

MYKOLO ROMERIO UNIVERSITETAS
SOCIALINĖS POLITIKOS FAKULTETAS
KOMUNIKACIJOS IR INFORMATIKOS INSTITUTAS

RAIMONDAS DRUSEIKIS

DUOMENŲ SAUGYKLOS ANALIZĖS PRIEMONIŲ
TAIKYMAS VERSLO ĮMONĖSE

Magistro baigiamasis darbas

Vadovė
Lekt. dr. Ramutė Naujikienė

Vilnius, 2013

MYKOLO ROMERIO UNIVERSITETAS
SOCIALINĖS POLITIKOS FAKULTETAS
KOMUNIKACIJOS IR INFORMATIKOS INSTITUTAS

RAIMONDAS DRUSEIKIS

DUOMENŲ SAUGYKLOS ANALIZĖS PRIEMONIŲ
TAIKYMAS VERSLO ĮMONĖSE

Elektroninio verslo vadybos magistro baigiamasis darbas
Studijų programa 621N20018

Recenzentas
2013 04 ...

Vadovė
Lekt. dr. R. Naujikiene
2013 04 ...

Atliko
EVV mis 1-01 gr. stud.
R. Druseikis
2013 04 ...

Vilnius, 2013

TURINYS

VARTOJAMŲ SANTRUMPŲ IR SĄVOKŲ SĄRAŠAS.....	5
LENTELIŲ SĄVADAS	6
PAVEIKSLŲ SĄVADAS	6
ĮVADAS	8
1. VERSLO ANALITIKOS SISTEMŲ ANALIZĖ IR APŽVALGA	11
1.1. VERSLO ANALITIKOS SISTEMOS.....	12
1.2. DUOMENŲ SAUGYKLOS SĄVOKA	15
1.3. OLAP DUOMENŲ KUBAI	17
1.4. PROGNOZUOJAMA VERSLO ANALITIKOS SISTEMŲ RAIDA.....	20
1 SKYRIAUS IŠVADOS	22
2. PARDAVIMŲ IR KLIENTŲ ANALIZĖS SISTEMOS PROJEKTAS	24
2.1. PARDAVIMŲ IR KLIENTŲ ANALIZĖS SISTEMOS PROJEKTO TIKSLAS	24
2.2. SISTEMOS REIKALAVIMŲ SPECIFIKACIJA.....	25
2.2.1. DUOMENŲ ANALIZĖS KUBO SPECIFIKACIJA	35
2.2.2. DETALI REIKALAVIMŲ SPECIFIKACIJA	37
2.3. PARDAVIMŲ IR KLIENTŲ ANALIZĖS SISTEMOS ARCHITEKTŪRA	39
2.3.1. VERSLO ANALITIKOS SISTEMOS ARCHITEKTŪRA.....	41
2.3.2. REIKALAVIMAI OLAP DUOMENŲ KUBO FUNKCIONAVIMO PALAIKYMUI.....	42
2.3.3. DUOMENŲ TRANSFORMACIJŲ PASLAUGOS	43
2.3.4. DUOMENŲ ATNAUJINIMO AUTOMATIZAVIMAS	49
2.3.5. DUOMENŲ ANALIZĖS KUBO FORMAVIMO PROCESAS	50
2.3.6. DUOMENŲ ANALIZĖS SISTEMOS FORMAVIMO MODELIS	52
3. OLAP KUBŲ ARCHITEKTŪROS LYGINAMOJI ANALIZĖ.....	54
3.1. „ŽVAIGŽDĖS“ IR „SNAIGĖS“ TIPO DUOMENŲ SCHEMAS	55
3.2. MOLAP ARCHITEKTŪRA.....	57
3.3. MOLAP PRANAŠUMAI IR TRŪKUMAI.....	59
3.4. ROLAP ARCHITEKTŪRA	59
3.5. ROLAP PRANAŠUMAI IR TRŪKUMAI.....	61
3.6. HOLAP ARCHITEKTŪRA	61

3 SKYRIAUS IŠVADOS	62
IŠVADOS	63
LITERATŪROS SĄRAŠAS	65
SANTRAUKA	68
SUMMARY	69

VARTOJAMŲ SANTRUMPŲ IR SĄVOKŲ SĄRAŠAS

DWH – Duomenų saugykla (angl. A Data Warehouse).

BI – verslo analitika arba verslo įžvalgos.

UML – modeliavimo kalba, kuri turtinga grafiniu atvaizdavimu ir elementų gausa.

BIS – verslo įžvalgų arba verslo analitikos sistemos.

IT – informacinės technologijos.

OLAP (Interaktyvus analitinis duomenų apdorojimas) – priemonės duomenų analitiniam apdorojimui realiu laiku.

ETL – ištraukti, transformuoti, importuoti (angl. extract, transform, load).

MOLAP (angl. Multidimensional OLAP)– daugiamatis OLAP. Duomenys yra saugomi daugiamatžio kubo struktūroje.

ROLAP (angl. Relational OLAP)– sąryšinė OLAP sistemos architektūra. Duomenys yra saugomi reliacinėse duomenų bazėse.

HOLAP (angl. Hybrid OLAP) – hibridinė OLAP architektūra. Agregacijos yra saugomos daugiamatžio kubo struktūroje, o detalesni duomenys – reliacinėse lentelėse.

CRM (angl. Customer Relationship Management) – santykių su klientais valdymo sistema.

ERP (angl. Enterprise Resource Planning) – verslo valdymo sistema.

SCM (angl. Supply chain management) – tiekimo grandinės valdymo sistemos.

LENTELIŲ SAŲVADAS

1 lentelė.	Panaudos atvejo „Analizuoti pardavimų ir klientų duomenis“ specifikacija.....	26
2 lentelė.	Panaudos atvejo „Pagal regionus“ specifikacija.....	28
3 lentelė.	Panaudos atvejo „Pagal akcijas“ specifikacija.....	29
4 lentelė.	Panaudos atvejo „Pagal paslaugas“ specifikacija.....	31
5 lentelė.	Panaudos atvejo „Pagal pardavimo kanalus“ specifikacija.....	32
6 lentelė.	Panaudos atvejo „Formuoti ataskaitas“ specifikacija.....	33
7 lentelė.	Pardavimų ir klientų duomenų kubo struktūra.....	35
8 lentelė.	Analizės kubo sąryšiai su pradiniais duomenų šaltiniais.....	37

PAVEIKSLŲ SAŲVADAS

1 pav.	Verslo analitikos platformų kūrėjų populiarumas pagal GARTNER.....	13
2 pav.	Duomenų saugyklos komponentai.....	16
3 pav.	Trijų dimensijų kubas su pardavimų informacija.....	19
4 pav.	Sistemos panaudos atvejų diagrama.....	26
5 pav.	Panaudos atvejo „Analizuoti pardavimų ir klientų duomenis“ sekų diagrama.....	27
6 pav.	Panaudos atvejo „Pagal regionus“ sekų diagrama.....	29
7 pav.	Panaudos atvejo „Pagal akcijas“ sekų diagrama.....	30
8 pav.	Panaudos atvejo „Pagal paslaugas“ sekų diagrama.....	32
9 pav.	Panaudos atvejo „Pagal pardavimų kanalus“ sekų diagrama.....	33
10 pav.	Panaudos atvejo „Formuoti ataskaitas“ sekų diagrama.....	34
11 pav.	Žvaigždės tipo schema.....	39
12 pav.	Snaigės tipo schema.....	40
13 pav.	Kuriamos sistemos OLAP duomenų schema.....	41
14 pav.	Verslo analitikos sistemos architektūra.....	42
15 pav.	DTS Designer langas.....	44
16 pav.	Vienos lentelės duomenų atnaujinimo DTS paketas.....	45
17 pav.	Įrašomas įrašas į veiksmų fiksavimo lentą apie darbų pradžia.....	45

18 pav.	DB SQL dialogo langas laikinos lentelės projektui.....	46
19 pav.	Duomenų kopijavimo ir transformavimo žingsniai.....	46
20 pav.	DB SQL dialogo langas laikinos lentelės pervardinimo į naują lentelę projektui...	47
21 pav.	Duomenų atnaujinimo proceso užbaigimo fiksavimas, įrašomas į veiksmų fiksavimo lentelę.....	48
22 pav.	DTS paleidimas per SQL SERVER Agent job.....	48
23 pav.	SQL SERVER Agent job pranešimų konfigūravimo langas.....	49
24 pav.	SQL SERVER Agent užduoties vykdymo ir jos parametrų parinkimo langas.....	50
25 pav.	Kubo formavimo procesas.....	51
26 pav.	Analizės kubo ir dimensijų atnaujinimo užduotys.....	52
27 pav.	Duomenų analizės sistemos formavimo modelis.....	53
28 pav.	Analysis Services įrankio kubų architektūros pasirinkimo vedlys.....	55
29 pav.	Žvaigždės tipo duomenų schema.....	56
30 pav.	Snaigės tipo duomenų schema.....	57
31 pav.	MOLAP architektūros charakteristikos, kai naudojama „žvaigždės“ tipo duomenų schema.....	58
32 pav.	MOLAP architektūros charakteristikos, kai naudojama „Snaigės“ tipo duomenų schema.....	58
33 pav.	ROLAP architektūros charakteristikos, kai naudojama „žvaigždės“ tipo duomenų schema.....	60
34 pav.	ROLAP architektūros charakteristikos, kai naudojama „snaigės“ tipo duomenų schema.....	60
35 pav.	HOLAP architektūros charakteristikos, kai naudojama „snaigės“ tipo duomenų schema.....	62

IVADAS

Šiandieninių įmonių veikla jau nebeįsivaizduojama be kompiuterizuotos apskaitos. Jau bene kiekvienoje, net ir mažiausioje įmonėje rasime bent vieną kompiuterį, kuriame vedama įmonės buhalterinė apskaita, finansai, rengiamos ataskaitos. Per pastaruosius metus įmonėse sparčiai išaugo prieinamos informacijos kiekiai, kompanijos kiekvieną dieną sukaupia milžiniškus duomenų kiekius. Tačiau rinka diktuoja savas sąlygas ir tempą. Atsižvelgiant į tai, kad kompanijos susiduria su didėjančiais reikalavimais aktyviai bei sparčiai reaguoti į rinkos jėgas ir tendencijas, duomenų panaudojimo efektyvumas tampa labai svarbiu veiksniumi. Visiems organizacijos darbuotojams reikia priėjimo prie informacijos, kuri padėtų jiems priimti pagrįstus sprendimus, imtis koreguojančių veiklos valdymo priemonių bei patobulinti verslo procesus. Vadovai bei kiti specialistai neturėtų būti priklausomi nuo IT išteklių ir specialistų sugebėjimų padėti. Informacija turi būti prieinama visoje kompanijoje tuo metu, kai tik jos prireikia, užtikrinant, kad kiekvienas turėtų jam reikalingus duomenis.

Verslo pasaulyje yra labai svarbu kuo greičiau sugebėti priimti reikalingus sprendimus, gauti norimą informaciją, rasti atsakymus į iškilusius klausimus. Yra svarbu apie savo verslą gauti visą informaciją – sekti klientų įpročius, jų poreikius, analizuoti pardavimų, klientų informaciją. Dažnu atveju įvairių ir didelių duomenų analizė tampa ypatingai sudėtingu procesu, kuris reikalauja daug laiko, resursų ir piniginių sąnaudų. Tačiau pasinaudojant verslo išvalgos (angl. Business intelligence) įrankiais galima sukurti sistemą, kuri leistų apjungti duomenis iš įvairių šaltinių į vieningą platformą – duomenų saugyklą. Pasinaudojus duomenų analizei skirtais įrankiais būtų galima kurti įvairias ataskaitas bei duomenis analizuoti įvairiais pjūviais.

Ideali situacija yra tada kai kiekvienas darbuotojas gali matyti sprendimą visos organizacijos mastu ir kiekvienas sprendimas dėsningas tikslams ir strategijoms. Verslo analitikos sistemos yra vienas iš elementų išplėsti prieinamumą ir kompiuterizuotų sistemų funkcijas, kurios palaiko sprendimų priėmimo procesus. Vykstantys pokyčiai technologijų srityje remiasi daugiausia duomenų saugyklomis (angl. Data Warehouse), analitinėmis sistemomis ir sprendimais naudojamais duomenims vizualizuoti. Platesnis naudojimas intelektualių verslo sistemų atneša skirtingą požiūrį į problemas susijusias su pasiektais sprendimais, ieškoma naujų sprendimų, naujų technologijų. Verslo analitikos įrankiai padeda analizuoti tendencijas, prognozuoti pokyčius. Jie

gali atlikti įvairių padalinių, produktų ar klientų analizę. Ši analizė padeda pamatyti koks yra organizacijos išsivystymo lygis, būklė ir kryptis, kur link organizacija eina. Didelės apimties ataskaitos, didelis duomenų kiekis padeda priimti sprendimus, tačiau sulėtina visą procesą. Verslo analitikos sistemos turi padėti rasti geriausią sprendimą. Pasieltas sprendimas turi būti nedviprasmiškas, pateiktas reikiamu laiku ir būtinai remiantis surinkta informacija. Verslo analitikos sistemos palaiko verslo procesus reikiamais momentais, kad sumažinti organizacijos riziką operatyviniu ir strateginiu lygiu.

Kuriant ir projektuojant verslo analitikos sistemas, duomenų saugyklas, duomenų analizės įrankius susiduriama su įvairiomis problemomis. Šiame darbe analizuojamos problemos, su kuriomis susiduriama, kuriant ir projektuojant duomenų analizei skirtus įrankius. Ieškomas tinkamiausias sprendimas, kuris padėtų kurti efektyvius įmonės klientų analizės modelius.

Įmonėms kasdien susiduriant su vis didėjančia konkurencija, tokie įmonės uždaviniai kaip marketingo, pardavimų valdymas, klientų aptarnavimas ir palaikymas iškelia tokias pagrindines problemas:

- pagerinti pardavimų darbuotojų sugebėjimus taip, kad tai leistų padidinti pardavimų apimtį;
- efektyviai pritraukti daugiau klientų;
- greitai ir efektyviai nukreipti produktus ir paslaugas, kuo efektyviau panaudojant turimas lėšas;
- mažinti pelningų klientų praradimo rizikos veiksnius;
- pagerinti klientų aptarnavimo rodiklius taip, kad klientai gautų veiksmingą, greitą ir precizišką aptarnavimą, taip padidinant klientų lojalumą.

Tyrimo objektas

Priemonės ir būdai įgyvendinti efektyvų pardavimų ir klientų duomenų analizės modelį, minimaliai įtakojant pagrindinių sistemų darbą.

Mokslinio tyrimo tikslas

Pateikti ir įgyvendinti duomenų saugyklos analizės priemonių taikymo modelį, kuris leistų užtikrinti pagrindinius įmonės veiklos procesus.

Mokslinio tyrimo uždaviniai:

1. Išnagrinėti naujausią mokslinę literatūrą susisteminant rodiklius darančius įtaką verslui.
2. Atlikti būsimų, kuriamos sistemos vartotojų poreikių analizę.
3. Suprojektuoti sistemos modelį, kuris atitiktų vartotojų lūkesčius.
4. Suprojektuoti ETL (angl. Extract, Transform, Load) procesus duomenų paėmimui iš darbinės sistemos, parinkti optimalias duomenų saugyklos schemas kuriamai sistemai.
5. Sukurti vartotojų reikalavimus atitinkantį įmonės klientų duomenų analizės kubą (angl. OLAP).
6. Atlikti OLAP architektūrinių sprendimų lyginamąją analizę, įvertinant geriausius verslo sprendimus priklausomus nuo kintamų parametrų.

Tyrimo uždavinių įgyvendinimo metodai:

- *mokslinės literatūros analizė* – atliekama siekiant atskleisti ir susisteminti pagrindinius rodiklius, darančius įtaką verslui;
- *lyginamoji analizė* – taikoma siekiant išanalizuoti OLAP architektūros skirtumus, privalumus ir trūkumus.

Darbo struktūra

Darbą sudaro įvadas, trys skyriai, išvados, literatūros sąrašas, santrauka lietuvių ir anglų kalbomis.

Pirmame skyriuje analizuojami verslo analitikos sistemų, OLAP kubų, duomenų sandėlio apibrėžimai. Nagrinėjama mokslinė literatūra apie verslo analitikos sistemų struktūrą ir OLAP kubo panaudojimo galimybes formuojant kubą iš įmonės turimų duomenų. Naudojantis moksline literatūra, išskiriamos pagrindinės BI sistemų raidos kryptys.

Antrame skyriuje pateikiamas siūlomo duomenų analizės sprendimo įgyvendinimo modelis. Aprašomas OLAP duomenų kubo analizės sprendimo ir įgyvendinimo projekto tikslas, nustatomi reikalavimai duomenų analizės sistemai. Įdiegiama siūloma duomenų analizės sistema su realiais įmonės duomenimis, detalizuojant kiekvieną sistemos kūrimo proceso žingsnį. Pateikiamos trumpos išvados.

Trečiame skyriuje pateikiamas siūlomo sprendimo skirtingų architektūrų lyginamoji analizė, kuri priklauso nuo kintamų parametrų. Kintami parametrai tai įmonės turimų duomenų kiekis ir struktūra. Toliau analizuojamos iškilusios problemos, ateities verslo vizija įmonėje, pateikiamos darbo išvados.

1. VERSLO ANALITIKOS SISTEMŲ ANALIZĖ IR APŽVALGA

Vienas iš svarbiausių veiklos procesų, verslo įmonėse yra klientų valdymas. Klientų valdymas – tai veiklos sritis, skirta supaprastinti santykius tarp įmonių ir klientų ir palengvinti su klientais susijusių rinkodaros, pardavimo sričių veiklas. Visa tai sudaro trys pagrindinės taisyklės:

1. Gerinti ir plėsti pardavimo kanalus.

Įmonės gali naudoti tiek tradicinius, tiek elektroninius (pavyzdžiui internetą) pardavimo kanalus, siekiant pritraukti daugiau klientų, naudotis įmonės produktais ar paslaugomis, padarant produktus ir paslaugas pasiekiamus virtualiai bet kada ir bet kur. Taip įmonės gali labiau patenkinti klientų poreikius, didinti jų pasitenkinimą bei ilgiau išlaikyti juos įmonės klientais.

2. Greitas, tikslus ir patikimas bendravimas su klientais.

Jei įmonė yra pilnai pajėgi greitai, tiksliai ir patikimai dirbti su klientų poreikiais susijusiose srityse, tokiose, kaip marketingas, pardavimai, paslaugos, lojalumo programos ir t.t., tuomet įmonė tampa į klientus orientuota įmone. Pasitelkiant verslo procesų kontrolę, įtraukiant sukauptą klientų, pardavimų ir paslaugų informaciją, įmonės derina savo veiklą, kad būtų patenkinti klientų poreikiai.

3. Gerinti santykius su klientais, pasitelkiant verslo įžvalgos sistemas.

Įmonės bendradarbiavimo su klientais metu, sukaupti dideli kiekiai vertingos informacijos, kurią galima apdoroti, analizuoti. Taip galima gauti vertingos informacijos, kuri gali būti pagrindu įmonės veiklos planavimui ir veiklos sprendimų priėmimui. Elektroninė komercija šiuo metu ypač plėtojasi ir vystosi, tapdama nauju verslo veiklos modeliavimo ir vykdymo instrumentu. Įmonės elektroninės komercijos sistemose, prekyba ir klientai yra pagrindas, o tvirti, tvarūs ir ilgalaikiai santykiai su klientais yra labai svarbus kriterijus sėkmingam įmonės vystymuisi. Dėl visu šių priežasčių, į klientus orientuotos, klientų valdymo sistemos pamažu tampa vienomis pagrindinių įmonių elektroninės komercijos sistemomis. Dažniausiai, įmonių klientų valdymo sistemos susideda iš keturių pagrindinių dalių – tai pardavimų, marketingo, produktų platinimo ir paslaugų valdymo. Daugumoje vakarų šalių įmonių, klientų valdymo sistemos padėjo įmonėms gauti naudą, pritraukiant naujų klientų, išlaikant esamus, taupant laiką, padidinant darbo efektyvumą bei didinant pelną. Tai ypač pastebima tose įmonėse, kurių veikla labiau susijusi su tiesioginiais

santykiais su klientais (pavyzdžiui, bankai, draudimo bendrovės, nekilnojamojo turto agentūros, telekomunikacijų ir ryšių įmonės, aviakompanijos, transporto įmonės, saugos tarnybos).

1.1. VERSLO ANALITIKOS SISTEMOS

Atsiradusios pigios duomenų saugojimo technologijos ir platus prieinamumas prie interneto suteikė galimybę organizacijoms pasiekti ir kaupti didelius kiekius duomenų. Duomenys dažniausiai būna įvairiarūšiai, kurie susideda iš komercinių, finansinių ir administracinių transakcijų. Taip pat duomenys gali būti iš elektroninių laiškų, tekstų, didelių tekstų ir kita. Šių duomenų prieinamumas atvėrė naujas perspektyvas ir galimybes, taip pat iškelė svarbų klausimą: ar yra įmanoma konvertuoti duomenis į informaciją ir žinias, kurios padėtų pagerinti organizacijų valdymą (Carlo Verzellis, 2009)?

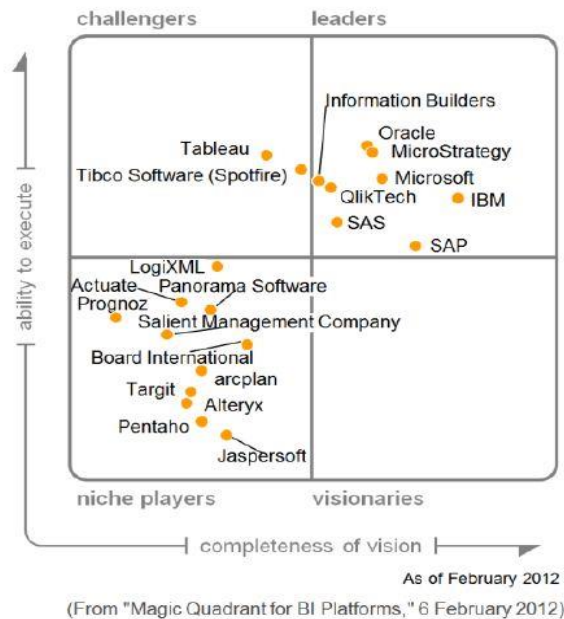
Visa tai susiję su terminu verslo analitika arba verslo įžvalgos. Kaip teigiama straipsnyje „Business Intelligence: Data mining and optimization for decision making“ (Carlo Verzellis, 2009) verslo analitika gali būti apibūdinama kaip kompletas matematinių modelių ir analizės metodologijų, kurios naudojamos prieinamų duomenų generavimui į žinias ir informaciją naudojamus sprendimo priėmimo procese. Taigi, verslo analitikos sistemas galima apibūdinti kaip šių savybių kombinaciją:

- duomenų rinkimas;
- duomenų saugojimas;
- žinių valdymas.

Skirtinguose šaltiniuose galima rasti skirtingus verslo analitikos sistemų apibrėžimus. Verslo analitikos sistemos (angl. *Business Intelligence* arba *BI*) sąvoka pirma kartą buvo paminėta 1958 m. Hans Peter Luhn IBM žurnale publikuotame straipsnyje „A Business Intelligence System“ (Business..., 2013). 1989 m. Gartner Group (IT tyrimų ir konsultavimo įmonė, Konektikutas, JAV) mokslininko Howard Dresner straipsnyje buvo pateiktas terminas „Business Intelligence“ su rinkiniu metodų ir sąvokų. Jų tikslas buvo pagerinti verslo sprendimų priėmimo procesus naudojant įvairius duomenų išteklius. Nuo to laiko verslo analitikos sistemos plėtojasi siūlydamos vis naujesnius ir įvairiau pritaikomus įrankius verslui.

