

Univerzitet u Beogradu
Matematički fakultet

Nataša V. Pucanović

**PRINCIPI I METODE INTELIGENTNOG POSLOVANJA –
PRIMENA U NARODNOJ BANCI SRBIJE**

magistarski rad

mentor: dr. Nenad Mitić

Beograd
2010.

PREDGOVOR

Ovaj rad je nastao posle višegodišnjeg rada na projektima razvoja skladišta podataka i sistema poslovne inteligencije uopšte. Cilj rada je da se sistematizuju i prikažu teorijske osnove, koncepti, principi i metode projektovanja sistema poslovne inteligencije, a pre svega da se sistematizuju, objedine i na neki način rezimiraju rezultati koji su u oblasti poslovne inteligencije do sada postignuti u Narodnoj banci Srbije. Rad se sastoji iz tri dela:

U prvom delu izložene su teorijske osnove, tj. osnovni koncepti i metode koji se koriste pri razvoju i implementaciji sistema poslovne inteligencije, odnosno skladišta podataka kao osnove na kojoj se ovi sistemi zasnivaju. Analizirani su glavni pravci koji su danas aktuelni u razvoju skladišta podataka, sa svojim prednostima i nedostacima. Pored toga, dat je presek arhitekture sistema za skladištenje podataka, kao i osnovne faze razvoja.

U drugom delu opisana su, na jednom opštem nivou, tri sistema razvijena u Narodnoj banci Srbije korišćenjem principa i metodologije skladištenja podataka. Ova tri sistema razvijana su u vremenskom periodu od nekoliko godina i sva tri su i danas u upotrebi. Među opisanim sistemima je i najstarije skladište podataka razvijeno u banci, ali i skladište koje je među poslednjima razvijeno i implementirano. U ovom delu rada povučene su neke paralele i izvedene zajedničke karakteristike sva tri skladišta. Objedinjena su i predstavljena iskustva iz kojih su vremenom nastali standardi koji se danas poštuju kada se ulazi u projekte iz oblasti skladištenja podataka.

U prilogima na kraju rada dat je detaljan uvid u tehničku implementaciju svakog od tri skladišta podataka, sa opisima dimenzija, tabela činjenica i poslovnih mera. Na kraju su priloženi i primeri izveštaja uradjeni u odabranim korisničkim alatima. Kada se govori o tehničkom aspektu razvoja skladišta ni u jednom momentu se ne pominju platforme na kojima se radi u Narodnoj banci. Cilj rada je da se pokaže na koji način su primenjeni osnovni BI koncepti u praksi, da se rezimiraju postignuti rezultati i da se naznače pravci kojima se ide dalje, nezavisno od proizvoda koji se nude na tržištu.

U razvoju i implementaciji ovih skladišta podataka bile su, pored autora rada, angažovane kolege zaposlene u IT sektoru, ali i poslovni korisnici bez čijeg aktivnog uključivanja ovi projekti nikada ne bi bili realizovani. Poseban doprinos celom poduhvatu uvođenja skladišta podataka dalo je rukovodstvo. Činjenica je da je upravo rukovodstvo imalo pravu viziju i dovoljno sluha za nove tehnologije i pokrenulo ideju o aktivnijem razvoju skladišta podataka, posebno od 2005. godine. Svaki od ovih projekata je, kao i svi ozbiljni IT poduhvati, rezultat timskog rada. Zato se ovom prilikom zahvaljujem na saradnji svim kolegama koji su učestvovali u razvoju opisanih skladišta podataka i na neki način pomogli da ovaj rad nastane.

Sadržaj

I OSNOVE SKLADIŠTENJA PODATAKA	5
1 UVOD U SISTEME ZA SKLADIŠTENJE PODATAKA.....	6
2 UPOREDNI PREGLED TRANSAKCIONIH SISTEMA I SISTEMA ZA SKLADIŠTENJE .	9
2.1 Transakcioni sistemi i analitičko izveštavanje	9
2.2 Nova generacija – sistemi za skladištenje podataka.....	11
3 ARHITEKTURA SISTEMA ZA SKLADIŠTENJE PODATAKA.....	14
3.1 Pravci razvoja arhitekture	14
3.1.1 CIF arhitektura	14
3.1.2 BUS arhitektura	15
3.1.3 Poredjenje CIF i BUS arhitektura.....	16
3.2 Elementi arhitekture	17
3.2.1 Izvori podataka.....	17
3.2.2 ETL procesi	17
3.2.3 Pripremna zona	19
3.2.4 Dimenzioni prezentacioni sloj	20
3.2.5 Data Mart u CIF i BUS arhitekturi.....	20
3.2.6 Korisnički alati za pristup podacima.....	22
3.3 OLAP	22
4 PROJEKTOVANJE SKLADIŠTA PODATAKA.....	26
4.1 Iterativni razvoj sistema za skladištenje podataka.....	26
4.2 Definisane logičkog modela.....	27
4.2.1 Analiza poslovnog sistema. Izbor poslovnih procesa.....	28
4.2.2 Definisane nivoa detaljnosti. Granularnost	29
4.2.3 Identifikacija poslovnih dimenzija	31
4.2.4 Identifikacija poslovnih mera.....	32
4.2.5 Princip magistrale. Jednobraznost dimenzija i činjenica	32
4.3 Dimenziono modeliranje	35
4.3.1 Zvezdasto dimenziono modeliranje.....	35
4.3.1.1 Tabele činjenica.....	36
4.3.1.2 Tabele dimenzija	38
4.3.2 Model pahuljice	40
4.3.3 Hijerarhije. Tehnike „bušenja“ podataka.....	42
5 AGREGATI U SKLADIŠTIMA PODATAKA.....	44
6 EKSTERNI PODACI U SKLADIŠTIMA PODATAKA.....	48
7 FIZIČKI MODEL SKLADIŠTA PODATAKA	50
7.1 Partitionisanje.....	50
7.2 Indeksiranje	51
7.3 Paralelizam.....	52
8 SKLADIŠTA PODATAKA I TEHNOLOGIJE	53
9 METAPODACI.....	56

II PRIMENE KONCEPATA SKLADIŠTENJA PODATAKA U NARODNOJ BANCII SRBIJE ..59	
1 UVODNE NAPOMENE	60
2 OPIS POSLOVNIH SISTEMA. OBUHVATNOST SKLADIŠTA.....	62
2.1 Penzijski fondovi.....	62
2.2 Kreditni odnosi sa inostranstvom	63
2.3 Bilansni podaci poslovnih banaka	63
3 OKRUŽENJE. ULAZNI PODACI. NEDOSTACI POSTOJEĆIH SISTEMA	65
3.1 Penzijski fondovi.....	65
3.2 Kreditni odnosi sa inostranstvom	66
3.3 Bilansni podaci poslovnih banaka	68
4 ZAJEDNIČKI CILJEVI I MOTIVI ZA UVODJENJE SKLADIŠTA PODATAKA	70
5 PROJEKTOVANJE SKLADIŠTA PODATAKA.....	72
5.1 Fazni razvoj sistema	72
5.2 Pripremna zona	74
5.3 Ažuriranje skladišta podataka. ETL strategije	75
5.3.1 Penzijski fondovi.....	76
5.3.2 Kreditni odnosi sa inostranstvom	77
5.3.3 Bilansni podaci poslovnih banaka	78
5.4 Fizička implementacija. Statistika rasta baza	80
5.5 Prezentacioni sloj.....	81
5.6 Završne napomene	82
6 KORISNICI SKLADIŠTA PODATAKA.....	84
7 DALJI PLANOVI	87
8 ZAKLJUČAK	89
III PRILOZI: PRIMERI IMPLEMENTACIJE	90
A. Penzijski fondovi.....	91
A.1. Izvori podataka	92
A.2. Opis poslovnih funkcija	93
A.3. Fazni razvoj sistema	94
A.4. Opis poslovnih dimenzija	96
A.4.1. Vremenska dimenzija	96
A.4.2. Dimenzija poslovnih banaka.....	98
A.4.3. Dimenzija penzijskih fondova	99
A.4.4. Dimenzija hartija od vrednosti.....	101
A.4.5. Dimenzija hijerarhijskog šifriranja podataka.....	102
A.4.6. Ostale dimenzije	104
A.5. Opis tabela činjenica	106
A.5.1. I faza	106
A.5.2. II faza	110
A.5.3. III faza	112
A.5.4. IV faza.....	113
A.5.5. V faza	115
A.6. Primeri izveštavanja.....	120
B. Kreditni odnosi sa inostranstvom	124
B.1. Projektovanje skladišta podataka.....	125
B.1.1. Dimenzije u podsistemima kreditnih zaduženja i kreditnih odobrenja.....	125
B.1.2. Poslovne mere u podsistemu kreditnih zaduženja.....	129
B.1.3. Poslovne mere u podsistemu kreditnih odobrenja	132
B.1.4. Tabela međjuvalutnih odnosa.....	133

B.2. Primeri izveštavanja.....	134
C. Bilansni podaci poslovnih banaka.....	137
C.1. Projektovanje skladišta podataka.....	138
C.1.1. Dimenzije u skladištu podataka.....	138
C.1.2. Tabele činjenica u skladištu podataka	141
C.2. Primeri izveštavanja.....	147
LITERATURA.....	150

I OSNOVE SKLADIŠTENJA PODATAKA

1 UVOD U SISTEME ZA SKLADIŠTENJE PODATAKA

Postoji više definicija skladišta podataka. Brojni analitičari iz ove oblasti, kao i veliki softverski proizvođači, dali su svoje definicije pojma i koncepta skladišta, odnosno skladištenja podataka, ali se ovaj rad zasniva na materijalima koje su objavili William Inmon [1] i Ralph Kimball [2] sa saradnicima, kao i manjim delom na knjigama i radovima nekoliko pratećih autora [4][6].

Neke od definicija skladišta i skladištenja podataka su:

- Skladište podataka je poslovno orijentisan, integrisan, nepromenljiv i vremenski odredjen skup podataka koji predstavlja centralnu bazu podataka za sisteme poslovne inteligencije [1].
- Skladište podataka je kopija transakcionih podataka specijalno strukturiranih za upite i analize [2].
- Skladište podataka je strukturiran repozitorijum subjekt orijentisanih, vremenski odredjenih, istorijskih podataka celokupnog preduzeća, iz kojih se izvode poslovne informacije koje omogućavaju donošenje poslovnih odluka [5].
- Skladištenje podataka predstavlja infrastrukturu i arhitekturu koja obezbedjuje uspešan, kvalitetan i, sa stanovišta preduzeća, sveobuhvatan sistem poslovne inteligencije. Dakle, skladištenje podataka nije nova tehnologija, već arhitektura, kojoj je neka od postojećih tehnologija neophodna da bi bila implementirana [1].

