším z předpokladů je, aby informace obsažené ve zdrojích, ze kterých autorka čerpá, byly důvěryhodné a správné, aby jej bylo možné použít pro definici pojmů uvedených v celé práci.

Jedním z omezení práce je ta skutečnost, že autorka nebude moci uvádět reálná data či informace a následně je využívat z důvodu bezpečnosti dat společnosti.

STRUKTURA PRÁCE

Práce je strukturována podle doporučení Katedry informačních technologií na Vysoké škole ekonomické v Praze. Práce je rozdělena do dvou ucelených částí, jež jsou teoretická a následně praktická část.

Po úvodu práce bude následovat část teoretická, kde budou vysvětleny pojmy, technologie a komponenty využívané při vypracování praktické části. Nejprve se autorka krátce zmíní o Business Intelligence. Poté bude následovat přiblížení Self Service Business Intelligence, kde se autorka zaměří především na nástroje Self Service Business Intelligence. Následně bude přiblíženo téma analytických služeb a Multidimenzionálních datových kostek, kde nejvíce pozornosti bude autorka věnovat tabulárnímu modelu datové kostky a jeho zpracování dat pomocí enginů. To povede k tématu Bezpečnosti datové kostky, kde autorka přiblíží čtenářům různé možnosti zabezpečení dat v tabulární kostce. Součástí teoretické

části bude v neposlední řadě zmínka také o sharepointu a o technologii použité pro získání dat.

Část praktická bude obsahovat analýzu podobného řešení zabezpečení dat reportů zmiňovaných v abstraktu práce. Po této analýze bude následovat analýza stávajícího řešení zabezpečení dat na starém serveru, na které bude dále navazovat workflow návrhu řešení a následný popis a implementace tohoto řešení na testovacích datech a testování jeho funkčnosti.

V závěru práce autorka zhodnotí úspěšnost či neúspěšnost zavedeného security modelu a dosažení stanovených cílů.

OČEKÁVANÉ PŘÍNOSY

Hlavním přínosem práce bude úspěšné dosažení cíle, neboli zavedení funkčního Security modelu na tabulární datovou kostku na testovací úrovni. To by dále mohlo vést k potřebným úpravám security modelu, například omezení přístupu k datům na nižších úrovních a následné zavedení tohoto zabezpečení také do produkce po konzultaci a odsouhlasení manažerů jednotlivých Business Unit.

Dalším z očekávaných přínosů, ač subjektivním, bude autorčino rozšíření znalostí v oblasti bezpečnosti dat a datových kostek. Autorka se stejně jako čtenáři této práce dozví o zabezpečení dat v podniku, především zabezpečení dat v tabulárním modelu datové kostky.

1 Rešerše informačních zdrojů

Tato kapitola pojednává o rešerši dostupných zdrojů informací, které byly použity při zpracování této bakalářské práce.

Již při výběru tématu bakalářské práce byly využity interní podnikové zdroje. Jelikož se jedná o podnikové zdroje, nejsou uvedeny názvy dokumentů, bude pouze nastíněn jejich užitečný obsah.

Jedním z interních podnikových zdrojů, který byl použit při zpracování této práce, je technická dokumentace podobného řešení implementovaného pro reporty. Tato dokumentace byla vytvořena společně s reporty před 3 roky, avšak bohužel nebyla aktualizována nově implementovanými úpravami řešení. I přes to tato dokumentace poskytla autorce základní informace o možném řešení zabezpečení za použití aplikace Microsoft Sharepoint 2016. Základními informacemi byla logika toku dat a informace o umístění SSIS balíčku na serveru používaném při prvotním řešení zabezpečení reportů.

Stěžejním typem informací pro zpracování teoretické i praktické části byly internetové články společnosti Microsoft, jelikož autorka používala při tvorbě bezpečnostního modelu aplikace vyvinuté společností Microsoft.

V práci byly převážně využity články zabývající se tématem SQL Server 2017. Pod tímto tématem lze dohledat články zabývající se tématikou Tabulárního datového modelu, které byly využívány nejvíce. Jednotlivé články především seznamují čtenáře s Tabulárním modelem a jeho komponenty a poskytují návod na vytváření a správu tabulárního modelu. Tyto články autorce pomohly k pochopení tématiky a při návrhu samotného security modelu.

Dalším významným internetovým zdrojem informací byly stránky podpory společnosti Microsoft, které autorce dopomohly k pochopení především technologií OLAP, ale také například k získání informací o nástrojích Self Service Business Intelligence.

Mezi významné knižní publikace, které byly autorkou využity při přípravě této práce či dokonce v práci samotné, patří především kniha Self Service Business Intelligence (Pour et al. 2018), která byla velmi významnou pro získání znalostí o technologii Self Service Business Intelligence a jejích nástrojích.

Kniha Podniková informatika (Gála et al. 2009) byla autorkou využita pro získání základních informacích o technologii Business Intelligence a jejích komponentech. Tato kniha byla použita například pro vymezení pojmu Business Intelligence.

Kniha OLAP (Lacko 2003) byla autorce nápomocna pro získání dalších informacích o multidimenzionálních modelech, jejich správě a použití.

Diplomová práce Možnosti analytických nástrojů v prostředí MS SQL Serveru Bc. Markéty Mazáčové (Mazáčová 2017) byla poměrně dosti nápomocná při získávání dalších informací

o tabulárním modelu, jeho engine a především analytických nástrojích v prostředí MS SQL Server, tedy pro získání jiného pohledu a jiného vysvětlení pojmů, které byly načteny z internetových zdrojů.

Zdroje uvedené v této kapitole se zabývají shodným tématem jako tato bakalářská práce. První zmíněná knižní publikace se zabývá problematikou Self Service Business Intelligence, která byla pro zpracování této práce bohatým zdrojem informací. Druhá zmíněná publikace doplnila informace o problematice a komponentech Business Intelligence. Třetí zmíněná kniha se zabývala tématem OLAP, avšak nezabývala se zcela zabezpečením těchto technologií. Zmíněná diplomová práce se zabývá analytickými nástroji v prostředí MS SQL Serveru, avšak nebyla zaměřena pouze na tabulární model a jeho zabezpečení.

2 Business Intelligence a komponenty

2.1 Business Intelligence

2.1.1 Vymezení pojmu

Cílem této kapitoly je přiblížit čtenářům pojem Business Intelligence. Existuje mnoho definic pojmu Business Intelligence, v této kapitole však bude uvedena pouze jedna z nich. Vymezení tohoto pojmu bylo čerpáno z knižní publikace Podniková informatika (Gála et al. 2009).

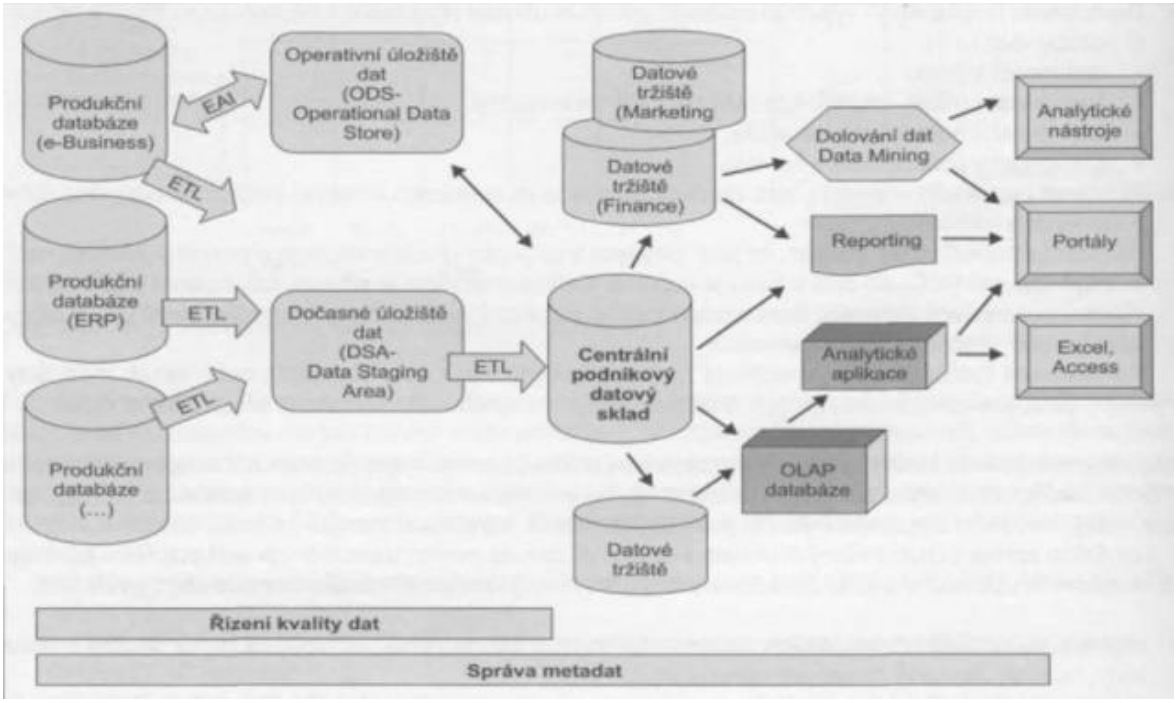
„Business Intelligence je sada procesů, know-how, aplikací a technologií, jejichž cílem je účinně a účelně podporovat řídící aktivity ve firmě. Podporují analytické, plánovací a rozhodovací činnosti organizací na všech úrovních a ve všech oblastech podnikového řízení, tj. prodeje, nákupu, marketingu, finančního řízení, controllingu, majetku, řízení lidských zdrojů, výroby a dalších.“

Záměrem BI je co nejefektivněji využít podniková data nejen pouze k jejich zpracování a realizaci běžných transakcí, ale především v rozhodování. Proto je tedy nutné nad těmito daty uvažovat multidimenzionálně (z více úhlů pohledu). Na základě těchto modelů jsou následně sestavovány statistické ukazatele, které pomáhají manažerům při samotném strategickém rozhodování. Správně použitá data poté mohou podnikům dokonce zajistit konkurenční výhodu. Díky BI aplikacím lze data využít ke snižování nákladů, zefektivnění podnikových procesů, a mnoho dalších. V současném řešení BI uživatelé přistupují k podnikovým datům přes reporty či dashboardy a pomocí nich mohou sledovat a hodnotit klíčové indikátory podniku či svého oddělení, analyzovat trendy a predikovat budoucí vývoj a řídit tak výkon firmy či svého oddělení.

Obecná architektura BI se skládá z následujících komponent:

- Produkční systémy – zdrojové aplikace pro BI
- ETL komponenty – extrakce, transformace a nahrávání dat do analytických databází
- Enterprise Application Integration (EAI)
- Data Staging Area – dočasné úložiště dat
- Operativní úložiště dat (ODS)
- Data Warehouse (DWH) – datové sklady
- Data Mart – datové tržiště, obsahují data podle záměru reportingu
- OLAP (Online Analytical Processing) databáze
- Reporting
- Analytické aplikace
- Data Mining – dolování dat
- Nástroje pro řízení kvality dat a správu metadat

Jak lze vidět na obrázku 1, BI získává data z produkčních databází, což jsou systémy podniku (například ERP, SCM), jejichž vlastností je architektura, která podporuje ukládání a modifikaci dat v reálném čase, nebo to mohou být také externí systémy. Tyto zdrojové systémy však nejsou určeny pro analytické úlohy. Data jsou z těchto produkčních databází získávána pomocí nástrojů ETL (Extract, Transform, Load), známé také jako datové pumpy. Úkolem ETL je vybrat a získat data ze zdrojových systémů, transformovat je do jednotného formátu a nahrát je do specifických datových struktur datového skladu, datového tržiště či OLAP databáze v dávkách po určitých časových intervalech.



Obrázek 1 Obecná architektura BI řešení (Gála et al. 2009)

Jednotlivé komponenty BI nebudou dále vysvětlovány v rámci této kapitoly. Komponenty, které jsou součástí řešení security modelu, budou vysvětleny v samostatných kapitolách.

2.1.2 Shrnutí

Cílem této kapitoly je přiblížit čtenářům pojem Business Intelligence a jeho komponent. Aplikace Business Intelligence nevytvářejí nová data, pouze používají data ze zdrojových systémů pro analýzy nad nimi.

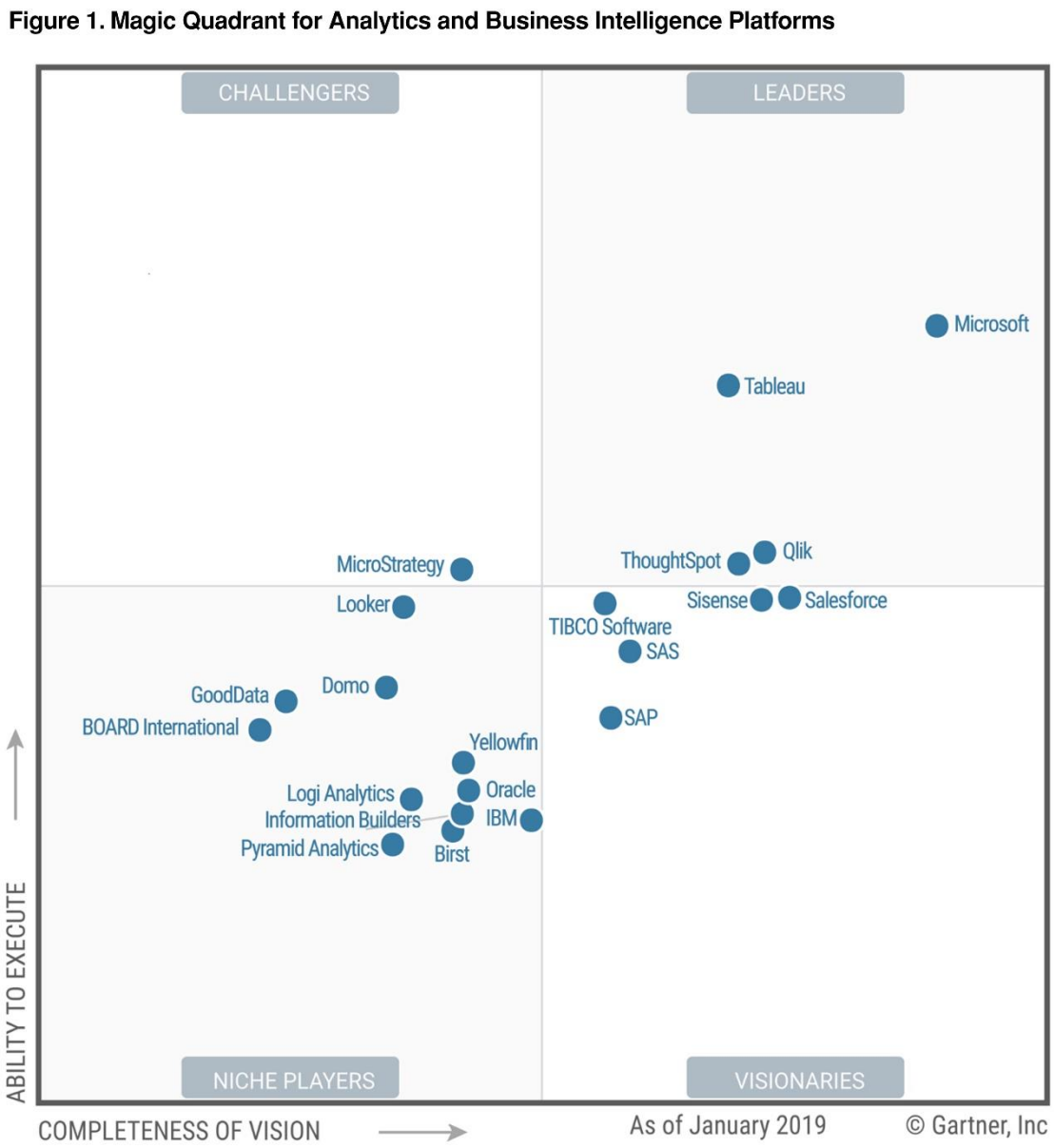
2.2 Self-Service Business Intelligence

2.2.1 Vymezení pojmu

Obecná definice pojmu Self Service Business Intelligence podle společnosti Gartner zní:“ *Self-Service Business Intelligence je zde definováno jako koncoví uživatelé navrhující a nasazující své vlastní reporty a analýzy v rámci schváleného a podporovaného portfolia architektury a nástrojů*“. (Gartner 2017)

Self-Service Business Intelligence je trend BI, jehož hlavním smyslem je soběstačnost uživatelů s nízkou technickou znalostí v oblasti zpracování podnikových dat, přístupu k datům, vytvářením si vlastních reportů a dashboardů. To je uživatelům umožněno pomocí několika nástrojů, mezi které patří například aplikace Microsoft Excel, PowerPivot, PowerBI, QlikView a další. (Pour et al. 2018)

Jak lze vidět na obrázku 2, podle společnosti Gartner je společnost Microsoft lídrem na trhu Analýzy a Business Intelligence.



Obrázek 2 Trh produktů BI (Gartner 2019)

2.2.2 Nástroje SSBI

Účelem této kapitoly je krajně čtenáře seznámit s nástroji Self Service Business Intelligence. Pro příklad budou uvedeny 3 nástroje.

Excel je aplikace, vyvinutá společností Microsoft pro vytváření tabulek a práce s nimi. Pro zpracování dat je v Excelu předdefinováno nespočet vzorců a funkcí, mezi které patří matematické funkce, logické funkce, vyhledávací a odkazovací funkce, textové funkce a převodní funkce. Excel umožňuje také přístup k databázím Analysis Services nebo import dat z různých zdrojů, jako například soubor csv. Do Excelu je možné nainstalovat několik rozšíření, ze kterých lze pro účely této práce vyzdvihnout SQL Server Master Data Services (MDS) a Power Pivot.

Rozšíření MDS umožňuje připojení se k MDS úložištím na databázových serverech, nahrání dat do Excelu, úprava či přidání nových záznamů, publikace nových dat z pracovního listu do MDS úložiště a další. Ke stažení dat do MDS úložiště je nutné vytvořit připojení k MDS databázi/ Master Data Manager webové aplikaci. (Guyer et al. 2017)

Jedním ze základních produktů a další doplňující komponentou, které lze využít v SSBI, je Power Pivot. V Power Pivot lze zobrazit a spravovat datový model, graficky znázorňovat vztahy mezi tabulkami, vytvářet složité výpočty, indikátory, vzorce a mnoho dalších. Díky sloupcově orientované databázi umožňuje efektivní komprimaci dat, tudíž lze pomocí Power Pivot pracovat s většími objemy dat, používá jazyk DAX, umožňuje integraci dat z různých datových zdrojů a dotazy vykonává v paměti. (MBI 2019)

Velmi známým nástrojem pro analýzu v SSBI je aplikace PowerBI, také od společnosti Microsoft. PowerBI je reportovací nástroj, který umožňuje připojení se k databázovým serverům či datovým kostkám. Aplikace umožňuje tvorbu datových modelů, reportů a dashboardů. (Stanovská nedatováno)

2.2.3 Shrnutí

SSBI je nástroj, který umožňuje uživatelům s menší znalostí IT využívat data pro ně potřebným a nejvhodnějším způsobem. Mohou si vytvářet své vlastní dashboardy, KPI a další potřebné komponenty v aplikacích SSBI. Aplikací podporující SSBI je na trhu dostatek a uživatelé si mohou vybrat dle jejich potřeb a představ. Lídrem na trhu Analytických nástrojů je společnosti Microsoft, která poskytuje hned několik nástrojů pro použití v SSBI.

Sama autorka bude využívat ke své práci aplikace od společnosti Microsoft především z toho důvodu, že splňují veškeré potřeby k dosažení cílů této práce. Pro účely této bakalářské práce bude k testování využit nástroj Power Pivot.

2.3 Analytické služby (Analysis Services)

2.3.1 Vymezení pojmu

Analysis services je služba v rámci databázového balíku MS SQL Server, která je používána pro podporu rozhodování a podnikové analýzy. Je to analytický datový engine, umožňující vytvářet analytické databáze, enterprise-grade podnikové modely sémantických dat pro použití v podnikových reportech a klientských aplikacích, jako jsou Power BI, Excel, reporty vytvořené pomocí služeb SQL Server Reporting Services a další nástroje pro vizualizaci dat. Je navržen tak, aby umožňoval získat co nejrychlejší odezvu z analytických procesů. (Duncan et al. 2018)

V analysis services je možné vytvořit 2 typy datových modelů, a to multidimenzionální a nebo tabulární. Tyto modely se vytvářejí jako projekt analytického datového modelu v aplikaci Visual Studio a je možné je nasadit se na server jako model analytické databáze. (Duncan et al. 2018)

2.3.2 Shrnutí

Služba SQL Server Analysis Services, vyvinuté společností Microsoft, slouží jako nástroj k analýze a pochopení podnikových dat, které mohou být rozděleny do více databází nebo do různých tabulek či souborů vzhledem k informacím, které nesou. Umožňuje vytváření analytických modelů nad daty a následně slouží jako zdroj pro klientské aplikace.