Verslo analitikos sistemos - tai taikomosios programos bei technologijos, naudojamos duomenų rinkimui, saugojimui, analizei bei priėjimui prie duomenų, kurie galėtų padėti organizacijoms greičiau priimti verslo sprendimus. BI apima veiklos sprendimų priėmimo sistemas, užklausų ir ataskaitų kūrimo internetu, analitinio apdorojimo (angl. *OLAP sistemos*), statistinės analizės, prognozavimo ir duomenų gavybos (angl. *data mining*) programas (Rossetti L., 2009).

UAB Softera Baltic tinklapyje ši sąvoka apibrėžiama taip „tai rinkinys metodikų, programinės įrangos ir technologijų skirtų duomenų iš skirtingų sistemų surinkimui, apjungimui ir pateikimui konkrečiam darbuotojui priimtinausiu būdu (ataskaitomis, įspėjimais, grafikais, tendencijomis, įžvalgomis) ir tinkamu metu. Verslo analitikos diegimas, tai jūsų veiklos informacijos sutvarkymas, valdymas, optimizavimas” (Kas yra verslo..., 2013). Tokia forma pateikta informacija dažnai naudojama, kai įmonė ketina priimti svarbų sprendimą ar tiesiog nori stebėti verslo pokyčius.



Šaltinis: (James Richardson, Gartner Symposium, 2012).

1 pav. verslo analitikos platformų kūrėjų populiarumas pagal GARTNER

Pasaulyje yra žinoma ne viena kompanija, kuri užsiima verslo analitinių sistemų bei platformų kūrimu ir platinimu. Tarp garsiausių kompanijų yra „Microsoft“, „SAS“, „IBM Cognos“, „WebFocus“, „MicroStrategy“, „Oracle“, „QlikTech“, „SAP“ ir daug kitų. Kompanijų

populiarumą ir aktualumą labai gerai parodo GARTER kompanijos atliktų tyrimų rezultatai. Buvo apklausta 1364 įvairiausi specialistai iš skirtingų kompanijų, kokius verslo analitikos sprendimus naudoja. Verslo analitikos platformų „magiškajame kvadrate“ (angl. Magic Quadrant) lyderio pozicijų neužleidžia „Oracle“, „Microsoft“, „MicroStrategy“ ir „IBM“ kompanijos (žr. 1 pav.) (Richardson, Symposium, 2012).

Informaciją apie verslo analitikos sistemas (angl. *Business Intelligence* arba *BI*) galima pastebėti ieškant internete, knygose, duomenų bazėse. Daugybėje įvairių šaltinių galima rasti informacijos apie BI svarbą, teikiamus privalumus. R.M. Bogza ir Dorin Zaharie straipsnyje „*Business intelligence as a competitive differentiator*“ (Bozga R.M, 2008) yra aiškinama, kokių konkurencinių pranašumų įmonei gali suteikti naudojamos BI technologijos. Straipsnio autoriai BI apibrėžia, kaip reikiamos informacijos pateikimą reikiamiems įmonės padaliniam reikiamu laiku. Teigiama, jog BICC (angl. *Business Intelligence Competency Center*) atskleidžia daugybę konkurencinių privalumų verčiant neapdorotus duomenis į žinias, kurios teigiamai veikia veiklos strategijas. BICC keliami iššūkiai straipsnyje skirstomi į 6 sritis:

1. Duomenų apdorojimo iššūkiai (greitas didelių ir nesutvarkytų duomenų apdorojimas, skirtingų formatų duomenų transformavimas į vieną formatą);
2. technologiniai iššūkiai (BI sistema turi į vieną visumą susieti visą įmonės vertės grandinėje atsirandančią informaciją, o tam dažnai prireikia itin sudėtingų technologinių sprendimų);
3. procesų iššūkiai (BI yra procesas, o visi įmonėse vykstantys procesai yra kintantys arba pasikartojantys. Bet kuriuo atveju, sėkmingos veiklos paslaptis yra įmonės. Kuo įmonėje vykstantys procesai bus artimesni įmonėms, aiškūs ir suprantami, tuo geresni bus veiklos rezultatai);
4. strateginiai iššūkiai (BI padeda pasirinkti tinkamą veiklos strategiją, tačiau dažnai atskiriems įmonės padaliniam reikalingos skirtingos strategijos. Taigi pagrindinis BI strateginis iššūkis yra kaip tinkamai pristatyti reikiamą informaciją reikiamiems įmonės padaliniam);
5. vartotojo iššūkiai (BI organizacijoje turi pateikti vartotojui tokią informaciją, kuri atitiktų jo pageidavimus, jo išprusimą, jo užimamas pareigas, bei padalinį kuriame jis dirba);
6. kultūriniai iššūkiai (BI pateikiama informacija turi atitikti įmonės vidinę kultūrą, jos veiklos standartus ir netrukdyti įprastinės jos veiklos).

Straipsnyje grafiškai pavaizduotos sritys, kurias apima bei kurioms naudą teikia BI. Teigiama, jog BICC įmonėje turėtų būti tam tikras centras apimantis skirtingas įgūdžių sritis (IT, analitinė veikla ir verslumas). Šio centro tikslas yra geriau suprasti veiklą, interpretuoti dabartinius veiklos rezultatus bei kuo tiksliau numatyti ateities perspektyvas (Bozga R.M, 2008).

Nors verslo analitikos sistemų panaudojimas paplitęs visame pasaulyje, taip pat ir Lietuvoje, tarp daugelio įmonių, tačiau nemaža dalis verslo įmonių susiduria su tam tikrais barjeriais ir sunkumais, žengdamos pirmus žingsnius link verslo analitikos sistemų panaudojimo. Theodoros Evgeniou ir Phillip Cartwright savo straipsnyje „*Barriers to Information Management*“ (Evgeniou, Cartwright, 2005) išskyrė pagrindinius barjerus, kodėl įmonėms yra sunku sklandžiai pradėti naudoti BI sistemas. Dažniausia pasitaikantys sunkumai nėra susiję su technologijomis, sunkiausia organizacijoms pakeisti savo kultūrą, atlikti struktūrinius ir infrastruktūros pokyčius. Dažnai verslo analitikos sistemos priverčia organizacijas keisti savo struktūras, o tai sudaro papildomų rūpesčių. Kitas svarbus aspektas yra įgūdžiai ir kompetencija. Dažnai įmonės neturi tinkamų specialistų arba jiems trūksta profesionalumo. Šiuo atveju labai svarbu išsiaiškinti ar darbuotojai, kuriems patikima BI sistemos diegimas ir priežiūra, išties yra kompetetingi. Kaip ir visuose projektuose svarbus vadovybės palaikymas, kitu atveju beveik visada projektas tampa nesėkmingu.

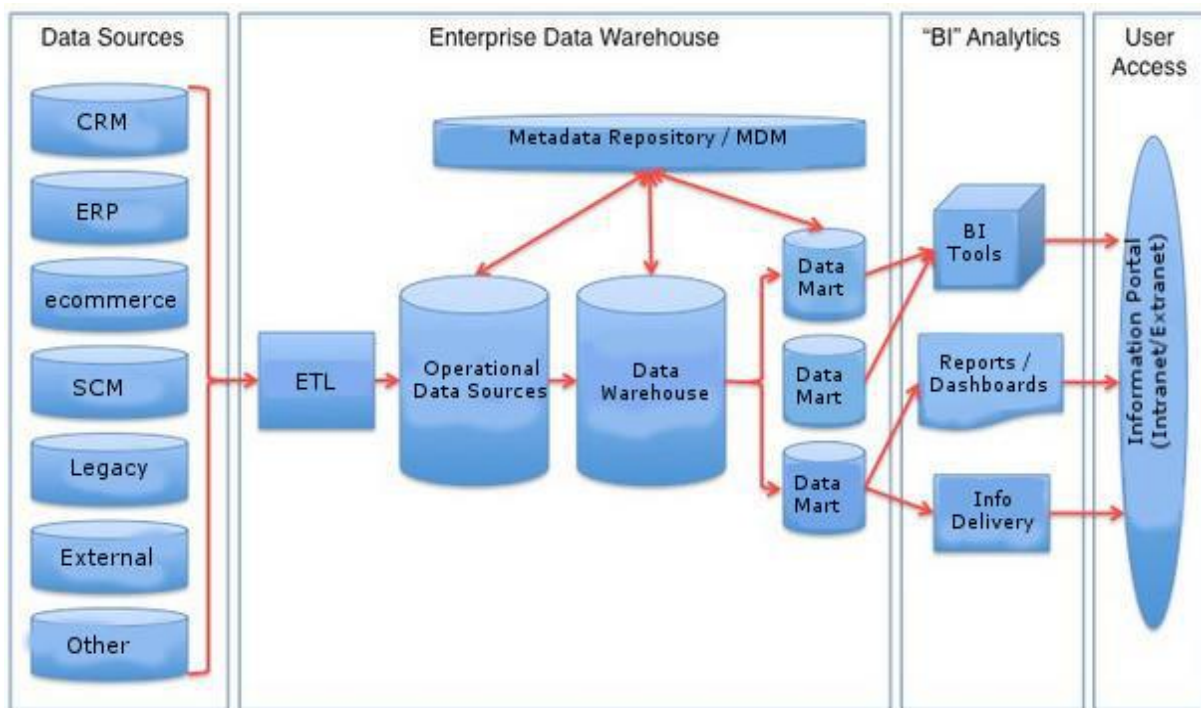
1.2. DUOMENŲ SAUGYKLOS SAŲOKA

Terminą „duomenų saugykla“ (angl. *Data Warehous* sutr. DWH) 1990 metais įvedė B. Inmon. Jis duomenų saugyklą apibrėžė: „Duomenų saugykla yra į veiklos sritis orientuotų, integruotų ir nekintančių, turinčių laiko matą duomenų rinkinys, naudojamas sprendimų priėmimo procese“. B. Inmon įvardino šiuos duomenų saugykloje saugomų duomenų požymius:

- **Į veiklos objektus orientuoti** : duomenys suteikiantys informacijos apie tam tikrą sistemos (įmonės, organizacijos) objektą, o ne apie sistemos vykdomas operacijas;
- **Integruoti** : duomenys surenkami iš įvairių šaltinių į vieną saugyklą, taip sudarydami duomenų visumą, kurie yra susiję tarpusavyje;
- **Turintys laiko matą** : saugykloje saugomos kintančių laike duomenų apibendrintos reikšmės, todėl duomenys turi laiko identifikatorių (pavyzdžiui periodą);

- **Nekintantys** : duomenys saugykloje nekinta. Gali būti įdedama naujų duomenų, tačiau saugomų duomenų faktai nemodifikuojami. Išlaikoma duomenų istorija.

Nors šis apibrėžimas paskelbtas maždaug prieš 20 metų, jis ir dabar gan tiksliai charakterizuoja duomenų saugyklą (žr. 2 pav.).



Šaltinis: sudaryta autoriaus pagal (BI/DW Architecture..., 2011).

2 pav. Duomenų saugyklos komponentai

Duomenų importavimas į duomenų saugyklą (angl. *Data Warehousing*) tai procesas reikalingas duomenų saugyklos sukūrimui. Jis susideda iš saugyklos kūrimo, pildymo ir užklausų vykdymo. Procese galima išskirti keletą žingsnių:

- **Duomenų šaltinio identifikavimas.** Norint sukurti duomenų saugyklą reikia turėti tam tinkamus duomenis. Dažniausiai imami kasdien kaupiami ir naudojami duomenys kurie yra verslo valdymo bei kitose informacinėse sistemose. Tokių duomenų išgavimas kartais gali būti labai sudėtingas ir reikalaujantis daug resursų.
- **Saugyklos projektavimas ir kūrimas.** Tai procesas, kurio metu kuriama saugykla, didžiausias dėmesys kreipiamas į tai kokios užklausos saugykloje bus vykdomos. Tam, kad etapas būtų sėkmingas reikalingas kuriamos duomenų struktūros supratimas ir nuolatinis bendravimas su galutiniu sistemos vartotoju. Dažniausiai šis žingsnis atliekamas iteracijomis.

Jis turi būti atliekamas itin kruopščiai. Vieną kartą sukūrus duomenų modelį ir jį užpildžius dideliais duomenų kiekiais vėliau būna labai sunku, o kartais ir neįmanoma tą modelį keisti.

- **Duomenų importavimas.** Tai duomenų perkėlimo procesas iš pradinio duomenų šaltinio į duomenų saugyklą. Perkeliant duomenis daromi tarpiniai žingsniai, kurie duomenis apdoroja, transformuoja į tinkamą formatą. Taip pat pritaikomos tam tikros verslo taisyklės, kurios leidžia duomenis vienareikšmiškai interpretuoti ir suprasti. Dažniausiai šie žingsniai yra brangiausi ir ilgiausiai trunkantys. Naudojamos taip vadinamos ETL (Extract/Transform/Load) (Išgaut/Tranformuok/Įdėk) programinės priemonės.
- **Duomenų pasikeitimų sekimas.** Periodinis saugyklos atnaujinimas duomenimis iš operacinės aplinkos. Problemos kyla sekant kuriuos duomenis reikia atnaujinti. Ne visose komercinėse sistemose ši problema sėkmingai išspręsta.
- **Duomenų valymas, auditavimas (cleaning).** Tai procesas vykdomas kartu su saugyklos užpildymu. Jo metu stengiamasi panaikinti neteisingus ar netikslus duomenis. Pavyzdžiui duomenys apie tą patį subjektą gauti iš skirtingų šaltinių gali sintaksiškai nesisieti arba dubliuotis, todėl šio proceso metu tokie įrašai yra tiesiog sutvarkomi ir išvalomi.
- **Duomenų agregavimas.** Saugykloje gali būti saugomi skirtingo detalumo duomenys. Labai retai prireikia operacinio lygio istorinių duomenų, todėl duomenys yra agreguojami. Taip taupoma vieta saugykloje. Kai kurios sumos gali būti iš anksto paskaičiuotos. Tada užklauskos su šiais duomenimis bus atliekamos žymiai greičiau.

Didelę reikšmę duomenų saugyklų kūrime turi metaduomenų apie operacinę (apskaitos) duomenų bazę ir duomenų saugyklą kaupimas. Metaduomenys – tai duomenys apie duomenis. Juose sukaupta informacija apie naudojamus duomenų tipus, tam tikrų įrašų fizinę ir loginę vietą, jų prasmę. Metaduomenys saugomi specialioje saugykloje (angl. *repository*). Jie padeda organizacijai sekti kur kokie duomenys kaupiami, keisti jų struktūrą, padeda išsiaiškinti kokios užklauskos duomenų saugykloje gali būti atliekamos.

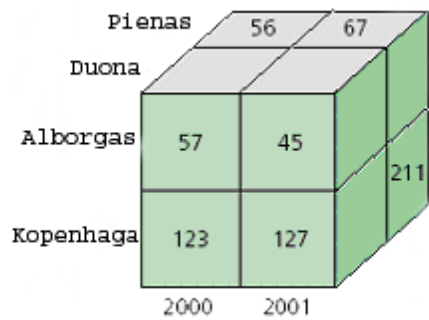
1.3. OLAP DUOMENŲ KUBAI

Codd savo knygoje pateikia tokį kubo apibrėžimą: kubas – tai duomenų struktūra, leidžianti greitai analizuoti duomenis (The Next Generation..., 2013), nes kubas agreguoja faktus kiekviename dimensijos lygmenyje. Kitais žodžiais OLAP apibrėžimas siejamas su OnLine

Analytical Processing sąvoka. OLAP naudojamas kai norima apibūdinti tam tikrus programinius produktus, kurie analizuoja realiu laiku tam tikrą verslo informaciją ir pateikia vartotojui jos apibendrinimą. Kaip galutinė informacija gali būti tam tikros lentelės, grafikai. Taip pat gali būti paprasti skaičiai. Informacijos gavimas priklauso nuo to, kokių pavidalu vartotojas ją nori gauti. OLAP produktų linija pasižymi paprasta ir suprantama vartotojui grafine aplinka. Taip pat sukurta taip, kad vartotojui, net ir nežinančiam kaip rašyti sudėtingą užklausą, būtų lengva gauti reikalingą verslui informaciją. OLAP sistemų duomenys yra saugomi duomenų bazėse kur vienu metu duomenimis gali naudotis ne vienas, o daug vartotojų ir gauti jiems rūpimą informaciją. Tokios programos dažniausiai numato ir įvairaus lygmens saugumo apribojimus vartotojams, turintiems skirtingus priėjimus prie duomenų (Evgeniou, Cartwright, 2005). OLAP sistemos pagrindinės savybės:

- Greita – užklausos rezultatai pateikiami per kelias ar kelias dešimtis sekundžių.
- Daugiamačiai duomenys – duomenys pateikiami hierarchiniu pavidalu kaip daugiamačis vaizdas.
- Analizė – sistema yra suderinama su verslo analitikos ir statistinės analizės sistemomis.

Duomenų kubai apibendrina duomenų peržiūros lenteles ir leidžia peržiūrėti duomenis daugeliu pjūvių (dimensijų). Duomenų kubai yra lengvai valdomi, į juos galima įdėti naujas dimensijas. Teoriškai jų skaičius gali būti neribotas, tačiau praktikoje dirbti su kubais didesniais nei sudarytais iš 15-20 dimensijų yra sudėtinga. Kubo dimensijų skaičius apsprendžia jo ląstelių skaičių. Paveiksle parodytas paprastas 3 dimensijų kubas su geografine, produktų ir laikotarpių informacija skirtingiems pardavimams (žr. 3 pav.). Kubo ląstelėse parodomas vienas matavimas - parduotas paslaugų kiekis. OLAP duomenų kubo užduotis yra pavaizduoti sukauptus duomenis įvairiomis dimensijomis. Taip vartotojui pateikiamas realaus pasaulio vaizdas, kuris jam priimtinas ir suprantamas. OLAP technologija neieško jokių paslėptų ryšių, tiesiog sugeba sukauptus duomenis atvaizduoti dinamiškai, su galimybe detalizuoti duomenis (angl. *drill down*). OLAP kubas puikiai tinka dideliems duomenų kiekiams atvaizduoti.



Šaltinis: sudaryta autoriaus.

3 pav. Trijų dimensijų kubas su pardavimų informacija

OLAP kubuose didelė tarpinių skaičiavimų dalis atliekama dar iki duomenų kubo panaudojimo, o vieną kartą atliktų tarpinių skaičiavimų rezultatais gali pasinaudoti visi prieigos teises turintys šios duomenų bazės vartotojai. Prieigos teises galima valdyti iki tam tikros kubo celės, t.y. dimensijos narių ir skaičiuojamųjų reikšmių. Kadangi OLAP duomenų bazės saugo tarpines agregacines reikšmes, pagal duomenų kubo įgyvendinimo būdą, duomenų pasikeitimai ar naujų duomenų įkėlimas į OLAP duomenų bazę gali pareikalauti atnaujinti ar pertvarkyti kubo informaciją (turi būti perskaičiuotos kai kurios agregacinės reikšmės).

Dimensija – tai tam tikros kategorijos lentelė, pagal kurią peržiūrimi susumuoti duomenys. Dimensijoms būdingos hierarchijos. Kiekviena dimensija saugo informaciją apie tą patį faktą skirtingame kontekste. Tai padeda gauti ar apskaičiuoti tam tikrą reikšmę tam tikrame detalizacijos lygyje. Dimensijos dažnai turi hierarchinę struktūrą. Pavyzdžiui, nagrinėjant tam tikro laikotarpio pardavimus, savaitės dimensijos jungiamos į mėnesių dimensijas, o šios savo ruožtu į metų ir taip toliau. Tam tikras duomenų lygis be jį sudarančių dimensijų gali turėti papildomų savybių, pavyzdžiui prekės matavimo vieneta, kuris tam tikroje dimensijoje nekinta ir nedidina dimensijų skaičiaus.

Faktas – tai lentelė, kurioje saugomi matavimai (faktai), t. y. skaičiuojamosios reikšmės ir išoriniai raktai, turintys ryšį su dimensijų pirminiais raktais. Faktą dažniausiai identifikuoja dimensijų kombinacija. Galutinai faktą identifikuoja dimensijų reikšmių kombinacija. Faktas egzistuoja tik tada, jei yra jį atitinkanti visų dimensijų reikšmių kombinacija. Dauguma daugiamačių duomenų modelių reikalauja, kad faktas turėtų bent jau žemiausio hierarchijos lygio dimensijų reikšmių kombinaciją (Aukštesnio lygio dimensijų reikšmės gali būti gaunamos

agreguojant žemesniųjų lygių reikšmes). Kiekvienas faktas turi savo detalumą. Kurį nulemia jį sudarančių dimensijų detalumas bei vieta dimensijų hierarchijoje. 3 paveikslėlio pavyzdyje fakto detalumas yra: metai pagal miestą ir produktą. Detalumai sudaryti iš aukštesnių arba žemesnių hierarchijų dimensijų yra detalesni (angl. *finer*) arba labiau apibendrinti (angl. *coarser*).

1.4. PROGNOZUOJAMA VERSLO ANALITIKOS SISTEMŲ RAIDA

Verslo analitikos sistemų svarba, kiekvienais metais tampa vis didesnė. Todėl nenuostabu, kad ir BI technologijos vystosi gana sparčiai. Pavyzdžiui, jeigu grįžtume keletu dešimtmečių atgal, niekas turbūt net nežinotų apie tokią duomenų analizės technologiją kaip OLAP kubai. Tačiau ši technologija šiandien jau nieko nestebina. Šiuo metu jau yra naudojamos daug pažangesnės ir naujesnės.

Lachlan James savo straipsnyje „*Top 14 Business Intelligence predictions for 2012*“ (James L., 2012) išskiria pagrindines verslo analitikos sistemų raidos kryptis:

1. Savitarnos analitika (angl. Self-Service "Business Intelligence")

Savitarnos analitika svarbi tuo, kad nereikia iš anksto žinoti ir numatyti kokie klausimai gali kilti verslui, į kuriuos atsakymo ieškome BI sistemose. Vartotojai patys gali kurti įvairius skydus ir ataskaitas paprasčiausias būdais. Savitarnos BI idėja, kad bet kokio verslo vartotojas galėtų analizuoti duomenis, tokiu būdu galėtų priimti geresnius verslo sprendimus (Software Tableau, 2012). Tai sumažina IT personalo įtraukimo poreikį ir iš esmės suteikia vartotojams galimybę būti savo BI sistemų kūrėjais ir savininkais.

Šiame darbe plačiau nagrinėjama būtent savitarnos analitikos sritis, kuriai priklauso ir OLAP duomenų kubai.

2. Verslo analitika „debesyse“ (angl. Cloud Business Intelligence)

"Gartner" ataskaitoje teigiama, kad BI debesyse vystosi gana pastoviu tempu, ir iki 2013 metų ši technologija sudarys 3% BI pajamų. Nors šios prognozės yra su didele perspektyva, manoma, kad verslo analitikos sistemų programinė įranga patirs stipresnę ekonomikos augimą. Debesis sumažina programinės ir aparatūrinės įrangos kaštus. Duomenys ir analizės rezultatai tampa pasiekiami iš bet kur ir bet kada. Šis sprendimas labiausia bus aktualus mažoms ir vidutinėms įmonėms.