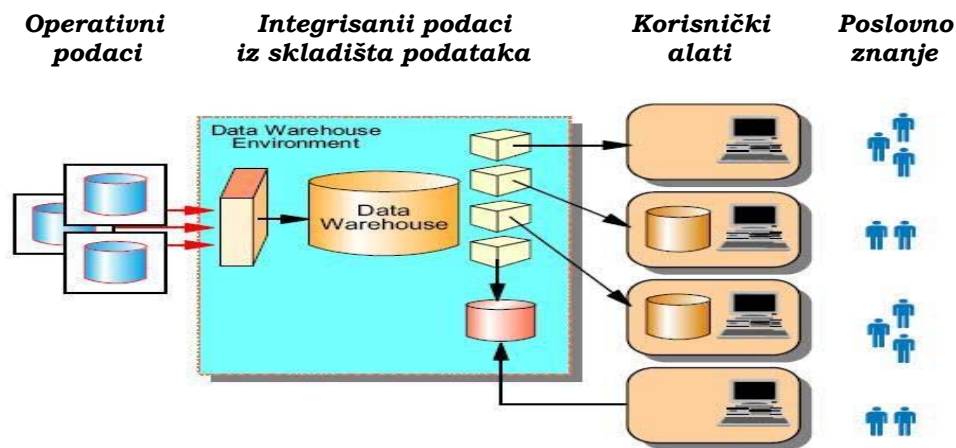
Osnovni zahtevi, prema [2], koji stoje pred sistemima za skladištenje podataka su:

- Poslovne informacije moraju biti lako dostupne. Sadržaj mora biti razumljiv, pregledan i intuitivan, kako po značenju tako i po imenovanju i to sa stanovišta krajnjih korisnika, a ne IT stručnjaka. Alati koji se krajnjim korisnicima stavljaju na raspolaganje u cilju kombinovanja podataka iz baze skladišta moraju biti laki za korišćenje. Cilj je da ti alati vraćaju rezultate upita korisnicima u što kraćem vremenu imajući u vidu masovnost podataka u bazama skladišta i veličinu tih baza koja se meri gigabajtima i terabajtima.
- Poslovne informacije moraju biti predstavljene konzistentno. Na primer, ako dva podatka u sistemu predstavljaju istu poslovnu meru moraju se imenovati identično. Obrnuto, ako dve stavke u sistemu ne označavaju isti poslovni pojam, ne smeju se isto imenovati. Konzistentnost podrazumeva da su svi predstavljeni podaci kompletni, pažljivo prikupljeni iz različitih izvora, prečišćeni i predstavljeni korisniku tek onda kada su spremni za upotrebu.

- Sistem mora biti adaptibilan i lako prilagodljiv promenama poslovnih procesa, potreba korisnika, uslova poslovanja, tehnologija i sl. Skladište podataka mora biti projektovano tako da može lako da uključi novonastale promene (nove podatke na zahtev korisnika) a ne naruši postojeće podatke ili aplikacije.
- Skladište podataka mora da predstavlja sigurnu zaštitu od neovlašćenog pristupa svim najznačajnijim i najpoverljivijim informacijama o poslovanju jedne organizacije (npr. šta se prodaje, kome i po kojoj ceni). Zato se od sistema za skladištenje očekuje kontrolisani pristup svim poverljivim informacijama.
- Skladište podataka mora da služi kao osnova za brže i kvalitetnije donošenje poslovnih odluka (koje proizilazi iz analize podataka koji se nalaze u bazi skladišta) primenom tehnika poslovne inteligencije kao što su OLAP, izveštavanje, istraživanje podataka (eng. *Data Mining*) i dr.

Dakle, sistem za skladištenje podataka predstavlja kolekciju integrisanih, detaljnih, sumarnih, kako tekućih tako i istorijskih podataka koji će biti osnovna podrška rukovodstvu u odlučivanju i omogućiti jednom preduzeću da ima sveobuhvatan uvid u svoje poslovanje.

Mesto skladišta podataka u svetu informacionih sistema prikazano je na slici 1.



Slika 1. Skladište podataka: od podatka do poslovnog znanja

Koncept inteligentnog poslovanja (poslovne inteligencije, eng. *Business Intelligence*) čini skup softverskih alata i metodologija koji omogućavaju korišćenje podataka iz skladišta podataka (eng. *Data Warehouse*) i njihovo pretvaranje u informacije potrebne za donošenje adekvatnih poslovnih odluka. Ovaj koncept postao je posebno zanimljiv u segmentima prodaje, marketinga i finansijskog poslovanja.

Prema jednoj od definicija analitičara Gartner grupe inteligentno poslovanje predstavlja proces prevodjenja (transformacije) podataka u poslovne informacije i dalje kroz proces iterativnog istraživanja transformaciju informacija u poslovno znanje.

Iz informatičkog ugla, inteligentno poslovanje obezbeđuje korisnicima analitički pristup podacima iz poslovanja preduzeća, tj. daje mogućnost analize podataka u cilju praćenja trendova u poslovanju, davanja odgovora na ključna poslovna pitanja i, u krajnjoj liniji, predstavlja podršku poslovnom odlučivanju. Sa poslovne tačke gledišta, inteligentno poslovanje predstavlja proces prikupljanja visoko-kvalitetnih informacija o poslovnim temama, koje analitičarima poslovanja omogućavaju donošenje zaključaka.

Akcent sistema za inteligentno poslovanje, pa i skladištenja podataka kao koncepta, tj. arhitekturnog okruženja na kome počivaju ovi sistemi nije na sirovom podatku, već na informaciji koja podrazumeva preradjene, prečišćene, objedinjene i organizovane podatke koji potiču iz različitih, nezavisnih izvora: postojećih transakcionih sistema, kao i eksternih izvora (dokumenata, poslovnih knjiga i sl.). Kvalitetna poslovna informacija mora da bude, po značenju i sadržaju, razumljiva i upotrebljiva grupama korisnika na različitim nivoima odlučivanja: operativnom, srednjem ili strateškom nivou odlučivanja.

2 UPOREDNI PREGLED TRANSAKCIONIH SISTEMA I SISTEMA ZA SKLADIŠTENJE

2.1 Transakcioni sistemi i analitičko izveštavanje

Transakcioni sistemi, poznati kao OLTP (eng. *On Line Transactional Processing*) sistemi, podržavaju izvršavanje poslovnih procesa i cilj im je da „uhvate“ svaki pojedinačni događaj, tj. svaku pojedinačnu transakciju tog poslovnog procesa i sve detalje vezane za tu transakciju. Primeri pojedinačnih transakcija su: unos porudžbenice, evidentiranje izvršenog plaćanja, unos izlazne fakture,... Izvršavanje svake pojedinačne transakcije rezultira formiranjem novih slogova podataka, čitanjem, izmenom ili brisanjem postojećih podataka. Zato, da bi obezbedili uspešno izvršavanje poslovnog procesa kroz evidentiranje njegovih pojedinačnih transakcija, transakcioni sistemi moraju da omoguće brze osnovne operacije kao što su: postavljanje upita, upis novih podataka, izmene i brisanje postojećih podataka. Upiti nad podacima transakcionih sistema odlikuju se većom predvidljivošću, lakše ih je unapred definisati, kako zbog strukture podataka tako i zbog obima podataka u sistemu i same prirode korisnika takvih sistema.

Transakcione sisteme odlikuju:

- Složene strukture podataka.
- Normalizovani podaci, predstavljeni, najčešće u trećoj normalnoj formi.
- Visoke performanse, pre svega brzina čitanja, upisa i izmene slogova.
- Kratko vreme odziva.
- Raspoloživost krajnjim korisnicima.
- Pouzdanost.

Pored brojnih prednosti, višegodišnje primene, iskustva u projektovanju i korišćenju kao i naprednih alata za projektovanje i korišćenje, transakcioni sistemi su se, uglavnom, pokazali nepogodni i malo upotrebljivi kada je reč o analizi kvaliteta poslovnih informacija, praćenju trendova, ad-hok izveštavanju ili donošenju poslovnih odluka.

Razlog se nalazi, pre svega, u činjenici da su podaci necentralizovani kao i da se čuvaju na različitim aplikativnim sistemima, često i na tehnološki različitim platformama. Osim toga, na razvoju svake transakcione aplikacije rade, uglavnom, različiti projektni timovi koji se bave samo domenom konkretne aplikacije ne sagledavajući celokupno poslovanje.

Sve to vodi redundansi podataka u transakcionim aplikacijama. Takve podatke je teško usaglasiti da bi se dobila jedna jedinstvena slika poslovnog sistema. U transakcionim sistemima podaci nisu uvek bliski po značenju krajnjim korisnicima, jer su ovi sistemi isključivo aplikativno, a ne poslovno orijentisani. Podaci su previše detaljni i kao takvi nisu pogodni za analitičko izveštavanje. Problem u sistemima koji se aplikativno razvijaju često predstavlja i neusaglašenost podataka u različitim aplikacijama po tipu ili formatu. Tako, na primer, telefonski broj je u jednoj aplikaciji dužine 15, a u

drugoj dužine 20 karaktera; jedna ista poslovna mera u jednoj aplikaciji može biti predstavljena u centimetrima, u drugoj u inčima, u trećoj u nekoj trećoj mernoj jedinici; pol zaposlenog u jednoj aplikaciji je vrednost iz domena 1 ili 0, u drugoj aplikaciji iz domena m ili f, u trećoj ž ili m.

Prema [1], glavni nedostaci klasičnih transakcionih sistema su:

- Kredibilitet podataka. Kredibilitet podataka dobijenih iz različitih transakcionih sistema jednog preduzeća dovodi se u pitanje posebno u većim preduzećima koja se sastoje iz više organizacionih delova u kojima se poslovni procesi prepliću i koja dele iste poslovne informacije. Česta pojava u praksi je da dva organizaciona dela jedne firme dostave izveštaje rukovodstvu o nekom segmentu poslovanja sa potpuno različitim rezultatima. Bitni pokazatelji poslovanja mogu biti toliko neusaglašeni izmedju organizacionih delova da to potpuno onemogućava svako donošenje poslovnih odluka na osnovu dobijenih izveštaja. Razlozi ovakve nesinhronizacije izmedju organizacionih delova mogu biti sledeći:
 - Podaci su izvedeni iz transakcionih sistema u različitim vremenskim trenucima. Kako se podaci u transakcionim sistemima neprestano menjaju sa svakom novom transakcijom, to su i dobijene vrednosti istih poslovnih mera različite.
 - Algoritmi kojima se podaci izvlače iz transakcionih sistema se razlikuju od jednog do drugog organizacionog dela.
 - Postoji uvek problem eksternih podataka sakupljenih iz različitih časopisa, sa sajtova ili dobijenih putem elektronske pošte, koji na različite načine mogu biti obradjeni u različitim odeljenjima.
 - Ne postoje uvek zajednički izvori podataka: na primer, jedan organizacioni deo svoje podatke čuva u datotekama dok drugi organizacioni deo podatke čuva u svojoj bazi podataka. Izmedju ovih heterogenih izvora ne postoji mogućnost deljenja zajedničkih podataka.
- Produktivnost. Problem produktivnosti transakcionih sistema posebno dolazi do izražaja u situacijama kada treba analizirati veće količine podataka i to ne samo iz jednog organizacionog dela, već celog preduzeća. Problem je, najpre, locirati ove podatke pretražujući datoteke, tabele ili izveštaje po svim organizacionim delovima. Sledi upoređivanje podataka izmedju organizacionih delova: utvrditi da li su isti poslovni pojmovi različito imenovani, kao i da li su različiti poslovni pojmovi identično imenovani. Dakle, treba napraviti detaljnu uporednu analizu svih podataka koje firma poseduje. Izvori za dobijanje sveobuhvatnog poslovnog izveštaja su brojni i potrebno je napisati isto toliko programa za izvlačenje podataka iz tih izvora. U realnom poslovnom sistemu zahtevi za ovakvim, objedinjenim izveštajima postaju sve češći i skoro nikada nisu jednokratni. Novi zahtev iziskuje

svaki put jednako veliki napor zaposlenih, objedinjavanje podataka uvek traje dugo, efikasnost i produktivnost se ozbiljno narušavaju.