2.4 Multidimenzionální datové kostky

2.4.1 Multidimenzionální databáze

Multidimenzionální databáze slouží jako podklad pro získání sumarizovaných a agregovaných údajů za použití nenormalizovaných tabulek, a to tabulky faktů a dimenzí. Do multidimenzionálních databází se ukládají upravené a vyčištěné údaje. Hlavní výhody multidimenzionálních databází jsou:

- Rychlý komplexní přístup k velkému objemu dat
- Přístup k multidimenzionálním a relačním datovým strukturám
- Možnost komplexních analýz
- Silné schopnosti pro modelování a prognózy

Multidimenzionální databáze mají také své nevýhody oproti relačním databázím, a to jsou například:

- vyšší nároky na kapacitu úložiště
- problémy při změně dimenzí bez přizpůsobení časové dimenze
- a další

Analytické databáze se označují pojmem OLAP. (Lacko 2003)

2.4.2 OLAP (Online Analytical Processing) databáze

OLAP je technologie, která umožňuje uspořádání rozsáhlých obchodních databází a podporuje analytické nástroje. Databáze OLAP jsou rozděleny do jedné nebo více krychlí a jednotlivé krychle jsou uspořádány a navrženy správcem krychlí tak, aby vyhovovaly způsobu načítání a analýzy dat a usnadňovaly tak vytváření potřebných kontingenčních tabulek a grafů. (Support Microsoft 2019)

Typické využití systémů OLAP je pro analýzu velkého množství údajů. Výsledkem analýzy jsou souhrny a reporty, které slouží manažerům jako podklady pro jejich rozhodování. Pro výpočet krychlí OLAP je potřebné vykonat velké množství výpočtů a agregací, a to skoro v reálném čase. (Lacko 2003)

Narozdíl od datových skladů obsahují předzpracované agregace dat podle definovaných hierarchických struktur dimenzí a jejich kombinací.

Základním principem technologie OLAP je několika dimenzionální tabulka umožňující velmi rychle a pružně měnit jednotlivé dimenze a nabízet tak uživateli různé pohledy na modelovanou ekonomickou realitu. Tabulka zároveň nabízí možnost drill down (prokliknutí do detailu), drill up (prokliknutí z detailu zpět na obecná data) nebo slice and dice pro výběr („vykrojení“) dat z OLAP kostky. OLAP kostky takto zahrnují předzpracované agregace dat podle výše uvedených hierarchických struktur dimenzí a jejich kombinací. Jde tak v podstatě o princip „n-dimenzionální Rubikovy kostky“ naplněné nejdůležitějšími podnikovými daty. (Gála et al. 2009)

Použití technologie OLAP se jeví jako vhodné pokud:

- Je nutné provádět komplexní analýzy a ad hoc dotazy rychle, avšak s negativním dopadem na systémy OLTP (Online Transactional Processing)
- Pro poskytnutí firemním uživatelům jednoduchý způsob vygenerování reportů nad daty
- Pro poskytnutí celé řady agregací, které uživatelům umožní získat rychlé a konzistentní výsledky. (Tejada 2018)

Technologie OLAP se může objevovat v několika variantách:

- ROLAP (Relational OLAP) – využívá technologie relačních databází.
- MOLAP (Multidimensional OLAP) data jsou uložena v multidimenzionálních – binárních OLAP kostkách.
- HOLAP (Hybrid OLAP) – kombinace MOLAP a ROLAP, kdy detailní data jsou uložena v relační databázi a agregované hodnoty jsou uloženy v binárních OLAP kostkách.
- DOLAP (Desktop OLAP) – DOLAP umožňuje připojení k centrálnímu úložišti dat a stažení určité části kostky na lokální počítač či mobilní zařízení. Analytické operace jsou poté prováděny na lokálním zařízení, není tedy nutné připojení k serveru. (Gála et al. 2009)

2.4.3 Shrnutí

Cílem této kapitoly bylo seznámit čtenáře se základy technologie OLAP a zmínit také její další varianty.

2.5 Tabulární model a jeho engine

2.5.1 Představení Tabulárního modelu

Tabulární modely jsou analytické databáze v Analysis Services, které pro ukládání dat mohou používat 2 druhy engine. In-memory, což znamená, že vykonávají dotazy k datům v paměti aplikace, nebo v DirectQuery, což znamená, že data nejsou kopírována do aplikace,

přes kterou k nim přistupujeme (například Power BI Desktop), ale dotazy jsou vykonávány přímo do databází. Tabulární modely kombinují technologii MOLAP a relačních databází. (Rabeler et al. 2018)

Zatímco in-memory modely jsou výchozí, DirectQuery je alternativní režim dotazů pro modely, které jsou buď příliš velké na to, aby se vešly do paměti, nebo když volatilita dat (rychlost změny hodnot uložených dat za určité časové období) vylučuje rozumnou strategii zpracování. DirectQuery dosahuje shody s in-memory modely prostřednictvím podpory široké škály datových zdrojů, schopností zpracovávat kalkulované tabulky a sloupce v DirectQuery modelu, zabezpečení na úrovni řádků prostřednictvím výrazů DAX, které se dostanou do databáze typu back-end, a optimalizace dotazů, jejichž výsledkem je rychlejší průchodnost(propustnost). (Rabeler et al. 2018)

Tabulární modely jsou vytvářeny v aplikaci SQL Server Data Tools (SSDT) použitím projektové šablony Tabular model. Tato šablona poskytuje podklady pro vytváření objektů sémantických modelů jako jsou tabulky, partitions (oddíly), vztahy, hierarchie, míry a KPI. (Rabeler et al. 2018)

Tabulární modely mohou být nasazeny do služby Azure Analysis Services nebo na instance služby SQL Server Analysis Services (SSAS) nakonfigurované na režim Tabular server. Nasazené tabulární modely je možné spravovat přes SQL Server Management Studio. (Rabeler et al. 2018)

Nástroje tabulárního modelu podporují pouze vazby 1:1 nebo 1:N s jednosměrnou filtrací. Pro vytvoření vazby M:N je nutné použít DAX funkce nebo obousměrné křížové filtry. (Rabeler et al. 2018)

2.5.2 Engine

Tato kapitola pojednává o výpočetních engine, které jsou součástí tabulárního modelu. Dotazy na tabulární databáze jsou v SSAS 2016 zpracovávány ve 2 vrstvách výpočetních engine, formula engine a storage engine. Existují 2 druhy storage engine, VertiPaq a DirectQuery. Tabulární model, který je využit v této práci, používá VertiPaq engine, a z toho důvodu bude v této kapitole tomuto engine věnováno více pozornosti.

Postup zpracování dotazů k datům je následující: Klient pošle MDX či DAX dotaz, SSAS Tabular ho rozebere, transformuje ho do plánu dotazů, který je dál zpracován pomocí Formula engine. (NisalBI nedatováno)

Dle typu dotazu formula engine provádí dotazy na storage engine. Tabulární model obsahuje 2 typy storage engine, a to in-memory analytický engine, známý jako Vertipaq engine, a DirectQuery engine. Typ storage engine se definuje již při vytváření Tabulárního modelu. DirectQuery převádí dotazy na SQL dotazy do zdrojové databáze. Vertipaq engine zpracovává data v lokální RAM paměti díky tomu, že při každém obnovení modelu si vytvoří

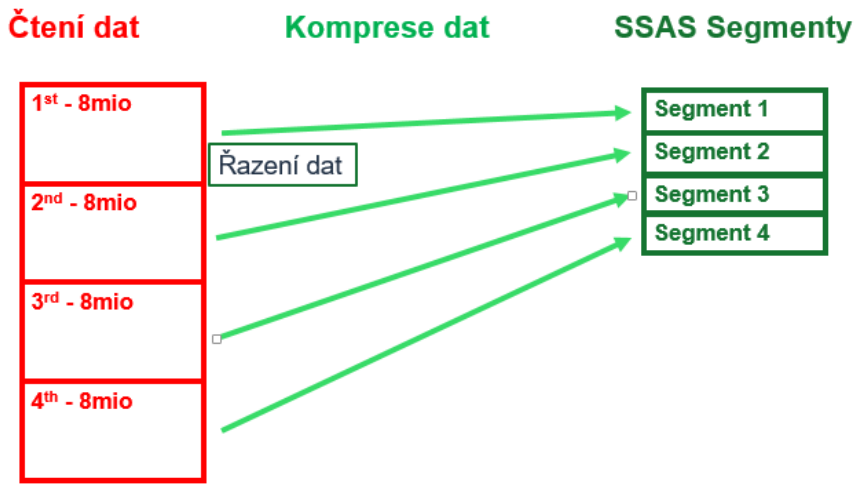
a uloží kopii dat. Zpracované požadavky na data obdrží v interní binární struktuře, externě popsané ve formátu xmSQL. (Mazáčová 2017)

1.5.2.1 Vertipaq engine

Stávající oficiální název Vertipaq engine je xVelocity in-memory engine, avšak z toho důvodu, že tento in-memory engine byl dlouhou dobu pojmenován jako Vertipaq, je známější pod tímto názvem. Navíc interně je stále zobrazován název Vertipaq a dotazovací engine vykonává Vertipaq dotazy. (Haman 2018)

Vertipaq engine při načítání dat nejprve data přečte a následně je transformuje do sloupcové struktury VertiPaq, tedy tabulku si rozdělí na sloupce. Pro jednotlivé sloupce tabulky si vytvoří slovník jedinečných hodnot, které dále nahrazuje indexy. Sloupce tak obsahují pouze číselné datové typy, konkrétně integer. (Haman 2018)

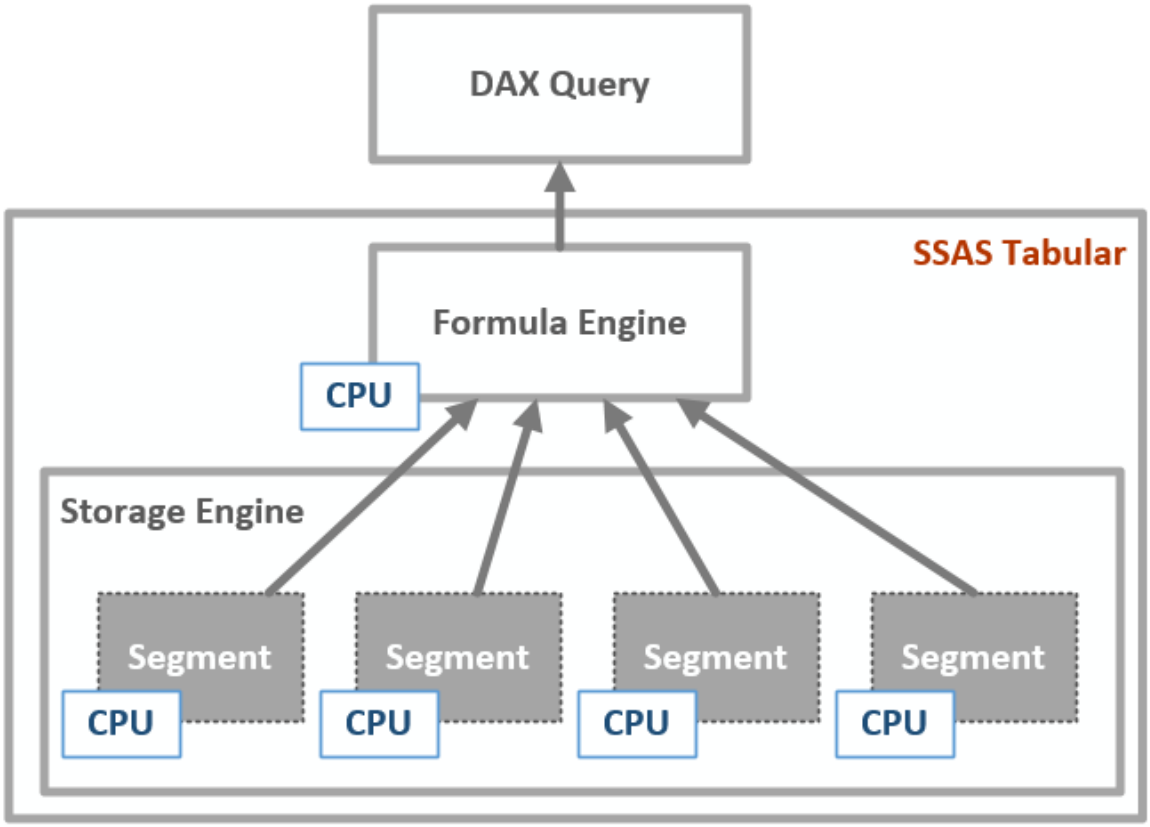
V dalším kroku si VertiPaq engine seřadí duplicitní výskyty indexů a zpracuje je podle algoritmu RLE (= Run Length Encoding). RLE je algoritmus pro komprimaci dat. Aby bylo dosaženo lepšího kompresního poměru, RLE tedy data nejprve seřadí tak, aby řádky následující ve sloupci obsahovaly stejnou hodnotu. Data jsou čtena a zpracovávána po dávkách, v angličtině *batch*. Jeden batch může obsahovat až 8 milionů řádků. Data jsou tedy rozdělena do dávek, zkomprimována a uložena do segmentů, jak je znázorněno níže na obrázku 3. Výsledné řešení je závislé na poměru počtu hodnot a počtu unikátních hodnot, ale i přesto je paměťově úspornější a rychleji prohledatelné, neboť je uložen pouze počet opakování stejných hodnot. (Joyful Craftsmen 2018)



Obrázek 3 Komprese dat ve VertiPaq engine (Joyful Craftsmen 2018)

Komplexnější kalkulace vykonává formula engine po získání výsledků z VertiPaq engine. (NisalBI nedatováno)

Na obrázku 4 je nastíněno schéma zpracování dotazů, kde lze vidět, že Formula engine zpracovává dotazy od jednoho uživatele pouze na jednom vlákně, zatímco VertiPaq engine používá více vláken paralelně.



Obrázek 4 Schéma zpracování dotazů v SSAS Tabular (Joyful Craftsmen 2018)

Pro in-memory engine je nutné, aby RAM na serveru, kde je datový model umístěn, byla dvakrát až třikrát větší než je velikost datové kostky.

1.5.2.2 DirectQuery engine

DirectQuery engine se dotazuje na data přímo na zdroji, tedy posílá SQL dotazy do zdrojové relační databáze, které SQL podporují. Je vhodnější použít DirectQuery pro velké datové objemy jelikož ukládá pouze metadata a navíc poskytuje nejaktuálnější data. (Duncan 2019)

Typ engine je možné změnit i po nasazení datové kostky do produkce.

2.5.3 Shrnutí

V tabulárním modelu datové kostky, nad kterým autorka práce vytváří security model, využívá právě VertiPaq storage engine. Tento typ engine je využíván proto, aby se zabránilo

přetížení výkonu testovacího serveru, k čemuž by mohlo dojít například v situaci, kdy by se současně k datovému modelu připojilo mnoho uživatelů a současně by poslali do databáze náročnější dotazy. In-memory engine uchovává data v lokální paměti, což ve skutečnosti je analytická databáze tabulárního modelu .

2.6 Perspektivy

Kapitola věnující se perspektivám v tabulárním modelu vychází z dokumentace na internetových stránkách společnosti Microsoft (Rabeler et al. 2018).

2.6.1 Vymezení pojmu

Perspektivy v tabulárních modelech umožňují definovat podmnožiny modelu, které budou zobrazené specifikovaným uživatelům či skupinám uživatelů, a zobrazují různé pohledy na konkrétní model tabulární databáze. Databáze modelu by měla být zabezpečena před přístupem k objektům v modelu pomocí perspektiv, jelikož veškeré zabezpečení pro konkrétní perspektivu je zděděno ze zabezpečení modelu. Avšak primární záměr vytváření perspektiv není zabezpečení tabulárního modelu, nýbrž usnadnění přístupu k datům pro uživatele. Není možné poskytnout uživateli pomocí perspektiv přístup k objektům modelu, ke kterým přístup nemá.

Perspektivy je možné vytvářet v návrháři modelu pomocí dialogového okna Perspectives a následně jsou zobrazovány v aplikacích využívající model s vytvořenými perspektivami. Například v excelu si uživatel může vybrat perspektivy, které mu jsou nabízeny, tedy k nimž má přístup.

2.6.2 Shrnutí

Z této kapitoly autorka usoudila, že není vhodné použít perspektivy jako druh zabezpečení tabulární datové kostky, ke které bude přistupovat velké množství uživatelů, jelikož by proces vytváření perspektiv zvlášť pro každého uživatele, případně skupinu uživatelů, byl zdlouhavý.

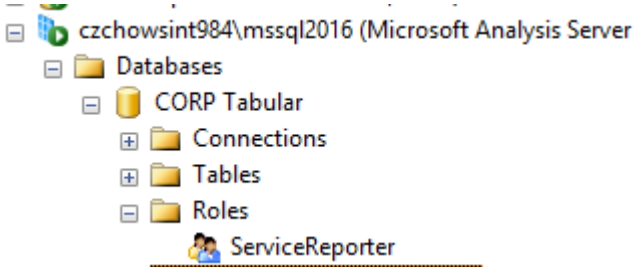
2.7 Security handling v tabulárním modelu

Pro zabezpečení tabulárního modelu se používá definice rolí uživatelů v databázi. Zabezpečení modelu pomocí rolí lze řešit až na úroveň řádků, které již vyžaduje použití jazyku DAX.

2.7.1 Databázové Role

Jak lze vidět na obrázku 5, zabezpečení analytické databáze je obdobné jako zabezpečení relační databáze. V databázi nebo přímo v tabulárním modelu jsou vytvářeny role, pod které

jsou přidáváni uživatelé. Těmto rolím jsou pak udělována práva, a to buď plná kontrola nad databází (Administrator), práva na procesování databáze (Process database), práva pouze na čtení dat z databáze (Read) či práva na čtení a procesování databáze (Read and Process).



Obrázek 5 Pohled na role v databázi (vlastní zpracování)

Vytvořené role lze zobrazovat a vytvářet filtry jak v aplikaci SSDT, tak v aplikaci SSMS. Filtry lze vytvářet pomocí výrazů v jazyku DAX.

K této úrovni zabezpečení lze také využít možnosti proměnných. Lze nastavit podmínky s využitím proměnných, jako je například Username(), která vrací uživatelské jméno přistupujícího uživatele ze služby Active Directory.

2.7.2 Shrnutí

Kapitola slouží k získání informací o základních možnostech zabezpečení tabulárního modelu. V kapitole bylo zmíněno, že tabulární model lze zabezpečit na 2 úrovních, a to na databázové úrovni a na úrovni řádků. V této práci bude využito zabezpečení na úrovni řádků.

2.8 Active Directory

2.8.1 Vymezení pojmu

Definice Active directory podle slovníku Gartner zní:*"Active Directory is the "directory service" portion of the Windows 2000 operating system. Active Directory manages the identities and relationships of the distributed resources that make up a network environment. It stores information about network-based entities (e.g., applications, files, printers and people) and provides a consistent way to name, describe, locate, access, manage and secure information about these resources. It the central authority that*

manages the identities and brokers the relationships between these distributed resources, enabling them to work together.“ (Gartner 2012)

V podstatě Active directory je služba, která sbírá a uchovává informace z aplikací, ke kterým je služba připojena. Tyto informace jsou organizovány do individuálních objektů a každý objekt má set atributů. Je to systém pro použití výhradně v prostředí Microsoft Windows.

2.8.2 Shrnutí

Například ve společnosti DHL služba Active Directory zajišťuje ucelený seznam uživatelů v podniku, který může být připojen k několika aplikacím, aby mohl získávat data ze všech těchto aplikací. Tato služba je následně využívána v rámci zabezpečení, ať už celé počítačové sítě, tak i například přístupu k aplikacím.

2.9 Sharepoint

Sharepoint je aplikace vyvinutá společností Microsoft. Informace uvedené v této kapitole byly převzaté z dokumentace Sharepointu (Microsoft 2019)

2.9.1 Vymezení pojmu

Sharepoint slouží primárně k vytváření webů organizací. Je možné ho používat jako bezpečné místo pro ukládání, uspořádání a sdílení informací a přístup k nim z libovolného zařízení. Tyto weby lze používat nejen interně, ale také ke komunikaci s partnery a zákazníky podniku.