„TechTarget“ 2012 metų tyrimas, per kurį apklausta daugiau nei 2600 IT profesionalų ir verslo analitikų, nustatė, kad vis daugiau įmonių planuoja padidinti savo išlaidas paslaugoms „debesyse“ per artimiausius metus (James L. 2012).

3. Didelių duomenų analizė (angl. Big Data Analysis)

Didelių duomenų analizės technologijos ir toliau įgauna didelę svarbą ir vietą rinkoje. Bet kodėl didelių duomenų analizės tema tampa svarbi? Atsakymas labai paprastas. Pirma, didelių duomenų (angl. Big Data) analizės produktai ir paslaugos tapo prieinamos organizacijoms. Tai suteikia galimybę ir gebėjimus surinkti daugiau informacijos nei bet kada anksčiau. Pavyzdžiui naujos kartos BI ir Google Analytics - labai reikšmingas pavyzdys, kaip tokie pokyčiai pakeitė įmonių duomenų analizės metodus ir mąstyseną: "tarkime, jūs turite gamybos įmonę, kuri gamina kūdikių buteliukus. 2001 metais, jūs tikriausiai orientavotės tik į du dalykus : kiek buteliukų pagaminama, kiek pinigų išleidžiame gamybai ir kiek gauname pardavę. Tačiau šiandien, jūs gali stebėti, kiek buteliukų jūs gaminate , kiek pinigų jums reikės padaryti taip, kad būtų patenkinti jūsų klientai, koks procentas nuolatinių klientų, ir tai, ką žmonės sako apie buteliukus socialiniuose tinkluose.

Didelių duomenų analizė leidžia įmonėms galvoti apie duomenų šaltinių plėtrą bei naujų kūrimą. Žinomiausios socialinių platformų terpės sukūrė didžiulį kiekį potencialių klientų duomenų. Tai duomenys, kurie yra prieinami analizei!

Neseniai "Gartner" atliktas tyrimas parodo milžiniškas duomenų augimo tendencijas, nurodoma, kad duomenų, gautų vien tik 2009 m. apimtis buvo didesnė nei ankstesnių 5000 metų kartu. "Gartner" spėja, kad pasaulio įmonių duomenys augs apie 650 procentų iki 2014 metų pabaigos (James L. 2012).

“Tableau“ kompanijos atliktas tyrimas rodo, kad vis dažniau įmonėse priimami sprendimai naudojant didelių ir detalių duomenų analizės rezultatus (Software Tableau, 2012).

4. Verslo įžvalgos mobiliuosiuose įrenginiuose (angl. Mobile Business Intelligence)

BI technologijos mobiliuosiuose įrenginiuose ir toliau auga. Mobilųjų BI augimas priklauso nuo atpigusių mobiliųjų įrenginių ir bevielio ryšio populiarumo augimo. Mobilųjų BI pagerina veiklos ir valdymo pajėgumus, vartotojai gali gauti informaciją per savo iPhone, iPad, planšetinius kompiuterius ir kitus išmaniuosius prietaisus. (Software Tableau, 2012). Verslo įžvalgos mobiliuosiuose įrenginiuose, leidžia operatyviai ir nepriklausomai nuo vietos gauti bet kokią atsakymą, kokį tik gali suteikti BI sistema. Visa tai nusveria saugumo ir integracijos problemas.

Pagal TechTarget atliktą tyrimą prognozuojama, kad beveik 50 procentų įmonių turės mobilius BI sprendimus dar 2013m. Tyrimo rezultatai iš SearchBusinessAnalytics.com teigia kad, 30% iš 249 tyrimo dalyvių teigė, kad jie ketina įdiegti mobilias BI sistemas per ateinančius metus. "Gartner" atliktas tyrimas rodo, kad mobilioji analitika išlieka kaip svarbiausia tendencija ir prioritetas, maždaug iki 2015 metų pabaigos.

5. Socialinė analitika (angl. Social Business Intelligence)

Kartu su socialinių tinklų evoliucija išaugo ir socialinės BI poreikis. Socialinis intelektas įtakoja BI pasaulį, naudojant socialinės žiniasklaidos priemones leidžiama įmonėms įgyti žinių apie įvairias vartotojų grupes. Dauguma kompanijų gan atsargiai žiūri į socialinę ir bendradarbiavimo analitikas. Dažnai tai būna dėl duomenų tikslumo ir painiavos kada duomenys gaunami ne iš pradinio duomenų šaltinio, bet yra bendrinami tarp draugų ir bendraamžių socialiniuose tiksluose. Tačiau tokių socialinių tinklų kaip "Facebook" interaktyvūs įrankiai išsprendžia šias problemas. Pagrindinis organizacijų uždavinys kaip išnaudoti šias galimybes ir apie klientus sužinoti dar daugiau informacijos, kurią vėliau galima bus panaudoti verslo tikslams (James L. 2012).

Kituose šaltiniuose taip pat išskiriamos ir kitos svarbios BI tendencijos, kurios vystosi labai dideliu greičiu: analitika operatyvioje atmintyje (angl. In-Memory Analytics), tekstinių duomenų gavyba (angl. Text Mining), vaizduojamoji analitika (angl. Visual Analytics), nuspėjamoji analitika (angl. Predictive Analytics).

Kaip matome iš visų šių apžvelgtų tyrimų, verslo įžvalgų sistemos yra nuolat tobulėjanti ir besiplečianti sritis. Vis daugiau didelių, vidutinių ir net smulkių įmonių pripažįsta BI sistemų naudą, vis dažniau jas diegia savo kompanijose siekdami efektyvesnės veiklos.

1 SKYRIAUS IŠVADOS

1. Dideli, informacinių sistemų sukaupiami duomenų kiekiai, negali būti efektyviai apdorojami be specialiai tam skirtų priemonių: lentelių transformavimo, perkėlimo į duomenų saugyklą, duomenų transformacijų ir agregavimo, daugiamačių kubų formavimo ir jų peržiūros. Šios priemonės ir aptartos analitinėje dalyje.
2. Didelių duomenų kiekių analizei geriausiai tinka daugiamatis duomenų modelis, kuris skirtingai nei reliaciniai modeliai, yra optimizuotas analitinių užklausų vykdymui. Šio modelio

esmė yra duomenų kubai, sudaryti iš faktų ir dimensijų lentelių bei turintys tam tikrus matavimus.

3. Transformuoti ir agreguoti reliaciniai duomenys saugomi duomenų saugykloje. Analizės dalyje nagrinėtos duomenų saugyklos savybės .
4. Duomenų saugyklos gali būti naudojamos ne tik duomenų analizei, bet ir duomenų, ateinančių iš skirtingų šaltinių, surinkimui bei integracijai.
5. Apžvelgti duomenų saugyklų principai bei OLAP priemonių naudojimas jose. Šios priemonės yra tinkamos ir verslo įmonių pardavimų ir klientų analizei skirtų duomenų analitiniam apdorojimui.
6. Darbe detaliau bus tiriamos OLAP priemonės MSSQL 2012 duomenų bazių valdymo sistemose.

2. PARDAVIMŲ IR KLIENTŲ ANALIZĖS SISTEMOS PROJEKTAS

Vienas svarbiausių duomenų analizės sistemos projekto etapų – tai pačios sistemos projektas. Tik gerai parengus sistemos projektą ir atsižvelgus į vartotojų reikalavimus bei poreikius, galima sukurti naudingą ir efektyvų produktą, kuris teiks naudą vartotojams. Sistemos projektavimas gali būti atliekamas dviem būdais: iš viršaus į apačią arba iš apačios į viršų. Dažniausiai sistemos projekto struktūra yra sudaroma abiem būdais. Valdymo lygmens projektavimo lygiai nustatomi iš viršaus į apačią, o techninio lygmens projektavimo lygiai - iš apačios į viršų. Projekto pagrindinis tikslas išskaidomas iki smulkesnių uždavinių, o po to, numatomi darbai, kurie gali būti lengvai įvertinti kaštų ar laiko rodikliais ir būtini sprendžiant prieš tai išskirtus uždavinius. Rengiant šį projektą, naudojama „iš viršaus žemyn“ metodika – t.y. pradžioje specifikuojamas galutinis rezultatas (norimas duomenų kubas), o toliau detalizuojama, kaip ir iš ko jis bus sudarytas.

2.1. PARDAVIMŲ IR KLIENTŲ ANALIZĖS SISTEMOS PROJEKTO TIKSLAS

Projekto įgyvendinimui buvo pasirinkta žinoma įmonė TEO LT, AB. Įmonė yra didžiausia integruotų telekomunikacijų, IT ir televizijos paslaugų teikėja Lietuvoje. Šiandien Bendrovė užima lyderio pozicijas interneto, duomenų perdavimo bei fiksuotojo telefono ryšio rinkose. Esami Bendrovės pajėgumai leidžia pasiūlyti klientams pačias moderniausias technologijas, nenusileidžiant pažangiausiems pasaulio telekomunikacijų rinkos žaidėjams. TEO LT, AB vizija – būti geriausiu partneriu bendraujant su nuolat kintančiu pasauliu. Pasitelkdama moderniausias technologijas, Bendrovė suteikia savo klientams galimybę pasiekti žmones, žinias ir pramogas (Trumpai apie TEO, 2012). Įmonės tinklas ir paslaugos nuolat sparčiai keičiasi ir plečiasi. Norint pasiūlyti klientams tai, ko jiems reikia, yra svarbu tirti rinką, analizuoti klientų poreikius.

Lietuvoje žmonių kiekis palyginus su kitomis Europos valstybėmis nėra didelis, tačiau telekomunikacinių įmonių teikiančių interneto, televizijos ar balso paslaugas yra labai daug. Todėl pritraukti naujus klientus yra labai sunku, praktiškai rinka jau yra pasidalinta. Visiškai naujų klientų, kurie anksčiau nesinaudojo jokiais paslaugomis, praktiškai nebelikę. Telekomunikacijų

įmonės kovoja tarpusavyje dėl kiekvieno kliento, bandydamos juos pervilioti pas save. Tam reikalinga įvairi klientų, pardavimų analizė, kuri padeda tai įgyvendinti. Kiekviena įmonė kaupia apie klientus įvairius duomenis. Sukaupus pardavimų duomenis galima panaudoti, ir gana efektyviai – išanalizavus pardavimus, galima nustatyti, kokias paslaugas dažniausiai perka klientai, kokie klientai labiau linkę įsigyti vieną ar kitą paslaugą, kokie pardavimų kanalai dirba efektyviausiai. Kur geriausi klientai susitelkę, geografiniu atžvilgiu. Visa ši informacija leidžia efektyviai nustatyti klientų poreikius, pasiūlyti jiems reikalingų paslaugų bei pritraukti daugiau klientų - tuo pačiu ir padidinti pardavimus, pajamas, klientų bazę.

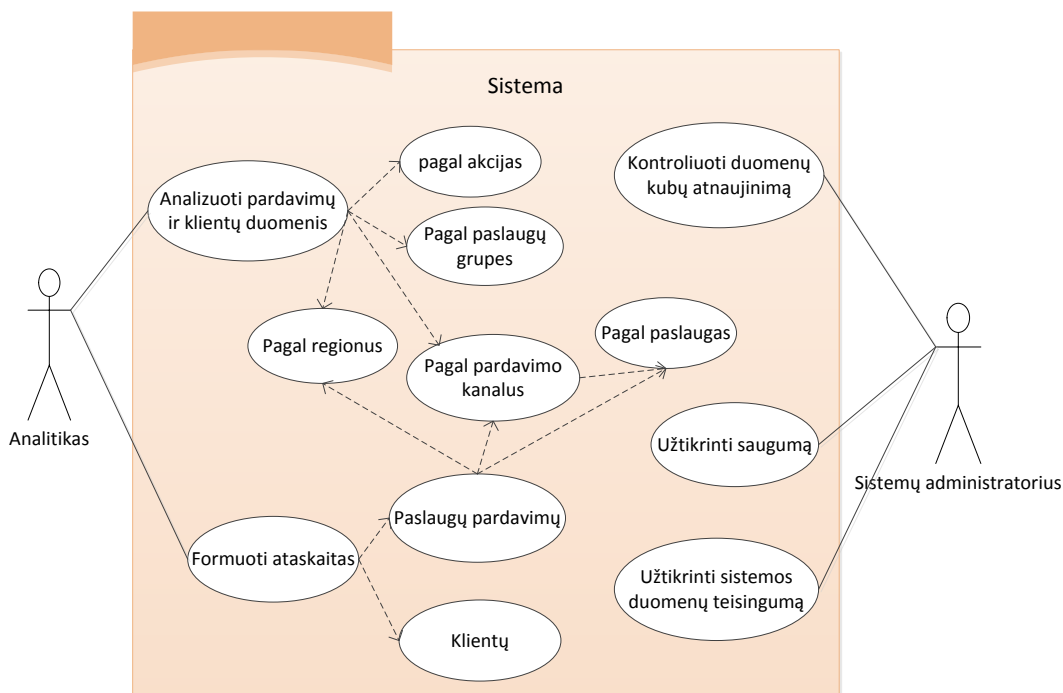
Taigi, atsirado poreikis analizuoti klientų ir paslaugų pardavimų duomenis. Tam buvo nuspręsta sukurti įmonės klientų duomenų analizės sistemą, kuri leistų atlikti įvairių pardavimų ir klientų duomenų analizę. Kadangi turimų klientų bazė yra didelė, per 500,000 tūkstančių klientų, o pardavimų informacijos sukaupiama dideli kiekiai, tai tokio pobūdžio duomenų analizei tinkamiausia yra sukurti duomenų analizės kubą. Jo pagalba būtų galima greitai ir efektyviai gauti norimus duomenis ir rezultatus.

Taigi, nuspręsta integruoti į jau esamą duomenų saugyklą duomenis apie įmonės padalinių pardavimus, klientus ir kitą susijusią informaciją.

2.2. SISTEMOS REIKALAVIMŲ SPECIFIKACIJA

Sistemos panaudos atvejų diagramos ir jų specifikacijos leidžia tiksliai identifikuoti būsimus sistemos vartotojus, jų poreikius bei atliekamus veiksmus. Visiems sistemos vartotojams sudaromos panaudos atvejų diagramos, jų specifikacijos bei tų vartotojų panaudos atvejų sekų, veiklos diagramos.

Sistemos vartotojai – tai įmonės darbuotojai, kurie analizuoja įmonės klientus, paslaugų pardavimus. Analitiko panaudos atvejų diagrama pavaizduota žemiau esančiame paveikslėlyje (žr. 4 pav.).



Šaltinis: sudaryta autoriaus.

4 pav. Sistemos panaudos atvejų diagrama

Sistemos administratoriaus funkcijos – tai prižiūrėti bendrą sistemos darbą, užtikrinti jos saugumą. Administratorius atsakingas už duomenų ir sistemos atnaujinimo procesus, jų korektiškumą.

Analitikas analizuoja sistemos pateikiamus duomenis ir priima strateginius sprendimus. Analitikas, sistemos pagalba, gali analizuoti pardavimų duomenis bei formuoti įvairias ataskaitas. Toliau pateikiamos kiekvieno panaudos atvejo specifikacijos.

- Panaudos atvejis „Analizuoti pardavimų ir klientų duomenis“

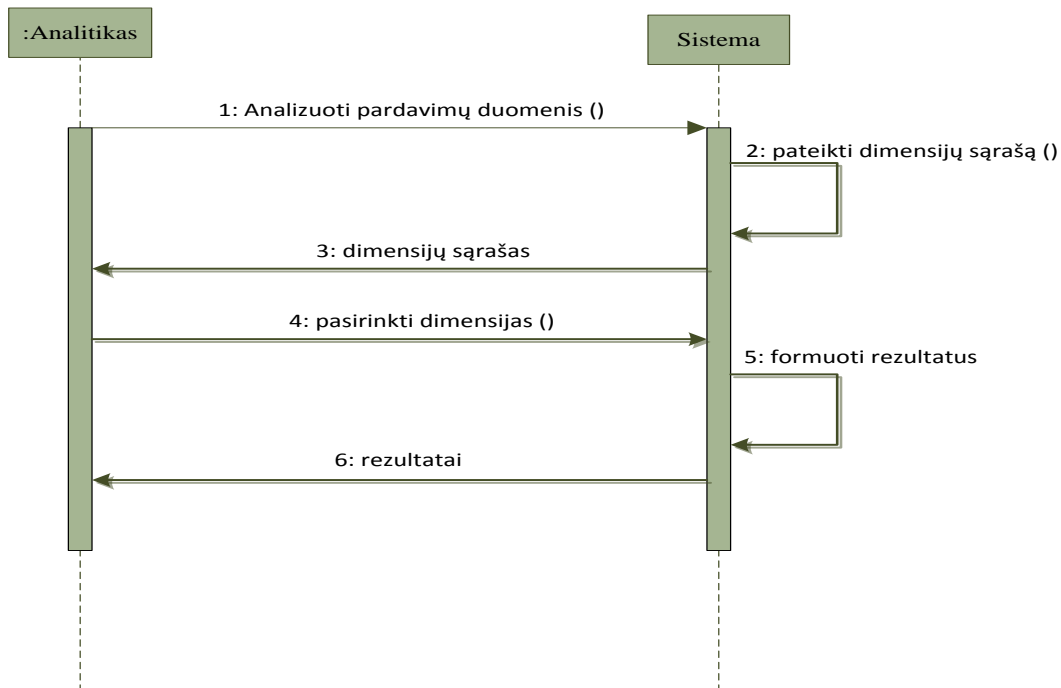
Panaudos atvejis „Analizuoti pardavimų ir klientų duomenis“ apima duomenų analizę įvairiais pjūviais (pardavimo kanalų, geografinio pasiskirstymo, paslaugų grupių, paslaugų, akcijų). Panaudos atvejo specifikacija pateikta žemiau (žr. 1 lentelę).

1 lentelė. Panaudos atvejo „Analizuoti pardavimų ir klientų duomenis“ specifikacija

Panaudos atvejis "Analizuoti pardavimų ir klientų duomenis"	
Aktorius	Analitikas
Prieš sąlyga	

Sužadinimo sąlyga		Vartotojas nori peržiūrėti sukauptus duomenis apie pardavimus ir klientus
Susiję panaudos atvejai	Išplečia PA	
	Apima PA	
	Specializuoja PA	
Pagrindinis įvykių srautas		Sistemos reakcija ir sprendimai
1. Vartotojas pasirenka duomenų analizę		1. Sistema pateikia dimensijų sąrašą
Po sąlyga		Vartotojas peržiūri duomenis norimais pjūviais
Alternatyvūs scenarijai		
1. Nepavyko prisijungti prie duomenų saugyklos		Sistema suformuoja klaidos pranešimą.

Panaudos atvejo „Analizuoti pardavimų ir klientų duomenis“ sekų diagrama pavaizduota žemiau esančiame paveikslėlyje (žr. 5 pav.).



Šaltinis: sudaryta autoriaus.

5 pav. Panaudos atvejo „Analizuoti pardavimų ir klientų duomenis“ sekų diagrama

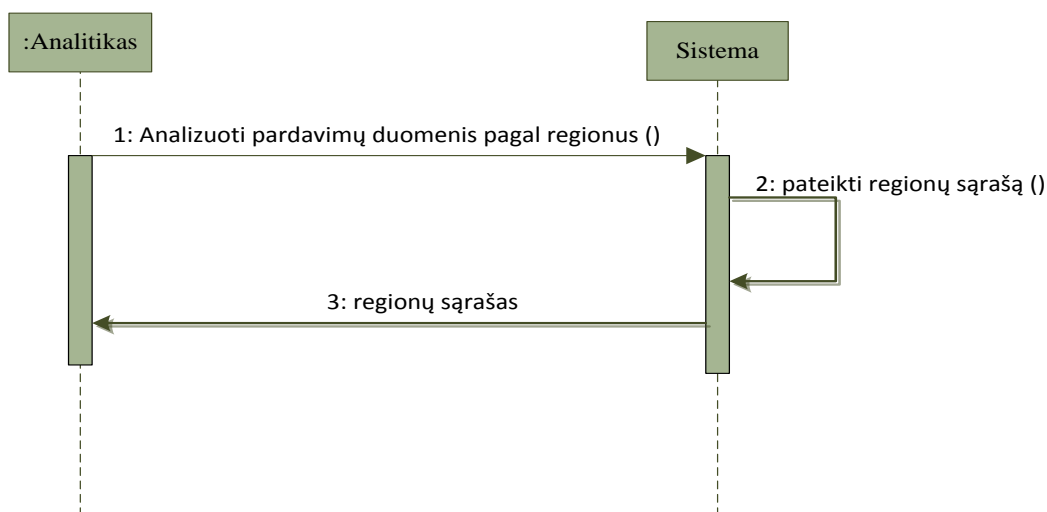
- Panaudos atvejis „Pagal regionus“

Panaudos atvejis „Pagal regionus“ išplečia panaudos atvejį „Analizuoti pardavimų ir klientų duomenis“. Sistemos vartotojui pasirinkus panaudos atvejį „Analizuoti pardavimų ir klientų duomenis“, vartotojas gali pasirinkti, pagal kokius kriterijus tuos duomenis analizuoti. Pasirinkus „Pagal regionus“, sistema suformuoja regionų sąrašą.

2 lentelė. Panaudos atvejo „Pagal regionus“ specifikacija

Panaudos atvejis "Pagal regionus"		
Aktorius		Analitikas
Prieš sąlyga		
Sužadinimo sąlyga		Vartotojas nori peržiūrėti sukauptus duomenis apie pardavimus ir klientus pagal regionus
Susiję panaudos atvejai	Išplečia PA	Analizuoti pardavimų ir klientų duomenis
	Apima PA	
	Specializuoja PA	
Pagrindinis įvykių srautas		Sistemos reakcija ir sprendimai
1. Vartotojas pasirenka duomenų analizę 2. Pasirenkama duomenų analizė pagal regionus		1. Sistema pateikia dimensijų sąrašą 2. Sistema suformuoja regionų sąrašą
Po sąlyga		Vartotojas peržiūri duomenis norimais pjūviais
Alternatyvūs scenarijai		
1. Nepavyko prisijungti prie duomenų saugyklos		Sistema suformuoja klaidos pranešimą.

Panaudos atvejo „Pagal regionus“ sekų diagrama pavaizduota žemiau esančiame paveikslėlyje (žr. 6 pav.).



Šaltinis: sudaryta autoriaus.