- Nemogućnost transformacije podatka u informaciju. Ovo je jedan od ključnih nedostataka klasičnih transakcionih sistema. Dobiti odgovor na uobičajeno bankarsko pitanje: "Koliko se aktivnost određenog računa razlikuje ove godine u odnosu na prethodnih pet godina" je gotovo nemoguće iz klasičnih sistema zato što su aplikativno, a ne subjekt orijentisani, neintegrisani i ne sadrže istorijske podatke.

Dakle, transakcioni sistemi rade na principu tekućih, trenutno važećih podataka nad tzv. primitivnim (atomskim) podacima koji imaju sledeće karakteristike:

- Primitivni podaci se menjaju kako promena nastane i uvek sadrže tekuću vrednost (npr. kada kupac promeni adresu stara više nije neophodna i ona se zamenjuje novom).
- Oni mogu da se obradjuju na predefinisani, predvidiv način, programima i procedurama koje se ponavljaju.
- Nema potrebe za istorijskim podacima, već se tekući podaci kada istekne neko vreme njihovog korišćenja brišu iz sistema ili stavljaju u arhivu (npr. podaci o prodaji se nakon izdavanja finansijskog izveštaja sklanjaju iz transakcionog sistema).
- Prilagodjeni su poslovnim funkcijama operativaca, a ne analitičara poslovanja i rukovodećih struktura.

2.2 Nova generacija – sistemi za skladištenje podataka

Sistemi za skladištenje podataka imaju fundamentalno različitu namenu, aspekt korišćenja, metode modeliranja i projektovanja, kao i tehnološki aspekt u odnosu na klasične transakcione sisteme.

Sistemi za skladištenje podataka se mogu okarakterisati kao:

- Poslovno, tj. subjekt orijentisani. Za razliku od klasičnih transakcionih, aplikativno orijentisanih sistema u sistemima za skladištenje podataka podaci se logički organizuju oko jednog poslovnog subjekta. Tipični subjekti poslovanja su: kupac, proizvod, račun, osiguravajuća polisa i sl. U ovakvoj organizaciji podataka deo ključa je uvek vremenska komponenta, tako da su podaci o subjektu uvek vremenski određeni. To mogu biti detaljni, dnevni podaci kao i sumarni mesečni, kvartalni ili godišnji podaci.
- Integrisani. Ovo je najvažnija od svih karakteristika koncepta skladištenja podataka. Podaci se prikupljaju iz različitih homogenih ili heterogenih izvora, konvertuju, reformatiraju, sumiraju, sortiraju kako

bi se neutralisale sve nekonzistentnosti koje su postojale u izvoru podataka (u transakcionoj aplikaciji, ili nekom drugom spoljnom izvoru podataka: poslovnoj knjizi, datoteci, dokumentu itd.). Proces transformacije i integracije zahteva vreme, ima svoju cenu i zahteva podršku i saglasnost svih organizacionih delova, posebno najvišeg rukovodstva, koje donosi odluke, raspolaže resursima i budžetom.

- Nepromenljivi. Podaci u sistemima za skladištenje podataka se, uglavnom, ne brišu i ne menjaju, već se periodično, u ciklusima dodaju nove serije podataka. Nove serije podataka obuhvataju nova pojavljivanja i, eventualno, korekcije postojećih podataka u skladištu. Ciklusi osvežavanja baze podataka, odnosno upis novih serija podataka, uslovljeni su potrebama korisnika, definiše ih projektni tim i ne zavise od nivoa detaljnosti na kome se nalaze sami podaci. Na primer, podaci u skladištu podataka mogu biti čuvani na dnevnom nivou, a ciklus upisa novih serija može biti jednom nedeljno, ili jednom mesečno.
- Vremenski odredjeni. Svaki slog u skladištu podataka je vremenski označen, odnosno zna se vremenski trenutak u kome je taj slog bio važeći. Osnovna komponenta ključa u sistemima za skladištenje je vremenska komponenta, koja u klasičnim, transakcionim sistemima ne mora da postoji. Vremenska odredjenost podataka je važna ako se uzme u obzir da je u sistemima za skladištenje podataka akcenat na istorijskim podacima, a manje na operativnim, dnevnim podacima. Istorijski podaci mogu biti starosti do deset godina, u nekim slučajevima i duže. To je jedna od razlika između sistema za skladištenje podataka i transakcionih sistema, koji se zasnivaju na tekućim, trenutno važećim podacima.

Neke bitne razlike između sistema za skladištenje podataka, tj. analitički orijentisanih sistema i transakcionih sistema su:

- Akcenat u sistemima za skladištenje nije na izvršavanju pojedinačne transakcije, već oceni, vrednovanju, odnosno merenju poslovnih procesa sagledavanjem rezultata svih izvršenih transakcija.
- Interakcija sa podacima u sistemima za skladištenje odvija se, isključivo, kroz postavljanje upita nad podacima.
- Upiti za pristup podacima, uglavnom, nisu unapred definisani, pa se kaže da su zahtevi za podacima nepredvidivi, promenljivi i odredjeni potrebama krajnjih korisnika, tj. potrebama posla.
- Skup podataka u skladištima podataka nisu samo trenutno važeći, tekući podaci u poslovnim sistemima, već i istorijski podaci.
- Sa aspekta projektovanja, treća normalna forma zamenjena je dimenzionalnim modelom (zvezdastom šemom ili nekom njenom modifikacijom).

Uporedni pregled transakcionih sistema i sistema za skladištenje podataka prikazan je u Tabeli 1.

OSOBINE	TRANSAKCIONI SISTEMI	SKLADIŠTA PODATAKA
Vreme odziva sistema	Milisekunde i sekunde	Sekunde, minuti, sati
Operacije nad bazom podataka	Upis, izmene, brisanje, čitanje podataka	Uglavnom čitanje podataka
Priroda podataka	Tekući, trenutno važeći podaci. Obično su stari do 30 dana	Upisuju se serije podataka u vremenskim razmacima. Istorijski podaci mogu da budu stari i po nekoliko godina
Organizacija podataka	Aplikativno orijentisani	Subjekt orijentisani, vezani za poslovni pojam i sa ugradjenom vremenskom komponentom
Obim podataka	Mali do velikih	Veliki i veoma veliki. Baze podataka mere se gigabajtima i terabajtima
Aktivnosti u sistemu	Transakciona obrada	Analitičke aktivnosti u cilju donošenja poslovnih odluka
Korisnici	Operativci	Svi nivoi odlučivanja u jednom preduzeću: operativni, srednji, najviši nivo

Tabela 1. Uporedni pregled karakteristika transakcionih sistema i sistema za skladištenje podataka

Još jedna razlika između standardnih transakcionih sistema i sistema za skladištenje podataka je u hardverskim zahtevima svakog od ovih okruženja. Smatra se da su tokom životnog ciklusa jednog transakcionog sistema zahtevi za hardverskim resursima relativno konstantni, nepromenljivi i predvidivi, sa manjim ili većim skokovima. Kada je reč o sistemima za skladištenje podataka hardver je, zbog samog načina kako se baza skladišta puni podacima i kako im se pristupa, ili potpuno zauzet, ili nije uopšte zauzet. Pri tome, pokušaji da se izračuna procenat zauzetosti, ili predvide vremenski momenti u kojima bi to zauzeće bilo najveće nisu se pokazali efikasnim i korisnim. Razlika u hardverskim zahtevima između transakcionih sistema i sistema za skladištenje podataka je razlog zbog koga dva ovakva sistema ne treba da budu na istoj mašini. Pravilo je da se mašina optimizuje i parametrizuje ili za transakciono okruženje ili za skladište podataka, nikada za oba.

3 ARHITEKTURA SISTEMA ZA SKLADIŠTENJE PODATAKA

U ovom poglavlju biće detaljnije opisani elementi koji su obavezan deo arhitekture svakog sistema za skladištenje podataka. Takođe, biće pomeniti i neobavezni elementi arhitekture na kojima posebno insistiraju neki analitičari sistema za skladištenje podataka i koji su se u praksi pokazali veoma korisnim.

3.1 Pravci razvoja arhitekture

Prema [6] postoje dva dominantna pravca kada je reč o arhitekturi sistema za skladištenje podataka. Prvi pravac propagira B. Inmon, koji se smatra „ocem“ skladištenja podataka. Drugi pravac propagira R. Kimball, danas zasigurno najuticajnija osoba u oblasti skladištenja podataka. Ova dva dominantna pristupa razvoju arhitekture skladišta podataka poznata su pod nazivom CIF i BUS arhitektura.

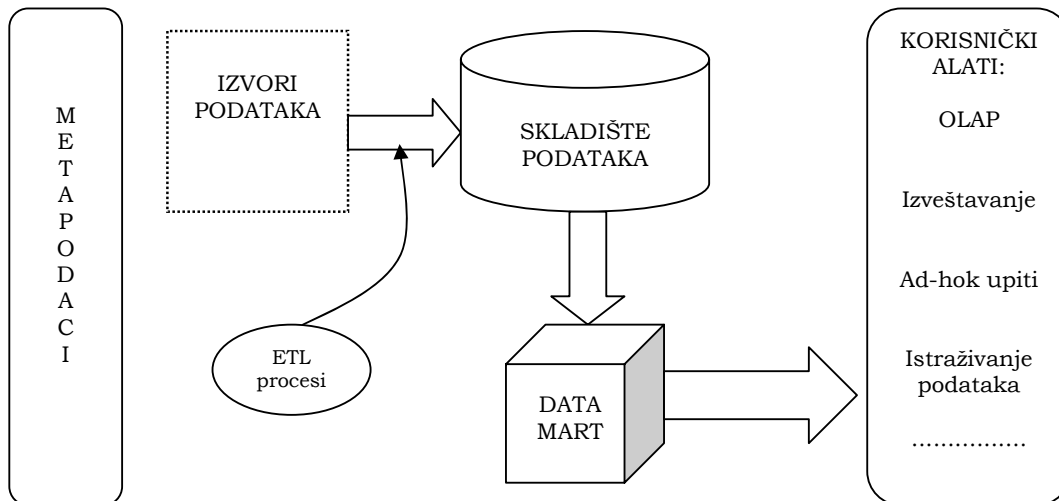
3.1.1 CIF arhitektura

U CIF (eng. *Corporate Information Factory*) arhitekturi skladište podataka je normalizovano i sadrži atomske podatke. Takve podatke kasnije će koristiti tzv. „ispostave podataka“, odnosno Data Mart¹ podsistemi.