Služba Sharepoint nabízí další služby a technologie. Jednou z nich je například Sharepoint Online, což je cloudová služba pro podniky. Dalšími ze služeb jsou Sharepoint Server, Sharepoint Foundation, Sharepoint Designer 2013 nebo OneDrive for Business sync.

Součástí stránek mohou být vložené aplikace. Seznam dostupných aplikací může být jednak seznam aplikací vytvořených či upravených v rámci organizace, ve které se služba používá, nebo seznam dostupných aplikací, které Sharepoint nabízí. Jednou z takových aplikací je například Custom List, což slouží jako seznam. Sloupce v tomto seznamu je možné dále upravovat dle požadavků.

2.9.2 Shrnutí

Pro vytváření bezpečnostního modelu budou využity právě seznamy na sharepointové stránce vytvořené speciálně pro projekt zabezpečení tabulárního modelu, jelikož se tím tým developerů zbaví zodpovědnosti za udělování přístupu uživatelům k datům obsaženým

v datové kostce. Sharepoint byl vytvořen, aby umožňoval sdílení seznamů například mezi zaměstnanci společnosti, čehož bude využito na zabezpečení datového modelu.

2.10 ACL tabulky

Tato kapitola má za cíl seznámit čtenáře s pojmem Access Control List. Zdrojem pro informace uvedené v této kapitole je dokumentace Microsoft (Kennedy et al. 2018)

2.10.1 Vymezení pojmu

Access Control List (ACL), v překladu seznam řízení přístupu, je seznam obsahující informace o oprávnění a přístupech k objektu, k němuž se váže. Jednotlivé záznamy v ACL seznamu lze označit jako Access Control Entries (ACE), v překladu položky řízení přístupu, kdy jednotlivé položky označují jednotlivé uživatele či skupiny uživatelů a jejich oprávnění, která se mohou týkat nejen přístupu, ale také práce se soubory či daty. Atributy, které se mohou v seznamu ACL objevit, se odvozují od objektu, ke kterému je ACL vázán.

Mohou být 2 typy ACL pro jakýkoliv zabezpečitelný objekt, a to Discretionary Access Control List (DACL) a System Access Control List (SACL).

DACL, v překladu volitelný seznam řízení přístupu, pracuje se zabezpečením tak, že jakmile se proces pokusí o přístup k objektu, systém nejprve zkontroluje, zda tento obejkt má DACL. Pokud objekt DACL nemá, přístup k objektu je tomuto procesu v rámci tohoto typu zabezpečení udělen v plné míře. Pokud však objekt má DACL, systém v sekvenci prochází veškeré záznamy, tedy ACE, v tomto seznamu, dokud nenajde alespoň jeden záznam ACE nebo pokud neprojde celý seznam. Pokud je DACL objektu prázdný, tedy neobsahuje žádné záznamy, systém odepře přístup veškerým procesům. Pokud systém nalezne v seznamu hledané ACE, zachová se následně dle obsahu tohoto záznamu v seznamu.

SACL, v překladu seznam řízení přístupu systému, slouží k zaznamenávání pokusů o přístup k objektu, ke kterému je SACL vázáno. Pokusy o přístup jsou zaznamenávány do Security event logu, tedy do protokolu událostí zabezpečení, kam je možné zaznamenat veškeré přístupy.

2.10.2 Shrnutí

Seznamy ACL lze použít pro zabezpečení několika typů objektů, od souborů po databáze. Tyto seznamy lze navrhnout podle své vlastní potřeby. Není nutné mít na objekt navázány oba typy ACL. Ve vytvářeném bezpečnostním modelu bude této technologie využito.

2.11 Integration services

Tato kapitola pojednává o komponentě Microsoft SQL Server, která slouží k přípravě datových pump. Informace v této kapitole byly převzaty z dokumentace Microsoft (Zhang 2018)

2.11.1 Vymezení pojmu

Integration services je nástroj v rámci databázového balíku MS SQL Server používající se pro získávání a transformaci dat do databáze. Integration services umožňují extrahovat a transformovat data ze širokého spektra zdrojů, jako jsou datové soubory typu excel, csv, xml, flat files, zdroje z odlišných databází či zdroje z odlišných serverů a načíst je do jedné či více cílových databází, souborů či jiných umístění.

Integration services nabízejí rozsáhlou sadu úloh a transformací, grafické nástroje pro vytváření balíčků a databázi katalogu integračních služeb, kde jsou balíčky ukládány, spouštěny a spravovány.

Grafické nástroje integration services umožňují vytvářet řešení bez vytváření kódu a dokonce i vytváření hierarchických balíčků, tedy parent-child balíčků, kdy parent balíček spouští child balíčky.

2.11.2 Shrnutí

Integration services je služba, která je velmi významnou součástí ETL v rámci řešení MS SQL Server. Integration Services zajišťuje sběr dat a jejich následné nahrávání do cílových destinací. Obvykle se Integration Services používá k nahrávání dat do Data Stage.

3 Návrh a implementace security modelu

Vytvářené řešení bezpečnostního modelu je vyvíjeno v rámci projektů v nadnárodní společnosti. Důvodem výběru řešení zabezpečení tabulární kostky je jednak existence podobného řešení pro existující reporty, což umožnilo autorce inspirovat se tímto řešením, analyzovat jej a po potřebné modifikaci jej zavést jako řešení bezpečnosti vytvořené tabulární kostky, ale také automatizace zabezpečování reportů, tedy podpora správy zabezpečení dat ze strany uživatelů. Jelikož se jedná o citlivá data, která jsou ve stávajících reportech na starém serveru zabezpečena pomocí speciálních tabulek, je potřeba tato data zabezpečit také pro nové řešení v podobě datové kostky. Dalším důvodem výběru tohoto typu zabezpečení řešeného přes sharepoint je také ta skutečnost, že tým, který spravuje datovou kostku, se nebude muset zabývat přidělováním či odebíráním práv uživatelům, jelikož to zodpovědní manažeři budou spravovat sami. Autorka tímto řešením tedy podpoří samostatnost uživatelů nejen z pohledu správy práv, ale také z pohledu práce s daty.

Datový model ve společnosti využívá metodiku podle R. Kimballa přesně podle její definice. Kimballova metodika popisuje datová tržiště jako: *„decentralizované datové sklady, které se pak mohou postupně integrovat do celopodnikového řešení.“* Tento přístup se nazývá také jako bottom-up přístup. (Pour et al. 2018)

3.1 Analýza podobného řešení

3.1.1 Popis

Podobné řešení zabezpečení dat, které je cílem této bakalářské práce, je nyní implementováno pro speciálně vytvořené reporty pro jednoho klienta, jímž je jedna nejmenovaná podniková část. Tyto reporty zobrazují informace o vybraných službách využívaných v rámci klientovi podnikové části.

Tato podniková část se dále rozděluje na dílčí části, které budou dále nazývány clustery.

Výše zmíněný klient žádal o vytvoření reportů, které budou sledovat data pouze pro danou podnikovou část a její clustery. Uživatelům spadajícím pod tyto clustery bude umožněn přístup pouze k datům vztahujícím se pro cluster, jehož jsou členy. Manažerům podnikových částí a dalším vybraným pověřeným osobám bude umožněn přístup ke všem datům.

Pro tento projekt byla vytvořena stránka na interním sharepointu, na kterou mají přístup pouze manažeři této podnikové části a tým developerů, kteří tento projekt vyvíjeli a podporují. Součástí této stránky jsou 4 implementované aplikace. V těchto seznamech lze nalézt seznam clusterů, seznam regionů, seznam zemí a souhrnný seznam uživatelů obsahující informace o uživatelském přihlašovacím jménu, regionu a clusteru. Někteří

uživatelé mohou mít přístup k datům z více clusterů, regionů a zemí. Tyto seznamy jsou spravovány skupinou pověřených uživatelů, nikoliv týmem developerů.

Data uložená na sharepointové stránce jsou stahována a ukládána do ACL tabulek v interní databázi pomocí ETL datové pumpy. Tato datová pumpa je vytvořena pomocí technologie SSIS, která získává data uvedená na sharepointu a ukládá je do tabulek v relační databázi.

V databázi bylo vytvořeno databázové schéma z důvodu jednoduchého odlišení projektu. Všechny související tabulky a procedury jsou vytvořeny pod tímto schématem. V databázi byly vytvořeny dimenze regionů, dimenze zemí, dimenze clusterů a souhrnná tabulka se seznamem přidělených práv.

Vytvořené tabulky obsahují stejná data jako filtr v SSRS report manageru při zobrazování a generování reportu. Tudíž uživateli, který přistupuje k reportu, jsou ve filtru nabízeny pouze možnosti, ke kterým má udělený přístup na sharepointové stránce.

Pro tento projekt byly vytvořeny reporty na bázi SSRS, uložené ve formátu rdl. V reportech a subreportech jsou vytvořené speciální datasety, které nabízejí pouze takové parametry filtrování, ke kterým má přistupující uživatel práva. Tyto speciální datasety obsahují dotazy do tabulek obsahující data pro danou podnikovou část s inner joiny na ACL tabulky nesoucí informace o právech konkrétního uživatele, kde je použit parametr uživatelského jména. Uživatelské jméno přistupujícího uživatele je získáváno z Active Directory, které se spolu s ostatními informacemi, jako je například čas spuštění reportu, nachází v tzv. Built-in-fileds, zobrazitelná v okně Report Data v Microsoft Visual Studio.

3.1.2 Shrnutí

Záměrem této kapitoly je přiblížení projektu s podobným řešením zabezpečení, jaké je implementováno do tabulární datové kostky. Podobné řešení zabezpečuje pouze reporty, tedy BI část Reporting.

3.2 Analýza stávajícího řešení

3.2.1 Popis

Ve stávajícím řešení celého projektu jsou manažeři podnikových částí schvalují udělení přístupu žádajícím uživatelům na reportovaná data zobrazená v několika reportech. Toto řešení nadále zůstane pro data reportovaná ve vytvořených SSRS reportech.

Stávající tabulární model obsahuje data o uživatelích a různých placených i neplacených službách, které využívají. Současně jsou taková data manažerům poskytována ve formě manuálně vytvořeného reportu předávaného prostřednictvím emailu a sharepointové stránky v souboru aplikace excel. Tým developerů za účelem automatizace vytvořil tento

tabulární model na účely reportování právě těchto dat v rámci automatizace manuálního vytváření reportů.

Data o uživatelích a jednotlivých službách jsou nahrávána každý den do tabulek v Data Stage pomocí ETL balíčků Integration Services, z toho jedna ze služeb je nahrávána pomocí skriptu vytvořeného v powershell. Z důvodu velkého množství služeb různého typu mají vytvořené SSIS balíčky stromovou hierarchii, kdy jeden hlavní balíček spouští několik dalších rodičovských balíčků jednotlivých služeb, rozdělených podle názvu služeb. Tyto balíčky už spouštějí jednotlivé balíčky, které již nahrávají data do Data Stage. Hlavní rodičovský balíček je spouštěn pomocí vytvořeného tasku v SQL Task scheduler, kde je nastaveno spouštění balíčku každý den. Výsledky běhu tohoto tasku jsou zaznamenávány do tabulky typu view každý den, kam jsou zapisovány výsledky běhu všech jednotlivých spouštěných balíčků v rámci hlavního parent balíčku.

Dalším z důležitých částí celého ETL procesu je task, jehož úkolem je přesun dat z Data Stage do Data Martu pomocí spouštění jednotlivých SQL stored procedures a dalším krokem tohoto tasku je procesování změn do analytické databáze tabulárního datového modelu.

3.2.2 Shrnutí

Přístup k datům v tabulárním datovém modelu má prozatím jen tým developerů spravující tabulární datovou kostku. Model obsahuje data všech podnikových částí. Před tím, než model bude uveden do produkce a bude moci být využíván dalšími uživateli ke sledování různých metrik, je nutné datový model zabezpečit.

3.3 Workflow návrhu

1. Vytvoření Sharepoint subsite
2. Úprava sharepointové subsite
3. Vytvoření schématu databází
4. Vytvoření potřebných tabulek – ACL tabulky
5. Vytvořit SSIS získávající data ze sharepointu
6. Vytvoření stored procedure získávající data z raw tabulky a vkládající je do ACL tabulek
7. Import ACL tabulek do tabulárního modelu
8. Vhodné použití a nastavení ACL tabulek v tabulárním modelu
9. Testování

Prerekvizity

K vytvoření security modelu je využíváno především 2 programů, a to SQL Server Management Studio (SSMS) a Microsoft Visual Studio 2015. SSMS je připojeno k serveru přes databázový engine, na kterém se nachází tabulární model a na kterém bude probíhat vývoj tabulek a procedur. Databáze, ve které je security model vytvářen, se nazývá CORP DM. Pro vývoj ETL je použita aplikace Microsoft Visual Studio 2015 a vývoj probíhá přes připojení na vzdálenou plochu SQL Serveru z toho důvodu, že vytvořený balíček bude uložen a spouštěn pravidelně pomocí SSMS SQL Task Scheduler, který bude tento balíček spouštět z SQL Serveru, nikoliv z lokálního počítače, a také z toho důvodu, že verze této aplikace naisnstalovaná na serveru obsahuje další rozšíření využitá v této práci. K vytvoření balíčku Integration Services je nutnou podmínkou nainstalovaná šablona Integration Services v aplikaci Microsoft Visual Studio.

Jelikož se jedná o projekt v nadnárodní společnosti, veškeré aplikace, programy a názvy objektů jsou uváděny v anglickém jazyce.

3.3.1 Vytvoření sharepoint subsite

Prvním krokem při tvorbě bezpečnostního modelu pro tabulární datový model je vytvoření sharepoint subsite, která bude vytvořena jako subsite týmové sharepoint stránky.

Vytvoření sharepoint subsite je jednoduché. V pravém rohu horní lišty se nachází ikona Nastavení. V Nastavení je možnost vytvoření nové stránky. Po kliknutí na tuto možnost se zobrazí vyskakovací okno, které obsahuje textové pole pro zadání názvu nově vytvářené subsite. Pod tímto textovým polem je zobrazen vygenerovaný odkaz na nově vytvářenou subsite. Po odkliknutí tlačítka vytvořit se následně přesměruje na nově vytvořenou stránku v režimu úprav.

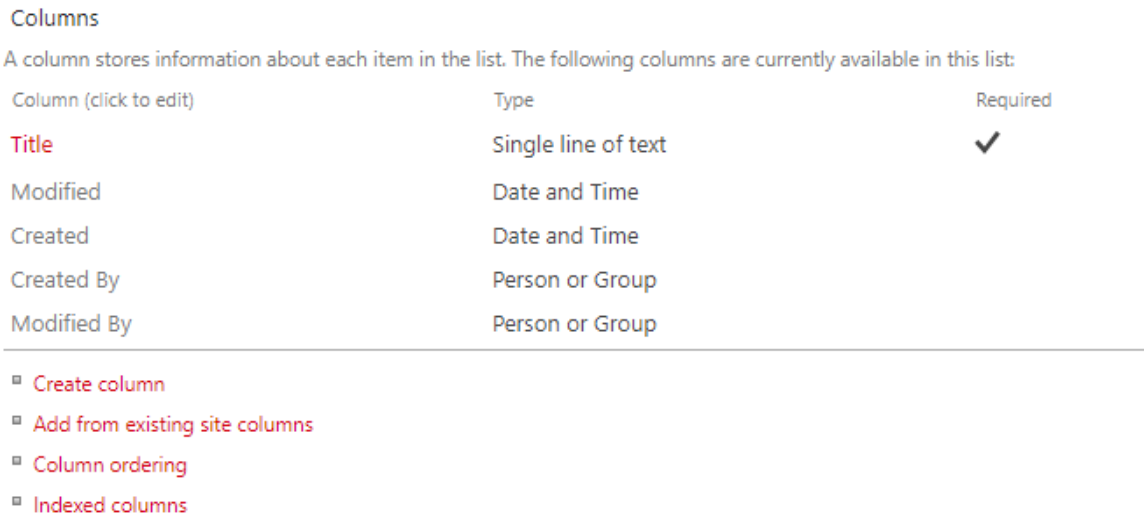
3.3.2 Návrh sharepointové stránky

V této kapitole jsou popsány pouze základní úpravy sharepointové stránky, které jsou nutné pro vytvoření praktické části této práce. Aplikace Sharepoint umožňuje vložit na vytvořenou stránku aplikace, například Custom List, který jsou pro účely bakalářské práce využity. Custom list je jednoduchý seznam, který je ve společnosti hojně využíván a z nabízených aplikací nejlépe vyhovuje požadavkům, a proto byl zvolen jako součást řešení zabezpečení.

Pro dosažení cíle práce je nutné vytvořit 2 seznamy, a to seznam podnikových částí a seznam uživatelů s jim přidělenými právy.

Při rozbalení možností nastavení stránky lze tedy vybrat možnost Site Contents či rovnou možnost Add an app. Z nabízených aplikací je vybrán Custom List. K přidání aplikace je povinné vyplnit název seznamu a při kliknutí na pokročilá nastavení lze přidat i popis aplikace. Ve vytvořeném listu lze přímo přidávat a odebírat sloupce a záznamy, avšak přehlednější správa sloupců seznamu se nabízí v horní záložce List v možnosti List Settings. Jak lze vidět na Obrázku 6, v této záložce je zobrazen seznam sloupců, jejich datový typ a

zda jsou povinné. V možnostech úprav se nachází možnosti tvorby nového sloupce a také možnosti pro jejich řazení.

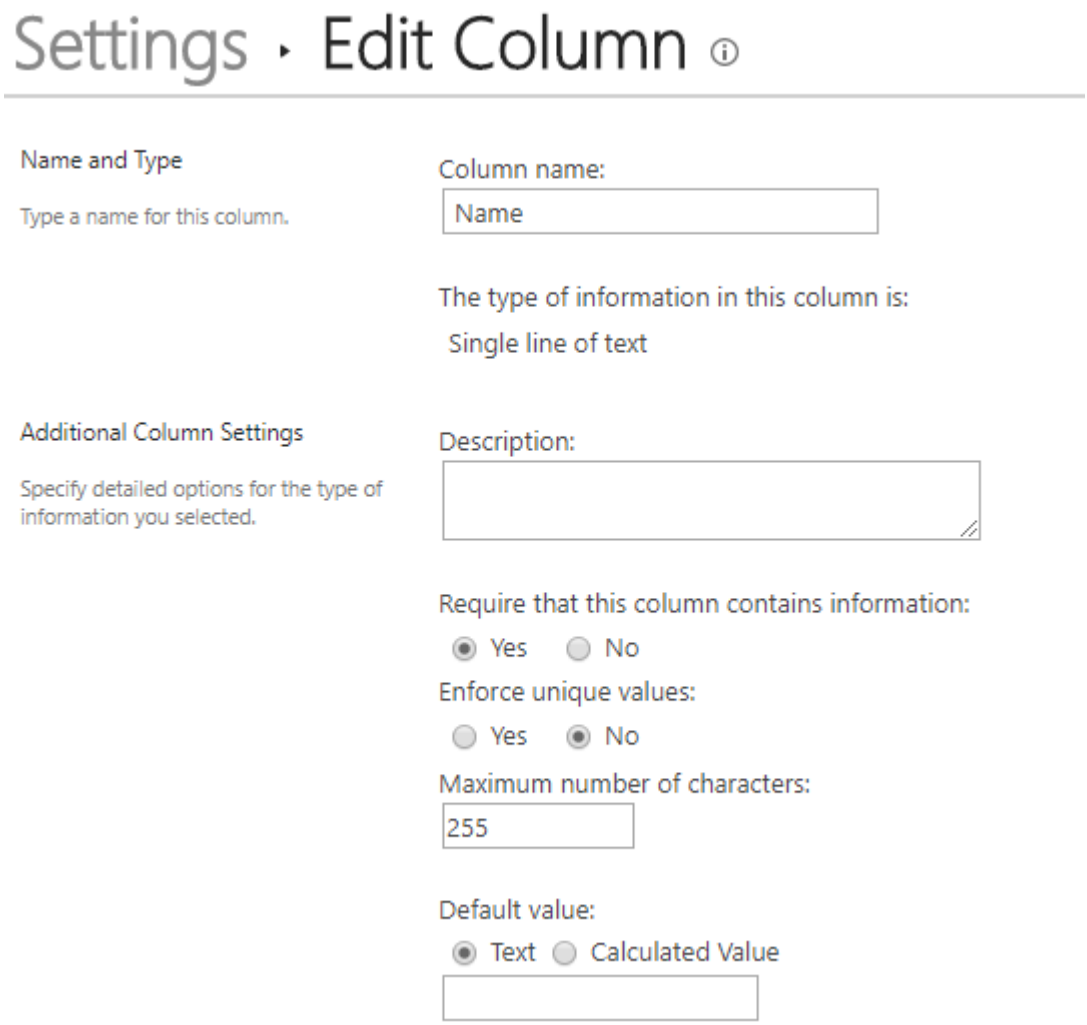


Obrázek 6 Pohled na správu sloupců v aplikaci Sharepoint (vlastní zpracování)

Úpravy již stávajících sloupců jsou prováděny po kliknutí na název sloupce ve výčtu sloupců, stejně tak, jako jejich odstraňování. V této části lze také vidět sloupce, které nejsou zobrazeny ve výchozím vytvořeném view listu. Tyto sloupce obsahují metadata, jako například datum vytvoření, datum úpravy, uživatelské jméno tvůrce záznamu, a podobné.