6 pav. Panaudos atvejo „Pagal regionus“ sekų diagrama

- Panaudos atvejis „Pagal akcijas“

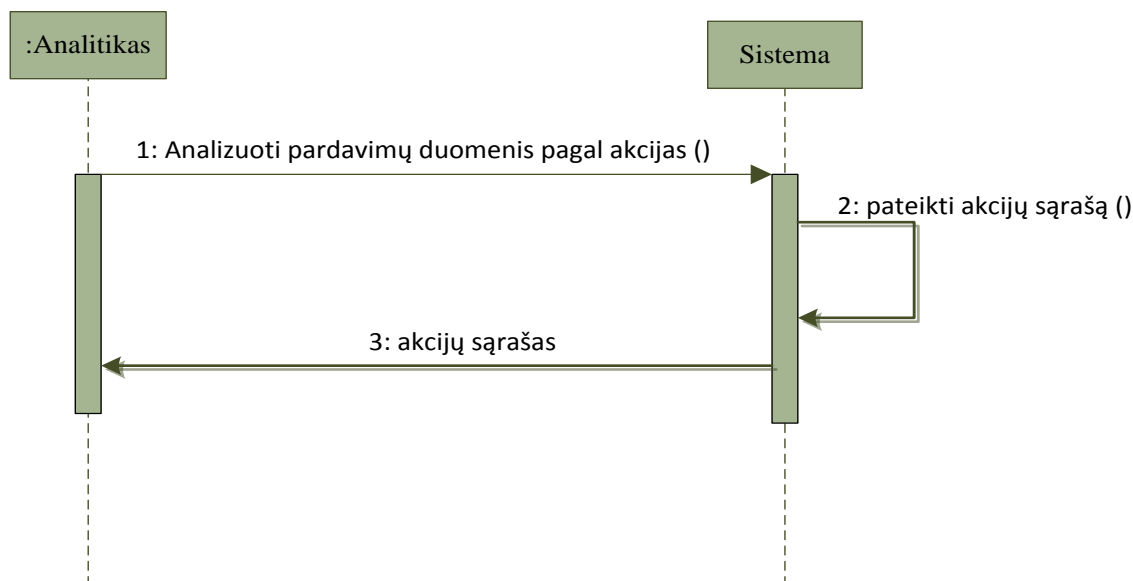
Panaudos atvejis „Pagal akcijas“ išplečia panaudos atvejį „Analizuoti pardavimų ir klientų duomenis“. Sistemos vartotojui pasirinkus panaudos atvejį „Analizuoti pardavimų ir klientų duomenis“, vartotojas gali pasirinkti, pagal kokius kriterijus tuos duomenis analizuoti. Pasirinkus „Pagal akcijas“, sistema suformuoja pardavimų akcijų sąrašą.

3 lentelė. Panaudos atvejo „Pagal akcijas“ specifikacija

Panaudos atvejis "Pagal akcijas"		
Aktorius		Analitikas
Prieš sąlyga		
Sužadinimo sąlyga		Vartotojas nori peržiūrėti sukauptus duomenis apie pardavimus ir klientus pagal pardavimų akcijas
Susiję panaudos atvejai	Išplečia PA	Analizuoti pardavimų ir klientų duomenis
	Apima PA	
	Specializuoja PA	
Pagrindinis įvykių srautas		Sistemos reakcija ir sprendimai

1. Vartotojas pasirenka duomenų analizę 2. Pasirenkama duomenų analizė pagal akcijas	1. Sistema pateikia dimensijų sąrašą 2. Sistema suformuoja akcijų sąrašą
Po sąlyga	Vartotojas peržiūri duomenis norimais pjūviais
Alternatyvūs scenarijai	
1. Nepavyko prisijungti prie duomenų saugyklos	Sistema suformuoja klaidos pranešimą.

Panaudos atvejo „Pagal akcijas“ sekų diagrama pavaizduota žemiau esančiame paveikslėlyje (žr. 7 pav.).



Šaltinis: sudaryta autoriaus.

7 pav. Panaudos atvejo „Pagal akcijas“ sekų diagrama

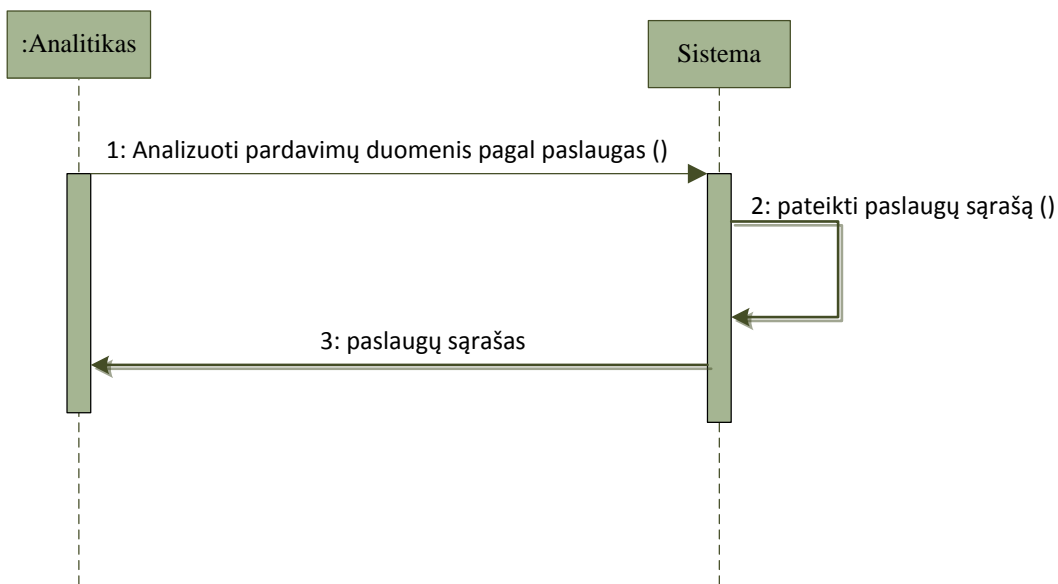
- Panaudos atvejis „Pagal paslaugas“

Panaudos atvejis „Pagal paslaugas“ išplečia panaudos atvejį „Analizuoti pardavimų ir klientų duomenis“. Sistemos vartotojui pasirinkus panaudos atvejį „Analizuoti pardavimų ir klientų duomenis“, vartotojas gali pasirinkti, pagal kokius kriterijus tuos duomenis analizuoti. Pasirinkus „Pagal paslaugas“, sistema suformuoja paslaugų sąrašą, pagal kurias galima analizuoti klientų ir pardavimų duomenis.

4 lentelė. Panaudos atvejo „Pagal paslaugas“ specifikacija

Panaudos atvejis "Pagal paslaugas"		
Aktorius		Analitikas
Prieš sąlyga		
Sužadinimo sąlyga		Vartotojas nori peržiūrėti sukauptus duomenis apie pardavimus ir klientus pagal paslaugas
Susiję panaudos atvejai	Išplečia PA	Analizuoti pardavimų ir klientų duomenis
	Apima PA	
	Specializuoja PA	
Pagrindinis įvykių srautas		Sistemos reakcija ir sprendimai
1. Vartotojas pasirenka duomenų analizę 2. Pasirenkama duomenų analizė pagal paslaugas		1. Sistema pateikia dimensijų sąrašą 2. Sistema suformuoja paslaugų sąrašą
Po sąlyga		Vartotojas peržiūri duomenis norimais pjūviais
Alternatyvūs scenarijai		
1. Nepavyko prisijungti prie duomenų saugyklos		Sistema suformuoja klaidos pranešimą.

Panaudos atvejo „Pagal paslaugas“ sekų diagrama (žr. 8 pav.).



Šaltinis: sudaryta autoriaus.

8 pav. Panaudos atvejo „Pagal paslaugas“ sekų diagrama

- Panaudos atvejis „Pagal pardavimo kanalus“

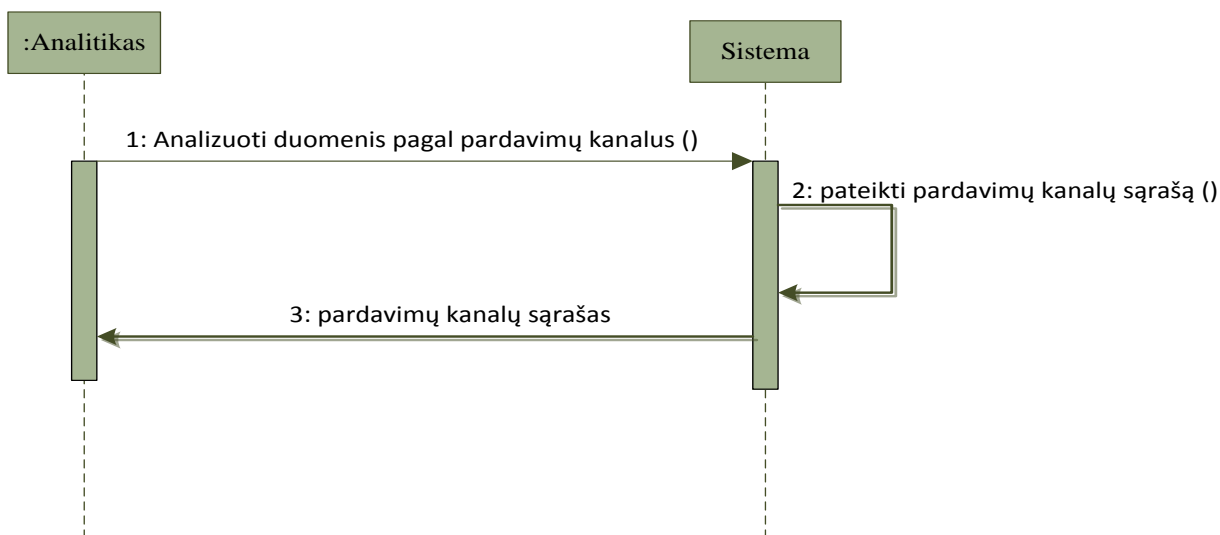
Panaudos atvejis „Pagal pardavimo kanalus“ išplečia panaudos atvejį „Analizuoti pardavimų ir klientų duomenis“. Pasirinkus „Pagal pardavimo kanalus“ panaudos atvejį, sistema suformuoja pardavimo kanalų sąrašą, pagal kuriuos galima analizuoti atskirų pardavimų kanalų pardavimus.

5 lentelė. Panaudos atvejo „Pagal pardavimo kanalus“ specifikacija

Panaudos atvejis "Pagal pardavimo kanalus"		
Aktorius	Analitikas	
Prieš sąlyga		
Sužadinimo sąlyga	Vartotojas nori peržiūrėti sukauptus duomenis apie pardavimus pagal pardavimų kanalus	
Susiję panaudos atvejai	Išplečia PA	Analizuoti pardavimų ir klientų duomenis
	Apima PA	
	Specializuoja PA	
Pagrindinis įvykių srautas		Sistemos reakcija ir sprendimai

1. Vartotojas pasirenka duomenų analizę 2. Pasirenkama duomenų analizė pagal pardavimų kanalus	1. Sistema pateikia dimensijų sąrašą 2. Sistema suformuoja pardavimo kanalų sąrašą
Po sąlyga	Vartotojas peržiūri duomenis norimais pjūviais
Alternatyvūs scenarijai	
1. Nepavyko prisijungti prie duomenų saugyklos	Sistema suformuoja klaidos pranešimą.

Panaudos atvejo „Pagal pardavimų kanalus“ sekų diagrama pavaizduota žemiau esančiame paveikslėlyje (žr. 9 pav.).



Šaltinis: sudaryta autoriaus.

9 pav. Panaudos atvejo „Pagal pardavimų kanalus“ sekų diagrama

- Panaudos atvejis „Formuoti ataskaitas“

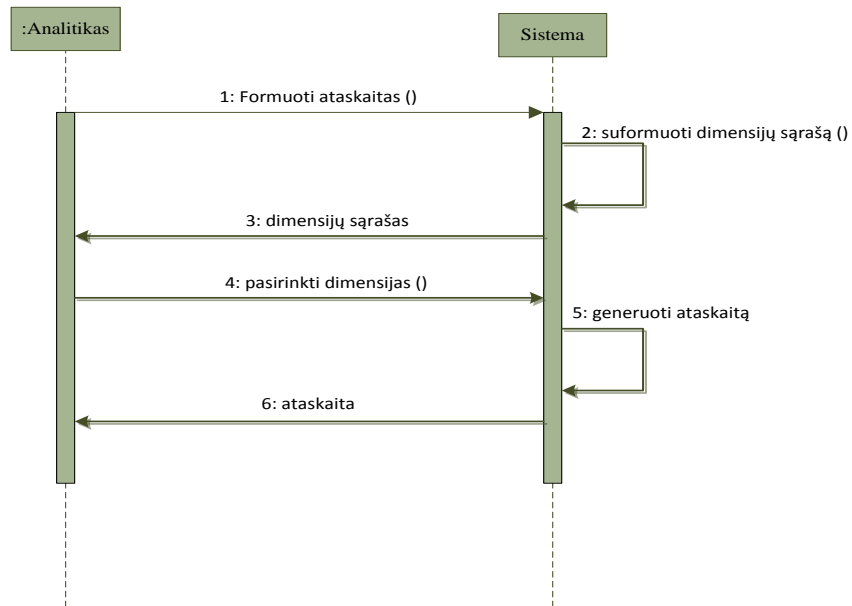
Panaudos atvejis „Formuoti ataskaitas“ apima įvairių analitinių ataskaitų formavimą. Analitikui, analizuojančiam sukauptus pardavimų ir klientų duomenis, yra svarbu sukurti įvairias ataskaitas, kuriose būtų galima matyti duomenis norimais pjūviais, diagramomis.

6 lentelė. Panaudos atvejo „Formuoti ataskaitas“ specifikacija

Panaudos atvejis "Formuoti ataskaitas"

Aktorius		Analitikas
Prieš sąlyga		
Sužadinimo sąlyga		Vartotojas nori sukurti analizuojamų duomenų ataskaitas
Susiję panaudos atvejai	Išplečia PA	
	Apima PA	
	Specializuoja PA	
Pagrindinis įvykių srautas		Sistemos reakcija ir sprendimai
1. Vartotojas pasirenka ataskaitų formavimą		1. Sistema pateikia dimensijų sąrašą
Po sąlyga		Vartotojas suformuoja norimas ataskaitas
Alternatyvūs scenarijai		
1. Nepavyko prisijungti prie duomenų saugyklos		Sistema suformuoja klaidos pranešimą.

Panaudos atvejo „Formuoti ataskaitas“ sekų diagrama pavaizduota žemiau esančiame paveikslėlyje (žr. 10 pav.).



Šaltinis: sudaryta autoriaus.

10 pav. Panaudos atvejo „Formuoti ataskaitas“ sekų diagrama

2.2.1. DUOMENŲ ANALIZĖS KUBO SPECIFIKACIJA

Sukauptų duomenų analizei naudojamas duomenų analizės kubas PARDAVIMAI. Šis kubas skirtas analizuoti informacija apie įmonės klientus, jų perkamas paslaugas ir produktus, analizuoti klientų lojalumą, pardavimus, pardavimų kanalus.

Busimo pardavimų analizės kubo struktūra (žr. 7 lentelę):

7 lentelė. Pardavimų ir klientų duomenų kubo struktūra

Dimensijos	
Laikotarpis	Ataskaitinis laikotarpis
Metai	Ataskaitinio laikotarpio metai
Ketvirtis	Ataskaitinio laikotarpio ketvirtis
Mėnuo	Ataskaitinio laikotarpio mėnuo
Savaitė	Ataskaitinio laikotarpio savaitė
Nuolaida	Paslaugos mokesčiui suteikta nuolaida
Nuolaidos pavadinimas	Paslaugos mokesčiui suteiktos nuolaidos pavadinimas
Nuolaidos procentas	Paslaugos mokesčiui suteiktos nuolaidos dydis procentais
Akcija	Paslaugos pardavimo akcija
Akcijos tipas	Paslaugos pardavimo akcijos tipas
Akcijos grupė	Paslaugos pardavimo akcijos grupė
Akcijos pavadinimas	Paslaugos pardavimo akcijos pavadinimas
Regionas	Vietovė, kurioje yra teikiamos paslaugos klientui
Regionas	Regionas, kuriame teikiama paslauga
Miestas kaimas	Miestas arba kaimas, kuriame teikiama paslauga
Tipas	Pardavimo veiksmo tipas: įrengimas arba atjungimas
Pardavimo kanalas	Pardavimo padalinys kuris pardavė paslaugą
Įmonė	Darbuotojo, kuris pardavė paslaugą, įmonė
Tarnyba	Darbuotojo, kuris pardavė paslaugą, tarnyba
Departamentas	Darbuotojo, kuris pardavė paslaugą, departamentas
Skyrius	Darbuotojo, kuris pardavė paslaugą, skyrius

Sektorius	Darbuotojo, kuris pardavė paslaugą, sektorius
Grupė	Darbuotojo, kuris pardavė paslaugą, grupė
Darbuotojas	Darbuotojo, kuris pardavė paslaugą, vardas ir pavardė
Paslaugos	Paslaugos informacija
Paslaugos tipas	Paslaugos tipo pavadinimas
Paslauga	Paslaugos pavadinimas
Technologija	paslaugos technologijos pavadinimas
Paslaugų gavėjas	Paslaugų gavėjo informacija
Amžius	Paslaugų gavėjo amžius suskirstytas režiais pagal metus
DVBT įsipareigojimai	Kliento turimi įsipareigojimai DVBT paslaugai režiais
IPTV įsipareigojimai	Kliento turimi įsipareigojimai IPTV paslaugai režiais
VOX įsipareigojimai	Kliento turimi įsipareigojimai VOX paslaugai režiais
ZEBRA įsipareigojimai	Kliento turimi įsipareigojimai ZEBRA paslaugai režiais
Kalba	Kliento fiskuotos telefono linijos naudojimo požymis
Vardas Pavardė	Kliento vardas ir pavardė
Paslaugų gavėjo numeris	Kliento paslaugų gavėjo numeris
Paslaugų gavėjo kategorija	Paslaugų gavėjo segmento pavadinimas
Turimų paslaugų rinkinys	Kliento turimų paslaugų pavadinimas
Statusas	Kliento statuso pavadinimas
Skaičiuojamieji laukai	
Klientų kiekis (DISTINCT COUNT)	Įmonės klientų kiekis
Paslaugų kiekis (DISTINCT COUNT)	Parduotų paslaugų kiekis
ARPU (SUM)	Vidutinės pajamos per paslaugą
Pajamos (SUM)	Pajamos gautos už suteiktas paslaugas

2.2.2. DETALI REIKALAVIMŲ SPECIFIKACIJA

Surinkus reikalavimus iš busimų sistemos vartotojų ir aprašius būsimo duomenų kubo struktūrą, toliau reikia tiksliai nustatyti, iš kur kiekviena dimensija ir skaičiuojamasis matas įgaus reikšmes. Iš kokių sistemų bus paimti duomenys. Toliau reikalinga nustatyti reikalingas informacines sistemas, duomenų bazines, laukus, lenteles, iš kurių duomenys bus naudojami duomenų kubui ir duomenų analizei. Reikės sukurti duomenų paėmimo, transformavimo, įkėlimo servisus (angl. DTS – Data Transformation Services), kurių pagalba bus perkeliama ir apdorojami reikalingi duomenys iš operacinių sistemų į duomenų saugyklą. Duomenų saugykloje esančius sutvarkytus ir paruoštus duomenis bus galima naudoti duomenų kube. Lentelėje (žr. 8 lentelę) surašyti visi projektuojamo analizės kubo laukai ir lentelės su jų laukais, iš kurių paimama informacija į duomenų analizės kubą.

8 lentelė. Analizės kubo sąryšiai su pradiniais duomenų šaltiniais

Dimensija	Duomenų bazės lentelė	Laukas
Laikotarpis	Laikotarpis_dim	laikotarpis_dwh_id
Metai	Laikotarpis_dim	Metai
Ketvirtis	Laikotarpis_dim	Ketvircio_pavadinimas
Mėnuo	Laikotarpis_dim	Menesio_pavadinimas
Savaitė	Laikotarpis_dim	Savaites_pavadinimas
Nuolaida	Nuolaidos_dim	Nuolaidos_dwh_id
Nuolaidos pavadinimas	Nuolaidos_dim	Nuolaidos_pavadinimas
Nuolaidos procentas	Nuolaidos_dim	Nuolaidos_procentas
Akcija	Akcijos_dim	Akcijos_dwh_id
Akcijos tipas	Akcijos_dim	Akcijos_Tipas
Akcijos grupė	Akcijos_dim	Akcijos_bendras_pavadinimas
Akcijos pavadinimas	Akcijos_dim	Akcijos_pavadinimas
Regionas	Adresas_dim	Adreso_dwh_id
Regionas	Adresas_dim	Regionas
Miestas kaimas	Adresas_dim	Gyvenvieta_kaimas
Tipas	Akciju_kubas_fct	Tipas
Pardavimo kanalas	Struktura_dim	Struktura_dwh_id
Įmonė	Struktura_dim	Imone
Tarnyba	Struktura_dim	Tarnyba
Departamentas	Struktura_dim	Departamentas

Skyrius	Struktura_dim	Skyrius
Sektorius	Struktura_dim	Sektorius
Grupė	Struktura_dim	Grupe
Darbuotojas	Struktura_dim	Vadybininkas
Paslaugos	Paslaugos_dim	Paslaugos_dwh_id
Paslaugos tipas	Paslaugos_dim	Paslaugos_tipas
Paslauga	Paslaugos_dim	Paslauga
Technologija	Paslaugos_dim	Technologija
Paslaugų gavėjas	PG_dim	PG_dwh_id
Amzius	PG_dim	amzius
DVBT įsipareigojimai	PG_dim	DVBT_isipareigojimai
IPTV įsipareigojimai	PG_dim	IPTV_isipareigojimai
VOX įsipareigojimai	PG_dim	VOX_isipareigojimai
ZEBRA įsipareigojimai	PG_dim	ZEBRA_isipareigojimai
Kalba	PG_dim	kalba
Vardas Pavardė	PG_dim	vardas + pavarde
Paslaugų gavėjo numeris	PG_dim	PG_nr
Paslaugų gavėjo kategorija	PG_dim	PG_kategorija
Turimų paslaugų rinkinys	PG_dim	Paslaugu_rinkiniai
Statusas	PG_dim	PG_statusas
Skaičiuojamieji laukai		Formulė
Klientų kiekis (distinct count)	Akciju_kubas_fct	DISTINCT COUNT (Pg_nr)
Paslaugų kiekis (distinct count)	Akciju_kubas_fct	DISTINCT COUNT (Paslaugos_ID)
ARPU (SUM)	Akciju_kubas_fct	SUM (ARPU)
Pajamos (SUM)	Akciju_kubas_fct	SUM (Pajamos)

Kadangi duomenys jau yra paimti iš pradinių duomenų šaltinių, tokių kaip CRM sistema, vartotojų apskaitos (angl. Billing) sistema, ir apdoroti bei patalpinti duomenų saugykloje, tai duomenų importavimo iš pradinių duomenų šaltinių į duomenų sandėlį proceso ir modelio nenagrinėsime detaliau. Visi analizės kubo duomenys yra paimami iš duomenų lentelių esančių duomenų saugykloje.