U ovoj arhitekturi, podaci se iz operacionalnih, izvornih sistema prevode, transformišu i sakupljaju u sveobuhvatno, korporativno skladište podataka. Skladište podataka sadrži istorijske podatke koji se ne ažuriraju, kao i neke izvedene podatke i služi kao izvor svakom pojedinačnom Data Mart-u. S obzirom da skladište predstavlja spremište podataka na nivou čitave kompanije, ono mora da sadrži atomske podatke u trećoj normalnoj formi. S druge strane, Data Mart podsistemi su dimenziono struktuirani i sadrže samo izvedene, agregirane podatke. Implementirani su tako da odgovaraju specifičnim potrebama odeljenja koje ga koristi. Korisnici CIF arhitekture pristupaju Data Mart-ovima preko korisničkih alata i podatke prilagođavaju svojim potrebama.

Na slici 2. dat je uprošćeni prikaz CIF arhitekture sistema za skladištenje podataka.

¹ Pojam Data Mart-a biće detaljnije opisan u nastavku rada



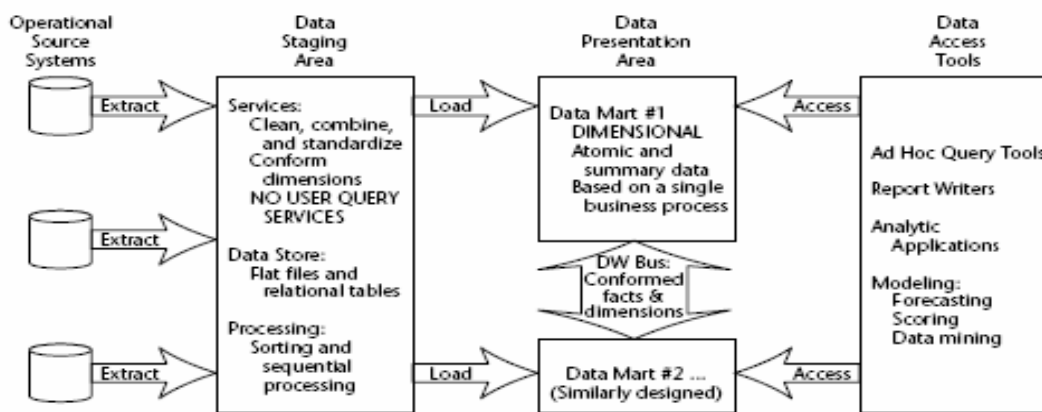
Slika 2. Osnovne komponente CIF arhitekture sistema za skladištenje podataka

3.1.2 BUS arhitektura

Po ovom pristupu skladište podataka se posmatra kao skup različitih Data Mart podsistema, koji poseduju konformisane² (usaglašene) dimenzije. Ovakvi Data Mart-ovi se grade tako što se direktno iz operacionálnih sistema izdvajaju podaci koji su značajni za pojedini poslovni proces. U ovoj arhitekturi skladište podataka je okrenuto poslovnim procesima, a ne pojedinim organizacionim jedinicama. Autor ovog modela, R. Kimbal, insistira da se pri tome koristi višedimenzionalni model podataka (i to striktno zvezdasta šema³) i takođe ističe da Data Mart-ovi ne smeju da sadrže samo agregirane podatke, već moraju sadržati i atomske podatke. Samo na ovaj način se može izaći u susret najrazličitijim zahtevima korisnika. Stoga je, po njemu, neprihvatljiva opcija da se u Data Mart-ovima čuvaju samo agregirani podaci, dok atomski ostaju “zarobljeni” u okviru normalizovanih modela. Kimball smatra da se Data Mart-ovi moraju graditi tako da zadovoljavaju princip konformnosti dimenzija i činjenica, kako bi se omogućila njihova integracija. Na taj način dobijeni podsistemi prodaje, nabavke, finansija, magacina dele zajedničke informacije izmedju sebe, jer su povezani skupom zajedničkih ili usaglašanih dimenzija. Za takve podsisteme kaže se da se nalaze na zajedničkoj „magistrali“ (eng. *Bus*) i svi zajedno čine deo arhitekture označen kao prezentacioni sloj na slici 3.

² Pojam konformnosti biće detaljnije opisan u nastavku rada

³ Pojam zvezdaste šeme biće detaljnije opisan u nastavku rada



Slika 3. Osnovne komponente BUS arhitekture sistema za skladištenje podataka

Još jedna razlika u odnosu na arhitekturu sistema prikazanu na slici 2. je postojanje pripremljene zone za prihvatanje podataka iz izvornih sistema. Kimball posebno insistira na postojanju prihvatne zone podataka. Svi elementi oba tipa arhitekture, prikazani na slikama 2. i 3., biće opisani u poglavlju 3.2.

3.1.3 Poredjenje CIF i BUS arhitekture

Osnovna razlika između CIF i BUS arhitekture, dakle, jeste u tome što prvo navedena podrazumeva da se pre učitavanja dimenzionalnih modela izvrši normalizacija podataka. Druga razlika se ogleda u različitim tretiranjima atomskih podataka. CIF arhitektura podrazumeva da se atomski podaci skladište u okviru normalizovanog skladišta podataka, dok BUS arhitektura insistira na tome da i atomski podaci moraju imati dimenzionalnu strukturu. Iako oba koncepta smatraju da se mora postići izvestan stepen integracije i koordinacije na nivou kompanije, ovaj se zahtev različito realizuje. Tako se kod CIF arhitekture ovaj zadatak rešava uvođenjem normalizovanog skladišta podataka, dok BUS arhitektura integraciju postiže uvođenjem konformisanih dimenzija.

Moguće je primeniti i hibridni pristup, tj. usvojiti najbolje karakteristike obe arhitekture. Iz CIF-a se preuzima normalizovano skladište podataka, ali mu se pridodaje dimenzionalno skladište koje je prisutno kod BUS arhitekture, koje će skladištiti atomske i sumarne podatke. Osnovni nedostatak ovog pristupa leži u redundantnosti atomskih podataka. Zbog toga ga uglavnom koriste kompanije koje su već implementirale normalizovano skladište podataka, ali žele i da koriste dobre strane BUS arhitekture.

Izbor najpodesnije arhitekture zavisi od više faktora. Ako je organizaciona struktura kompanije takva da postoji potreba za uskom povezanošću organizacionih jedinica, i za deljenjem informacija među njima, onda je bolje prihvatiti BUS arhitekturu. Ovaj izbor je poželjniji i u slučaju da kompanija želi brzo da vidi prve rezultate ulaganja u izgradnju skladišta podataka, s obzirom da centralizovano skladište, koje propagira Inmon, zahteva značajno vreme za implementaciju. Ne treba zaboraviti i da ovakav sveobuhvatan

projekat zahteva i značajna novčana ulaganja. Ovo je značajan razlog zbog kojeg se sve više kompanija opredeljuje za Kimball-ov pristup. Ovo, sa stanovišta troškova, predstavlja poželjnije rešenje, jer u slučaju da kompanija nema dovoljno novca da odmah krene sa izgradnjom celokupnog skladišta podataka, ona može najpre ulagati u dizajn i implementaciju podsistema koji će podržati najznačajnije poslovne procese.

3.2 Elementi arhitekture

U ovom poglavlju biće detaljnije opisani svi pomenuti elementi oba tipa arhitekture sistema za skladištenje podataka.

3.2.1 Izvori podataka

Sa aspekta sistema za skladištenje podataka izvori podataka mogu biti:

- Transakcioni sistemi koji podržavaju izvršenje svake pojedinačne transakcije. Transakcioni sistemi rade nad bazama podataka koje mogu biti: relacione, hijerarhijske, mrežne i sl., smeštene na različitim tehnološkim platformama. Ove baze podataka ne sadrže istorijske podatke.
- Arhivski podaci. Ovo su podaci iz transakcionih sistema koji su, nakon dostizanja određene starosti, arhivirani i smešteni u arhivske baze.
- Interne datoteke koje se razmenjuju u okviru organizacionih delova preduzeća.
- Eksterni podaci koji dolaze u preduzeće putem elektronske pošte, skidaju se sa Veb sajtova ili iz časopisa.

3.2.2 ETL procesi

Upis podataka iz izvornih sistema u skladište podataka vrši se u tri koraka, i to:

Korak 1: Izdvajanje (eng. *Extract*) podataka je proces izbora i izvlačenja operativnih podataka, interesantnih sa aspekta skladištenja iz različitih izvornih sistema. U arhitekturama u kojima postoji pripremna zona ovaj korak podrazumeva prebacivanje tako izdvojenih podataka u pripremnu zonu radi dalje obrade.

Korak 2: Transformacija (eng. *Transform*) podataka je proces integrisanja izdvojenih podataka i njihove dalje provere, prečišćavanja, reformatiranja i usaglašavanja pre upisa u samu bazu skladišta.

Korak 3: Upis (eng. *Load*) tako pripremljenih podataka u ciljno skladište, odnosno punjenje skladišta podacima. U arhitekturama u kojima postoji

pripremna zona ovaj korak podrazumeva punjenje skladišta transformisanim podacima koji su iz izvornih sistema izdvojeni i prebačeni u pripremnu zonu.

Izdvajanje, transformacija i punjenje skladišta implementiraju se kroz seriju programskih procedura koje imaju za cilj:

- Automatizovano izvlačenje podataka iz izvornih sistema kojih može biti puno i koji mogu da budu nehomogeni. U takvim sistemima često se nalaze podaci u različitim formatima, neusaglašeni po značenju i imenovanju.
- Transformaciju podataka u formate koji su potrebni za skladištenje podataka.
- Izvodjenje sumarnih podataka u cilju poboljšanja performansi upita koji će se postavljati nad podacima iz skladišta. Ukoliko u sistemu postoji pripremna zona, izvodjenje sumarnih podataka može da se vrši u pripremljenoj zoni pre upisa u samo skladište, ili nakon punjenja skladišta u samoj ciljnoj bazi.
- Inicijalno punjenje skladišta se vrši jednokratno podacima iz arhivskih baza. Reč je o masovnom punjenju velikih količina podataka. Inicijalno napunjeni podaci poslužiće kao istorijski podaci i imaju poseban značaj u analizi dužih vremenskih serija. Arhivski podaci su jednokratni izvor podataka i neće se koristiti u daljim osvežavanjima skladišta podataka.
- Otkrivanje promena nastalih nad izvornim podacima.
- Osvežavanje baze skladišta podataka.

Posebnu važnost, kada je reč o ETL (eng. *Extract Transform Load*) procesima, ima faza transformacije. Transformacija podrazumeva prečišćavanje izvornih podataka, eliminisanje nekonzistentnosti u njima, dodavanje vremenske komponente (ukoliko ne postoji), spajanje izvornih podataka i njihovo integrisanje u jedinstvene formate. Transformacija se sprovodi pre upisa u ciljnu bazu skladišta ili paralelno sa upisom, posebno kod većih baza podataka u cilju uštede vremena. Preporuka je da se transformacije ne rade u bazi skladišta da se ne bi narušavale performanse upita koji se izvršavaju. Potrebe za transformacijom izvornih podataka proizilaze iz same činjenice da su, sa aspekta skladišta podataka, najčešći izvorni sistemi nesinhronizovani, nezavisno aplikativno razvijani transakcioni sistemi, kao i heterogeni eksterni izvori podataka. U takvim izvornim sistemima čest je slučaj da se atributi koji se odnose na isti poslovni pojam različito imenuju, da se isto imenuju poslovni pojmovi različitog značenja. Takođe, slobodno definisani pojmovi kao što su telefonski brojevi, imena, adrese i sl. se najčešće različito zapisuju u različitim sistemima. Sve te nekonzistentnosti treba eliminisati, formate uskladiti i identifikovati iste poslovne pojmove. Uskladjivanje je neophodno i

ukoliko je u izvornim sistemima narušen referencijalni integritet ili nema jedinstvenih ključeva.