Seznam podnikových částí je nutné vytvořit z důvodu snadnějšího vytváření zabezpečení v rámci tabulární datové kostky. Seznam je pojmenován jako Tab_Business_Units. V tomto seznamu je vytvořen pouze sloupec s názvem podnikové části. Jelikož ve vytvořeném listu je předvytvořený sloupec s názvem „Title“ a v tomto seznamu bude použit pouze jeden sloupec, nabízí se možnost tento sloupec pouze upravit. Sloupec bude vhodnější přejmenovat na Name z důvodu lepšího odlišení v následné práci s tímto sloupcem, a ostatní nastavené hodnoty lze zanechat, jelikož splňují požadavky na požadovaný sloupec,

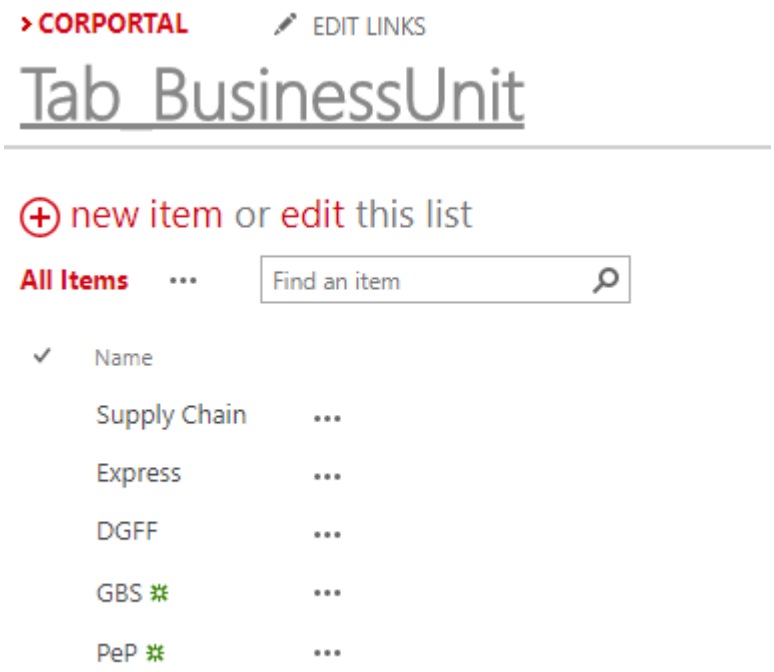
tedy je povinný a je datového typu single line of text. Nastavení tohoto sloupce lze vidět na Obrázku 7.



Obrázek 7 Úprava předvytvořeného sloupce v aplikaci Sharepoint (vlastní zpracování)

Jako další je vytvořen seznam uživatelů a jejich práv s názvem Tab_AccessControlList. V tomto seznamu je nutné vytvořit sloupec UserName typu single line of text obsahující uživatelské jméno, DisplayName typu Person or Group, kde je následně umožněn výběr ze seznamu aktivních uživatelů, který je napojen na Active Directory, a Business Unit typu Lookup. Sloupec UserName lze vytvořit z předem vytvořeného sloupce Title, kde je nutné změnit pouze název sloupce, jelikož ostatní kritéria, která pro požadovaný sloupec jsou, splňuje. Jelikož je třeba zajistit, aby se do cílové relační databáze dostaly hodnoty, kde vztah uživatele a podnikové části je 1:N, musí sloupec Business Unit být typu Lookup. Aby se však mohl použít typ tohoto sloupce, je nutné vytvořit další subsite, kde bude výčet hodnot, které budou následně použity pro výběr hodnot ve sloupci Business Unit. Tedy stačí vytvořit subsite pro Business Units, kde je povinný pouze 1 sloupec, jelikož více informací v tomto seznamu není třeba jak bude dále definováno, lze ponechat tedy předvytvořený sloupec

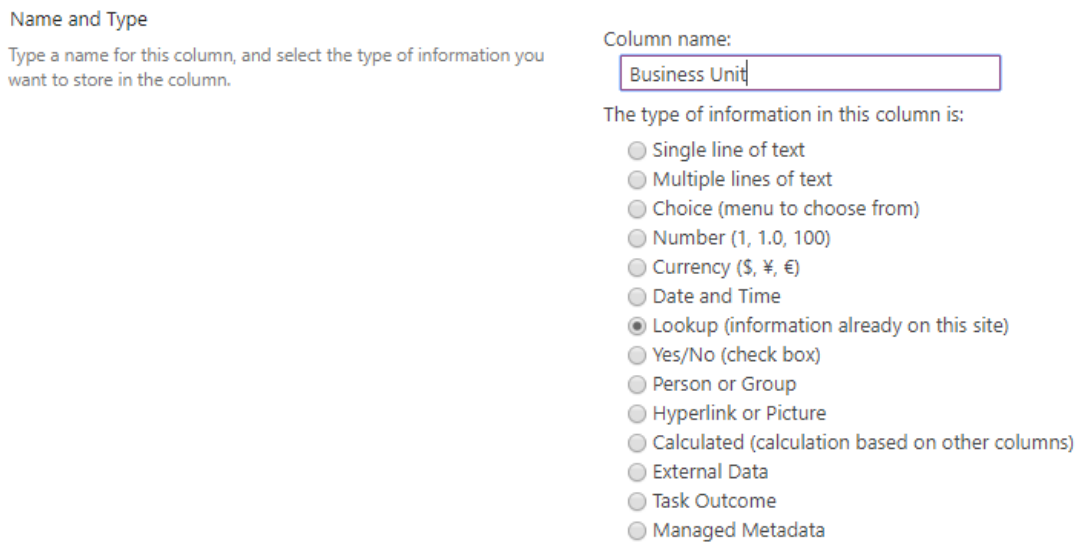
Title, typu single text, a vložit seznam podnikových částí jako jednotlivé záznamy, jak lze vidět na obrázku 8.



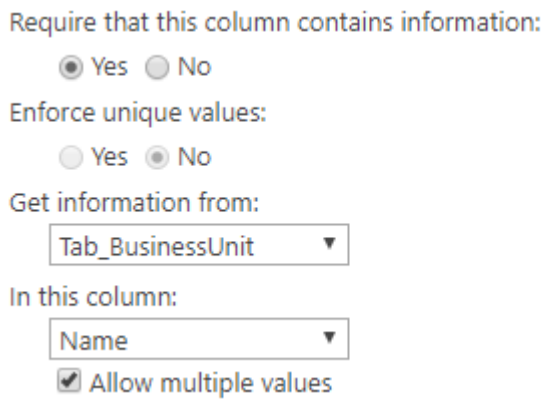
Obrázek 8 Seznam podnikových částí vytvořený v aplikaci Sharepoint (vlastní zpracování)

Nyní je možné vytvořit sloupec Business Unit, kde nejdůležitějšími parametry sloupce jsou typ sloupce, jímž je Lookup, povinnost sloupce, která je vyžadována, zdroj hodnot tohoto

sloupce, což je seznam Tab_BusinessUnit, název sloupce v tomto seznamu, kde se hodnoty nacházejí, a povolit mnohonásobné hodnoty, viz obrázek 9 a obrázek 10.



Obrázek 9 Nastavení datového typu při tvorbě sloupce Business Unit v aplikaci Sharepoint (vlastní zpracování)



Obrázek 10 Nastavení potřebných atributů při tvorbě sloupce Business Unit (vlastní zpracování)

Stránku je nutné zabezpečit, aby přístup měli pouze manažeři podnikových částí. Pro účely této práce prozatím přístup na sharepoint stránku dostanou členové týmu developerů, aby mohlo být zabezpečení následně testováno.

Práva na zabezpečení listů budou přidělena skupinám uživatelů. V první řadě bude vytvořena sharepoint skupina uživatelů, kterou budou tvořit manažeři podnikových částí. Skupiny se vytvářejí v nastavení stránky v možnosti People and Groups ve skupině možností Users and Permissions. Při vytváření nové skupiny lze nastavit vlastníka skupiny, práva na

zobrazení a práva na úpravu členství ve skupině, automatické přijímání členství ve skupině, práva na samotnou sharepoint stránku a například také emailovou schránku této skupiny.

Je nutné tedy vytvořit skupinu vlastníků projektu, kteří budou mít právo na zobrazení stránky a veškerých jejích komponent a také na správu listu s právy uživatelů. V rámci vytváření a testování bezpečnostního modelu pouze v týmu vývojářů, nebude prozatím skupina vlastníků projektu vytvářena. Práva na vytváření záznamů v seznamu ACL_Users budou udělena pouze týmu vývojářů k testování datového toku.

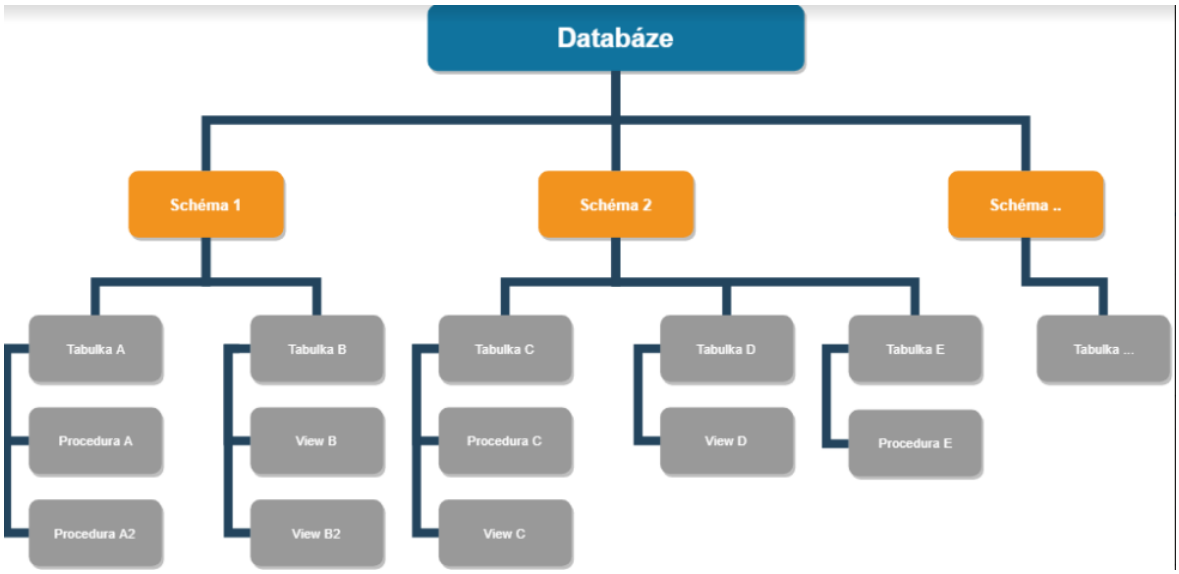
Další skupinou bude skupina správců stránky, v tomto případě tým developerů podílející se na vývoji tabulární datové kostky, a tedy i tohoto security modelu. Tato skupina bude mít plnou kontrolu nad sharepointovou stránkou a všemi jejími komponenty.

3.3.3 Vytvoření schématu databází

Schéma v MS SQL, které je v práci využito, se může chovat jako uživatel, schéma může mít omezená práva stejně jako uživatelé.

Formální definice databázového schématu je soubor vzorců nazývaných jako omezení integrity uložené v databázi.

Zařazení schématu lze vidět na Obrázku 11.



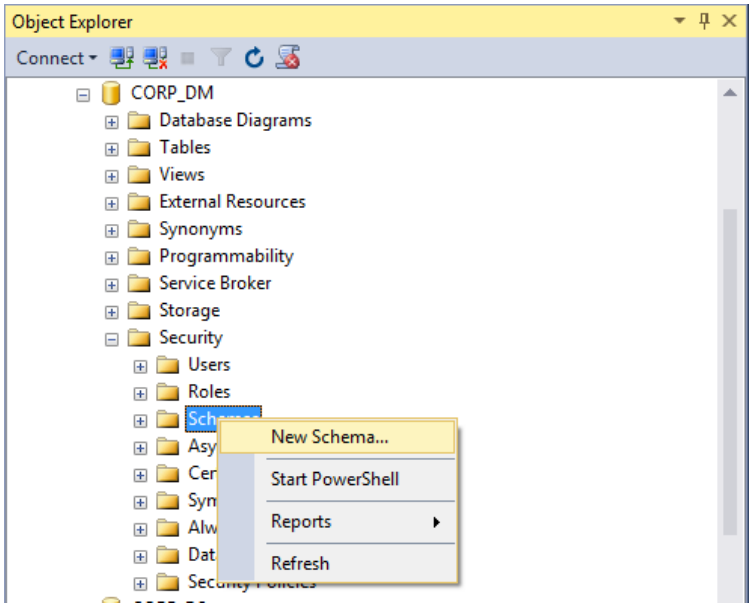
Obrázek 11 Zařazení schéma v databázovém stromu (Vlastní zpracování)

Schéma databází není nutnou podmínkou při tvorbě tabulek, avšak jelikož ve stávající databázi jsou tabulky rozlišovány schématem podle různých kritérií, bude vytvořeno nové schéma i pro tabulky související se zabezpečením datové kostky pro odlišení jeho použití.

Schéma je možné vytvořit 2 způsoby.

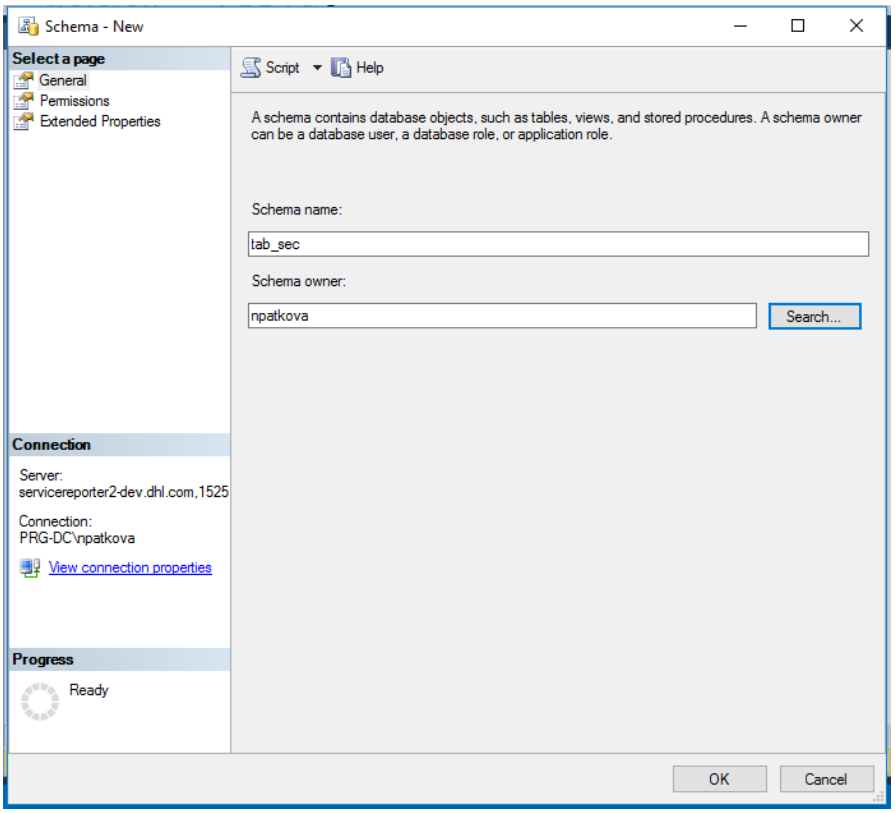
Jedním ze způsobů je vytvoření schématu manuálně v databázi přes dialogová okna.

V záložce Object Explorer je po rozbalení databáze seznam objektů nacházející se v této databázi. Pod záložkou Security se nachází záložka Schemas. Pravým kliknutím na tuto záložku se objeví možnosti vyobrazené na Obrázku 12.



Obrázek 12 Vytváření nového schéma (vlastní zpracování)

Po kliknutí na možnost Nové schéma se objeví okno vyobrazené na Obrázku 13. V tomto okně je nutné vyplnit název vytvářeného schématu a jeho vlastníka.



Obrázek 13 Dialogové okno tvorby schématu (vlastní zpracování)

Jak lze vidět na Obrázku 13, SSMS nabízí možnost vygenerovat script pro tvorbu předpřipraveného nového schématu.

Druhou možností tvorby schématu je pomocí příkazu. Tento příkaz lze vygenerovat i v rámci vytváření schématu prvním způsobem, viz Obrázek 13, tlačítko Script.

Příkaz obsahuje jméno databáze, pro kterou schéma bude vytvořeno, jméno vytvářeného schéma a vlastníka schématu.

Příkaz je následující:

```
USE [CORP_DM]
GO
CREATE SCHEMA [tab_sec] AUTHORIZATION [npatkova]
GO
```

Stejně tak je nutné schéma vytvořit v Data Stage, což je možné vytvářet stejnými způsoby.

Pro účely této bakalářské práce není nutné specifikovat další vlastnosti vytvořeného schématu.

3.3.4 Vytvoření potřebných tabulek

Pro zabezpečení tabulárního modelu je potřeba vytvořit ACL tabulky, které budou následně v modelu využity.

V první řadě bude vytvořena jedna souhrnná tabulka obsahující veškerá data ze souhrnného seznamu, který bude následně vytvořen na sharepointové stránce, odkud budou data čerpána. Každý uživatel bude mít v tabulce 1 záznam i přesto že může být součástí několika podnikových částí. Tabulka bude vytvořena v dočasném úložišti dat, tedy Data Stage, aby bylo dodrženo celkové schéma datového skladu.

Souhrnná tabulka bude pojmenována ACL_Ishare_Raw a vytvořena pomocí následujícího příkazu:

```
CREATE TABLE [tab_sec].[ACL_iShare_raw](
       [ID] [int] NOT NULL,
       [User] [nvarchar](4000) NOT NULL,
       [Business Unit Lookup Values] [nvarchar](4000) NOT NULL
) ON [PRIMARY]
```

Následující tabulky budou vytvořeny v rámci datového tržiště, tedy Data Martu.

Další potřebnou tabulkou obsahující data ze sharepointu je tabulka Business Unit, která bude následně využívána především ve stored procedure, která bude popsána a vysvětlena v další kapitole. Tato tabulka je pojmenována BU a obsahuje pouze 2 sloupce, a to OID typu integer, které bude uměle vytvářeno příkazem identity, a BU typu nvarchar, kde jsou hodnoty ze sharepointu. Tabulka je vytvořena následujícím příkazem:

```
CREATE TABLE [tab_sec].[BU](
       [OID] [int] IDENTITY(1,1) NOT NULL,
       [BU] [nvarchar](255) NOT NULL
) ON [PRIMARY]
```

Dále budou vytvořeny ACL tabulky, které již budou obsahovat data získaná ze sharepointu rozdělena na seznam uživatelů a seznam podnikových částí. Tyto tabulky budou pro účely bakalářské práce dostačující. Tabulka s raw daty musí být rozdělena, jelikož vztah uživatele

k podnikové části může být 1:N, což by výrazně ztěžovalo další použití těchto dat a především z toho důvodu, aby byla dodržena 3. normální forma.

Tabulka uživatelů obsahuje uživatelské jméno a ID jako primární klíč. Uživatelské jméno je typu nvarchar a ID jako integer převzatý ze sharepointu. Všechny sloupce jsou nenulové. Příkaz pro vytvoření této tabulky je následující:

```
CREATE TABLE [tab_sec].[ACL_User](
       [ID] [int] NOT NULL,
       [User] [nvarchar](50) NOT NULL
) ON [PRIMARY]
```

Tabulka podnikových částí obsahuje ID uživatele a k němu přiřazenou podnikovou část, kdy vztah uživatele a podnikové části může být 1:N, čili 1 ID uživatele může být v tabulce obsaženo vícekrát pro různé podnikové části. ID uživatele bude stejného typu jako v tabulce

uživatelů z důvodu následného propojování tabulek, tudíž integer, a podniková část je typu nvarchar. Příkaz pro vytvoření této tabulky je tedy následující:

```
CREATE TABLE [tab_sec].[ACL_BU](
       [UID] int NOT NULL,
       [BU] [nvarchar](255) NOT NULL
) ON [PRIMARY]
```

V rámci této kapitoly zabývající se tvorbou tabulek lze zahrnout také tvorbu view, které bude využíváno v tabulárním modelu. View v databázi je dynamický Select. Slouží jako jakýsi pohled na data, který může obsahovat data spojená z několika tabulek.