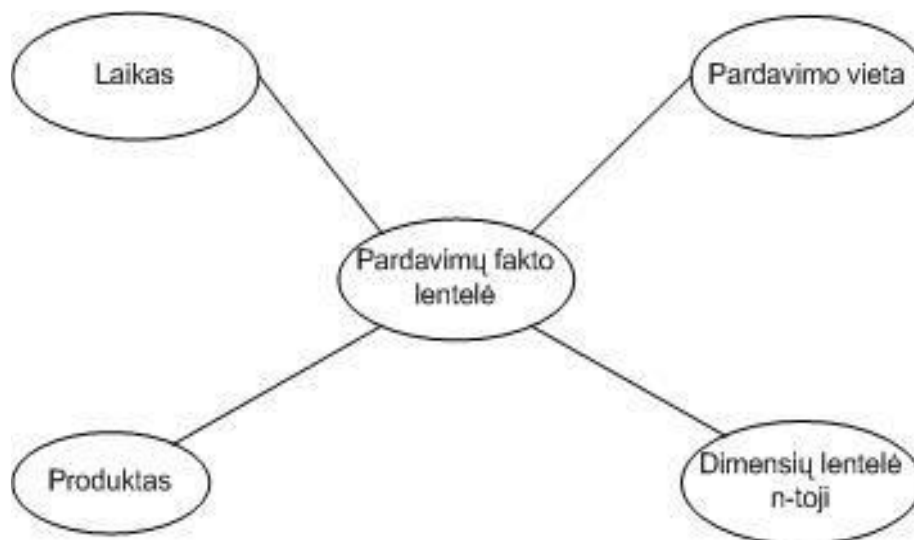
2.3. PARDAVIMŲ IR KLIENTŲ ANALIZĖS SISTEMOS ARCHITEKTŪRA

Gerai parengtas ir su būsimos sistemos vartotojais suderintas sistemos projektas leidžia sumažinti projekto įgyvendinimo klaidų tikimybę, leidžia užtikrinti būsimų sistemos vartotojų reikalavimų ir lūkesčių išpildymą. Tinkamai parengtas sistemos projektas leidžia numatyti reikalingus būsimos sistemos elementus, jų tarpusavio sąryšius ir sujungimo būdus, veiksmų sekas.

Rengiant šį projektą, panaudota „iš viršaus žemyn“ projektavimo technologija – t.y. pradžioje specifikuojamas galutinis rezultatas (duomenų analizės kubas), o toliau detalizuojama, kaip ir iš ko jis bus gaunamas, kokių elementų reikės, kaip jie turi būti kuriami ir pan.

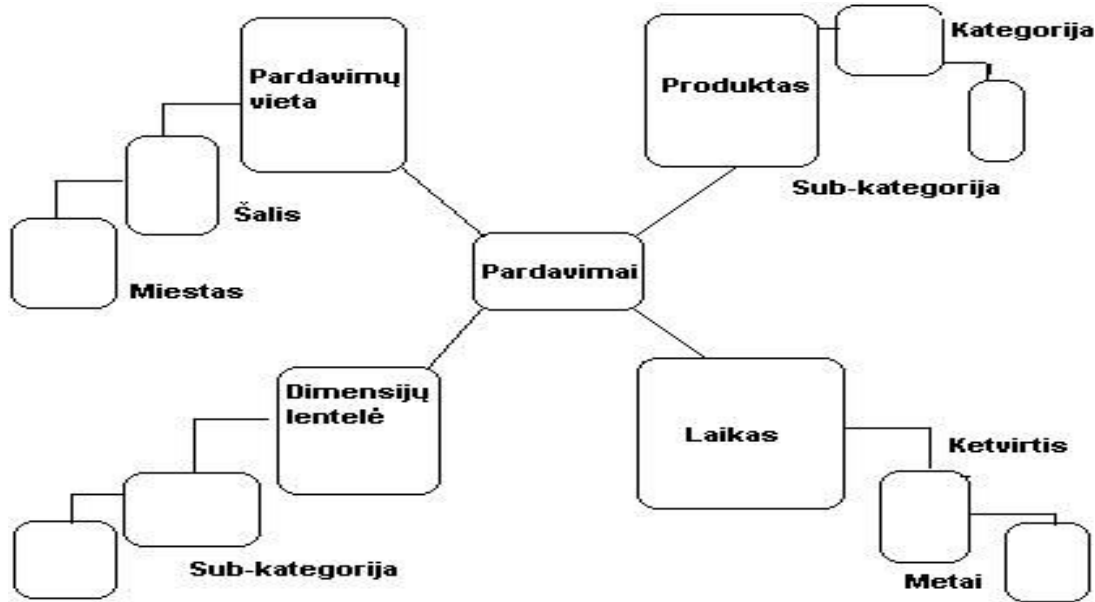
Kiekvienai verslo funkcijai (pardavimai, paslaugos, finansai ir kt.) reikalinga atskira lentelė, kurioje saugojami konkretūs verslo proceso matai. Priklausomai nuo verslo proceso atitinkamai fakto lentelė gali būti sujungta su dimensijų lentelėmis, kurios yra sujungtos ryšiais su kitomis fakto lentelėmis, be to gali naudoti ir unikalias dimensijų lenteles, kurios tiks tik vienintelei fakto lentelei.

Duomenų saugyklų kūrimo dimensinių modelių schemas yra dviejų tipų: žvaigždės tipo schema, snaigės tipo schema (Data warehouse..., 2013). Schemų principiniai paveikslėliai pateikti 11 ir 12 paveikslėliuose (žr. 11,12 pav.).



Šaltinis: sudaryta autoriaus.

11 pav. Žvaigždės tipo schema

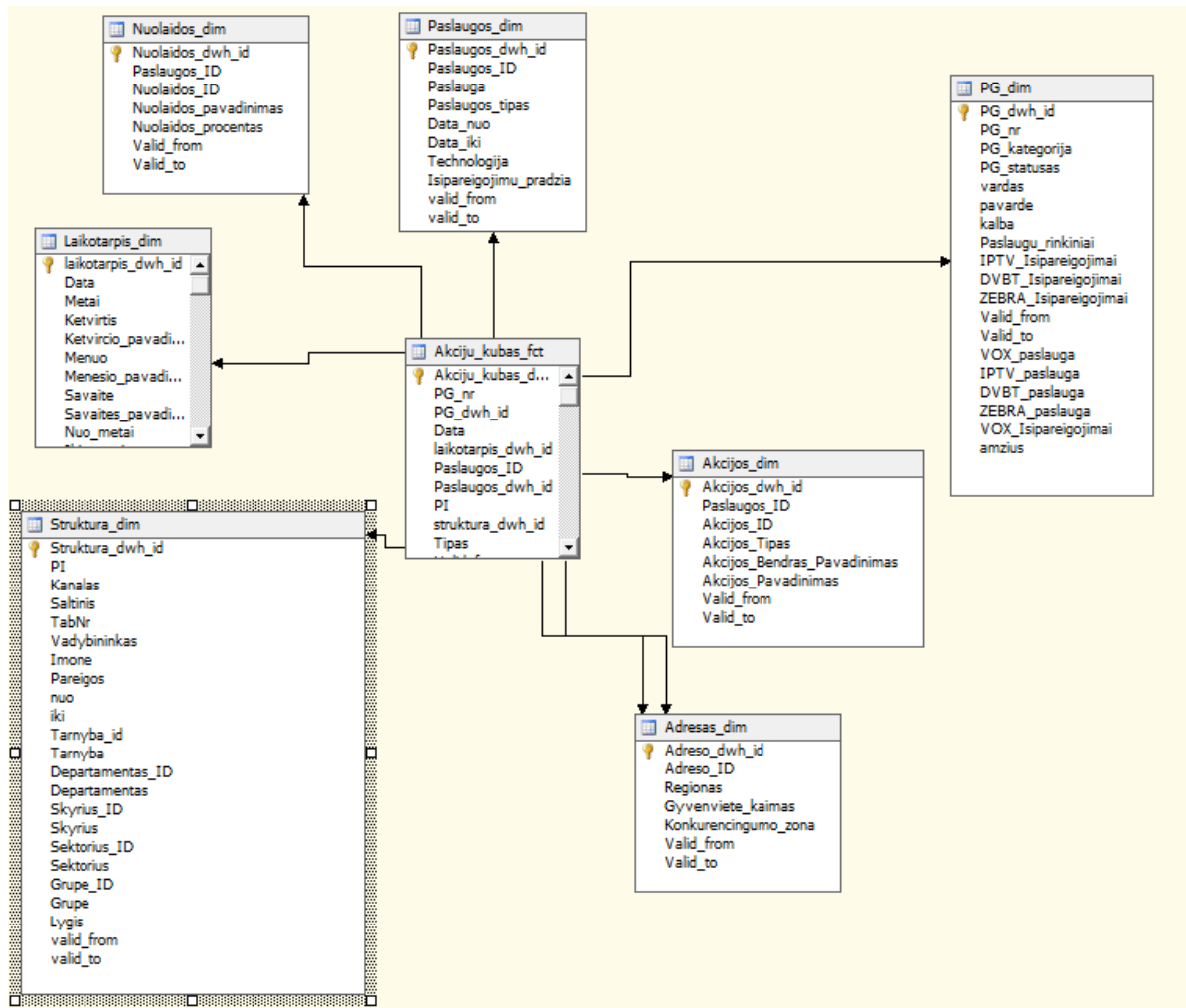


Šaltinis: sudaryta autoriaus.

12 pav. Snaigės tipo schema

Kiekviena iš schemų turi tam tikrų privalumų ir trūkumų, į kuriuos reikia atsižvelgti projektuojant duomenų saugyklą. Žvaigždės tipo schema yra nenormalizuota, tačiau lentelių skaičius yra mažesnis, įrašai gali pasikartoti. Tuo tarpu snaigės tipo schema yra hierarchinio pobūdžio, joje galima atlikti sudėtingas užklausas, taip pat išsprendžiamas duomenų integralumas. Snaigės duomenų saugyklos schemas greitaveika yra didesnė negu žvaigždės. Detaliau apie žvaigždės ir snaigės privalumus ir trūkumus atliktas eksperimentinis tyrimas kitame šio darbo skyriuje. Šiam atvejui buvo pasirinkta žvaigždės tipo architektūra.

Suprojektuoto duomenų analizės kubo schema (žr. 13 pav.).



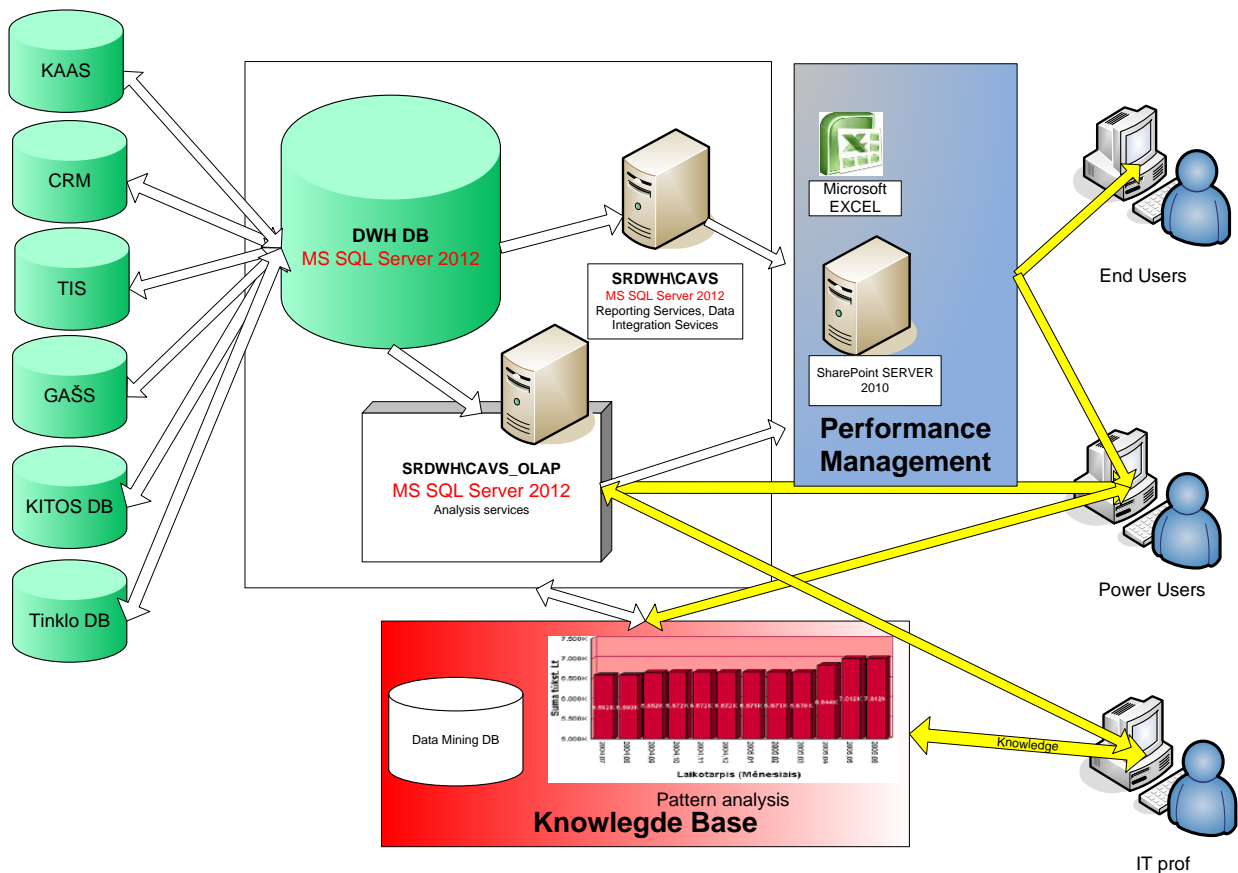
Šaltinis: sudaryta autoriaus.

13 pav. Kuriamos sistemos OLAP duomenų schema

2.3.1. VERSLO ANALITIKOS SISTEMOS ARCHITEKTŪRA

TEO LT, AB įmonė turi ir naudoja daugybę informacinių sistemų, kurios naudoja atskirus serverius. Kiekvienas serveris turi tam tikrus reikalavimus, priklausomai nuo sistemos, yra dubliuojamas, turi veidrodines kopijas. Visuose serveriuose saugoma atitinkamo lygio informacija apie verslą, klientus, ir t.t.. Duomenų analizei yra naudojamas atskiras duomenų bazių serveris, vadinamas duomenų saugykla. Į duomenų saugyklą automatiškai kiekvieną parą yra atnaujinami duomenys iš kitų sistemų. Duomenys yra transformuojami, apdorojami, sutvarkomi ir tinkamu ir suderintu formatu saugomi duomenų augykloje. Kuriant įmonės klientų analizės sistemą, vienas

svarbiausiu reikalavimų yra kuo mažiau apkrauti darbinės įmonės sistemas. Todėl duomenys, duomenų analizės sistemai, visada turi būti naudojami tik iš duomenų saugyklos. Toks sprendimas leidžia minimaliai apkrauti pagrindines įmonės sistemas (nes duomenų nukopijavimas reikalauja gerokai mažiau pagrindinių sistemų resursų, nei tų duomenų analizė tame pačiame darbiniame serveryje). Bendra sistemos architektūra (žr. 14 pav.).



Šaltinis: sudaryta autoriaus.

14 pav. Bendra sistemos architektūra

2.3.2. REIKALAVIMAI OLAP DUOMENŲ KUBO FUNKCIONAVIMO PALAIKYMUI

OLAP duomenų kubai saugomi analizės servisuose, kuriuose yra apdorojami ir atnaujinami. Analizės servisas yra tame pačiame serveryje kaip ir duomenų saugykla. Analizės servisams

keliami dideli techniniai reikalavimai, kadangi yra svarbu trumpas duomenų kubų atnaujinimo, užklausų apdorojimo laikas. Duomenų saugyklos serveris turi pakankamai techninių resursų, kad galima būtų užtikrinti sklandų ne tik duomenų saugyklos darbą, bet ir analizės servisų, kuriuose diegiami duomenų kubai.

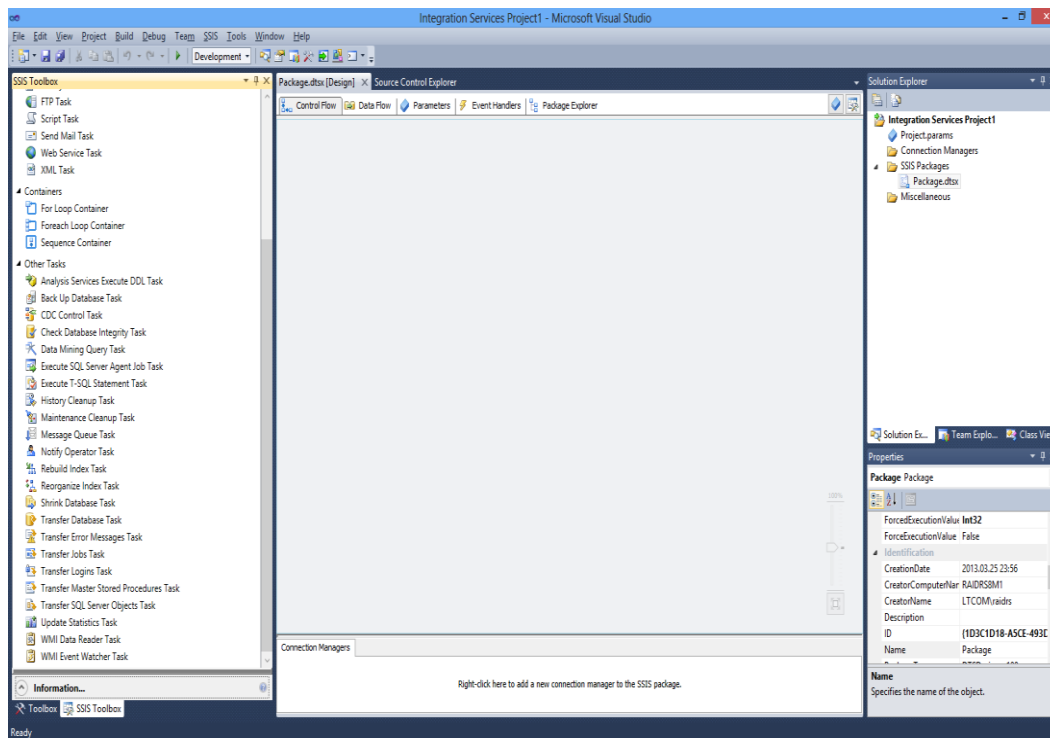
Be įprastų techninės įrangos parametrų, nemažiau svarbus ir ryšio tarp serverių parametrai. Tačiau duomenų pasikeitimai tarp analizės servisų ir duomenų saugyklos, vyksta tame pačiame serveryje todėl tinklui papildomi reikalavimai nekeliami.

2.3.3. DUOMENŲ TRANSFORMACIJŲ PASLAUGOS

Analizės etapo metu buvo nurodyta, jog vienas svarbiausių reikalavimų - tai kuo mažiau apkrauti pagrindines operacines sistemas. Dėl to visi analizei reikalingus duomenys saugomi duomenų saugykloje. Yra svarbu užtikrinti savalaikį ir patikimą duomenų atnaujinimo mechanizmą tarp pagrindinių įmonės sistemų ir duomenų saugyklos.

Duomenų atnaujinimui nuspręsta panaudoti DTS (angl. Data Transformation Service). Tai įrankis, kuris yra prie kiekvienos Microsoft SQL serverio licencijos ir duodamas nemokamai. DTS – tai įrankiu rinkinys, kuris naudojamas norint eksportuoti ar importuoti duomenis tarp skirtingų duomenų šaltinių (pvz.: Microsoft SQL server 2012, Microsoft Excel, Oracle DBVS, FLAT file, CSV ir pan. duomenų formatai). DTS leidžia duomenis paimti, transformuoti, apdoroti, apjungti ir importuoti į norimą vietą. Ryšiai tarp skirtingų duomenų šaltinių yra užtikrinami, pasitelkiant įvairias atviro standarto duomenų prieigas (pvz.: OLE DB, ODBC).

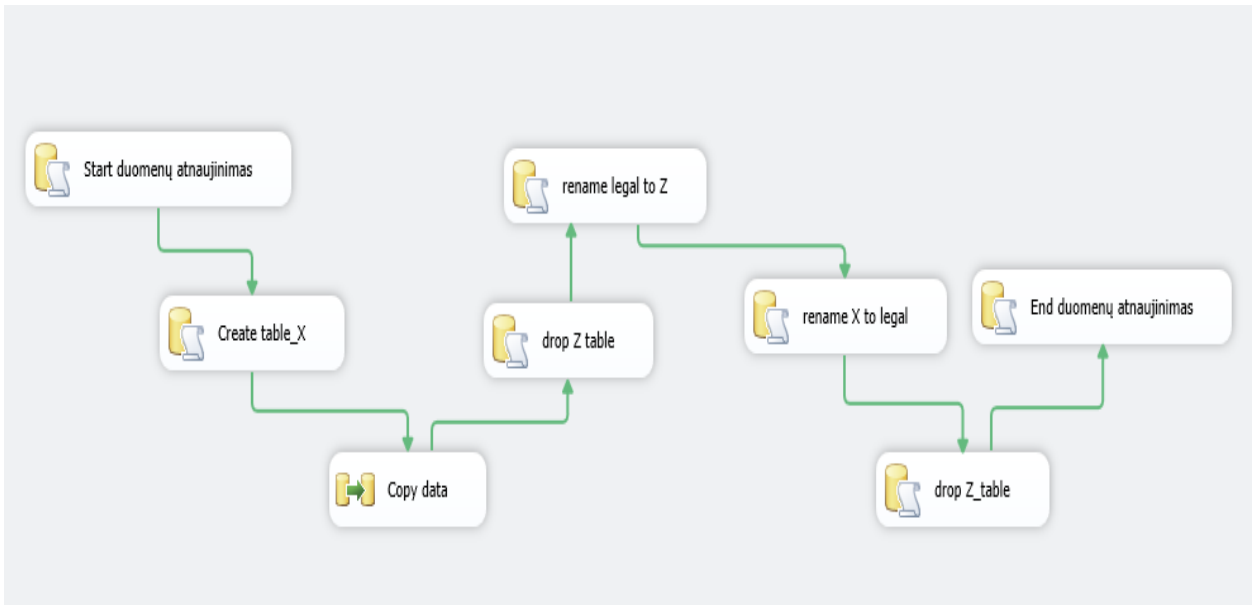
Kartu su Microsoft SQL Server 2012 yra naudojama Visual Studio programa, kurioje pateikiamas DTS dizaineris leidžia kurti DTS modelį, pasitelkiant grafinius elementus ir juos dėliojant DTS dizainerio (angl. DTS Designer) lange (žr. 15 pav.).



Šaltinis: sudaryta autoriaus.
15 pav. DTS Designer langas

Kairiajame DTS dizainerio lange matome prietaisų skydelį, iš kurio galime pasirinkti norimą veiksmą ir įtempti į pagrindinį langą. Apačioje „Connection Managers“ lange apibrėžiami prisijungimai prie duomenų šaltinių. DTS dizainerio pagalba, šiame projekte realizuotas reikalingų duomenų perkėlimas iš klientų valdymo sistemų (CRM ir apskaitos sistemos) duomenų bazių į duomenų saugyklą, iš kurio duomenis ima analizės kubas.

Bendruoju atveju, vienos lentelės perkėlimo metu, yra nuskaitymi reikiami duomenys iš pradinių duomenų šaltinių, transformuojami ir apdorojami ir galiausiai įrašomi į laikiną duomenų saugyklos duomenų lentelę „table_X“. Sena lentelė esanti duomenų saugykloje pervardinama į laikiną lentelę „table_Z“. Ir tik tada laikina lentelė „table_X“ su naujais duomenimis yra pervardinama į duomenų saugyklos duomenų lentelę, kuri vėliau panaudojama duomenų kubo kūrimo ir atnaujinimo procese. Vienos lentelės perkėlimo DTS paketas pavaizduotas žemiau esančiame paveikslėlyje (žr. 16 pav.).