Osvežavanje baze skladišta podataka je, po pravilu, proces ubacivanja novih serija podataka u određenim vremenskim intervalima. Taj proces retko dozvoljava izmene nad ranije ubačenim podacima. Izmene postojećih podataka nisu sasvim zabranjene u sistemima za skladištenje podataka, ali su vrlo retke i preporuka je izbegavati ih.

Kada se govori o osvežavanju skladišta podataka, jedno od ključnih pitanja je i na koji način identifikovati skup slogova iz izvornog sistema, koji će biti zahvaćeni u sledećem ciklusu osvežavanja. Efikasno rešenje je da se, okidačima ili korisničkim kodom, u tabelama činjenica evidentira vreme upisa podataka u bazu skladišta. Pri sledećem osvežavanju baze, upoređuje se vreme upisa u bazu skladišta sa vremenom izmene podataka u izvornom sistemu. Na taj način identifikuje se skup slogova koji su novi ili izmenjeni u izvornom sistemu i koji će biti kandidati za sledeći upis u bazu skladišta. Dakle, svako sledeće punjenje baze skladišta zahvata značajno manji skup izvornih podataka u odnosu na masovno, inicijalno punjenje.

Preporuka je da se ETL aktivnosti za svaku tabelu iz baze skladišta rade u zasebnim procesima tj. programima, zbog lakšeg održavanja ETL procedura. Za implementiranje ETL procesa može se koristiti neki od standardnih alata na tržištu, može se razvijati sopstveni ETL alat unutar samog preduzeća, ili se može praviti korisnički programski kod. Ciklusi u kojima će se ETL procesi odvijati zavise isključivo od samog poslovnog sistema i određeni su prirodom posla i zahtevima krajnjih korisnika.

Smatra se da efikasnost ETL procesa predstavlja jedan od ključnih faktora uspešnosti projekata izrade sistema za skladištenje podataka. Aktivnosti projektovanja i implementiranja ETL procesa troše i do 70 procenata vremena angažovanja na projektu.

3.2.3 Pripremna zona

Pripremna zona (eng. *Data Staging Area*) je neobavezan deo arhitekture i njeno postojanje zavisi, isključivo, od pojektanata skladišta podataka i prirode svakog pojedinačnog projekta. Pripremna zona je deo arhitekture namenjen smeštanju podataka prepisanih iz izvornih sistema, u kojoj se podaci prečišćavaju i pripremaju za upis u ciljno skladište. Skoro sva prečišćavanja i transformacije podataka odvijaju se u pripreмноj zoni. Osnovni zahtev arhitekture koja uključuje pripremnу zonu je da pripremnа zona ostane nedostupna krajnjim korisnicima jer je nepogodna za postavljanje upita i analitičko izveštavanje. Dakle, ovo je, sa aspekta korisnika, neprezentacioni sloj i služi isključivo tehničkom timu. Pripremnа zona se, najčešće, organizuje kao skup normalizovanih relacionih tabela. Cilj je da se u jednom relacionom okruženju, uz minimalne neophodne transformacije, objedine i usklade podaci iz heterogenih izvora i da se iz nje rade sve dalje ETL aktivnosti i punjenje ciljne baze skladišta. To je model koji neće biti dostupan korisnicima jer je poslovno nerazumljiv i analitički malo upotrebljiv. U praksi, pripremnа zona se pokazala kao koristan element arhitekture, posebno kada su izvori podataka brojni i podaci u njima

nesinhronizovani. Pripremna zona biće deo arhitekture koji će, uglavnom, biti zastupljen i u praktičnim primerima ovog rada.

3.2.4 Dimenzioni prezentacioni sloj

Na prezentacionom nivou (sa aspekta krajnjeg korisnika) podaci su prečišćeni, objedinjeni, transformisani i organizovani tako da postaju lako dostupni krajnjim korisnicima za postavljanje upita, izradu izveštaja i pristup kroz analitičke alate poslovne inteligencije. Insistira se na tome da na ovom nivou budu sasvim isključene normalizovane baze podataka i da podaci budu dimenzionalno struktuirani. Dimenzionalno modeliranje usvojeno je kao tehnika najpogodnija za predstavljanje podataka u skladištima podataka. Prezentacioni sloj skladišta podataka čini jedan, ili više integrisanih Data Mart podsistema.

3.2.5 Data Mart u CIF i BUS arhitekturi

Data Mart se može shvatiti kao podsistem sveobuhvatnog sistema za skladištenje podataka jednog preduzeća. Data Mart se formira na osnovu zahteva određene grupe korisnika i sadrži podatke koji po značenju, poslovnim potrebama i izgledu odgovaraju baš toj grupi korisnika. Ta grupa korisnika je, najčešće, jedan organizacioni deo u preduzeću, tj. neko njegovo odeljenje: finansije, prodaja, nabavka, magazin, ...

Prema B. Inmon-u, kao što je prethodno rečeno, projektuje se sveobuhvatno skladište podataka celog preduzeća. Izvor podataka za Data Mart treba da bude upravo ta zajednička baza opšte namene. Data Mart se, u tom slučaju, izvodi kao podsistem velikog sistema u kome će biti isključivo sumarni, podaci. Periodičnim osvežavanjem, podaci iz sveobuhvatnog sistema se prebacuju u Data Mart, analogno prebacivanju izvornih (transakcionih) podataka u sistem za skladištenje. Količina podataka u Data Mart podsistemu je mnogo manja nego u celom skladištu podataka. To su podaci koji će pokriti deo poslovnih procesa preduzeća i biće relevantni samo određenoj grupi korisnika.

Postoje zamerke druge grupe analitičara okupljenih oko R. Kimball-a, da je izgradnja sveobuhvatnog sistema za skladištenje skupa, dugo traje, zahteva ogromne resurse i rezultati se ne vide brzo. Teže se „opipava puls“ korisnika i koriguju zahtevi.

Prema prvoj grupi analitičara, dugoročno gledano, sveobuhvatno skladište podataka opravdava svoju cenu. Zagovornici ovog pristupa u projektovanju skladišta podataka tvrde da više Data Mart podsistema nikada ne može da zameni jedinstveno skladište podataka iz sledećih razloga:

- Struktura podataka u Data Mart sistemu odgovara potrebama samo jednog organizacionog dela.

- Može da dodje do neusaglašenosti između istih izveštaja iz različitih organizacionih delova.
- Zvezdaste šeme kojima su predstavljene strukture podataka u Data Mart sistemima su prilagodjene i optimalne samo jednoj grupi korisnika i nikome više. To su strukture iskoristljive samo sa stanovišta organizacionog dela za koji su optimizovane i ne mogu biti opšte namene za celo preduzeće.
- Jedan od mogućih problema može da bude preklapanje podataka, pa i preklapanja ETL procedura koje se rade za svaki Data Mart ponaosob. Tako je broj ETL procedura sa svakim novim sistemom sve veći, dolazi do dupliranja vremena i posla, ETL algoritmi nisu usaglašeni, hardverski zahtevi su veliki.
- Prisutan je problem održavanje ovakvih sistema. Mogućnost greške je veća, promene je neophodno sprovesti na više mesta. Sve to vodi većem utrošku vremena, novca, ljudskih resursa.
- Svaki novi Data Mart pravi se iz početka. Evidentna je slaba proširljivost ovakvih sistema.

U praksi su zastupljena oba pristupa, zavisno od poslovne problematike, obima podataka, resursa, zahteva korisnika i afiniteta projektnog tima.

Iz svega ovoga izvode se osnovne karakteristike Data Mart podsistema, u CIF i BUS arhitekturi. To su, ujedno, i razlike između Data Mart-a i korporativnog skladišta podataka koje propagira CIF arhitektura:

- Data Mart je vodjen zahtevima grupe ljudi i projektuje se prema njihovim potrebama. Tek na Data Mart nivou određeni poslovni entiteti i procesi dobijaju na većem značaju, favorizuju se, dok su ostali manje bitni ili nebitni. U korporativnom skladištu podataka svi entiteti i poslovni procesi su istog značaja.
- Dok skladište podataka odlikuju masovni podaci, Data Mart ima bar za red veličine manje podataka jer pokriva samo određene poslovne procese.
- Dok skladište podataka može da sadrži velike količine detaljnih, granularnih podataka, Data Mart bi trebalo da bude više sastavljen od sumarnih, agregiranih podataka. U CIF arhitekturi Data Mart sadrži isključivo agregirane podatke, dok u BUS arhitekturi delimično sadrži i atomske podatke zbog kvalitetnijih analiza.
- Data Mart je po nivou detaljnosti podataka, formi, prikazu prilagodjen analitičarima poslovanja i brzim analizama, a ne matematičarima i uopšte IT stručnjacima.

- Obrada Data Mart podataka je mnogo jeftinija sa stanovišta hardvera koji se angažuje, kako zbog količine tako i zbog struktuiranosti podataka koji su prilagodjeni brzim obradama.
- Data Mart podaci se obradjuju korišćenjem programskih alata za inteligentno poslovanje, uz primenu tehnika kakve su OLAP, istraživanje podataka i sl.
- U CIF arhitekturi preporuka je da skladište podataka sadrži istorijske podatke, stare i do 10 godina, a da Data Mart sadrži mnogo "mladje" podatke stare 12 do 15 meseci, zavisno od poslovnih potreba. Istorijski podaci će biti dostupni analitičarima, ukoliko budu potrebni, ali se neće plaćati cena njihovog čuvanja u radnom okruženju.

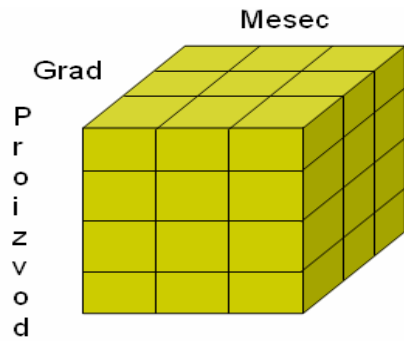
Jedno od pitanja koje se postavlja kada preduzeće raspolaže bazom za skladištenje podataka je: Kada je vreme za uvođenje novog Data Mart podsistema? Smatra se da je to upravo onaj trenutak kada je formirna dovoljno velika grupa ljudi koja na isti način tretira podatke i postavlja slične upite. Praćenje sistema za skladištenje, praćenje aktivnosti korisnika unutar njega, praćenje upita koji se postavljaju, podataka koji se tim upitima zahvataju, jedan je od načina da se identifikuju kandidati za odvojene Data Mart podsisteme.

3.2.6 Korisnički alati za pristup podacima

Korisnički alati (eng. *Data Access Tools*) predstavljaju alate pogodne za pristup podacima iz skladišta podataka, odnosno podacima iz Data Mart podsistema. Ovo su analitički alati projektovani da rade sa denormalizovanim bazama podataka. Služe postavljanju upita nad podacima u prezentacionom sloju. Korisnički alati mogu da budu vrlo jednostavni alati za postavljanje ad-hok upita, ali i veoma sofisticirani OLAP alati, alati za istraživanje podataka i alati za modeliranje i prognoze.