Vytvořené View spojuje tabulky ACL_BU a ACL_User, kdy vztah uživatele a podnikové části bude ponechán 1:N, kdy tedy 1 uživatel může být na několika řádcích, závisle na počtu podnikových částí, k nimž má právo přístupu.

Příkaz pro vytvoření View je následující:

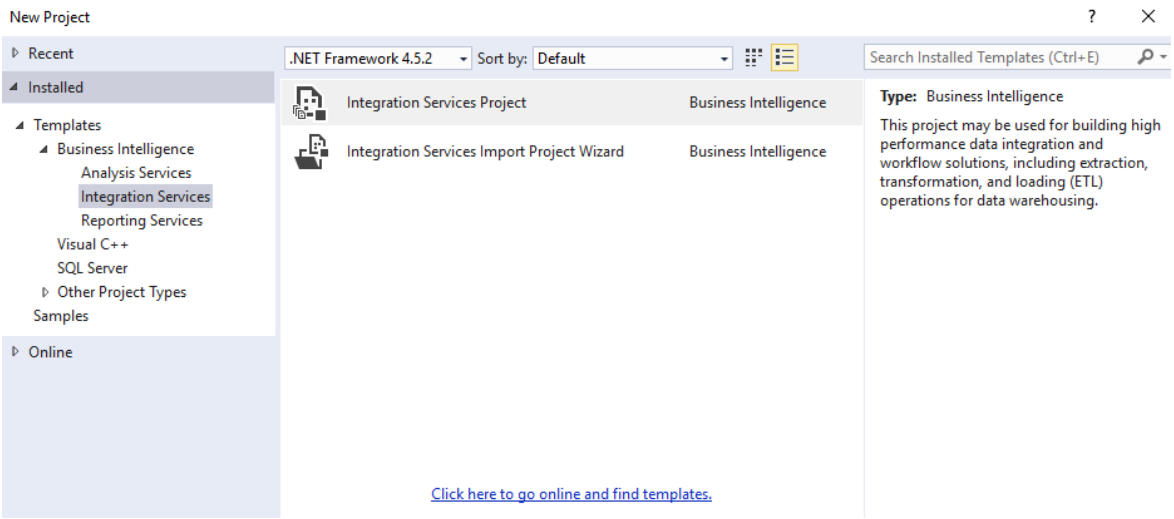
```
CREATE VIEW [tab_sec].[V_ACL_User_BU]
AS (

SELECT u.[ID], u.[USER] AS UserName, b.[BU]
FROM tab_sec.ACL_User u
LEFT JOIN tab_sec.ACL_BU b
ON  u.[ID]=b.[UID]

)
```

3.3.5 ACL SSIS

SSIS balíček je forma ETL, která tedy umožňuje získávání, transformaci a nahrávání dat do cílových destinací. V tomto případě je nutné vytvořit SSIS balíček, který bude získávat data ze sharepointu, nahrávat je do příslušných tabulek v relační databázi a následně spouštět vytvořené stored procedures. SSIS balíček bude nejprve získávat data týkající se

podnikových částí, jenž je krátký seznam existujících podnikových částí, a následně budou získávána data týkající se oprávnění přidělených specifickým uživatelům.

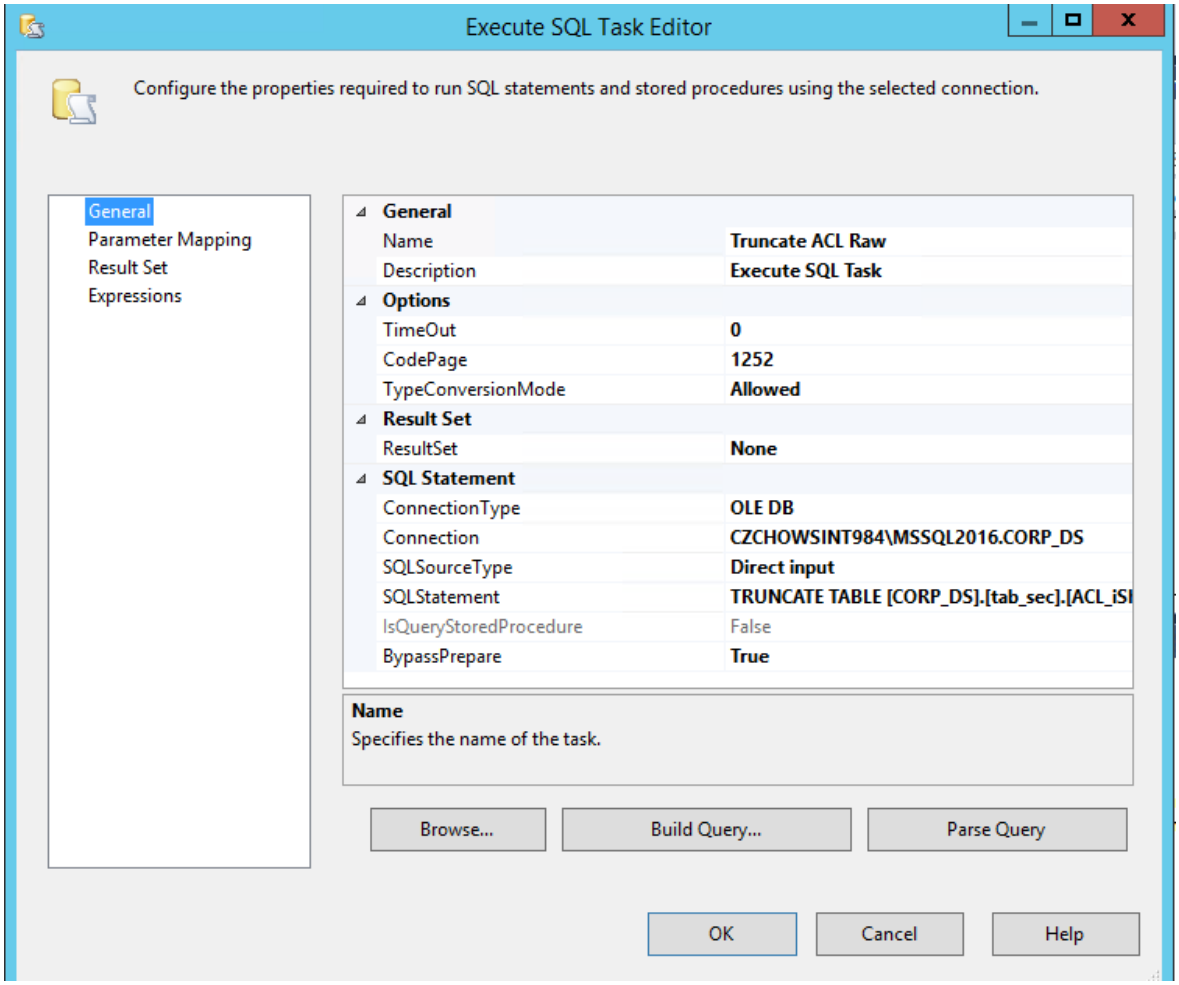
SSIS balíček je vytvářen v aplikaci Microsoft Visual studio. Při vytváření nového projektu v aplikaci Visual Studio je vybrána šablona Integration Services Project, jak je vyobrazeno na obrázku 14.



Obrázek 14 Tvorba Projektu SSIS balíčku (vlastní zpracování)

Prvním objektem ve vytvářeném balíčku, tedy i prvním krokem, bude Execute SQL Task, který bude mazat obsah tabulky obsahující souhrnná data ze sharepointu. V návrháři tohoto objektu bude upraven název objektu, nastaveno připojení a do SQLStatement bude vložen

SQL příkaz TRUNCATE TABLE s úplným názvem tabulky, stejně jako je tomu na obrázku 15. Důvod použití tohoto příkazu bude definován později.

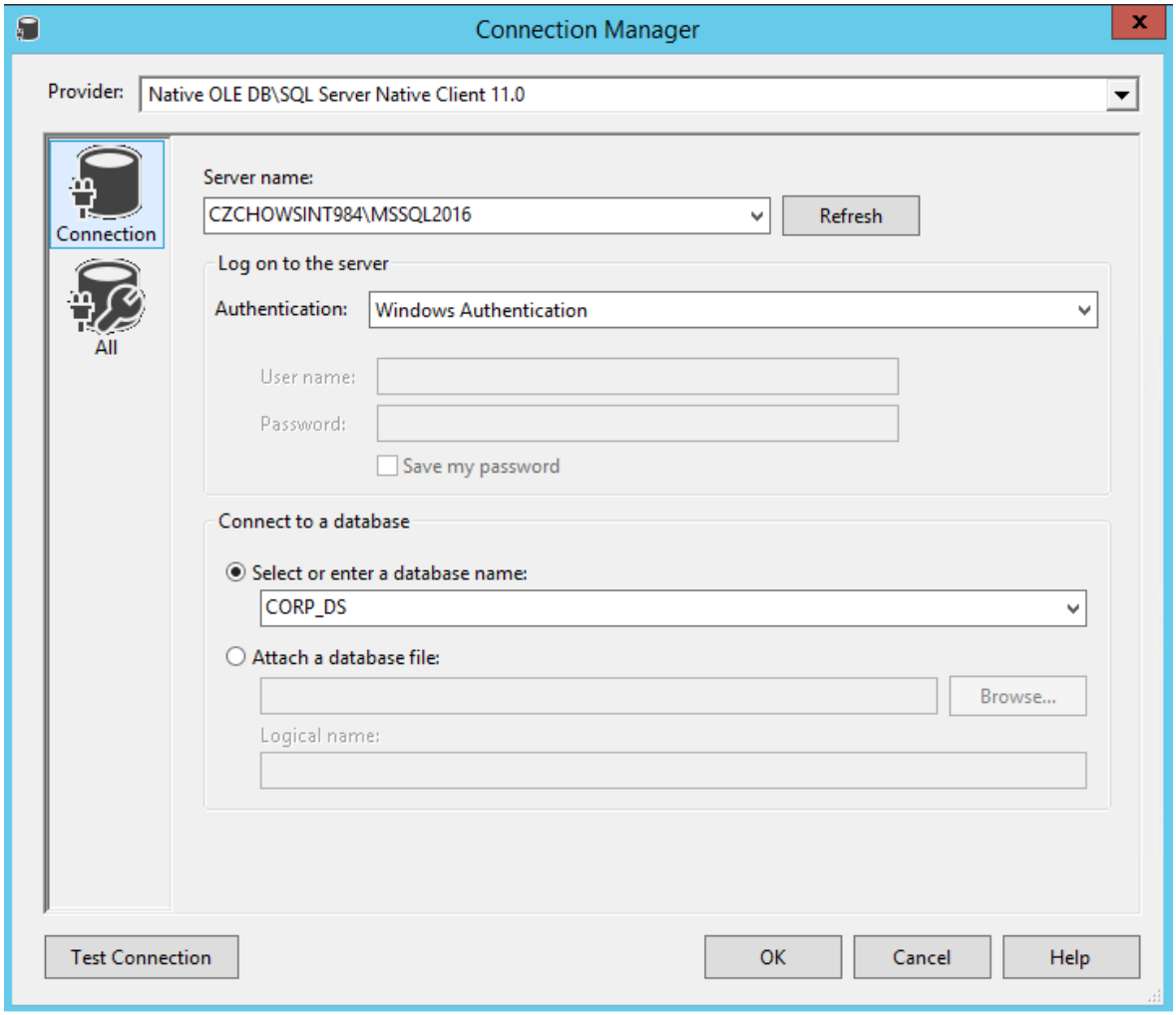


Obrázek 15 Execute SQL Task návrhář (Vlastní zpracování)

Následujícím vloženým objektem bude Data Flow Task. V tomto objektu budou vnořené 2 další. Prvním z nich bude Sharepoint List Source a druhým OLE DB Destination.

Pro správné nastavení této komponenty je nutné vytvořit připojení na cílovou databázi. Toto připojení bude vytvořeno jako OLE DB připojení. Nastavení OLE DB konektoru lze

vidět na obrázku 16, kde je nutné vyplnit název serveru, na kterém se databáze nachází, zvolit typ autorizace a zvolit název cílové databáze, ve které se cílová tabulka nachází.



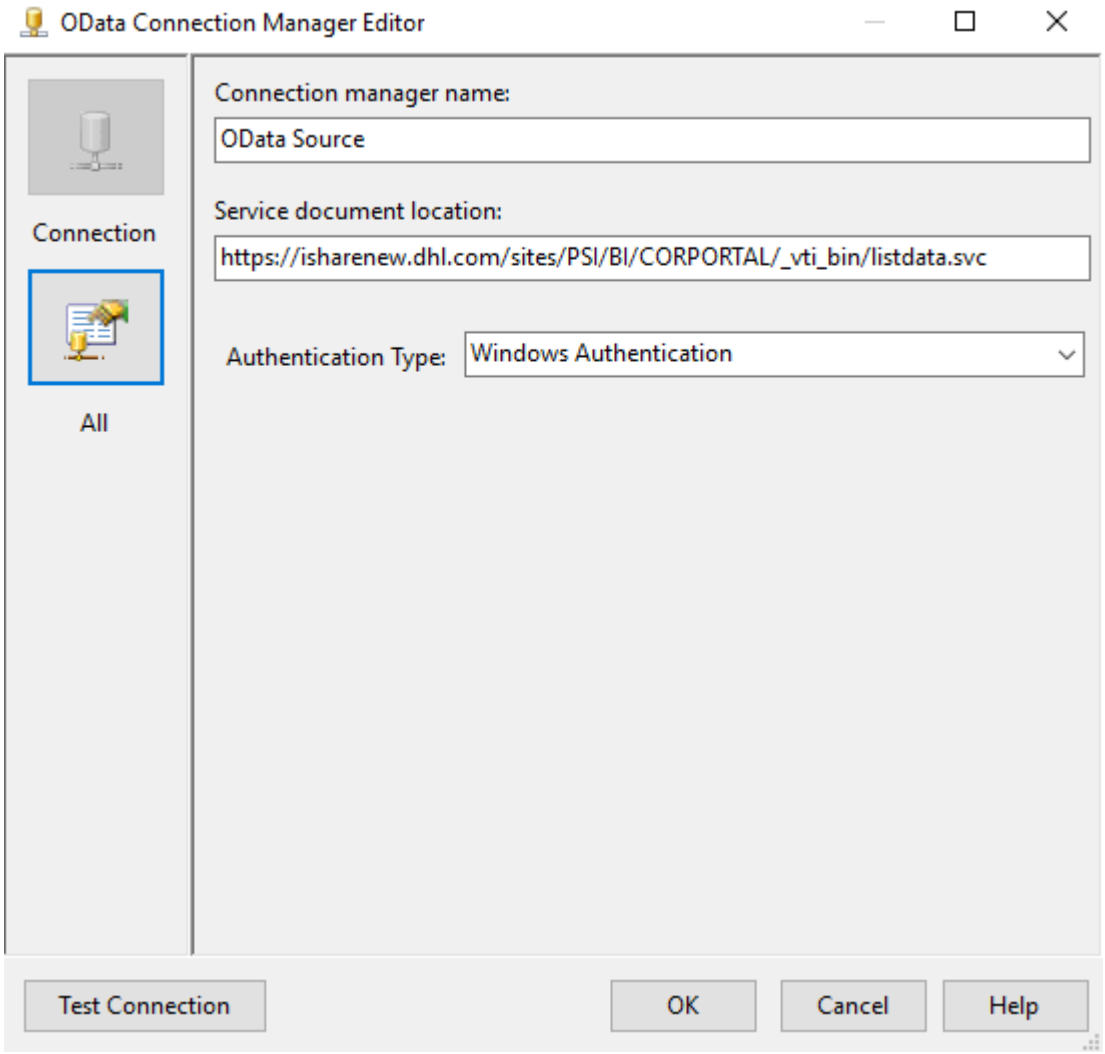
Obrázek 16 Nastavení OLE DB konektoru (Vlastní zpracování)

Před dokončením tvorby připojení se doporučuje toto připojení otestovat.

Pro vytváření následující části ETL je podmínkou mít nainstalovány v aplikaci Visual Studio adaptéry pro integraci se sharepointem, se kterým je pracováno. Tyto adaptéry jsou již nainstalovány v prostředí, ve kterém je balíček vyvíjen, a proto se touto instalací nebude tato práce zabývat.

V první řadě je nutné nastavit konektor na sharepoint. Typem konektoru bude OData Source. Při vytváření konektoru je nutné vyplnit Connection Manager Name a Service Document Location. Při odkazování se na Service Document Location je nutné dodržet formu odkazu na danou site. Odkaz je v následující formě: „https://**mycompany.sharepoint**.com/sites/**mysite**/_vti_bin/listdata.svc“, kde zvýrazněné části je nutné nahradit vlastní částí URL na sharepoint. V tomto případě bude odkaz vypadat následovně: „https://**isharenew.dhl**.com/sites/**PSI/BI/CORPORTAL**/_vti_bin/listdata.svc“.

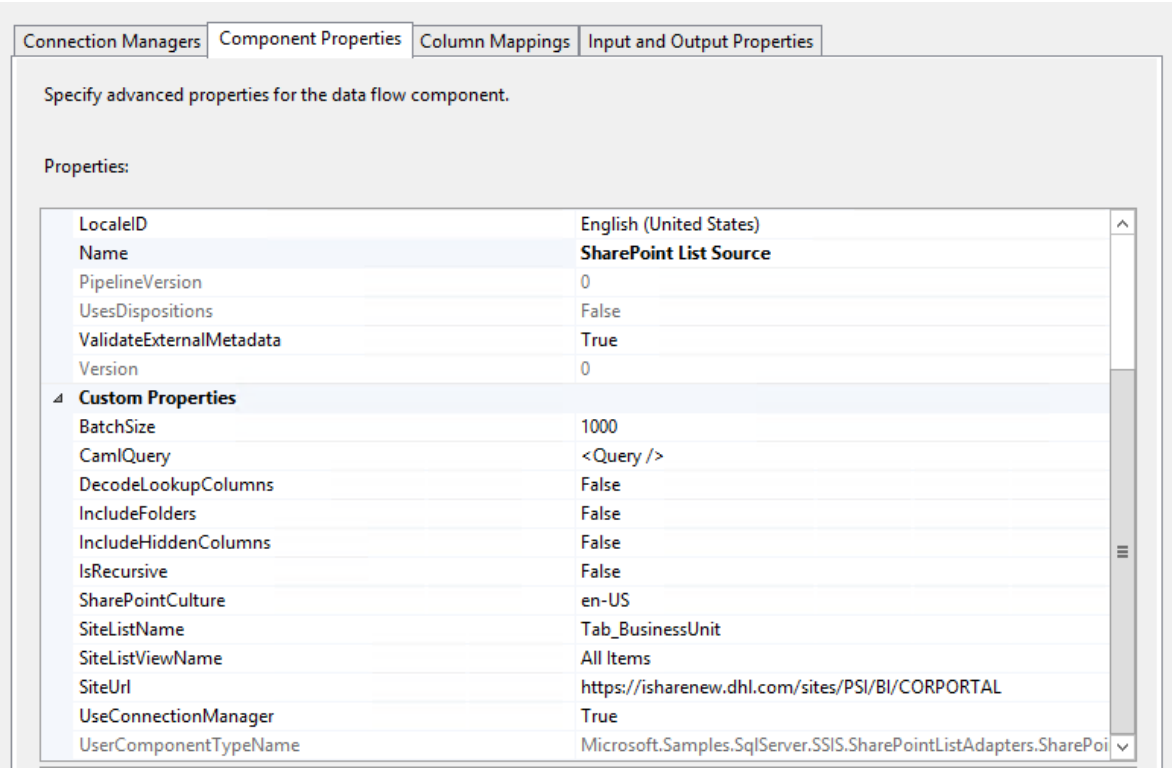
Vložená URL odkazuje na XML formát dané stránky. Při vložení URL odkazující na hlavní stránku dané stránky, test připojení proběhne úspěšně, avšak nebude možné nastavit funkční připojení na sharepoint v OData objektu. Authentication Type lze ponechat výchozí hodnotu, tedy Windows authentication, jako je vyobrazeno na obrázku 17, jelikož bude využíváno ověření uživatele ve windows, který balíček spouští.