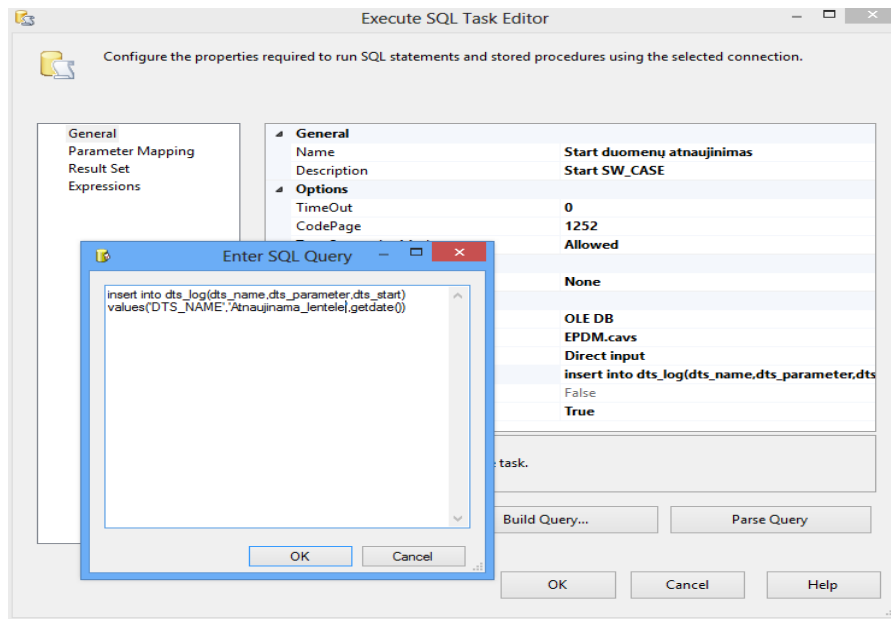


Šaltinis: sudaryta autoriaus.

16 pav. Vienos lentelės duomenų atnaujinimo DTS paketas

Vienos lentelės duomenų atnaujinimo procesą galima išskaidyti i kelis etapus:

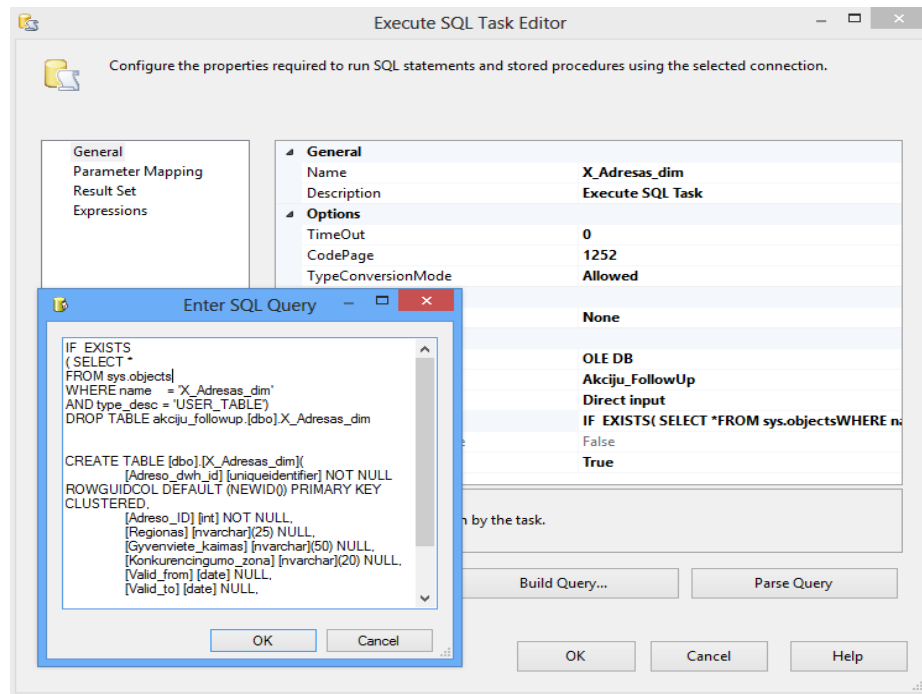
1. Įrašomas įrašas į veiksmų fiksavimo lentą, kad duomenų atnaujinimas tam tikros lentelės yra pradėtas. Nurodoma data ir laikas kada startavo duomenų atnaujinimas, lentos pavadinimas, kuri yra atnaujinama.



Šaltinis: sudaryta autoriaus.

17 pav. Įrašomas įrašas į veiksmų fiksavimo lentą apie darbų pradžią

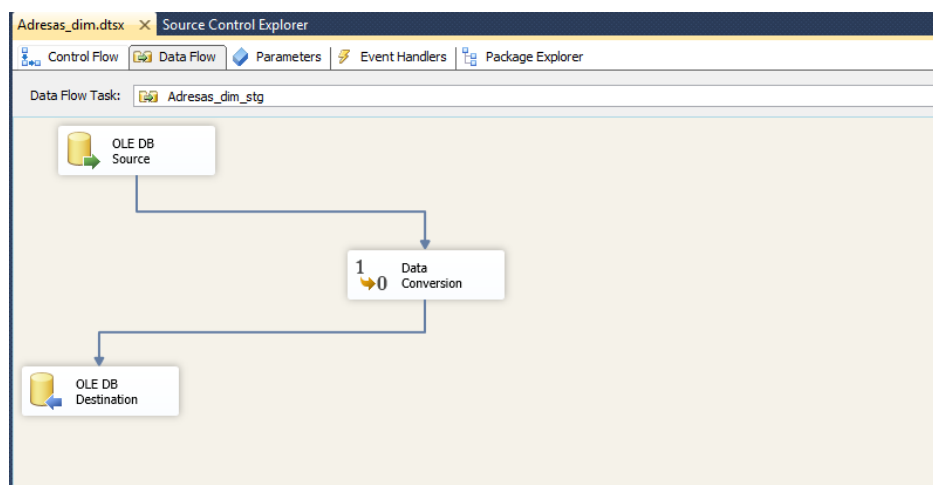
2. Sukuriama laikina lentelė į kurią bus talpinami nauji duomenys.



Šaltinis: sudaryta autoriaus.

18 pav. DB SQL dialogo langas laikinos lentelės projektui

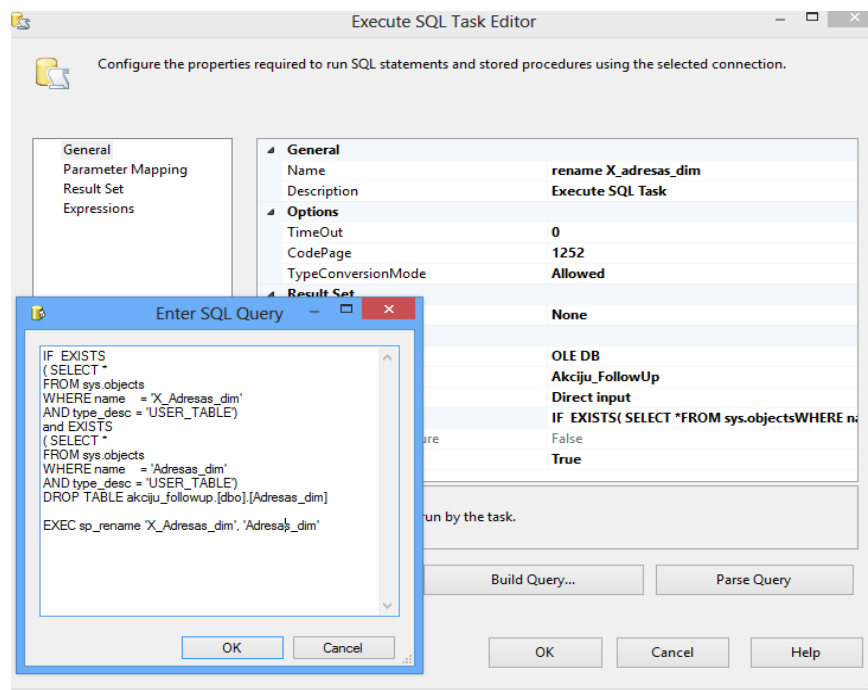
3. Kopijuojami duomenys iš pradinio duomenų šaltinio į duomenų saugyklos laikiną lentelę. Šis žingsnis susideda iš trijų: nuskaityti duomenys, transformuojami, įrašomi. Visi žingsniai pavaizduoti žemiau esančiame paveikslėlyje (žr. 19 pav.).



Šaltinis: sudaryta autoriaus.

19 pav. Duomenų kopijavimo ir transformavimo žingsniai

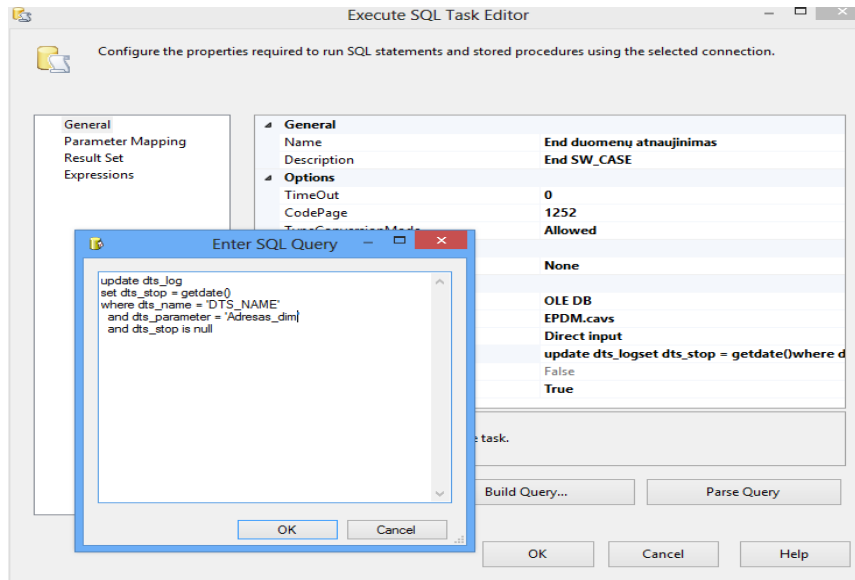
4. Toliau yra ištrinama laikina lentelė „Z“, tam kad į ją galėtume perkelti senus duomenis. Tai yra daroma savotiška duomenų atsarginė kopija, tam kad procesui neužsibaigus sėkmingai būtų galima atstatyti senus duomenis.
5. Seni duomenys patalpinami į laikinę „Z“ lentelę, tiesiog pervardinus lenteles.
6. Laikina lentelė „X“ yra pervardinama į duomenų lentelę, kurią vėliau galėsime naudoti analizės kubo kūrimo procese. Primenu, kad laikinoje lentelėje „X“ buvo pakrauti nauji duomenys, taigi ir pervardintoj lentelėje bus patys naujausi duomenys. Tai yra esminis žingsnis duomenų lentelės atnaujinime.



Šaltinis: sudaryta autoriaus.

20 pav. DB SQL dialogo langas laikinos lentelės pervardinimo į naują lentelę projektui

7. Ištrinti „Z“ lentelę su atsarginės kopijos duomenimis. Šis žingsnis yra atliekamas iškart po sėkmingai atlikto 6 žingsnio, kada esam tirki, kad naujausi duomenys atsinaujino sėkmingai.
8. Užbaigtas duomenų atnaujinimo procesą, pažymėdami veiksmų fiksavimo lentelėje.

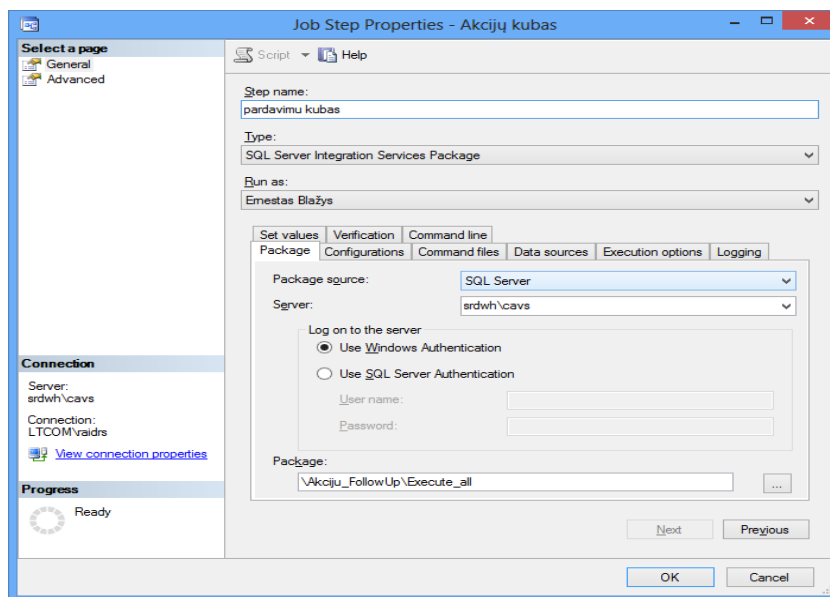


Šaltinis: sudaryta autoriaus.

21 pav. Duomenų atnaujinimo proceso užbaigimo fiksavimas, įrašomas į veiksmų fiksavimo lentelę

Visų duomenų saugyklos lentelių duomenų atnaujinimo procesas vyksta tokiu pačiu 8 žingsnių principu.

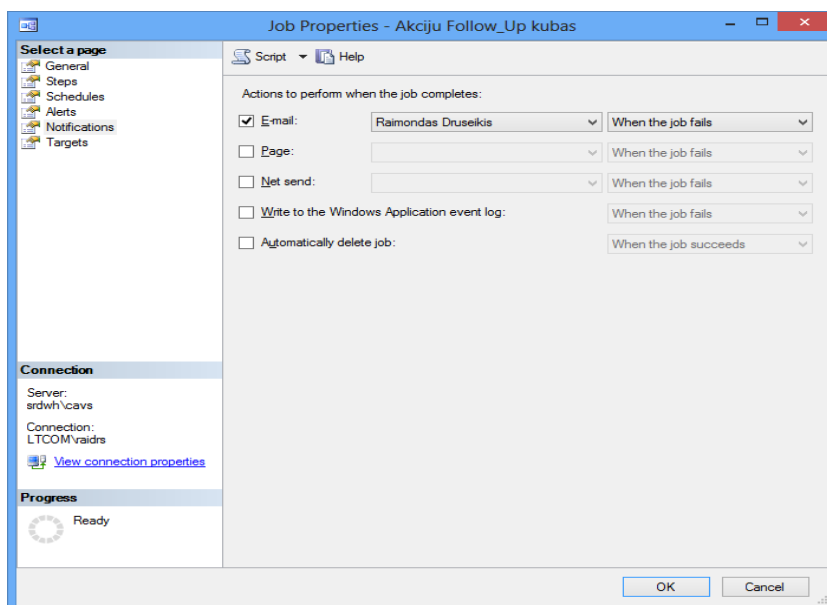
Sukurtas DTS paketas paleidžiamas per SQL SERVER Agent Job servisą. Tam reikia sukurti darbą kuris galėtų paleisti nurodytą DTS paketą. Paveiksle parodytas vienas iš darbo žingsnių, kuriame paleidžiamas DTS paketas (žr. 22 pav.).



Šaltinis: sudaryta autoriaus.

22 pav. DTS paleidimas per SQL SERVER Agent job

Duomenų atnaujinimo procese numatyta, jog įvykus klaidai, atnaujinimo procesas stabdomas ir administratoriui išsiunčiama elektroninė žinutė apie klaidą. Sėkmės atveju, sistemos administratorius apie tai taip pat informuojamas. Žinučių gavėją galima nurodyti SQL SERVER Agent darbe, arba galima pačiame DTS pakete sukurti užduotį, kuri vienu ar kitu atveju siųstu atitinkamą pranešimą.



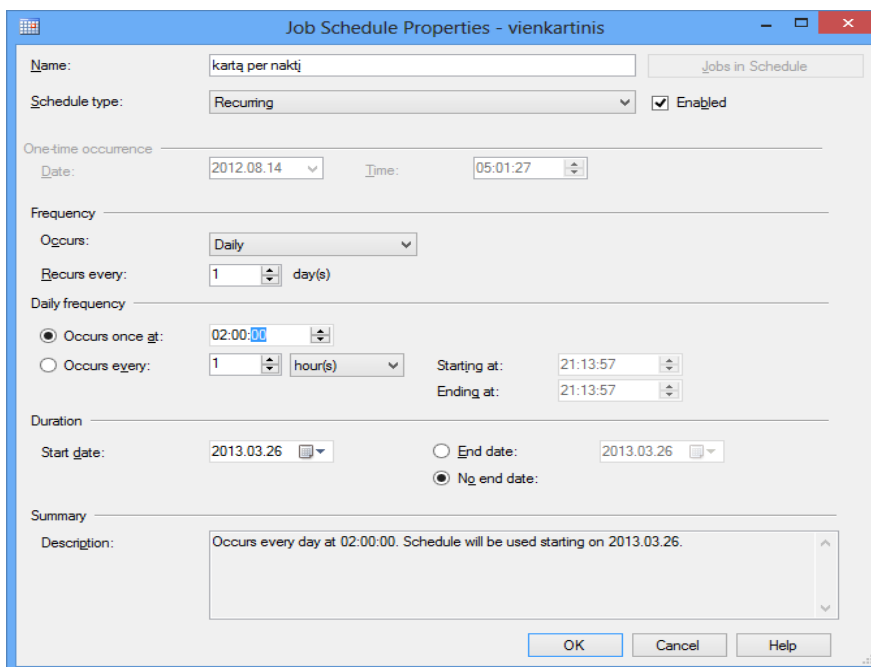
Šaltinis: sudaryta autoriaus.

23 pav. SQL SERVER Agent job pranešimų konfigūravimo langas

2.3.4. DUOMENŲ ATNAUJINIMO AUTOMATIZAVIMAS

Analizei skirtų duomenų savalaikis atnaujinimas yra svarbus procesas, užtikrinantis, jog vartotojui bus pateikta naujausia informacija. Projekto pradžioje buvo nustatyta, kad visiškai pakankamas duomenų atnaujinimo periodas – kartą į parą. Duomenis atnaujinant naktį, tokiu būdu yra mažiau apkraunamos informacinės sistemos kada su jomis dirbama t.y. darbo valandomis. Tam tikslui bus panaudojamas sukurtas DTS paketas, perkeliantis duomenis iš Pradinių duomenų šaltinių į duomenų saugyklą. Taip pat panaudojamas MS SQL SERVER 2012 Agent užduotis, paleidžianti DTS paketą. Taigi, sukurtai MS SQL SERVER 2012 Agent užduočiai sukuriame paleidimo planą. Nustatome periodiškumą, kuriuo laiku bus atnaujinami duomenų

saugyklos duomenys. Šiuo atveju kaip pavaizduota žemiau esančiame paveikslėlyje (žr. 24 pav.) duomenų atnaujinimas kiekvieną naktį 02:00 val..

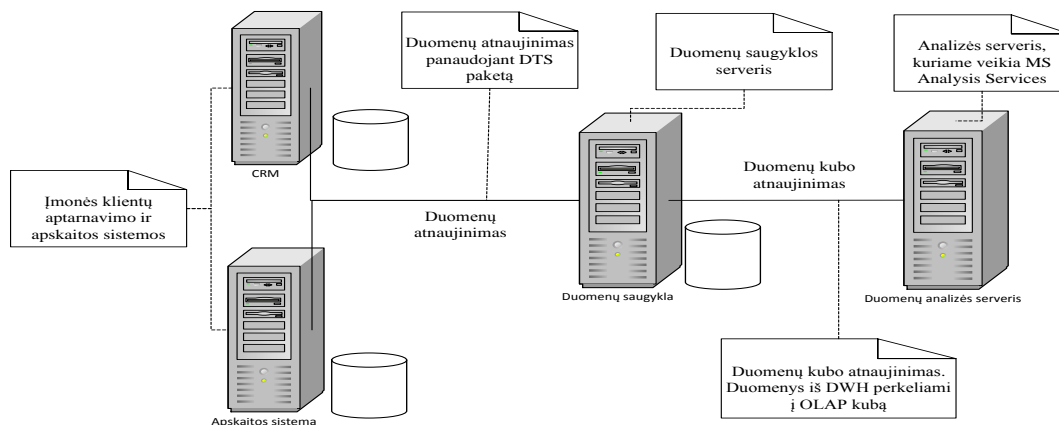


Šaltinis: sudaryta autoriaus.

24 pav. SQL SERVER Agent užduoties vykdymo ir jos parametrų parinkimo langas

2.3.5. DUOMENŲ ANALIZĖS KUBO FORMAVIMO PROCESAS

Duomenų kubas formuojamas keliais etapais. Tai tam tikra procesų seka. Kiekvieno etapo metu vyksta duomenų išgavimas iš pradinių šaltinių ir tų duomenų įkėlimas į kitą sistemą. Bendras duomenų kubo formavimo procesas pavaizduotas žemiau esančiame paveikslėlyje (žr. 25 pav.).



Šaltinis: sudaryta autoriaus.

25 pav. Kubo formavimo procesas

Analizės kubo formavimo procesas yra kiek sudėtingesnis, kadangi analizei skirti duomenys yra saugomi duomenų saugykloje. Pirmojo etapo metu, panaudojant DTS (ang. Data Transfer Services), analizei reikalingi duomenys iš darbinių sistemų yra nukopijuojami į analizės MS SQL duomenų bazės serverį:

- Paruošiamos naujos laikinos lentelės
- Nuskaitomi nauji duomenys iš klientų aptarnavimo ir apskaitos sistemų
- Nuskaityti duomenys įrašomi į naujai paruoštas lenteles
- Naujos lentelės pervardinamos į pastovias lenteles

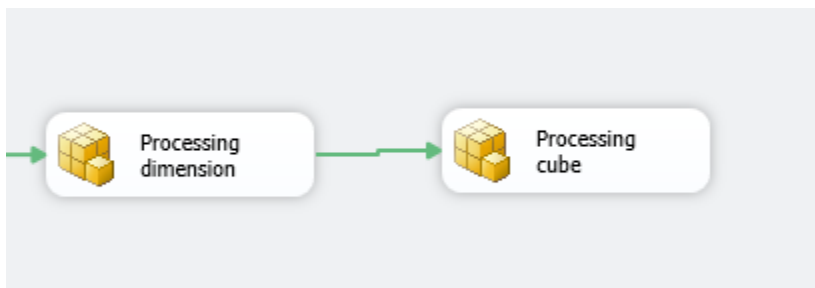
Sekantis žingsnis – tai faktų lentelės suformavimas:

- Ištrinama esama faktų lentelė
- Iš naujo suformuojama faktų lentelė
- Apskaičiuojamos tarpinės reikšmės
- Rezultatai išsaugomi analizės duomenų bazės serverio faktų lentelėje

Paskutinis analizės kubo formavimo etapas – tai kubo atnaujinimas:

- Atnaujinamos analizės kubo dimensijos
- Duomenys iš duomenų saugyklos serverio perkeliama į OLAP duomenų saugyklą
- Suskaičiuojamos agregacijos

Kubo atnaujinimo etapas yra realizuotas DTS pakete. Kubo ir dimensijų atnaujinimui yra skirtos užduotys iš DTS užduočių panelės, kaip pavaizduota žemiau esančiame paveikslėlyje (žr. 26 pav.).



Šaltinis: sudaryta autoriaus.