3.3 OLAP

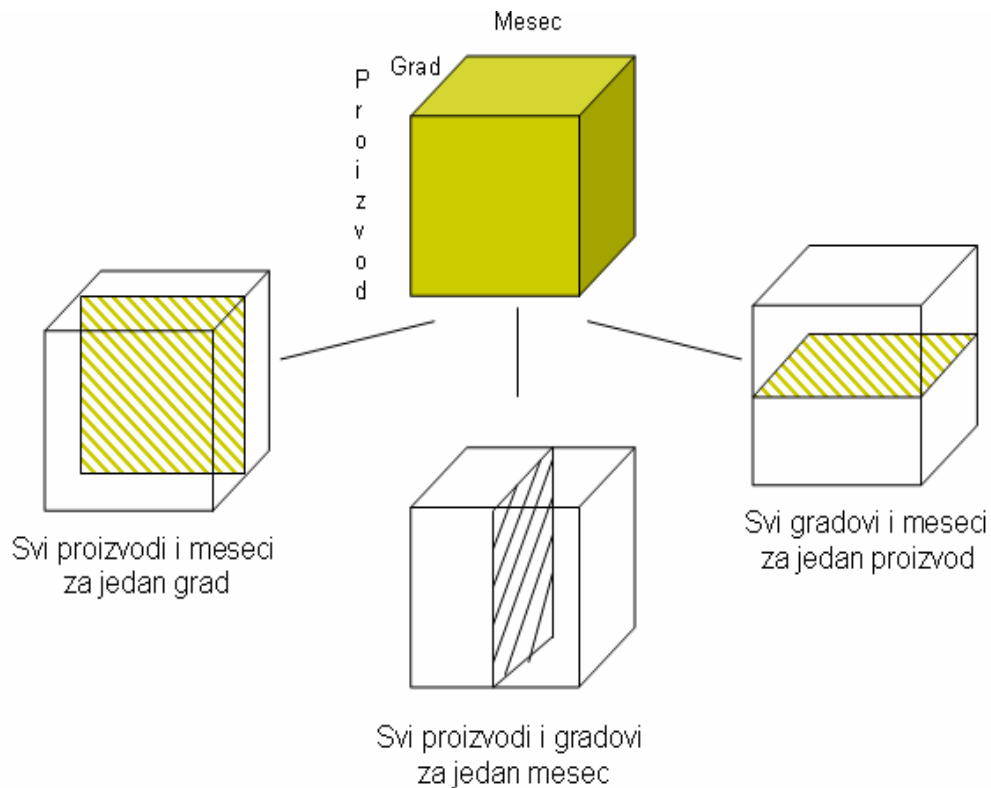
OLAP (eng. *On Line Analytical Processing*) tehnologija, bazirana na multidimenzionalnom SUBP pristupu, smatra se najpogodnijom za obradu Data Mart podataka. Ovo je pristup koji korisniku obezbeđuje multidimenzionalni pogled na podatke i omogućava „kretanje“ kroz dimenzije i hijerarhije u okviru tih dimenzija. Relacioni i multidimenzionalni pristup odlikuje isti logički model, zasnovan na dimenzijama i tabelama činjenica. Ova dva pristupa se razlikuju tek na nivou fizičke implementacije. U praksi, Data Mart podsistemi se, u mnogim slučajevima, i dalje implementiraju relaciono. Kada govorimo o skladištenju podataka i OLAP-u imamo u vidu da je osnovna razlika između OLAP-a i skladištenja podataka u tome što je OLAP tehnologija, a skladištenje podataka je koncept, odnosno metodologija i arhitektura. Osnovni elementi OLAP tehnologije su tzv. kocke (eng. *cube*) u kojima se nalaze Data Mart podaci. Primer multidimenzionalne OLAP kocke dat je na slici 4.



Slika 4. Multidimenzionalna OLAP kocka

Multidimenzioni SUBP obezbeđuju fleksibilan pristup podacima, mogućnost "sečenja" podataka duž jedne dimenzije ili u više pravaca duž više dimenzija, "bušenja" naviše i naniže duž hijerarhija u okviru dimenzija kocki i dinamičko povezivanje sumarnih sa detaljnim podacima. Smatra se da je upravo ovakva organizacija podataka najpogodnija za optimalne analize poslovanja.

Primer sečenja podataka duž dimenzija kocke dat je na slici 5.



Slika 5. Različiti pogledi na iste podatke u OLAP kocki

Višedimenzionalne kocke odlikuju sledeće karakteristike:

- Obezbeđuju brz pristup podacima.
- Optimizovane su za interaktivne analitičke upite.
- Strukture podatka mogu da se maksimalno optimizuju, ukoliko se poznaju modeli upita.
- Skalabilnost, dobre performanse, fleksibilnost.

Tabele činjenica i dimenzione tabele se „izvlače“ iz standardnih zvezdastih šema i čuvaju eksterno, u repozitorijumu. Savremeni korisnički alati za pristup podacima u kockama nisu SQL orijentisani, ali mogu biti konfigurisani tako da se, u slučaju da korisniku trebaju podaci koji nisu izdvojeni u kockama, omogući pristup i originalnim zvezdastim šemama. Dakle, cilj je korisnicima obezbediti dinamički prolaz do relaciono strukturiranih podataka, ukoliko su kocke izvorno izvedene iz tih relacionih podataka. Kocke se „pune“ podacima iz osnovnih tabela pomoću ETL procedura. Korak dalje u razvoju OLAP tehnologije predstavljaju alati koji obezbeđuju automatizovano „izvlačenje“ relaciono strukturiranih podataka u kocke, čime postaje nepotrebno pisanje ETL procedura za njihovo formiranje.

Multidimenzionalni pristup ima i svoje nedostatke. Neki od njih su:

- Multidimenzionalni formati ne mogu da obradjuju ni približno toliku količinu podataka kao standardni relacioni formati.
- Ovaj pristup ne podržava procese izmene podataka.
- Strukture su nefleksibilne ukoliko im se ne pristupa na predvidjeni način.
- Dinamičko povezivanje podataka je diskutabilno.
- Evidentna je ograničenost u tehnologijama koje mogu da omoguće dinamički prelazak sa ne-relacionih u relacione strukture i obrnuto. Razlike između SQL i OLAP pristupa su brojne, ali sve je više proizvoda na tržištu koji počinju da prevazilaze ove razlike. Danas je jedna od tendencija proizvođača pravljenje RSubP-a koji sadrži OLAP organizovane podatke i kojima se pristupa preko SQL interfejsa.

Skladište podataka se smatra izvorom pogodnih podataka za još neke napredne tehnike inteligentnog poslovanja, kao što je istraživanje podataka. Korišćenjem masovnih detaljnih, istorijskih podataka i uz podršku naprednih tehnologija i algoritama (neuronske mreže, klasteri, stabla odluka, diskriminacione analize itd.) identifikuju se složene, skrivene veze među podacima. Cilj je dobijanje modela koji obezbeđuju: praćenje trendova, analize rizika i sl.

Istraživanje podataka je pristup primenljiv u različitim granama poslovanja:

- Telekomunikaciona preduzeća ga koriste u cilju analize korišćenja svojih usluga.
- Osiguravajuća i berzanska društva u cilju smanjenja rizika poslovanja.
- Medicina za predviđanje efikasnosti hirurških intervencija.
- Preduzeća iz finansijskog sektora za efikasno analiziranje karakteristika tržišta ili performansi konkretnog preduzeća sa kojim nameravaju da posluju.

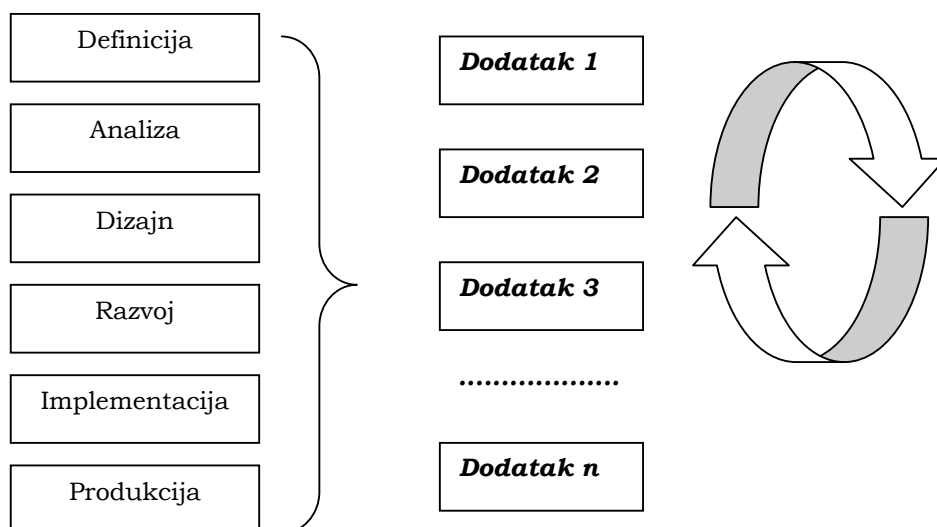
4 PROJEKTOVANJE SKLADIŠTA PODATAKA

4.1 Iterativni razvoj sistema za skladištenje podataka

Sistemi za skladištenje podataka su, prvenstveno, namenjeni analitičarima poslovanja. Njihov zadatak je analiza informacija neophodnih u korporativnom odlučivanju. Zato je, pri izradi ovakvih sistema, neophodno ući u način razmišljanja analitičara poslovanja, razumeti način na koji on doživljava skladište podataka i utvrditi šta on od skladišta podataka očekuje. Analitičar poslovanja je osoba koja funkcioniše po principu otkrića i tek kad vidi ekran ili izveštaj počinje da istražuje mogućnosti sistema koji je pred njim. U svetu analitičara poslovanja zahtevi dolaze na kraju „istraživačkog“ procesa. Pristup i potrebe poslovnih analitičara kao ključnih korisnika sistema za skladištenje podataka uzrokuju bitnu razliku između transakcionih sistema i sistema za skladištenje podataka i to u načinu projektovanja sistema.

U transakcionim sistemima jedna od metoda projektovanja koja se koristi je poznata pod nazivom „vodopad“. Osnovna odlika ovog pristupa je da se aktivnosti uzastopno odvijaju jedna za drugom tako da kraj jedne inicira početak sledeće. Aktivnosti su vodjene zahtevima, tj. zahtev korisnika inicira razvoj sistema. Za razliku od ovog pristupa koji polazi od zahteva, u sistemima za skladištenje podataka je prisutan potpuno suprotan princip koji polazi od podataka i nudi informacije koje će inicirati zahteve korisnika. Korisnici sistema za skladištenje podataka, uglavnom, funkcionišu po principu: „Pokaži mi šta imam na raspolaganju pa ću ti reći šta mi treba“. U skladu sa tim prihvaćeno je da je sistem za skladištenje podataka najbolje razvijati iterativno-inkrementalno. Razvoj kompletnog softvera, tj. automatizacija svih poslovnih procesa koji će biti obuhvaćeni skladištem podataka, odvija se fazno. Najpre se potpuno implementira inicijalni podskup skupa svih poslovnih procesa, a zatim se uzastopnim koracima implementiraju, kao nadogradnja prethodnog koraka, nove i složenije poslovne logičke celine. Svaka faza razvoja skladišta podataka uključivanjem softverskih dodataka dodaje postojećem sistemu pojedine nove funkcije, pri čemu se postojeće zadržavaju. Svaki takav softverski dodatak se naziva priraštaj (eng. *increment*). Na primer, jedna iteracija razvoja sistema može da obuhvati poslovne procese vezane za prodaju. U ovoj iteraciji prolazi se kroz sve faze razvoja softvera, počev od definicije, analize preko dizajna, razvoja, implementacije i eksploatacije u produkcionom okruženju od strane krajnjih korisnika. Sledeća iteracija razvoja obuhvata poslovne procese vezane za magacinsko poslovanje i njihova automatizacija prolazi kroz iste faze razvoja softvera kao prethodna. Sve zajedničke funkcionalnosti iz prve iteracije biće iskorišćene u drugoj uz neophodna proširenja. Tek nakon n ovakvih iteracija kompletno skladište podataka biće implementirano i stavljeno u produkciju.