Obrázek 17 Nastavení OData konektoru (vlastní zpracování)

Po vytvoření konektoru lze nastavit objekt Sharepoint List Source, jehož záměrem bude získávat data o podnikových částí z vytvořeného listu Tab_BusinessUnit. V první záložce je nutné vybrat konektor na sharepoint, kterým je předchozí vytvořený OData Source.
V následující záložce s názvem Component Properties je nutné nastavit jméno listu, view

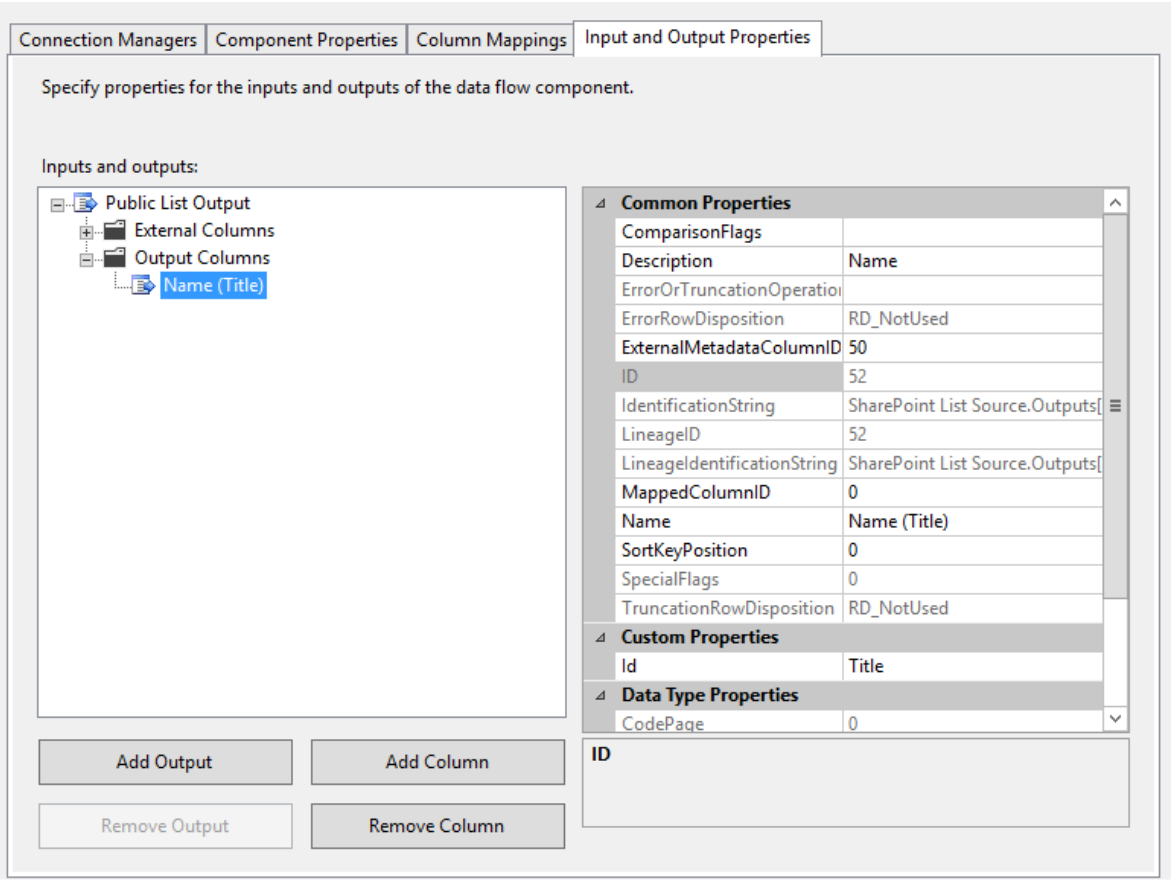
v tomto listu a URL stránky, ze které jsou data čerpána. Nastavení těchto hodnot lze vidět na obrázku 18.



Obrázek 18 Nastavení komponentu Sharepoint List Source (vlastní zpracování)

V záložce Input and Output Properties je možné pro snadnější mapování sloupců a pro přenos menšího obnosu dat lze vybrat sloupce, které budou následně nahrávány do

databáze. Ve složce Output columns jsou tedy nepotřebné sloupce, což je vyobrazeno na obrázku 19.

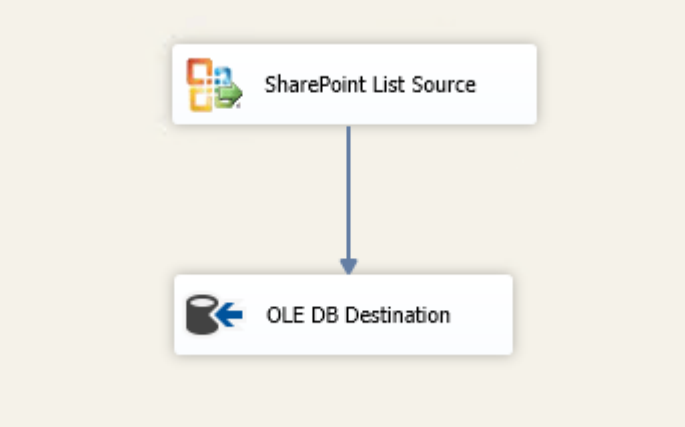


Obrázek 19 Nastavení sloupců komponentu Sharepoint List Source pro data podnikových částí (vlastní zpracování)

Jak lze vidět na obrázku 19, v záložce Input and Output Properties je možné upravovat datové typy přenášených sloupců. V této komponentě je přenášen jediný sloupec, a to sloupec Name (Title), kde (Title) je uvedeno z toho důvodu, že tento sloupec byl původně vytvořen jako Title v kapitole 3.3.2 Návrh sharepointové stránky.

Dalším objektem ve zpracovávané části Data Flow Task, je objekt OLE DB Destination, jehož cílem je nahrát data do cílové tabulky v databázi.

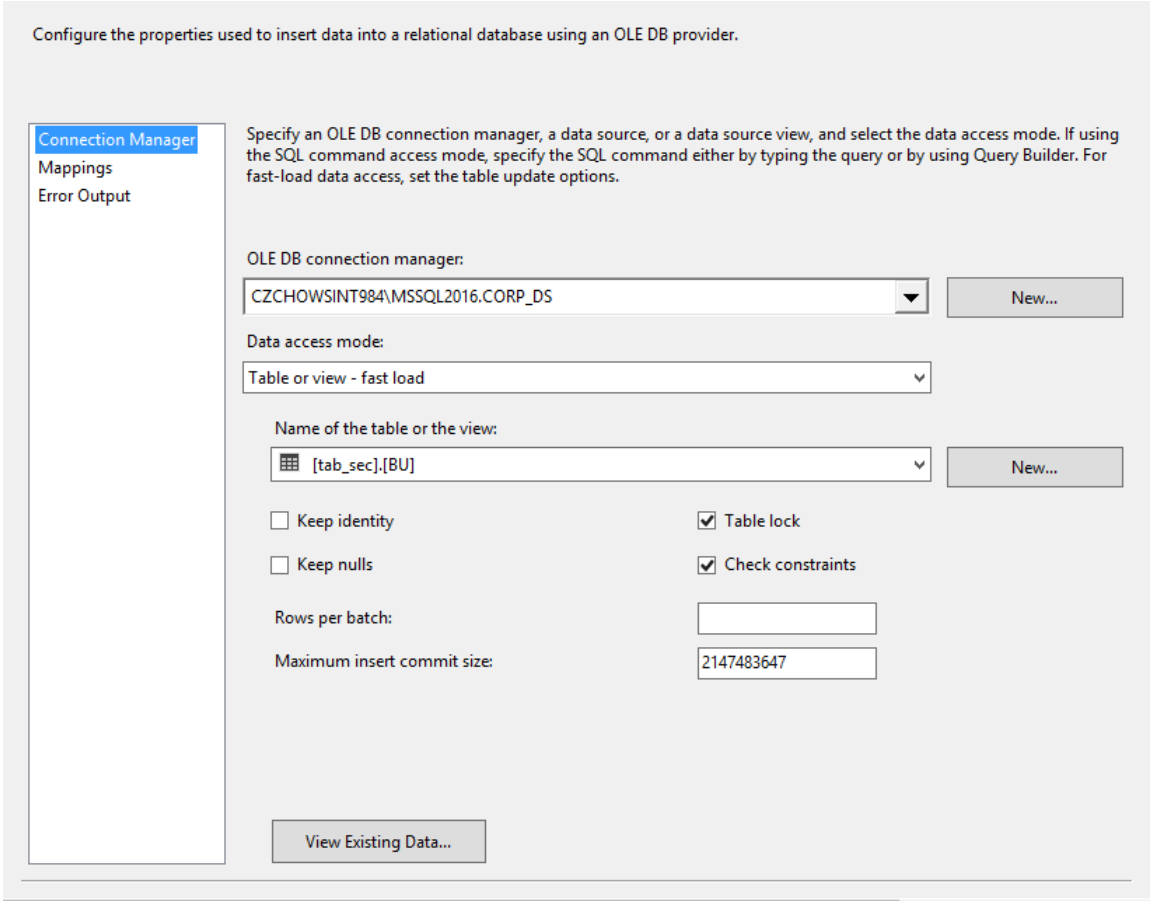
Po výběru konektoru je vhodné vložit objekt OLE DB Destination a napojit ho na předchozí objekt, kterým je Sharepoint List Source, aby bylo umožněno nastavení mapování sloupců. Toto napojení objektu je zachyceno na obrázku 20.



Obrázek 20 Pohled na Data Flow Task s nastavenými komponenty (vlastní zpracování)

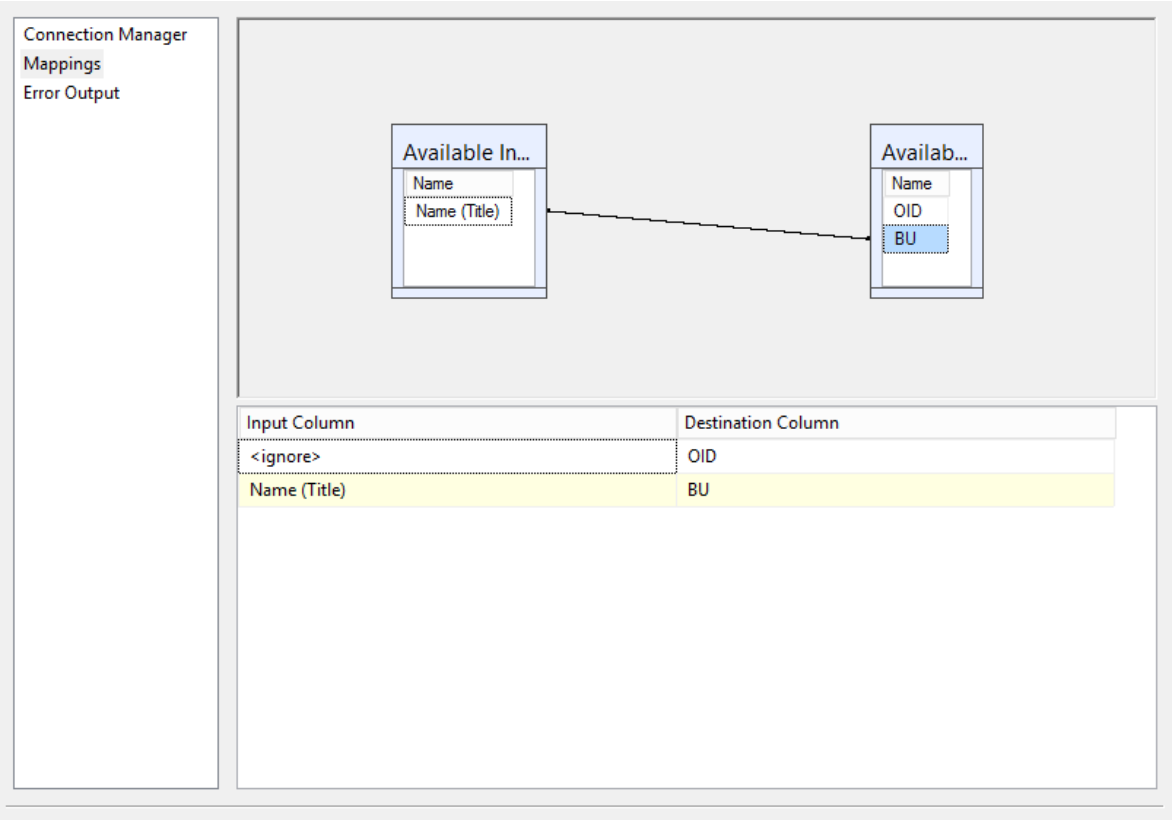
Napojení objektu na jiný je znázorněno pomocí Data Flow Path objektu ve tvaru šipky, který určuje směr toku dat a pořadí spouštění komponent balíčku.

V tomto vytvořeném objektu, v záložce Connection Manager, je vybrán vytvořený OLE DB konektor, v poli Data Access Mode je vybráno Table or view – fast load a v dalším poli bude specifikována tabulka, do které jsou data nahrávána, jako je na obrázku 21.



Obrázek 21 Nastavení připojení komponentu OLEDB Destination (vlastní zpracování)

V záložce Mappings bude nastaveno mapování sloupců vycházejících ze zdroje do tohoto objektu se sloupci nacházejícími se v cílové tabulce, viz obrázek 22.



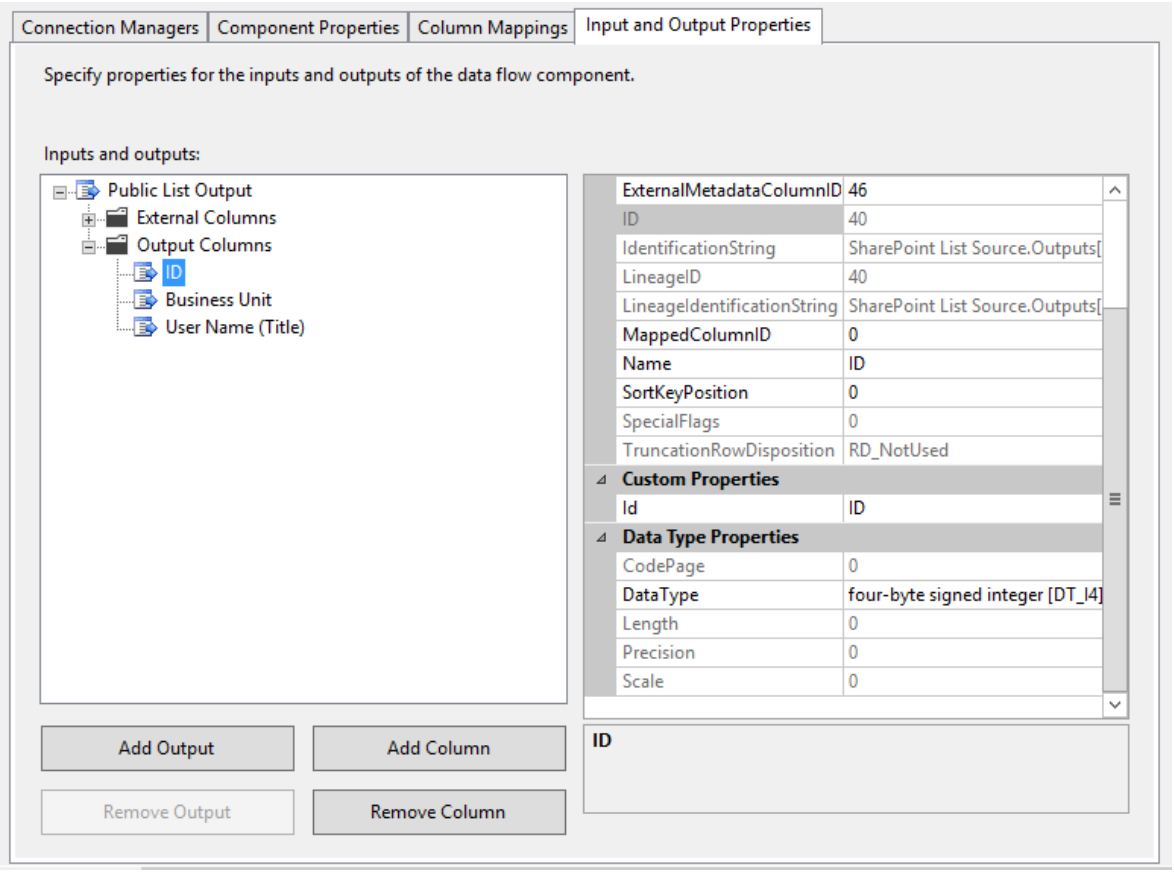
Obrázek 22 Nastavení mapování sloupců komponentu OLEDB Destination (vlastní zpracování)

Je vhodné zkontrolovat datové typy sloupců, které budou stahovány ze sharepointu objektem OData a formát sloupců, které budou nahrávány do databáze. Tyto sloupce musejí mít stejný datový typ a jelikož relační databáze podporují pouze UNICODE datové typy, tak se jako nejvhodnější datový typ pro textové hodnoty jeví Unicode string [DT_WSTR] datový typ. Toto nastavení lze zkontrolovat a případně změnit pravým kliknutím na objekty a výběru Show Advanced Editor, kde input columns jsou sloupce vcházející do objektu a output columns jsou sloupce z objektu vycházející. Je tedy zapotřebí v objektu OData zkontrolovat input columns a v objektu OLEDB Destination sloupce external columns.

Vytvořený Data Flow Task je vhodné přejmenovat z důvodu odlišení od dalšího vytvářeného pro lepší přehlednost. První Data Flow Task je tedy pojmenován Business Units.

Data Flow Task pro získání dat týkající se oprávnění v seznamu Tab_AccessControlList bude obdobný jako předchozí Data Flow task a je pojmenován jako Access Control List. Data Flow Task obsahuje stejné 2 komponenty na sebe navázané, kdy oba tyto komponenty využívají stejné konektory jako předchozí komponenty. V komponentě Sharepoint List

Source je tedy rozdíl pouze v Site List Name a v návrháři komponent a Output Columns jsou omezeny na 3 sloupce, které jsou vyobrazeny na obrázku 23.



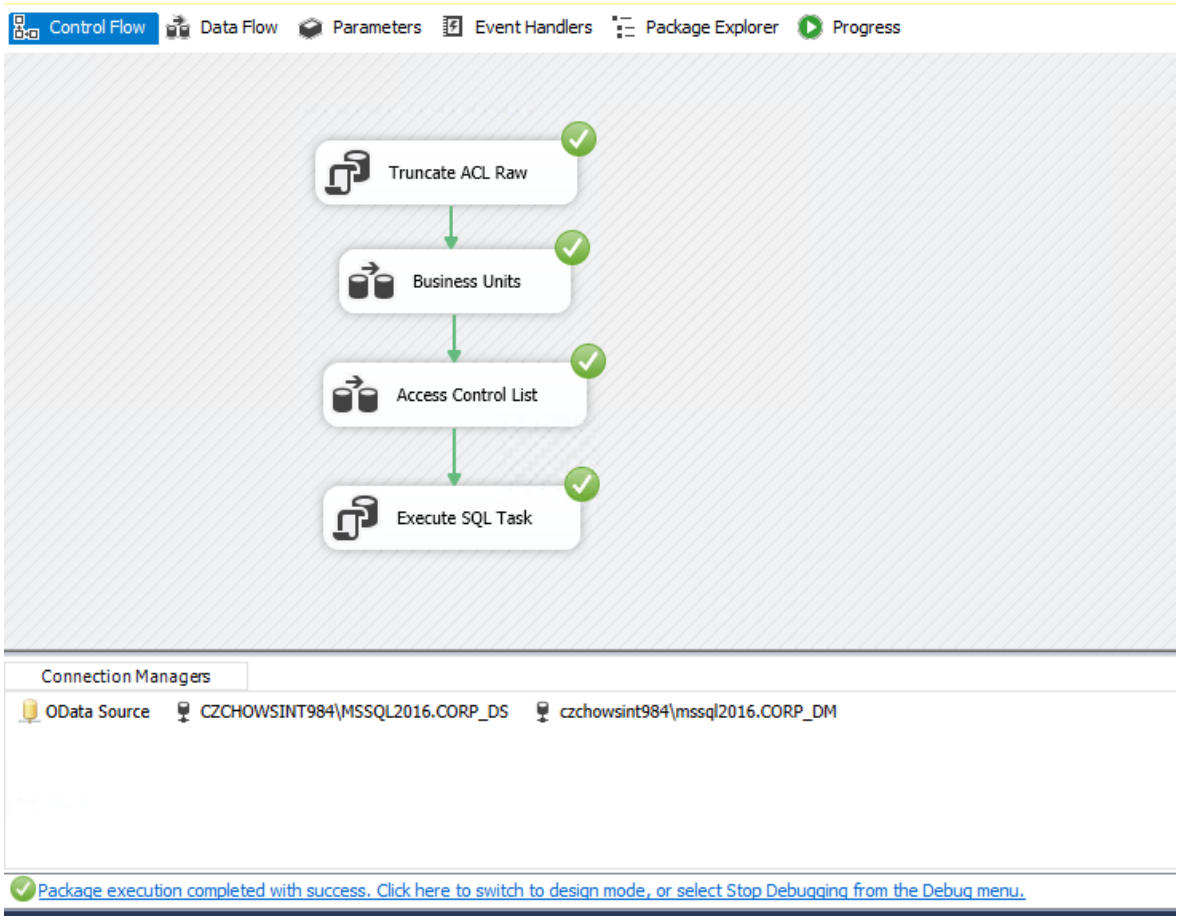
Obrázek 23 Nastavení výstupních sloupců komponentu Sharepoint List Source pro list Tab_AccessControlList (vlastní zpracování)

V této záložce lze zkontrolovat datové typy přenášených sloupců . V tomto případě datový typ sloupců Business Unit a User Name (Title) bude ponechán Unicode String (DT_WSTR) a pro sloupec ID to je four-byte signed integer [DT_I4].