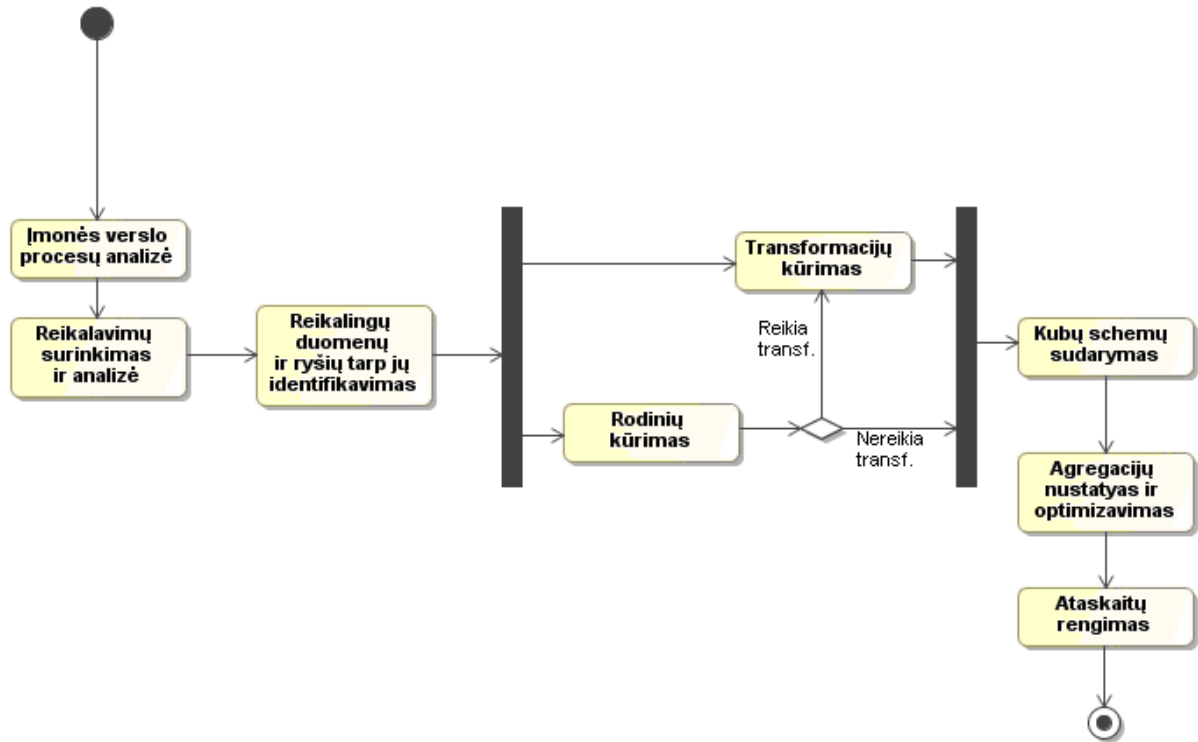
26 pav. Analizės kubo ir dimensijų atnaujinimo užduotys

Kubo atnaujinimo etapai čia atliekami automatiškai, paleidus kubo atnaujinimo procesą. Priklausomai nuo poreikių, galima atnaujinti kubą trejopai:

- Laipsniškas atnaujinimas (angl. Incremental update) – leidžia atnaujinti kubą, nekeičiant esamų kubo duomenų, o tiktai pridėdant prie kubo atsiradusius naujus duomenis. Šis kubo atnaujinimo procesas yra vienas paprasčiausių ir greičiausių – neperskaičiuojamos esančios agregacijos, pridėdamos tiktai naujos – tai sumažina kubo atnaujinimo laiką.
- Duomenų atnaujinimas (angl. Refresh data) – šio kubo atnaujinimo metu yra perskaičiuojamos visos kubo reikšmės ir atnaujinami duomenys. Šio kubo atnaujinimo būdo privalumas yra tai, jog vartotojai gali naudotis kubu netgi jo atnaujinimo metu. Taigi, esant poreikiui, bet kada galima atnaujinti kubo duomenis.
- Visiškas atnaujinimas (angl. Full Process) – tai pilnas kubo atnaujinimas. Šis kubo atnaujinimas reikalingas tada, kai pakeičiama pati kubo struktūra, sukuriamos arba panaikinamos tam tikros dimensijos, jų reikšmės ir pan. Šio atnaujinimo metu kubu naudotis vartotojai negali.

2.3.6. DUOMENŲ ANALIZĖS SISTEMOS FORMAVIMO MODELIS

Šio darbo tikslas yra ne tik sukurti konkrečią analizės sistemą ir ja integruoti į įmonės veiklos procesus, bet ir sukurti OLAP analizės kubų formavimo proceso modelį, kuris padėtų panašiuose projektuose, leistų lengviau numatyti galimus sunkumus, sprendimus ir technologijas. Supaprastintas analizės kubų formavimo proceso modelis, sukurtas remiantis literatūros šaltiniuose pateikiamais metodais, pateiktas žemiau esančiame paveikslėlyje (žr. 27 pav.).



Šaltinis: sudaryta autoriaus.

27 pav. Analizės kubo formavimo modelis

Kiekvienas šio proceso etapas gali būti detalizuotas išsamiai. Dalis šio proceso etapų yra detalizuoti šio projekto skyriuose.

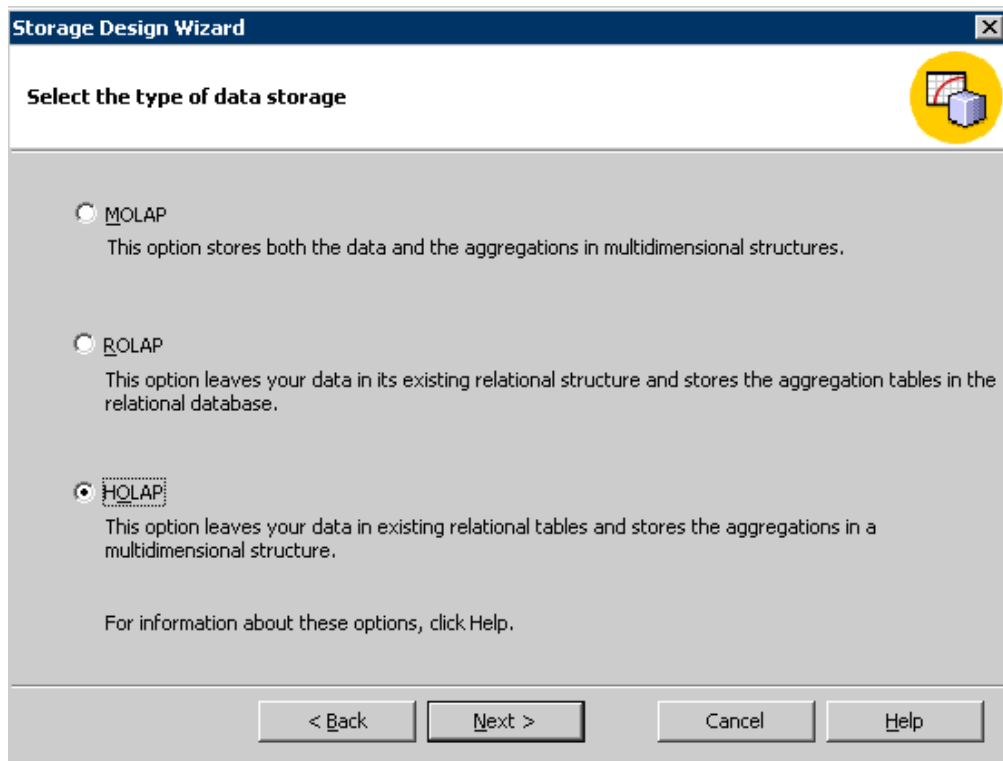
3. OLAP KUBŲ ARCHITEKTŪROS LYGINAMOJI ANALIZĖ

Naudojant MS Analysis Services įrankį buvo sudaryta skirtingų OLAP architektūrų kubai ir atliktas jų palyginamasis tyrimas. Analysis Services palaiko tris pagrindines duomenų saugojimo architektūras:

- MOLAP (angl. *Multidimensional Online Analytical Processing*) – ir detalūs ir konsoliduoti duomenys saugomi daugiamačiame duomenų struktūroje. Skaičiavimai saugomi kartu su duomenimis. Toks duomenų saugojimas leidžia potencialiai greičiausiai vykdyti užklausas. Užklausų vykdymo greitis priklauso tik nuo to, kiek procentų skaičiavimų atlikta iš anksto. MOLAP taikomas kubams, kurie dažnai peržiūrimi ir kur reikia greito atsako į užklausas. Šios architektūros tipui trūksta standartizavimo. Kitas MOLAP trūkumas yra jo apimtis, daugiamačiai duomenų modeliai tinkami darbui su suminiais duomenimis, tačiau jeigu reikia detalesnio lygmens analizės, duomenų bazės dydis greitai išauga.
- ROLAP (angl. *Relational Online Analytical Processing*) – konsoliduoti ir detalūs duomenys saugomi reliacinėje duomenų bazių struktūroje, be to detalūs duomenys gali būti išgaunami tiesiog iš pradinio duomenų šaltinio (duomenų sandėlio). Užklausos tokioje duomenų struktūroje vykdomos žymiai lėčiau. ROLAP naudojamas dideliems duomenų kiekiams arba tada, kai užklausos nėra dažnos. Šis panaudojimo būdas tinkamas analizuoti archyvinis duomenis.
- HOLAP (angl. *Hybrid Online Analytical Processing*) – detalūs duomenys saugomi reliacinėje duomenų saugykloje, o konsoliduoti duomenys – daugiamačiame kube. Šiuo atveju kubo užimama atmintis mažesnė nei MOLAP atveju, o užklausos aptarnaujamos greičiau nei ROLAP atveju. Užklausos, kurios išgauna susumuotus duomenis šioje struktūroje, įvykdomos taip pat greitai kaip ir MOLAP. Užklausos kurios naudoja bazinius duomenis, pavyzdžiui gilyn einančios iki tam tikro fakto, turi gauti duomenis iš reliacinės struktūros ir nėra tokios greitos. Kubai saugomi HOLAP yra mažesni, negu MOLAP struktūroje ir atsako į užklausas greičiau, nei kubai esantys ROLAP struktūrose, kai užklausiama išanksto susumuoti duomenys. HOLAP taikomas kai dirbama su kubais,

kuriuose reikia greito atsako į išanksto susumuotų duomenų užklausas dideliuose duomenų kiekiuose.

MS SQL 2012 Analysis Services - tai viena iš nedaugelio sistemų, kurios leidžia pasirinkti iš trijų duomenų saugojimo architektūrų (žr. 28 pav.):



Šaltinis: sudaryta autoriaus.

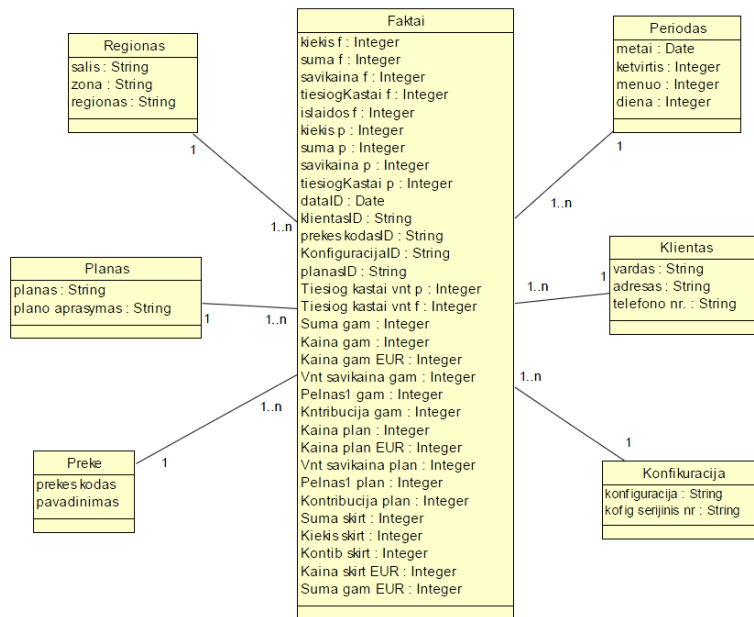
28 pav. Analysis Services įrankio kubų architektūros pasirinkimo vedlys

Tai įgalina naudoti Analysis Services įvairiose probleminėse srityse. Tačiau duomenys kubui išgaunami iš duomenų sandėlio, todėl norint užtikrinti korektišką kubo struktūrą, patys duomenų šaltinis turi būti teisingai suprojektuotas, priešingu atveju kubo projektavimas tampa sudėtingas. Duomenys iš tradicinės reliacinės DB yra transformuojami į „žvaigždės“ ar „snaigės“ tipo architektūrą.

3.1. „ŽVAIGŽDĖS“ IR „SNAIGĖS“ TIPO DUOMENŲ SCHEMOS

Žvaigždės schema - kai visos dimensijų lentelės tiesiogiai sujungtos su faktų lentele. Schema pavadinta dėl esybių išsidėstymo žvaigždės pavidalu. Žvaigždės schemas centras- faktų lentelė. Joje yra saugomi tikrieji duomenys (faktai) ir visos užklausos yra formuojamos būtent iš

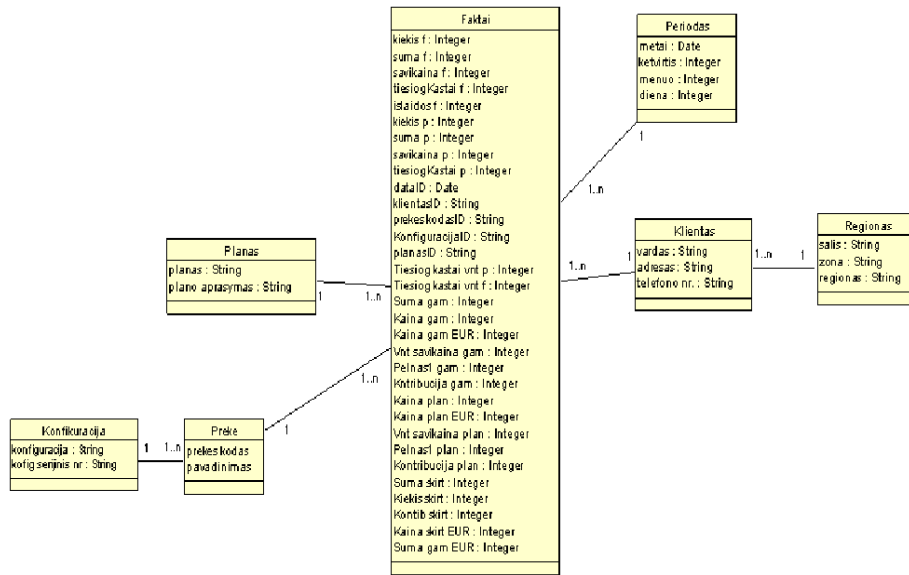
šios lentelės. Faktai dažniausia yra skaitmeniniai atributai tokie kaip kaina, kiekis, kartai, nuolaida, pajamos. Jie gali būti sumuojami, randamas jų vidurkis, maksimumas, minimumas, reikšmių kiekis, jie gali būti agreguojami, apdorojami įvairiomis matematinėmis ir statistinėmis funkcijomis. Faktai yra išmatuojamos skaitinės reikšmės apie daiktą kurį aprašo fakto lentelė. Sekančiame paveiksle yra parodyta „žvaigždės“ tipo duomenų schema iš kurios duomenis ima OLAP kubas (žr. 29 pav.).



Šaltinis: sudaryta autoriaus.

29 pav. Žvaigždės tipo duomenų schema

Snaigės schema - kai viena ar kelios dimensijų lentelės nėra tiesiogiai sujungtos su faktų lentele, o siejasi per kitas dimensijų lenteles. Snaigės schema papildo dimensijos lenteles hierarchine struktūra. Sekančiame paveiksle pavaizduota „snaigės“ tipo duomenų architektūra, kuri naudojama OLAP kube (žr. 30 pav.).



Šaltinis: sudaryta autoriaus.

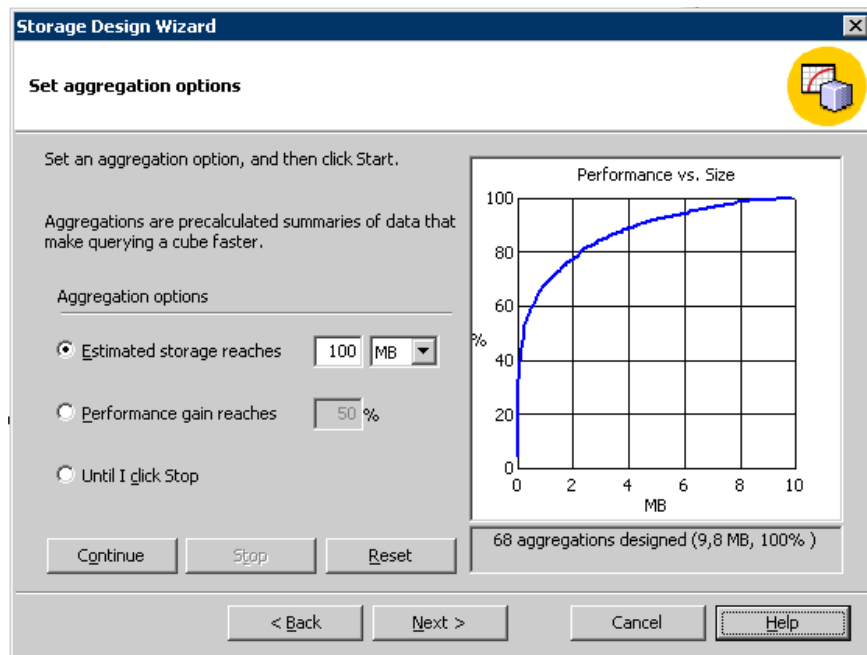
30 pav. Snaigės tipo duomenų schema

Kiekviena iš schemų turi tam tikrų privalumų ir trūkumų, į kuriuos reikia atsižvelgti projektuojant duomenų saugyklą. Žvaigždės tipo schema yra nenormalizuota, tačiau lentelių skaičius yra mažesnis, įrašai gali pasikartoti. Tuo tarpu snaigės tipo schema yra hierarchinio pobūdžio, joje galima atlikti sudėtingas užklausas, taip pat išsprendžiamas duomenų integralumas.

3.2. MOLAP ARCHITEKTŪRA

MOLAP (angl. *Multidimensional Online Analytical Processing*) - tai duomenų analizės priemonė paremta duomenų kubo modeliu. MOLAP saugo jau susistemintus kubo duomenis, darant užklausą nereikia kiekvieną kartą skaičiuoti kubo reikšmių, todėl MOLAP priemonės pasižymi greitu atsakymu į užklausą. Tačiau į MOLAP kubą sudėtingiau pridėti naujas dimensijas, be to, kube nesaugomi detalizuoti duomenys, todėl negalima pasinaudoti „Drill - In“ galimybėmis.

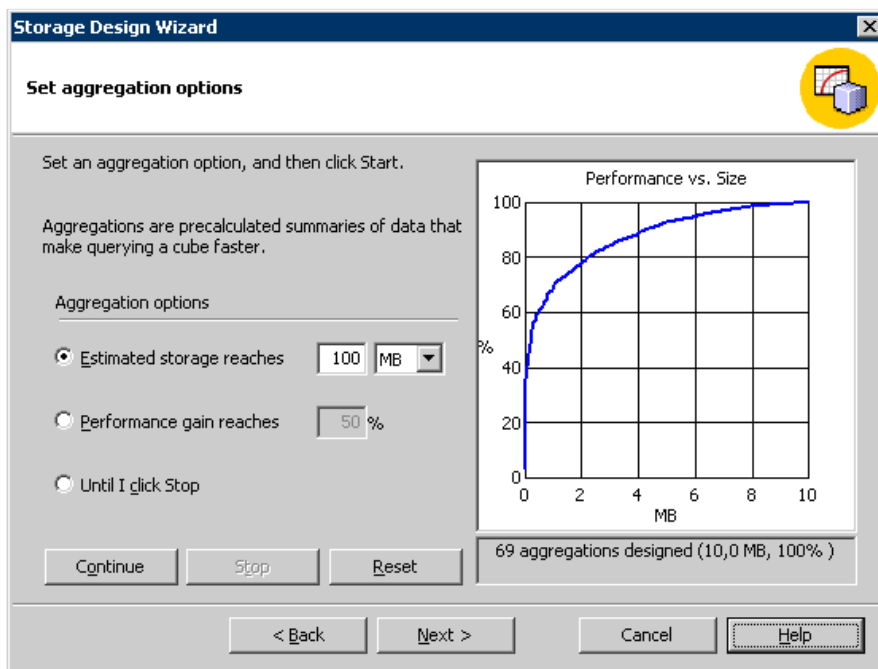
MOLAP architektūros kubo, kuris ima duomenis iš „žvaigždės“ tipo duomenų schemas, apdorojimo charakteristikos pavaizduotos sekančiame paveiksle (žr. 31 pav.).



Šaltinis: sudaryta autoriaus.

31 pav. MOLAP architektūros charakteristikos, kai naudojama „žvaigždės“ tipo duomenų schema

MOLAP architektūros kubo, kuris ima duomenis iš „snaigės“ tipo duomenų schemas, apdorojimo charakteristikos pavaizduotos sekančiame paveiksle (žr. 32 pav.).



Šaltinis: sudaryta autoriaus.

32 pav. MOLAP architektūros charakteristikos, kai naudojama „Snaigės“ tipo duomenų schema

3.3. MOLAP PRANAŠUMAI IR TRŪKUMAI

Pranašumai:

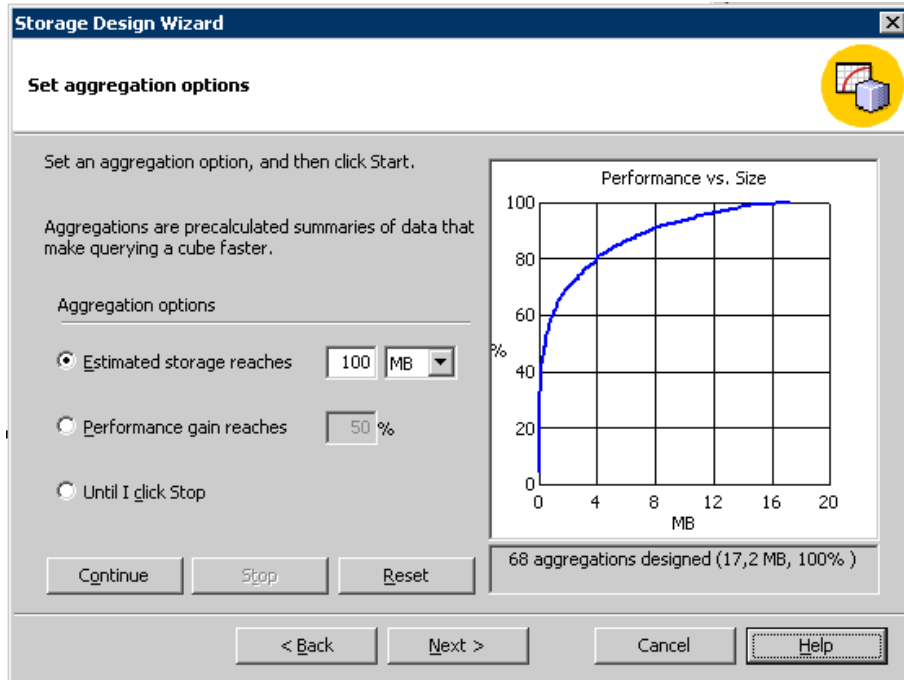
- Greitas užklausų vykdymas – MOLAP kubai greitai išgauna duomenis ir yra optimalūs informacijos išrinkimo ir pateikimo procese.
- Gali vykdyti daug sudėtingų skaičiavimų. Visi skaičiavimai yra saugomi pačiame OLAP kube.