V Control Flow části SSIS bude vložen poslední krok, kterým bude Execute SQL Task, jehož úkolem bude zavolat či spustit sql proceduru v databázi, která je popsána v další kapitole. Tento objekt bude nastaven podobně jako první objekt v tomto balíčku, avšak rozdíl bude v poli SQL Statement, kde bude EXEC sp_ACL_Raw a jelikož se tato procedura nachází

v datovém tržišti, je nutné vytvořit ještě připojení na tuto databázi, stejným postupem jako připojení na dočasné datové úložiště, a tento objekt připojit na tento konektor.

Po uložení všech změn lze balíček spustit a sledovat, zda-li proběhne vše v pořádku. Na obrázku 24 lze vidět, že všechny kroky proběhly úspěšně.



Obrázek 24 Úspěšné spuštění SSIS balíčku (vlastní zpracování)

V případě, že by některá z komponent neskončila úspěšně, ostatní komponenty by nebyly spuštěny a proces SSIS ETL by byl pozastaven. Lze nastavit ignorování vzniklého problému v záložce Event Handler, tudíž by další komponenty byly spouštěny i přes předchozí problémy.

3.3.6 ACL Stored Procedure

V tomto kroku je vytvořena stored procedure, jejímž záměrem je rozdělovat data z nenormalizované tabulky ACL_Ishare_Raw do normalizovaných tabulek ACL_BU a ACL_User.

Stored procedures jsou součástí procesu ETL, kdy stored procedures mohou obsahovat SQL příkazy, které mají za úkol data získávat, transformovat nebo přesouvat. Záměr vytvoření

stored procedures může být různý. V této práci bude vytvořena stored procedure se záměrem rozdělit nezpracovaná data do dimenzí aby byla dodržena 3. normální forma.

Pro snadnější práci s daty uloženými na sharepointu, je nutné získaná data rozdělit. Z tohoto důvodu byly vytvořeny tabulky zvlášť pro uživatele a zvlášť pro jejich podnikové části. Data ze sharepointu jsou dosud nahrávána do jedné souhrnné tabulky. Aby byla data rozdělena do separátních tabulek, bude vytvořena stored procedure. Z toho důvodu, že není očekávaná častá větší změna v datech a souhrnná tabulka bude denně přehrávána aktuálními daty, budou data v dimenzích mazána a nahrávána při každém spuštění této stored procedure.

```
CREATE PROCEDURE [tab_sec].[ACL_Procedure]
AS
BEGIN

	----INSERT INTO ACL USER
	TRUNCATE TABLE CORP_DM.tab_sec.ACL_User
	INSERT INTO CORP_DM.tab_sec.ACL_User
		SELECT ID
			,[user] AS username
		FROM CORP_DS.tab_sec.ACL_iShare_raw


	---- INSERT INTO ACL BU
	TRUNCATE TABLE CORP_DM.tab_sec.ACL_BU
	INSERT INTO CORP_DM.tab_sec.ACL_BU
		SELECT s1.ID
			,s1.Business_Unit
		FROM (
			SELECT  ID
				value AS Business_Unit
			FROM CORP_DS.tab_sec.ACL_iShare_raw t
			CROSS APPLY STRING_SPLIT(REPLACE([Business Unit Lookup
Values], ';#', '#'), '#')
			) s1
		INNER JOIN CORP_DS.tab_sec.BU b ON b.BU = s1.Business_Unit
END
```

Procedura je rozdělen na 2 části, kdy v první části bude plněna tabulka ACL_User. V příkazu výše lze vidět, že nejprve plní tabulku ACL_User a poté tabulku ACL_BU. Historizace těchto údajů není nutná a tak jsou tabulky nejprve odstraňovány a následně vytvářeny znovu.

Pro odstranění tabulek je použit příkaz Truncate, jelikož je rychlejší než příkaz Delete.

Do sloupce ID je nahráváno pole ID ze sharepointu bez dalších nutných transformací.

Sloupec uživatelské jméno nabývá takových hodnot, které obsahují uživatelovu lokaci a jeho uživatelské jméno, oddělené kombinací znaků ; a #. Tento tvar není potřeba nijak upravovat, jelikož hodnoty v tomto sloupci budou využívány právě v tomto tvaru.

Tabulku ACL_BU je nutné naplnit hodnotami ID uživatele a podnikových částí. Pro hodnoty ID uživatele není potřeba dalších transformací, a tak ji lze jednoduše vybrat z tabulky ACL_iShare_Raw. Komplikace nastává pro výrazy ve sloupci BU_Lookup_Values v tabulce ACL_iShare_Raw, které z důvodu přenosu mnohonásobných hodnot do jednoho řetězce jednotlivé hodnoty odděluje kombinací znaků, je nutné tyto hodnoty očistit do použitelné formy. Například mnohonásobná hodnota „GBS, PeP“ je v databázi uložena jako „4;#GBS;#5;#PeP“.

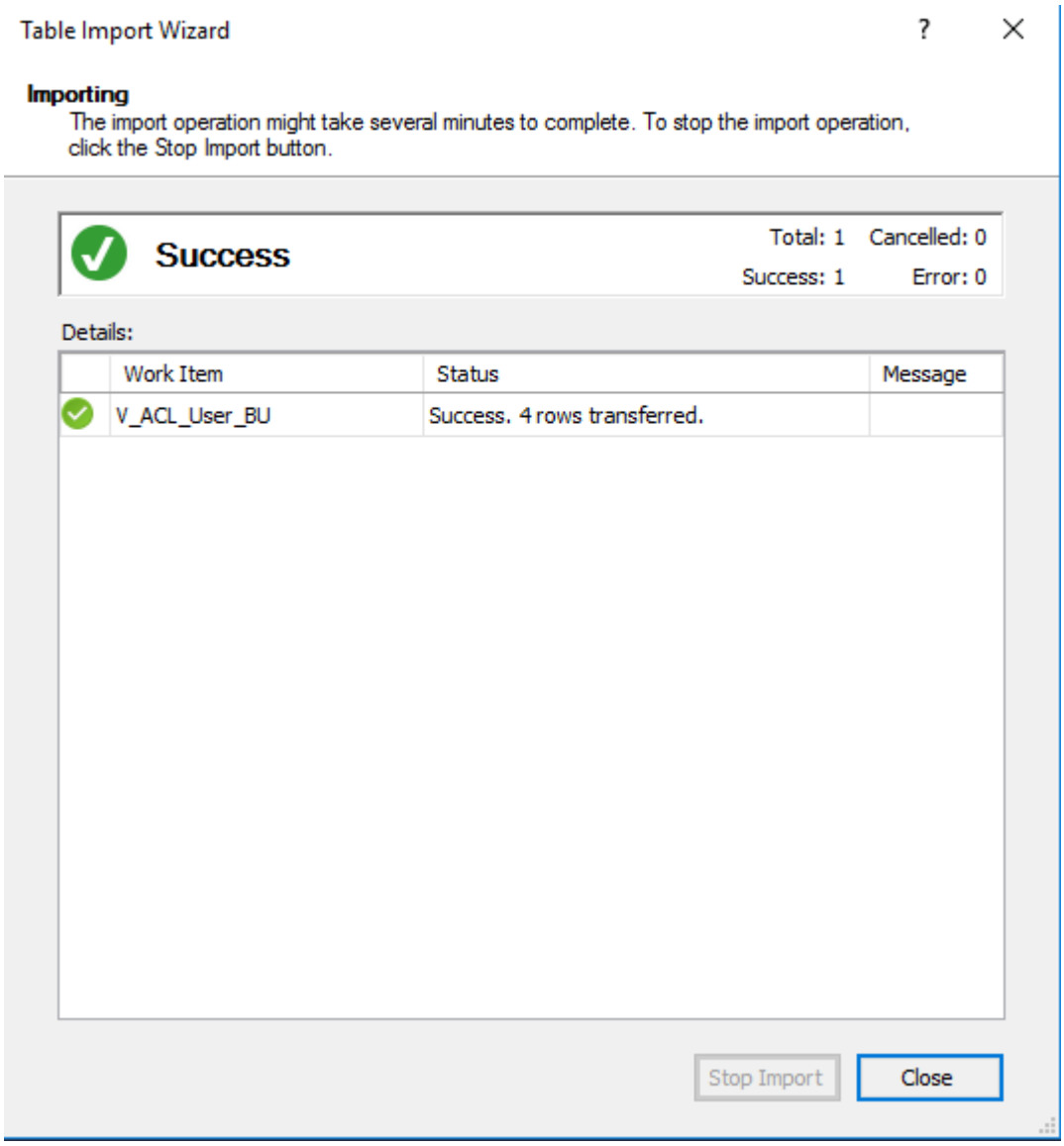
K tomu je využit vnořený select, nazývaný také jako subselect, vracející ID a hodnotu value označenou jako business_unit z tabulky ACL_iShare_Raw, za použití příkazu cross apply String_Split, který rozděluje hodnoty v jednom řetězci podle vybraného znaku do jednotlivých záznamů, tedy řádků. Výstupem vnořeného příkazu replace je nahrazení znaků ;# znakem # pro zjednodušení následného zpracování hodnoty.

Tento vnořený select, aby nevracel také zbylé znaky, které nebyly odstraněny předchozím příkazem, je nutné napojit pomocí Inner join na tabulku obsahující seznam podnikových částí. Výsledkem celého příkazu select je seznam ID a podnikových částí rozdělených do řádků ve vztahu 1:1.

3.3.7 Import ACL tabulek do tabulárního modelu

Data získaná z ACL tabulek je nutné naimportovat do tabulárního datového modelu, aby se s nimi mohlo dále pracovat.

Import tabulek je jednoduchý. V aplikaci Microsoft Visual Studio po připojení se k vytvořenému tabulárnímu modelu je zobrazena v horním panelu záložka Model, pod kterou lze nalézt možnost Existing Connections. V otevřeném dialogovém okně lze zvolit připojení ke zdroji, ve kterém se nachází tabulka, která bude přidána do modelu. Po odkliknutí tlačítka Open se zobrazí okno Select Tables and Views, kde lze již vybrat tabulku. V tomto okně lze také vybrat možnost Preview and Filter, kde lze aplikovat filtry na výběr dat, která mají být nahrána. Po odkliknutí se zobrazí další dialogové okno, kde lze vidět stav nahrávání tabulek. Po jejich úspěšném nahrání je vyobrazeno také počet přenesených řádků, viz obrázek 25.



Obrázek 25 Import tabulky do tabulárního modelu (vlastní zpracování)

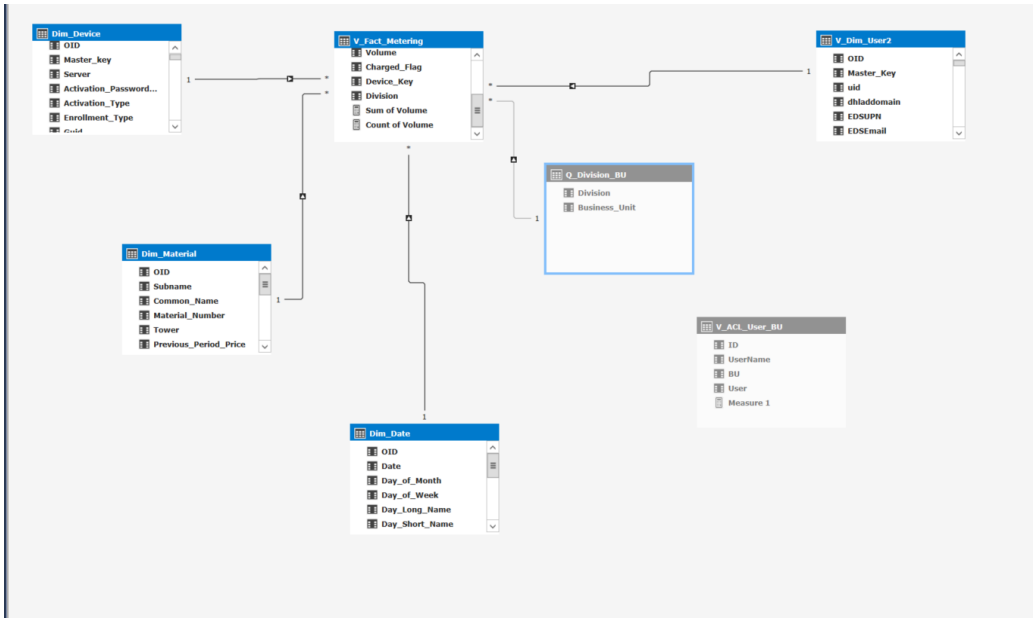
Tabulka jménem Q_Division_BU je přidána z toho důvodu, že ve faktové tabulce je granuralita dat nižší a data vztahující se k podnikovým částem jsou detailnější, jsou zobrazeny na úrovni divizí, kdy tedy vztah divize a podnikové části může být 1:N. Tabulka Q_Division_BU je použita jako jakési přemostění, kdy sloupce v této tabulce jsou BU značící podnikové části a Division značící divize. Tyto sloupce jsou získány z dimenzionální tabulky uložené v relační databázi, která nese data o podnikových částech. Abychom zamezili duplikovaným datům v této tabulce, je lepší použít možnost použít příkaz k importu tabulek. Příkaz je následující:

```
SELECT DISTINCT
  Dim_Crest_Division.Business_Unit
  ,Dim_Crest_Division.Division
FROM
  Dim_Crest_Division
WHERE
Dim_Crest_Division.Business_Unit IS NOT NULL
AND
Dim_Crest_Division.Division IS NOT NULL
```

V pravém spodním rohu okna lze příkaz validovat.

3.3.8 Vhodné použití a nastavení ACL tabulek v tabulárním modelu

Nově přidané tabulky je nutné v modelu napojit na tabulku faktů, aby bylo uskutečněno následné filtrování dat ovlivněné vztahem. Vytváření vztahů mezi tabulkami lze jednoduše v grafickém zobrazení modelu v aplikaci Visual Studio, které je vyobrazeno na obrázku 26. Přetažením sloupce z tabulky Q_Division_BU do tabulky v_Fact_Metering lze vytvořit vztah těchto tabulek založený na hodnotách ve sloupcích Division.



Obrázek 26 Grafické zobrazení tabulárního modelu (vlastní zpracování)

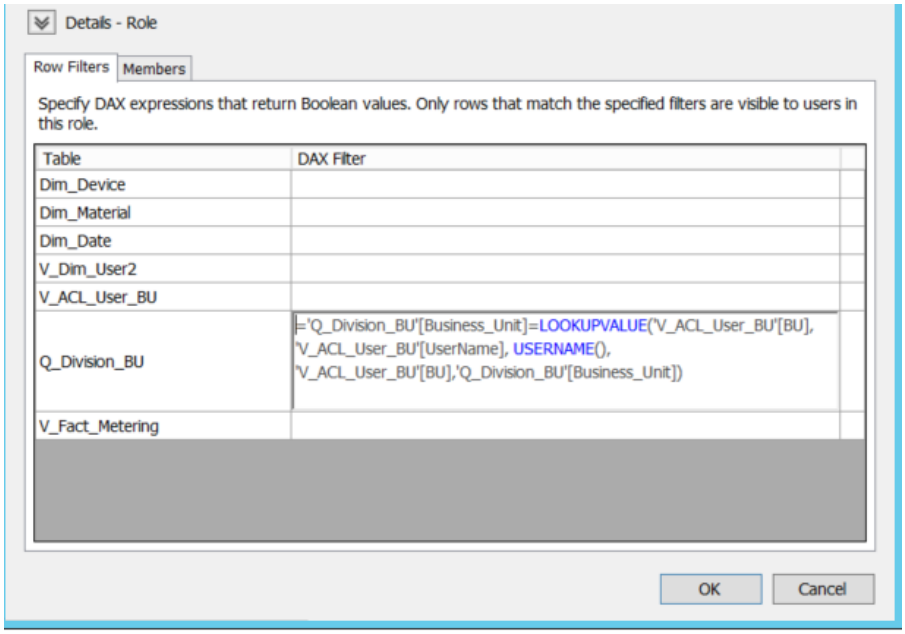
Doporučuje se tabulky týkající se oprávnění schovat před samotnými uživateli. Tento krok lze uskutečnit pravým kliknutím na tabulku v grafickém modelu a výběru možnosti Hide from Client Tools. Takto schované tabulky mají v grafickém zobrazení šedou barvu, a pro jejich zobrazení je nutné mít povoleno zobrazování schovaných položek. I přes to však mohou uživatelé posílat dotazy i na tyto tabulky. Této skutečnosti může být zamezeno použitím DAX příkazu „=FALSE()“ pro tyto tabulky.

Pro dosažení správného fungování bezpečnostního modelu je vztah vytvořen pouze mezi tabulkou Q_Division_BU a faktovou tabulkou v_Fact_Metering. Tabulku v_ACL_User_BU není nutné napojovat na Q_Division_BU tabulku, jelikož z ní budou získávány hodnoty, které budou dále využívány při filtrování dat. Navíc tabulku v_ACL_User_BU není možné napojit, jelikož tento vztah hlásí chybnou hlášku, ve které je uvedeno, že hodnoty v_ACL_User_BU obsahují duplikáty. Tato chybová hláška je uvedena z toho důvodu, že při napojování sloupců BU a BU by vyžadovalo vztah M:N.

Role lze vytvářet v SSDT. Pomocí jazyku DAX lze vytvářet filtry na data v tabulkách. V tomto případě poslouží příkaz Lookupvalue, který vyžaduje definici výsledné hodnoty a poté podmínky, které se musejí rovnat. Jazyk DAX vyžaduje název tabulky i sloupce, kdy název tabulky je uveden v jednoduchých uvozovkách a název sloupce v této tabulce je uveden v hranatých závorkách.

Je potřeba získat uživatelské jméno přistupujícího uživatele. Toho je docíleno příkazem USERNAME(). Získanou hodnotu je nutné nalézt v tabulce v_ACL_User_BU ve sloupci UserName. Z této tabulky je potřeba získat hodnotu BU. Z tabulky Q_Division_BU je poté nutné získat hodnotu Division k hodnotě BU. Nakonec z tabulky v_Metering je nutné vyfiltrovat data pouze pro vybrané divize.

Na obrázku 27 lze nalézt příkaz DAX pro tabulku Q_Division_BU.



Obrázek 27 DAX filtr vytvořené role (vlastní zpracování)

3.3.9 Testování

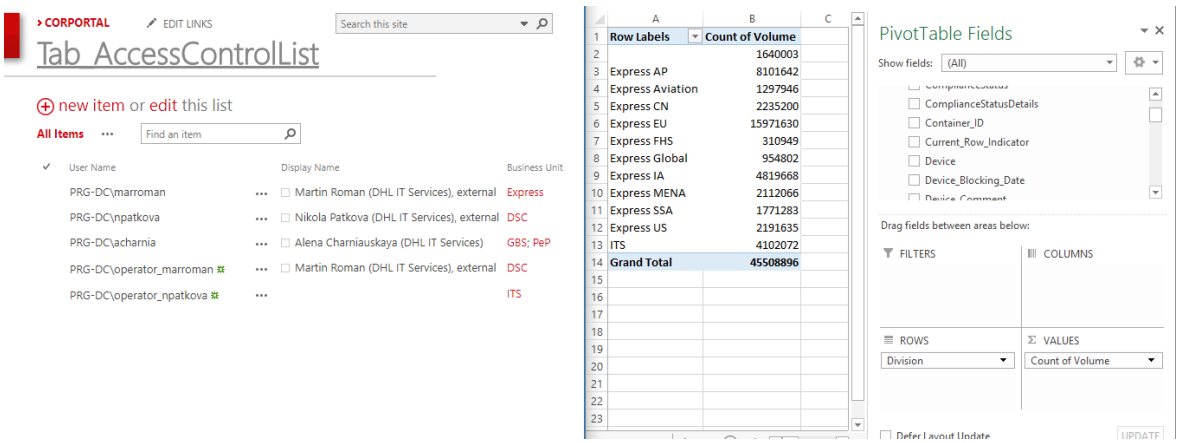
V této kapitole bude testován přístup k datovému modelu pomocí přidělených rolí.

Testování rolí proběhlo pomocí operátorského účtu, jemuž byla přidělena vytvořená role. Tento operátorský účet nemá přístup do analytické databáze na serveru, a tak bylo možné implementované řešení otestovat. Jelikož nebylo možné spustit aplikaci excel pomocí tohoto operátorského účtu, byla pod tímto účtem spuštěna aplikace Microsoft Visual Studio.

Aplikace Microsoft Visual Studio umožňuje analýzu rolí pomocí aplikace excel tlačítkem Analyze in excel. Po kliknutí na toto tlačítko je otevřena aplikace excel pod účtem, pod který je uživatel přihlášen v aplikaci Microsoft Visual Studio. V otevřené aplikaci excel je automaticky vytvořená tabulka typu Power Pivot.

V první řadě lze otestovat, zda jsou zobrazeny tabulky, které byly v SSDT schovány pro klientské aplikace. Na pravé straně okna se nachází seznam tabulek. Schované tabulky se v tomto listu nenacházejí.

Dále lze otestovat filtrování dat v tabulkách. V tabulárním modelu jsou vytvořená kalkulovaná pole. K otestování postačí kalkulovaný sloupec Count of Volume jako zobrazené hodnoty, a zobrazení divizí například jako řádky. Na obrázku 28 lze vidět výsledek tohoto testování. V tabulce se nachází pouze názvy divizí spadající pod podnikové části (vpravo), ke kterým byl uživateli přidělen přístup (vlevo).



Obrázek 28 Výsledek testování role (vlastní zpracování)

Prázdná hodnota v prvním řádku prvního sloupce je zapříčiněná tím, že historická data nenesou informaci o divizích.

Závěr

Cílem této práce byla tvorba, implementace a testování bezpečnostního modelu pro již vytvořený Tabulární model.

V první části této práce byly vysvětleny pojmy související nejen se samotným tématem této práce, ale také s použitými technologiemi a komponenty při vytváření tohoto bezpečnostního modelu. Pojmy byly rozděleny zvlášť do kapitol a pod každou kapitolou je uvedeno shrnutí kapitoly, kde je obsaženo autorčino shrnutí kapitoly a případně nějaký další komentář.

Po části teoretické následovala část praktická, ve které byl popsán současný stav Tabulárního modelu a jeho zabezpečení a popsáno bylo také řešení zebezpečení reportů, které bylo prvotní inspirací při návrhu tohoto modelu. Následně byl vytvořen návrh modelu a před samotným postupem tvorby modelu bylo zmíněno, které aplikace byly pro tvorbu použity. Dále již byly popsány kroky nezbytné pro vytvoření bezpečnostního modelu. Některé kroky nejsou zcela navazující, aby se zamezilo opakování kapitol. Postup pro vytvoření modelu byl rozdělen do devíti kroků včetně testování a byl dokumentován také pomocí obrázků a vložených příkazů.

Jednou z obtížnějších součástí při vytváření modelu bylo nastavení připojení komponenty SSIS ETL na sharepoint. Zprvu probíhala tvorba SSIS ETL balíčku na lokálním počítači, avšak jelikož byl jeden z přenášených sloupců typu multichoice, komponent OData Source již nestačil. Verze aplikace Microsoft Visual Studio s nainstalovaným rozšířením pro Sharepoint se nachází na vzdálené ploše, proto byla tvorba SSIS ETL balíčku přesunuta na vzdálenou plochu.

Další obtížnou součástí vytváření modelu byla tvorba návrhu použití vytvořených ACL tabulek v samotném Tabulárním modelu a následné vytváření a aplikace filtrů ve vytvořené roli. Především při testování vytvořené role bylo překážkou celopodnikové zabezpečení, jelikož autorka společně s ostatními vývojáři v týmu, má administrátorská práva na server, a tak nebylo zcela efektivní testovat role pomocí těchto účtů, jelikož se vytvořené filtrování neprojevilo a byla zobrazována veškerá data. Tato překážka byla vyřešena pomocí operátorského účtu, který již práva k přístupu do analytické databáze přidělena neměl.

I přes veškeré obtížnosti bylo cíle a dílčích cílů této práce úspěšně dosaženo. Byl navržen a implementován security model využívající ACL tabulky pro filtrování dat zobrazených v klientských aplikacích. Data v těchto tabulkách jsou vkládána do seznamu nacházejícím se na stránce sharepoint a pomocí vytvořeného balíčku SSIS, jehož součástí je i spouštění vytvořené stored procedures, jsou data zpracována a ukládána do těchto ACL tabulek v relační databázi. Tabulka byla též přidána do tabulárního modelu a byla vytvořena role s filtry podle podnikové části, využívající data v ACL tabulkách.

Řešení vytvářené v rámci této bakalářské práce bylo jedním z mnoha možných řešení. Dalším možným postupem vývoje tohoto bezpečnostního modelu je nastavení

automatického spouštění SSIS balíčku pomocí SQL Task Scheduleru. Jednou z možností je přidat tento balíček jako součást hlavního, rodičovského balíčku, případně přidat krok nový. Z důvodu nutnosti procesování tabulární kostky pro zanesení veškerých změn by nebylo vhodné řešení vytvářet pro tento krok zcela nový task.

Terminologický slovník

Termín	Význam
ACL	Access Control List – seznamy vázající se k objektům, které obsahují informace o oprávnění a přístupech k danému objektu
Analýza	Analýza - postup, jehož výstupem je logický model vytvářeného systému.
Aplikace	Aplikace - řešení řídících, finančních, obchodních, výrobních a dalších procesů a funkcí podniku pomocí prostředků informačních a komunikačních technologií, tj. aplikačního a základního software, technických a komunikačních prostředků a s nimi spojených služeb poskytovaných jejich uživatelům.
Atribut	Atribut - vlastnost entity nebo sloupec/pole u databázové tabulky nebo položka v datovém souboru. Datová část objektu, tj. proměnná, která je součástí objektu.
Business Intelligence (BI)	Business Intelligence - BI - souhrn technologií, a nástrojů, které umožňují na základě aplikace multidimenzionálního přístupu k měření vytvářet a optimalizovat aplikace na podporu analytických a plánovacích aktivit podniku
Datová pumpa (ETL)	Datová pumpa - ETL - Extract, Transform, Load - datová pumpa - software pro realizaci transformací dat mezi různými datovými zdroji
Datový Typ	Datový typ - označuje typ dat uložených v proměnné, např. číslo, text, apod. Určuje množinu přípustných hodnot a operace s těmito hodnotami
Dimenze	Dimenze - analytické hledisko pro identifikaci a hodnocení sledovaných ukazatelů a je tak součástí de facto každé metriky. Z informatického pohledu se jeví jako struktura dat, případně jako databázová tabulka obsahující záznamy o jednotlivých prvcích dimenze, tj. např. o jednotlivých službách informatiky, dodavatelích IT apod.
Dimenzionální tabulka	Dimenzionální tabulka - tabulka obsahující kontext k datům fakt tabulky, obsahuje obvykle textové popisné údaje.
Dotazovací jazyk	Dotazovací jazyk - umožňuje uživateli formulovat dotazy či získávat informace z databáze formulací požadavku podle definovaných standardů.

Engine	Engine – softwarový modul, používaný systémem pro správu databází k práci s daty, jako například tvorbě, čtení a aktualizaci dat z databáze
Faktová tabulka	Fakt tabulka - tabulka obsahující obchodní fakta a hodnoty, zpravidla číselné a aditivní.
Komponenta	Komponenta - stavební blok aplikace, samostatná binární část kódu
Metadata	Metadata - informace o datech, tj. názvy tabulek, názvy atributů, datové typy, primární klíče, vazby, komentáře, atd.
Multidimenzionální databáze	Multidimenzionální databáze - databáze, kde jsou data uložena na principu vícerozměrové matice. Hodnoty jsou přístupné přímo pro danou kombinaci prvků dimenzí.
OLAP	OLAP - Online Analytical Processing - informační technologie založená především na koncepci multidimenzionálních databází. Jejím hlavním principem je několikadimenzionální tabulka umožňující rychle a pružně měnit jednotlivé dimenze a měnit tak pohledy uživatele na modelovanou ekonomickou realitu.
Produkční server/databáze	Produkční databáze - databáze uchovávající data, které vytváří a spravují transakční informační systémy a představují hlavní vstup do řešení business intelligence.
Relační databáze	Relační databáze - databáze splňující dvě základní podmínky: databáze je chápána jako množina relací, jsou k dispozici minimálně operace selekce, projektování a spojení.
Report	Report - výstup ze software s výsledky hodnot ukazatelů. Může být vyjádřen tabulkou, grafem, atd.
Role	Role - představují typové skupiny pracovníků charakterizované vykonáváním obdobných činností. Jeden člověk může zastávat více rolí. Role mají definované odpovědnosti, které zahrnují provádění určitých činností.
Autorizace	Autorizace - oprávněnost určité entity (uživatele, programu) k přístupu k určitým zdrojům (např. serveru, souboru). Předpokladem autorizace entity je úspěšná autentizace
Data Mart	Data Mart - DMA datové tržiště, analytická databáze určená pouze pro určitou část podniku (útvar, oblast řízení apod.).
Databáze	Databáze - DB - data, které slouží více aplikacím, jsou v nich minimalizovány redundance dat a existuje

	vhodně centralizovaná správa těchto dat. Cílem databázového systému je uspořádat datové zdroje (datovou základnu) na počítači tak, aby tyto zdroje mohly být využívány.
Databázový server	Databázový server - Database Server, server orientovaný na zpracování dat v databázích.
Datové tržiště	Datové tržiště - Data Mart - DMA - analytická databáze určená pouze pro určitou část podniku (útvar, oblast řízení apod.).
Kostka	Kostka - označení pro multidimenzionální databázi (nebo její princip). Kostka je realizována buď prostřednictvím multidimenzionální databáze, nebo zvláštní struktury relační databáze.
Procedura	Procedura - ucelená, relativně samostatná jednotka programu
SQL	SQL - Structured Query Language – databázový dotazovací jazyk
3. NF (3. normální forma)	3. normální forma – metodika pro návrh datové struktury databáze, jehož dodržení vede k optimálnímu využití vlastností systému OLTP při tvorbě databázových aplikací

Použitá literatura

DUNCAN, Owen. 2019. *What is Azure Analysis Services?* [online]. Microsoft Docs. [cit. 2019-04-30]. Dostupné z: https://docs.microsoft.com/en-us/azure/analysis-services/analysis-services-overview

DUNCAN, Owen, HAMILTON, Bruce, CAI, Saisang, DUNN, James, KUMAR, Sudeep, MILENER, Gene a RABELER, Carl. 2018. *About SQL Server Analysis Services* [online]. Microsoft Docs. [cit. 2019-04-30]. Dostupné z: https://docs.microsoft.com/en-us/sql/analysis-services/analysis-services?view=sql-server-2017

DUNCAN, Owen, GUYER, Craig a RABELER, Carl. 2018. *Tabular model roles* [online]. Microsoft Docs. [cit. 2019-04-14]. Dostupné z: https://docs.microsoft.com/en-us/sql/analysis-services/tabular-models/tabular-model-roles-ssas-tabular?view=sql-server-2017

GÁLA, Libor, POUR, Jan a ŠEDIVÁ, Zuzana. 2009. *Podniková informatika*. 2. přepracované vydání. Praha: Grada Publishing. ISBN 978-80-247-2615-1.

GARTNER, 2017. *Gartner IT Glossary – Active Directory* [online]. Gartner. [cit. 2019-04-27]. Dostupné z: https://www.gartner.com/it-glossary/active-directory/

GARTNER, 2017. *Gartner IT Glossary - Self-Service Business Intelligence*. Gartner [online]. [cit. 2019-04-14]. Dostupné z: https://www.gartner.com/it-glossary/?s=self+service+business+intelligence

GUYER, Craig, MILENER, Gene, KUMAR, Sudeep, HAMILTON, Bruce, ROTH, Jason, LAUDENSCHLAGER, Douglas a RABELER, Carl. 2017. *Master data services Overview (MDS)* [online]. Microsoft Docs. [cit. 2019-04-30]. Dostupné z: https://docs.microsoft.com/en-us/sql/master-data-services/master-data-services-overview-mds?view=sql-server-2017

HAMAN, Martin. 2018. *Data Day 2018* [online]. Dostupné z: https://www.gopas.sk/getattachment/47219b58-3c18-427e-b304-c940604185a8/Martin_Haman_all.pdf.aspx;.php;.html;;

JOYFUL CRAFTSMAN. 2018. *Hitchhiker's Guide to the SSAS Tabular* [online]. [cit. 2019-04-14]. Dostupné z: https://www.sqlsaturday.com/779/Sessions/Details.aspx?sid=85530

KENNEDY, John a SATRAN, Michael. 2018. *Access Control Lists*. Microsoft Docs [online]. [cit. 2019-04-23]. Dostupné z: https://docs.microsoft.com/en-us/windows/desktop/secauthz/access-control-lists

LACKO, Luboslav. 2003. *Databáze: datové sklady, OLAP a dolování dat s příklady v Microsoft SQL Serveru a Oracle*. Vydání první. Brno: Computer Press. ISBN 80-7226-969-0.

MAZÁČOVÁ, Markéta. *Možnosti analytických nástrojů v prostředí MS SQL Serveru*. Brno, 2017. Diplomová práce. Mendelova univerzita v Brně. Vedoucí práce Jan Přichystal.

MBI tým. *AQ001A*: *Aplikace: Power Pivot* [online]. MBI. [cit. 2019-04-30]. Dostupné z: https://mbi.vse.cz/mbi/index.html#obj/APPLICATION-156

POUR, Jan, MARYŠKA, Miloš, STANOVSKÁ, Iva a ŠEDIVÁ, Zuzana. 2018. *Self Service Business Intelligence: Jak si vytvořit vlastní analytické, plánovací a reportingové aplikace*. Praha: Grada Publishing. ISBN 978-80-271-0616-5.

RABELER, Carl, GUYER, Craig a DUNCAN, Owen. 2018. *Tabular Models*. Microsoft Docs [online]. [cit. 2019-04-14]. Dostupné z: https://docs.microsoft.com/en-us/sql/analysis-services/tabular-models/tabular-models-ssas?view=sql-server-2017

RABELER, Carl, GUYER, Craig a DUNCAN, Owen. 2018. *Perspectives in tabular models* [online]. Microsoft Docs. [cit. 2019-04-29]. Dostupné z: https://docs.microsoft.com/en-us/sql/analysis-services/tabular-models/perspectives-ssas-tabular?view=sql-server-2017

ŘEHOŘ, F. *F455: BI: Self Service BI* [online]. MBI. [cit. 2019-04-14]. Dostupné z: https://mbi.vse.cz/mbi/index.html#obj/FACTOR-110

STANOVSKÁ, Iva. *AQ002A: Aplikace: Power BI – Principy, funkce* [online]. MBI. [cit. 2019-04-30]. Dostupné z: https://mbi.vse.cz/mbi/index.html#obj/APPLICATION-161

SUPPORT OFFICE, 2019. *Přehled technologie OLAP (Online Analytical Processing)* [online]. Support Office. [cit. 2019-04-14]. Dostupné z: https://support.office.com/cs-cz/article/p%C5%99ehled-technologie-olap-online-analytical-processing-15d2cdde-f70b-4277-b009-ed732b75fdd6#top

TEJADA, Zoiner. 2018. *Technologie OLAP (Online Analytical Processing)* [online]. Microsoft Docs. [cit. 2019-04-14]. Dostupné z: https://docs.microsoft.com/cs-cz/azure/architecture/data-guide/relational-data/online-analytical-processing

VertiPaq engine in SSAS Tabular Database [online]. 2017. [cit. 2019-04-30]. Dostupné z: http://nisalbi.blogspot.com/2017/11/vertipaq-engine-in-ssas-tabular-databse.html

What is SharePoint? [online]. Microsoft Support Office. [cit. 2019-04-30]. Dostupné z: https://support.office.com/en-us/article/what-is-sharepoint-97b915e6-651b-43b2-827d-fb25777f446f

ZHANG, Janine, GUYER, Craig, MILENER, Gene a MABEE, Dan. 2018. *SQL Server Integration Services* [online]. Microsoft Docs. [cit. 2019-04-30]. Dostupné z: https://docs.microsoft.com/en-us/sql/integration-services/sql-server-integration-services?view=sql-server-2017

GARTNER. 2019. *2019 Magic Quadrant for Analytics and Business Intelligence Platforms*[online]. Gartner. [cit. 2019-04-14]. Dostupné z: https://www.qlik.com/us/gartner-magic-quadrant-business-intelligence

Seznam obrázků

Obrázek 1 Obecná architektura BI řešení (Gála et al. 2009) 7
Obrázek 2 Trh produktů BI (Gartner, 2019) 9
Obrázek 3 Komprese dat ve VertiPaq engine (Joyful Craftsmen, 2018) 14
Obrázek 4 Schéma zpracování dotazů v SSAS Tabular (Joyful Craftsmen, 2018) 15
Obrázek 5 Pohled na role v databázi (vlastní zpracování) 17
Obrázek 6 Pohled na správu sloupců v aplikaci Sharepoint (vlastní zpracování) 25
Obrázek 7 Úprava předvytvořeného sloupce v aplikaci Sharepoint (vlastní zpracování) 26
Obrázek 8 Seznam podnikových částí vytvořený v aplikaci Sharepoint (vlastní zpracování) 27
Obrázek 9 Nastavení datového typu při tvorbě sloupce Business Unit v aplikaci Sharepoint (vlastní zpracování) 28
Obrázek 10 Nastavení potřebných atributů při tvorbě sloupce Business Unit (vlastní zpracování) 28
Obrázek 11 Zařazení schéma v databázovém stromu (Vlastní zpracování) 29
Obrázek 12 Vytváření nového schéma (vlastní zpracování) 30
Obrázek 13 Dialogové okno tvorby schématu (vlastní zpracování) 31
Obrázek 14 Tvorba Projektu SSIS balíčku (vlastní zpracování) 35
Obrázek 15 Execute SQL Task návrhář (Vlastní zpracování) 36
Obrázek 16 Nastavení OLE DB konektoru (Vlastní zpracování) 37
Obrázek 17 Nastavení OData konektoru (vlastní zpracování) 38
Obrázek 18 Nastavení komponentu Sharepoint List Source (vlastní zpracování) 39
Obrázek 19 Nastavení sloupců komponentu Sharepoint List Source pro data podnikových částí (vlastní zpracování) 40
Obrázek 20 Pohled na Data Flow Task s nastavenými komponenty (vlastní zpracování) 41
Obrázek 21 Nastavení připojení komponentu OLEDB Destination (vlastní zpracování) 42
Obrázek 22 Nastavení mapování sloupců komponentu OLEDB Destination (vlastní zpracování) 43
Obrázek 23 Nastavení výstupních sloupců komponentu Sharepoint List Source pro list Tab_AccessControlList (vlastní zpracování) 44
Obrázek 24 Úspěšné spuštění SSIS balíčku (vlastní zpracování) 45
Obrázek 25 Import tabulky do tabulárního modelu (vlastní zpracování) 48
Obrázek 26 Grafické zobrazení tabulárního modelu (vlastní zpracování) 49
Obrázek 27 DAX filtr vytvořené role (vlastní zpracování) 50
Obrázek 28 Výsledek testování role (vlastní zpracování) 51

Seznam zkratek

ACL	Access Control List
BI	Business Intelligence
ETL	Extract, Transform, Load
MS SQL	Microsoft SQL
OLAP	On Line Analytical Processes
SQL	Sequel Query Language
SSAS	SQL Server Analysis Services
SSBI	Self Service Business Intelligence
SSDT	SQL Server Data Tool
SSIS	SQL Server Integration Services
SSMS	SQL Server Management Studio