Trūkumai:

- Lėtai vykdo užklausas su detalesnio lygio duomenimis.
- Ribotas duomenų kiekis, kadangi visos įmanomas reikšmės jau yra apdorotos jo kūrimo procese ir saugomos kube.
- Užima pakankamai daug vietos

3.4. ROLAP ARCHITEKTŪRA

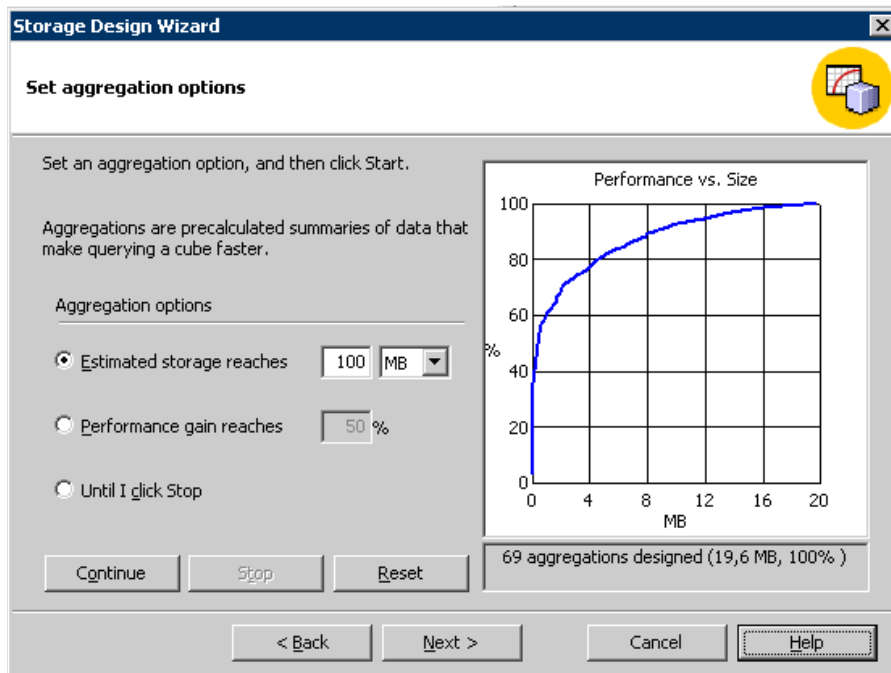
ROLAP (angl. *Relational Online Analytical Processing*) veikia reliacinės duomenų bazės pagrindu kur sukuriamos užklausos, formuojančios duomenų kubus. Kiekvieną kartą analizuojant kubo duomenis, kubo reikšmės perskaičiuojamos. ROLAP užklausos nėra tokios greitos kaip MOLAP, ypač dirbant su dideliais duomenų kubais, tačiau čia saugoma detali informacija, ją išgaunant galima pasinaudoti „Drill - In“ galimybėmis. ROLAP duomenų kubą paprasčiau papildyti naujomis dimensijomis. Greitesnis kubo atnaujinimo procesas. Sekančiame paveiksle pavaizduota ROLAP architektūros kubo vykdymo charakteristikos (žr. 33 pav.).



Šaltinis: sudaryta autoriaus.

33 pav. ROLAP architektūros charakteristikos, kai naudojama „žvaigždės“ tipo duomenų schema

ROLAP architektūros kubo, kuris ima duomenis iš „snaigės“ tipo duomenų schemas, apdorojimo charakteristikos pavaizduotos sekančiame paveiksle (žr. 34 pav.).



Šaltinis: sudaryta autoriaus.

34 pav. ROLAP architektūros charakteristikos, kai naudojama „snaigės“ tipo duomenų schema

3.5. ROLAP PRANAŠUMAI IR TRŪKUMAI

Pranašumai:

- Duomenys saugomi standartinėje reliacinėje duomenų bazėje todėl detalių duomenų išgavimui galima panaudoti ir kitus įrankius (įrankiai nebūtinai turi būti OLAP įrankiai);
- ROLAP priemonės leidžia lanksčiau valdyti didelius duomenų kiekius;
- ROLAP įrankiai geriau apdoroja neskaičiuojamas reikšmes (angl. non-aggregatable facts), pvz. tekstinius duomenis. MOLAP priemonės lėtai vykdo užklausas su tokiais duomenimis;

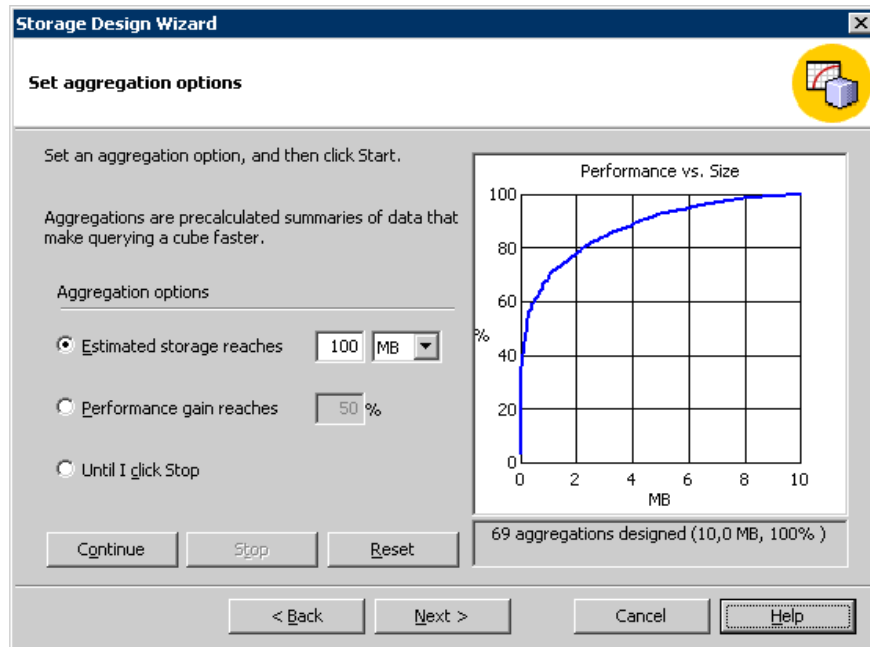
Trūkumai:

- ROLAP lėčiau veikia esant dideliame duomenų kiekiui;
- Dauguma ROLAP priemonių praleidžia skaičiuojamųjų reikšmių lentelių kūrimą, tada nukenčia užklausos vykdymo greitis nes reikia apdoroti didesnius duomenų kiekius. ROLAP naudoja bendros paskirties duomenų bazes, kurios neturi specialių MOLAP naudojamų priemonių (pvz. hierarchinis indeksavimas). Tačiau modernios ROLAP priemonės jau turi SQL kalboje numatytus specialius operatorius (angl. CUBE, ROLLUP).
- ROLAP naudoja SQL užklausas visiems skaičiavimams, todėl ROLAP įrankiai netinka sudėtingiems modeliams, kuriuos aprašyti SQL kalba labai sudėtinga.

3.6. HOLAP ARCHITEKTŪRA

HOLAP (angl. *Hybrid Online Analytical Processing*) technologijos, tai mėginimas suderinti MOLAP ir ROLAP privalumus. Suglausto tipo informacija ir HOLAP kubo įtakos technologija greitesniam veikimui. Kai reikalinga kažkokia detali informacija, HOLAP gali kreiptis iš kubo į bendros paskirties duomenų bazes ir išgauti reikiamą informaciją.

HOLAP architektūros kubo, kuris ima duomenis iš „snaigės“ tipo duomenų schemas, apdorojimo charakteristikos pavaizduotos sekančiame paveiksle (žr. 35 pav.).



Šaltinis: sudaryta autoriaus.

35 pav. HOLAP architektūros charakteristikos, kai naudojama „snaigės“ tipo duomenų schema

Taigi lyginant visus aprašytus OLAP kubų ir duomenų schemų trūkumus ir pranašumus, galime daryti išvadas, kad mažiausia vietas užima OLAP kubas naudojant ROLAP architektūrą bei pasirenkant žvaigždės tipo duomenų schemą. Greičiausia užklausas vykdo MOLAP duomenų analizės kubas.

3 SKYRIAUS IŠVADOS

Tyrimo metu buvo sudaryta „snaigės“ ir „žvaigždės“ tipo duomenų schemas ir pagal kiekvieną iš jų sukurti trijų skirtingų architektūrų OLAP kubai. Tarpusavyje palyginta MOLAP, ROLAP ir HOLAP architektūros duomenų apdorojimo priklausomybė nuo agreguotų duomenų kiekio atsižvelgiant į duomenų schemas struktūra, iš kurios imami kubo duomenys. Kai naudojama „Snaigės“ tipo duomenų schema, nepriklausomai nuo OLAP architektūros, agreguotų duomenų kiekis yra didesnis, nei naudojant „žvaigždės“ tipo schemą. Lyginat tarpusavyje OLAP architektūras MOLAP ir HOLAP architektūras agreguotų duomenų kiekis nesiskiria, tuo tarpu ROLAP architektūros duomenų kiekis beveik padvigubėja.

Remiantis atliktu tyrimu, galima teigti, kad geriausia naudoti „žvaigždės“ duomenų modelį. Renkantis OLAP kubo architektūrą patartina naudoti MOLAP arba HOLAP architektūras.

IŠVADOS

1. Atlikus verslo analitikos sistemų apžvalgą ir analizę nustatyta, kad verslo įmonės gali analizuoti savo klientų ir pardavimų duomenis, naudodamos OLAP duomenų analizės kubus. Taip pagerindamos verslo valdymo procesus ir pagrindinius veiklos valdymo rodiklius tokius kaip: klientų išlaikymas, pardavimų didinimas, naujų klientų pritraukimas.
2. Išanalizuota įmonės, kuri teikia telekomunikacines paslaugas, duomenų analizės sistemos reikalavimai, pateiktos pagrindinių panaudos atvejų sekų diagramos, bei duomenų saugyklos modelio detalus aprašymas ir realizacijos diagramos.
3. Ištirti įmonės veikloje dalyvaujantys elektroniniai duomenys ir nustatyta, kad duomenų analizei bus naudojami duomenys importuoti į duomenų saugyklą.
4. Atliktas duomenų integravimo procesas, pasinaudojant MS SQL Server 2012 integravimo paslaugomis, kurio dėka duomenys iš verslo valdymo sistemų (CRM, Apskaitos sistemų) DB yra išgaunami, atitinkamai pagal poreikius transformuojami bei užkraunami į duomenų saugyklos lenteles.
5. Remiantis pateiktu duomenų analizės įrankių kūrimo modeliu, sukurta ir įdiegta duomenų analizės sistema TEO, LT įmonėje.
6. Numatytas sistemos tobulinimo procesas, kuris gerina jos funkcionalumą:
 - reikalinga papildomai įtraukti šiuos įmonės verslo procesus – detaliau analizuoti paslaugų vartojamumą ir technologinius duomenis.
 - siūloma, kad integravimo projektas būti papildytas automatiniu kubų atnaujinimu po duomenų išsaugojimo į duomenų sandėlį etapo.
 - rekomenduojama pateikti duomenų audito ataskaitą, kurioje būtų informacija apie klaidingus duomenis.
 - siūloma suformuoti duomenų kubų ataskaitas, pasinaudojant MS SQL Server 2012 ataskaitų paslaugomis, kurios būtų integruojamos į įmonės portalą.
7. OLAP kubų architektūrų tyrimo dalyje suprojektuotos „snaigės“ ir „žvaigždės“ tipo duomenų schemas ir pagal kiekvieną iš jų MOLAP, ROLAP ir HOLAP architektūros kubai. Atlikta skirtingų architektūrų duomenų apdorojimo priklausomybė nuo duomenų kiekio, atsižvelgiant į duomenų schemas struktūrą, iš kurios imami kubo duomenys. Remiantis atlikto tyrimo

rezultatais, rekomenduojama „žvaigždės“ tipo duomenų schema ir MOLAP arba HOLAP duomenų kubo architektūra.

8. Nustatyta, kad MOLAP architektūra (palyginus tris OLAP architektūras ROLAP, MOLAP, HOLAP), geriausiai tinka didelės agregacijos kubams: MOLAP architektūra leidžia greitai pateikti informaciją užklausų metu, neilgas duomenų kubo kūrimo laikas. Mažesnės agregacijos kubams reiktų rinktis HOLAP architektūrą. Didelių duomenų analizei, kuri analizuojama retai, geriausia tinkanti ROLAP architektūrą, kuri yra lėtesnė, tačiau užima labai nedaug vietos laikmenose.

LITERATŪROS SĄRAŠAS

1. A. Rozeva. Dimensional Hierarchies – Implementation in Data Warehouse Logical Scheme Design. International Conference on Computer Systems and Technologies - CompSysTech'07., 2007. p. 6.
2. An Overview of Business Intelligence Technology. Communications of the ACM 54, no. 8 (August 2011): 88–98.
3. BI/DW Architecture Fundamentals [žiūrėta 2013.03.01].
<<http://www.dynacrongroup.com/2011/04/bi-dw-architecture-fundamentals/>>.
4. Borgza R.M., Zahari D. (2008) Business intelligence as a competitive differentiator. *IEEE Xplore*.
5. Business intelligence [žiūrėta 2013.03.10].
<http://en.wikipedia.org/wiki/Business_intelligence>.
6. Carlo Verzellis. Business Intelligence: Data mining and optimization for decision making. A John Wiley and Sons, Ltd., Publication. 2009. ISBN: 978-0-470-51138-1
7. Codd E.F., Codd S.B., Salley C.T.. Providing OLAP (On – line Analytical Processing) to User – Analysts: An IT Mandate. 1993
8. Čekavičius V.et. al. Statistika ir jos taikymai: vadovėlis, 2003. –238 p. ISBN 091001444122
9. D. Dzemydienė (2006). Intelektualizuotų informacinių sistemų projektavimas ir taikymas. Vilnius. Mykolo Riomerio universiteto leidybos centras.
10. Dapkevičius, V. (2006). Veiklos analizės IS smulkiojo verslo įmonėms-Business intelligence system for SB market. Lietuvos akademinių bibliotekų tinklas LABT-- Magistro darbas.
11. Data warehouse design onsiderations. [žiūrėta 2013.03.10].
<<http://technet.microsoft.com/library/Cc917644>>.
12. Davidavičienė, V., et al. (2009). Elektroninis verslas: vadovėlis. Vilnius: Technika.
13. James L. (2012, March 1). Top 14 Business Intelligence Predictions. [žiūrėta 2013.03.10]. <<http://smartdatacollective.com/yellowfin/47360/top-14-business-intelligence-predictions-2012>>.

14. James Richardson, Gartner Symposium/Itxpo konferencija, lapkričio 5-8, 2012.
Barcelona, Spain
15. Kas yra verslo analitika. [žiūrėta 2013.02.07]. <<http://www.softera.lt/lt/verslo-analitika/kas-yra-verslo-analitika/>>.
16. L.Kalinauskienė, OLAP sistemų projektavimas žinių modelio pagrindu: Informacinių sistemų projektavimo metodai ir technologijos, 146-150psl.
17. L.Kalinauskienė, OLAP sistemų projektavimas žinių modelio pagrindu: Informacinių sistemų projektavimo metodai ir technologijos, 146-150psl.
18. Liya Wu, Gilad Barash, Claudio Bartolini. A Service-oriented Architecture for Business Intelligence. IEEE International Conference on Service-Oriented Computing and Applications(SOCA'07). 0-7695-2861-9/07. 2007.
19. Marius Vilimas. Lyginamoji OLAP priemonių analizė ir taikymas interneto svetainėse. Straipsnis 2004.
20. MERKEVIČIUS, E. GARŠVA, G. CEPKOVA-TAJA (2005). Intelektualios sprendimų paramos sistemos struktūra kredito rizikos vertinimui. Kaunas.
21. Mikhail Bondarenko, Nikolay Slipchenko, Kateryna Solovyova ir kiti. Systemological business modeling for informatikon systems development. International Book Series "Information Science and Computing". 2009.
22. Morris, Langdon. 2009.Business Model Innovation The Strategy of Business Breakthroughs. International Journal of Innovation Science. Vol. 1 Issue 4.
23. MSDN Library [žiūrėta 2013.02.25]. <<http://msdn.microsoft.com/library/default.aspx>>
24. Mundy J., Thornthwaite W., Kimball R. The Microsoft Data Warehouse Toolkit: With SQL Server 2005 and Microsoft Business Intelligence Toolset. Indianapolis: Willey, 2006. 792p.
25. Mundy J., Thornthwaite W., Kimball R. The Microsoft Data Warehouse Toolkit: With SQL Server 2005 and Microsoft Business Intelligence Toolset. Indianapolis: Willey, 2006. 792p.
26. Regina Jasilionienė ir Rima Tamošiūnienė. Ryšių su klientais valdymo sistemų investicijų efektyvumo vertinimas: investicijų grąžos metodo taikymo teoriniai ir praktiniai aspektai. Verslas: teorija ir praktika. 9(3): 221–228. 2008. ISSN 1648-0627
27. Rukavičius, V., (2005, 2007). *OLAP sistemų analizė*. Kauno technologijos universitetas.

28. Seven Drawbacks of Traditional OLAP. [žiūrėta 2013.02.20].
<<http://datakeyword.blogspot.com/2013/01/seven-drawbacks-of-traditional-olap.html>>
29. Simon, A. R., & Shaffer, S. L. (2001). *Data Warehousing and Business Intelligence for E-Commerce*. Morgan Kaufmann.
30. Software, T. (2012). Top 10 BI Trends for 2013. [žiūrėta 2013.03.10].
<<http://www.slideshare.net/TableauSoftware/top-10-bi-trends-for-2013>>.
31. Songini, M. L. (2003). Real-time Data Warehousing Tests IT's Ability to Serve Businesses, Users Say. *Computerworld*, 37(34), 14.
32. Stock, T. (2011). Using a data warehouse to solve risk, performance, reporting and compliance-related issues. *Journal of securities operations & custody*. - London : Stewart, ISSN 1753-1802, ZDB-ID 24196204. - Vol. 3.2010/11, 4, p. 305-315.
33. The Microsoft Data Warehouse Toolkit, The Toolset, c 2006 by Wiley Publishing
34. The Next Generation of Business Intelligence. Comparing QlikView to Traditional OLAP Tools. [žiūrėta 2013.02.20].
<<http://www.brintex.com/ArticleAdmin/uploads/pressUploads/171/370/TheNextGenerationOfBusinessIntelligence.pdf>>.
35. Theodoros Evgeniou, Phillip Cartwright. Barriers to Information Management. *European Management Journal* Vol. 23, No. 3, pp. 293–299, 2005
36. Tomkūnaitė, L. Nemuraitė „Duomenų saugyklos konceptų modelis“, Kauno technologijos universitetas, 2005
37. Trumpai apie TEO. [žiūrėta 2013.03.10]. <<http://www.teo.lt/apieteo>>.
38. UML basics. (2008) An introduction to the Unified Modeling Language. UML Basics [žiūrėta 2013.03.01]. <http://en.wikipedia.org/wiki/Unified_Modeling_Language>.
39. W.H.Imon, OLAP and Data warehouse, 2000
40. Wu, F., Mahajan, V., & Balasubramanian, S. (2003). An analysis of e-business adoption and its impact on business performance. *Journal of the Academy of Marketing Science*, 31(4), 425–447. doi:10.1177/0092070303255379
41. Žemaitis, V., Milkintis, (2010). *E – verslo plėtros Lietuvoje ekonominė analizė ir įtakos šalies ūkio konkurencingumui tyrimas*. Lithuanian Academic Libraries Network (LABT) / Siauliai University.

Druseikis R. Duomenų saugyklos analizės priemonių taikymas verslo įmonėse / Elektroninio verslo vadybos magistro baigiamasis darbas. Vadovas lekt. dr. R. Naujikienė – Vilnius: Mykolo Romerio universitetas, Socialinės politikos fakultetas, Komunikacijos ir informatikos institutas, 2013. – p. 69

SANTRAUKA

Viena iš elektroninio verslo sandaros dalių yra verslo įžvalgos sistemos. Verslo įžvalgos sistemos apibūdina informaciją apie klientus, konkurentus, rinkas, verslo procesus ir kitus legaliais metodais visiems prieinamoje informacinėje erdvėje surinktus duomenis.

Šio darbo tikslas - pateikti ir įgyvendinti duomenų saugyklos analizės priemonių taikymo modelį, kuris leistų užtikrinti pagrindinius įmonės veiklos procesus. Nagrinėjamos priemonės ir būdai įgyvendinti efektyvų pardavimų ir klientų duomenų analizės modelį, minimaliai įtakojant pagrindinių sistemų darbą. Sukurtas ir realizuotas pavyzdinis projektas, kuris suteikia galimybes atlikti įvairių duomenų analizę.

Pirmoje darbo dalyje yra nagrinėjama mokslinė literatūra. Išanalizavus, kokie gali būti taikomi duomenų analizės sprendimai, pasirinkta OLAP duomenų analizės technologija kaip vienas geriausių sprendimo būdų pasirinktu atveju. Darbe taip pat nagrinėjama verslo įžvalgų, duomenų saugyklų apibrėžimai, paaiškinama architektūra.

Antrojoje darbo dalyje nagrinėjama duomenų analizės sistemų projektavimo metodika. Pagal pasirinktą metodiką projektuojama duomenų analizės sistema, atitinkanti keliamus organizacijos ir vartotojų lūkesčius. Analizuojama projektavimo struktūra.

Trečiojoje darbo dalyje yra atliekama lyginamoji OLAP technologijos architektūrų analizė. Išskiriami privalumai ir trūkumai. Realizavus duomenų analizės projektą skirtingais architektūriniais OLAP kubų sprendimais, gauti duomenys susisteminti ir pateiktos išvados.

Darbo išvadose pateikti gauti rezultatai, rekomendacijos, kokie tinkamiausi duomenų analizės įrankiai priklausomi nuo kintamų parametrų, kurie užtikrina efektyvius verslo procesus. Aprašyti projekto rezultatai, bei galimybės išplėsti ir pritaikyti projektą.

Druseikis R. Application of data warehouse analysis means in business enterprises / Masters's Work in E-business management. Supervisor lect. dr. R. Naujikienė – Vilnius: Mykolas Romeris University, 2013. – p. 69

SUMMARY

One of the building blocks of e-business is a business intelligence system. BI system describes the information about customers, competitors, markets, business processes and other the public information space of the data collected made by legal methods. .

The main goal of this work - to provide and implement the model of application of data warehouse tools for analyzing, which would allow to ensure core business processes of a company's . Relevant tools and techniques are enabling to implement effective sales and model of customer data analysis, which made a minimally influence to the major systems. Designed and implemented a model of project that offers opportunities for a wide range of data analysis.

In the first part of this work you can see the analysis of the scientific literature. After the analysis of the data analysis of solutions are selected OLAP data analysis technologies as one of the best solutions for the selected event. In this work also analyzed the business intelligence, the definitions of data warehousing and explaining architecture.

In the second part of the work were examining the methodology of design of a data analysis system. According to the chosen methodology are projected the data analysis system, which confirm requirements of the users and organization.

In the third part is a comparative analysis of the OLAP technology architectures. Is described advantages and disadvantages. The data are structure and made the conclusions when is realize the data analysis project with different architectural solutions of OLAP cubes,

In the end of the work are made results of the work, recommendations what are the most appropriate of data analysis tools depend on variable parameters, which ensure efficient business processes. Is described the results of the project and the ability to extend and customize the project.