Grafički prikaz iterativno-inkrementalnog razvoja sistema dat je na slici 6.



Slika 6. Iterativno-inkrementalni razvoj sistema za skladištenje podataka

Prednosti ovakvog razvoja su višestruke:

- Dodaci, odnosno nove funkcije se lakše razumeju i testiraju.
- Ovakav pristup ostavlja mogućnost da se bogato korisničko iskustvo iskoristi i ugradi u sledećim iteracijama. Već nakon prve iteracije moguće je dobiti odgovore krajnjih korisnika na softver koji imaju pred sobom. Praćenje rada korisnika, osluškivanje njihovih zamerki, sugestija i potreba, uočavanje tipova upita koji postavljaju nad tim malim delom budućeg skladišta u velikoj meri pomaže projektantima u sledećim iteracijama.
- Sve uočene projektantske i programerske greške mnogo je lakše i jeftinije korigovati na inicijalnom podskupu nego na celom softverskom proizvodu.
- Skraćuje se vreme uvođenja sistema za skladištenje u produkciju, odnosno brže se vide prvi rezultati razvoja novog sistema. Efikasnost u razvoju sistema pozitivno deluje na sve učesnike u projektu, a posebno na rukovodeće strukture u firmi koje donose konačne odluke o novcu i resursima koji će se izdvojiti za razvoj novih informacionih sistema.

Iterativno-inkrementalni pristup implementacije skladišta podataka biće korišćen u praktičnom delu ovog rada.

4.2 Definisanje logičkog modela

Definisanje logičkog modela skladišta podataka, prema [2], obuhvata sledeće korake:

4.2.1 Analiza poslovnog sistema. Izbor poslovnih procesa

Faza analize poslovnog sistema obuhvata, pre svega, identifikaciju poslovnih procesa i identifikaciju subjekata u poslovanju. Dalje, ova faza podrazumeva analizu procesa i subjekata. Na kraju faze analize vrši se izbor onih poslovnih procesa i subjekata, iz skupa svih identifikovanih, koji su najbitniji sa aspekta celokupnog poslovanja organizacije. Izabrani procesi i subjekti poslovanja biće dalje implementirani u sistemu za skladištenje podataka. Faza analize podrazumeva i određivanja prioriteta u implementaciji. Kupac, proizvod, nabavka, magacin, transport, naplata, glavna knjiga su neki primeri poslovnih pojmova, odnosno poslovnih entiteta i poslovnih procesa.

Pri izboru poslovnih procesa koji će biti implementirani neophodno je voditi računa o sledećem:

- Koji procesi posebno utiču na efikasnost poslovanja?
- Koji procesi obezbeđuju najveći profit preduzeću?
- Koji procesi imaju najpovoljniji odnos cene implementacije i koristi koju bi od te implementacije imalo preduzeće (eng. *cost-benefit*)? Ova analiza treba da obuhvati inicijalnu cenu implementacije poslovnog procesa, kao i cenu daljeg održavanja tog informacionog sistema. Kada se određuje cena održavanja jednog informacionog sistema uzima se u obzir: cena održavanja hardvera, cena održavanja softvera, ali i ulaganja u obuku i usavršavanje zaposlenih. Pri tome, uzima se u obzir i ušteda u vremenu koju novi softver donosi, odnosno smanjenje angažovanja pojedinih zaposlenih. Na kraju, u analizu se uključuje i veća efikasnost u donošenju poslovnih odluka primenom novog softvera. Analizom svih ovih parametara poslovanja dobija se ROI (eng. *Return On Investment*) pokazatelj isplativosti i efikasnosti uvođenja novog sistema.
- Koji procesi imaju najveći strateški značaj za firmu?
- Koje vrste poslovnih odluka mogu da se donesu na osnovu tih poslovnih procesa?
- Koji su procesi mogu najlakše implementirati?
- Koliko će dugo trajati implementacija svakog od tih procesa?

Iako će se skladište podataka implementirati inkrementalno, preporuka je da se u ovoj fazi sagleda celokupan poslovni sistem. Određivanjem koji će entiteti ući u opseg modela, a koji ne, postavljaju se „granice“ modela podataka, tzv. „opseg integracije“. Ove granice je neophodno postaviti na početku, pre izrade modela. Oko samih granica, tj. „opsega integracije“ moraju da se slože projektni tim, rukovodstvo firme i krajnji korisnici.

Najbolji način da se predstavi veza poslovnih procesa i poslovnih subjekata sistema je korišćenjem matrice u kojoj su redovi poslovni procesi, a kolone svi identifikovani poslovni subjekti. Redovi ove matrice predstavljaju potencijalne Data Mart podsisteme.

Primer jedne takve matrice veze u sistemima nabavke, skladištenja i prodaje dat je na slici 7.

BUSINESS PROCESSES	COMMON DIMENSIONS						
	Date	Product	Store	Promotion	Warehouse	Vendor	Contract Shipper
Retail Sales	X	X	X	X			
Retail Inventory	X	X	X				
Retail Deliveries	X	X	X				
Warehouse Inventory	X	X			X	X	
Warehouse Deliveries	X	X			X	X	
Purchase Orders	X	X			X	X	X X

Slika 7. Matrični prikaz veze poslovnih procesa i subjekata sistema

Redovi matrice su identifikovani poslovni procesi i oni nisu uvek grupisani po organizacionim delovima preduzeća. Redovi matrice mogu, prilikom implementacije, da se razbiju na detaljnije procese ukoliko su izvori podataka u nekom procesu različiti ili ukoliko jedan proces ne može da bude implementiran u jednoj iteraciji. Preporuka je da se prvo implementiraju Data Mart podsistemi prvog nivoa, tj. pojedinačni procesi koji mogu direktno da se izvedu iz izvornog, operativnog sistema. Njihovom implementacijom se, za relativno kratko vreme, krajnjim korisnicima daje na uvid odredjeni skup podataka, prate se aktivnosti i ponašanje korisnika i dobija se na vremenu. Pravu dobit za krajnje korisnike donose tek konsolidovani Data Mart podsistemi drugog nivoa. Oni su složeniji i kroz njih se implementira ukrštanje poslovnih procesa.

Matrica veza subjekata poslovanja i poslovnih procesa je jedan od najvažnijih rezultata analize poslovnog sistema i deo izlaznog dokumenta. Izlazni dokument iz ove faze izrade poslovnog modela ne treba da bude mnogo obiman i mora da bude napisan jezikom krajnjih korisnika, tj. poslovno a ne tehnički razumljivim jezikom.

4.2.2 Definisane nivoa detaljnosti. Granularnost

Ovo je trenutak u razvoju sistema za skladištenje kada definišemo najniži nivo detaljnosti podataka neophodan da bi se postigla dovoljna analitičnost. Nivo detaljnosti, odnosno sumarnosti podataka u skladištu podataka predstavlja njihovu granularnost.

Kada je nivo detaljnosi podataka u pitanju, u IT krugovima vladaju oprečna mišljenja. Jedna grupa analitičara smatra da u sistemu za skladištenje treba

sasvim izbegavati detaljne podatke i uključivati isključivo agregirane, sumarne podatke. Druga grupa je mišljenja da sistem za skladištenje podataka treba da sadrži i podatke na najnižem nivou detaljnosti, čak i podatke na nivou svake pojedinačne transakcije, kakva je na primer POS transakcija u sistemu prodaje. Druga grupa svoj stav pravda činjenicom da su detaljni podaci upravo ti koji obezbeđuju veću analitičnost krajnjim korisnicima i omogućavaju „bušenje“ podataka od najvišeg nivoa agregacije do najdetaljnijih podataka na nivou pojedinačne transakcije. Činjenica je da se iz detaljnih podataka uvek mogu izvesti sumarni podaci, dok obrnut proces nije moguć. Mišljenje druge grupe analitičara je da bi izostavljanjem tih najdimenzionalnijih, najdetaljnijih podataka korisnici bili uskraćeni za veliki broj najdetaljnijih, najfinijih analiza poslovnih procesa.

Pristup koji će biti prikazan u praktičnom delu ovog rada pokušava da nadje kompromis između oba pravca u modeliranju sistema za skladištenje podataka. Ono što je bio jedan od imperativa pri izradi ovih modela je struktura korisnika koji će ove sisteme koristiti i njihovi konkretni zahtevi.

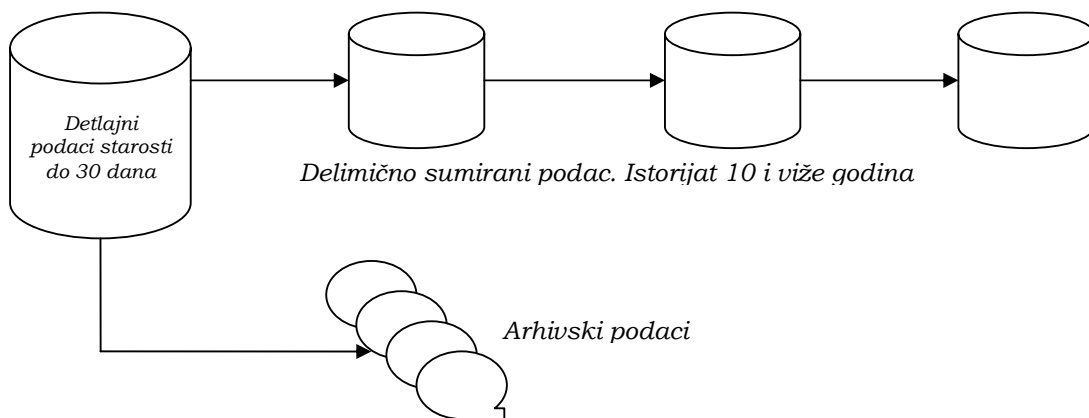
Smatra se da je pravilan izbor nivoa granularnosti podataka ključna projektantska odluka u celokupnom procesu projektovanja i implementacije skladišta podataka. Između ostalog i zbog toga što direktno utiče na količinu podataka u skladištu, kao i na tip upita na koji će sistem za skladištenje moći da odgovori. Što su podaci detaljniji nivo granularnosti je niži. Tako je, na primer, svaka pojedinačna prodajna transakcija najnižeg nivoa granularnosti, odnosno najveće detaljnosti. Nizak nivo granularnosti imaju i podaci o svakom pojedinačnom telefonskom razgovoru koji je kupac napravio u toku meseca. Sumarni podaci o transakcijama ili telefonskim razgovorima na mesečnom, kvartalnom ili godišnjem nivou su podaci sve veće granularnosti. Podaci niske granularnosti su detaljniji, ali i masovniji, tako da zauzimaju više prostora u odnosu na granulirane podatke. Ukoliko se u sistemu skladišti svaki pojedinačni telefonski poziv kupca, to na mesečnom nivou može iznositi i nekoliko hiljada slogova, odnosno na destine hiljada bajtova. Granulirani podaci o telefonskim razgovorima kupaca na mesečnom nivou u bazi čuvaju samo jedan slog sa sumiranim podacima, odnosno mogu da zauzmu par stotina bajtova. Granularnost je osnova za:

- Fleksibilnost sistema za skladištenje podataka.
- Iskoristivost sistema. Kada se govori o iskoristivosti sistema misli se na njegovu iskoristivost od strane različitih grupa korisnika. Npr. jedna grupa analitičara posmatra pomenute transakcije trgovanja na najdetaljnijem nivou, tj. na nivou dana, proizvoda i kupca. Druga grupa posmatra transakcije svih kupaca zbirno na mesečnom nivou po proizvodima. Treća grupa korisnika posmatra zbirne transakcije na mesečnom nivou po svakom pojedinačnom kupcu i to za sve proizvode. Pravilnim izborom nivoa granularnosti osnovnih podataka omogućava se svim ovim grupama korisnika da koristeći isti sistem dodju do rezultata.
- Adekvatan i brz odgovor na buduće, nepredvidive zahteve za podacima. Kada se, pojave zahtevi za novim informacijama sistem za

skladištenje podataka bi trebalo da u najkraćem roku odgovori takvim zahtevima.

Projektanti skladišta podataka često pribegavaju uvođenju dva ili više nivoa granularnosti podataka. Podaci prvog, najnižeg nivoa granularnosti su detaljni i masovni. Kao takvi čuvaju se u sistemu ograničeno vreme (npr. 30 dana), nakon čega se delimično sumiraju i tako sumirani ostaju u bazi skladišta u mnogo manjem obimu i duže vreme. Ovako sumirani podaci se dalje koriste za analize. Količina sumarnih podataka je manja čime se poboljšavaju performanse upita, ali postoji ograničenje u detaljnosti analiza koje je moguće izvesti nad ovakvim podacima. Osim kratkotrajnih detaljnih i sumarnih podataka, ovaj pristup zahteva i postojanje "arhivskih" detaljnih podataka. U arhivske podatke smeštaju se svi podaci na najdetaljnijem nivou, sačuvani na magnetnoj traci ili nekom sličnom masovnom mediju, nakon isteka odredjenog životnog veka ovih podataka. Princip postojanja više nivoa granularnosti podataka u sistemu poznat je pod nazivom "dvojna (dualna) granularnost".

Grafički prikaz dualne granularnosti dat je na slici 8.



Slika 8. Dualna granularnost u sistemima za skladištenje podataka

Smatra se da je zbog cene, efikasnosti, lakoće pristupa i mogućnosti da se odgovori na skoro svaki postavljeni upit dualna granularnost najbolji izbor za većinu skladišta podataka u kojima se obradjuju velike količine podataka.

4.2.3 Identifikacija poslovnih dimenzija

Dimenzioni podaci predstavljaju attribute poslovnog procesa u kontekstu kojih se vrši analiza i vrednovanje poslovnih mera, tj. činjenica. Podaci koji su kandidati za dimenzije modela identifikuju se odgovorom na pitanje kako se u poslovnom sistemu prate poslovne mere. Za takve attribute se kaže da sirovom podatku daju značenje. Neki primeri poslovnih dimenzija su: kupac, proizvod, račun i magacin. U ovoj fazi razvoja sistema za skladištenje potrebno je odgovoriti na pitanja: koji su dimenzioni atributi interesantni sa aspekta poslovnog sistema koji se analizira, kako se grupišu ti atributi i

koliko često se menja svaki od ovih atributa. Određivanjem nivoa detaljnosti neke poslovne mere određuje se osnovna dimenzionalnost tabele činjenica u kojoj se nalazi podatak o toj poslovnoj meri. Ukoliko se u ovoj fazi ne sagledaju sve poslovne dimenzije dolazi se, najčešće, u situaciju da se kasnijim dodavanjem novih dimenzija povećava nivo detaljnosti podataka. Promena nivoa detaljnosti podataka sa sobom povlači korigovanje programskih modula za punjenje tabela činjenica i, najčešće, potpuno pražnjenje i inicijalno punjenje tabela činjenica svim podacima.

4.2.4 Identifikacija poslovnih mera

Poslovne mere ili činjenični podaci (činjenice, eng. *facts*) su numeričke vrednosti, tj. iznosi kojima se izražavaju rezultati poslovnih procesa i kroz koje svaka organizacija meri uspeh svog poslovanja. Ova vrsta podataka u sistemima za skladištenje se identifikuje kroz odgovor na pitanje šta se meri nekim poslovnim procesom. Primeri poslovnih mera su: iznos prodaje u dolarima, cena proizvoda, količina prodatih proizvoda ili ukupan iznos porudžbine. Za činjenične podatke kaže se da su, uglavnom, aditivni jer su nad njima primenljive operacije agregiranja. Činjenice mogu biti sasvim aditivne, tj. aditivne sa aspekta svih dimenzija koje su povezane sa tim činjeničnim podatkom. Kada je agregiranje izvodljivo samo sa aspekta nekih, ne svih, dimenzije modela koje su povezane sa činjeničnim podatkom, te činjenice su delimično aditivne.

Za činjenične podatke kao što je iznos prodaje ili cena proizvoda neophodno je da se zna u kojoj su jedinici mere izraženi (milijarde, milioni, hiljade) i u kojoj valuti. Tabele činjenica uvek su povezane spoljnim ključevima sa primarnim ključem svake od njenih dimenzija.

U ovoj fazi izrade modela poželjno je da se utvrdi inicijalni broj slogova u tabelama činjenica, kao i (orijentacioni) broj slogova koji će se u tabele činjenica dodavati pri svakom narednom punjenju.

4.2.5 Princip magistrale. Jednobraznost dimenzija i činjenica

Pojam magistrale [2] aktuelan je u sistemima za skladištenje podataka koji se razvijaju kao skup komplementarnih, konzistentnih, nezavisno implementiranih Data Mart podsistema. Da bi ovako koncipiran sistem za skladištenje podataka bio dugoročno uspešan potrebno je da se razvija inkrementalno i to u strogo definisanoj i uredjenoj arhitekturi. Arhitektura sistema o kojoj je ovde reč predstavlja organizovano okruženje tj. pravila, standarde i okvire u kojima će se graditi celokupan sistem za skladištenje. Osnovne principe ove arhitekture mora da ispoštuje svaki nezavisno razvijan Data Mart podsistem. Ovaj koncept je u informatičkom rečniku poznat pod nazivom arhitektura magistrale (eng. *bus architecture*). Reč „magistrala“ potiče iz elektrotehnike i označava sve one zajedničke strukture sistema sa kojima je sve ostalo u sistemu povezano i iz kojih sve ostalo u sistemu crpi energiju. To su, dakle, pokretači, odnosno ključni elementi sistema.

U fazama analize i definicije logičkog modela identifikovani su poslovni procesi koji će biti implementirani, kao i dimenzije i tabele činjenica preko kojih će se ti procesi implementirati. Svi identifikovani procesi u sistemu dele

neke zajedničke dimenzije. Taj skup zajedničkih dimenzija predstavlja okvire celog sistema za skladištenje podataka i implementacija svakog pojedinačnog Data Mart-a će se odvijati u tako postavljenim okvirima. Na ovaj način svi Data Mart podsistemi postaju međusobno komplementarni kao delovi jedne jedinstvene slike. Da bi se potpuno odredio skup dimenzija na kome će se razvijati svaki novi podsistem, neophodno je u fazi analize sagledati sve poslovne procese jednog preduzeća. Dakle, bez obzira što se sistem za skladištenje razvija inkrementalno kao skup Data Mart podsistema, neophodno je da se sveobuhvatno sagleda celokupno poslovanje preduzeća. I to na početku, u fazi analize poslovnog sistema.

Primer deljenja dimenzija u podsistemima nabavke, skladištenja i prodaje dat je na slici 9.



Slika 9. Zajedničke dimenzije – arhitektura magistrale

Arhitektura magistrale je nezavisna od baze podataka i platforme na kojoj je sistem za skladištenje implementiran.

Osnovni princip koji mora da zadovolji sistem za skladištenje koji se implementira kao skup nezavisnih Data Mart podsistema je princip uskladjenosti, tj. jednobraznosti dimenzija. Ukoliko se dimenzije kao što je vremenska dimenzija, dimenzija kupaca ili dimenzija proizvoda implementiraju u Data Mart podsistemima neuskladjeno, takvi podsistemi neće moći da se ukrštaju i razmenjuju podatke izmedju sebe. Svaki pokušaj kombinovanja podataka iz ovih podsistema davaće pogrešne rezultate upita.

Postoji nekoliko osnovnih pravila postizanja jednobraznosti dimenzija:

- Dve dimenzije su jednobrazne ukoliko su identične i na identičan način su povezane sa tabelama činjenica. Npr. vremenska dimenzija u sistemu prodaje može da bude identična vremenskoj dimenziji u sistemu nabavke. To su dimenzije koje mogu da budu fizički implementirane kroz jednu tabelu. Ako su fizički razdvojene, one imaju iste ključeve, broj slogova, opise, sadržaje i nazive atributa.
- Dve dimenzije su jednobrazne ukoliko su izvedene agregiranjem iz osnovne dimenzije najvišeg nivoa detaljnosti (eng. *roll-up*). Osnovna vremenska dimenzija čuva podatke na nivou pojedinačnog dana,

osnovna dimenzija banaka sadrži podatke o svakoj pojedinačnoj banci, osnovna dimenzija proizvoda sadrži podatke o svakom pojedinačnom proizvodu. Izvedena dimenzija nedelje ili meseca mora da bude jednobrazna sa osnovnom vremenskom dimenzijom. Ona ne sadrži atribute osnovne dimenzije koji su karakteristični za dane (npr. naziv dana, broj dana u nedelji ili oznaku da li je dan radni ili neradni). Ali, svi zajednički atributi osnovne i izvedene dimenzije moraju da budu usaglašeni po opisu, sadržaju i imenovanju.

- Dimenzija je jednobrazna osnovnoj dimenziji ukoliko sadrži podskup njenih redova. Primer su dimenzije izvedene iz dimenzije svih proizvoda izborom određenih proizvodnih linija. Tako se dobijaju usaglašene dimenzije u kojima su razdvojeni podaci o prehrambenim, hemijskim ili tehničkim proizvodima. Ovakve dimenzije se formiraju prema potrebama određene grupe krajnjih korisnika. Usporedne analize mogu da se obavljaju samo nad presekom ovih skupova podataka, odnosno po proizvodima koji su zajednički u proizvodnim linijama koje se prate.
- Dimenzija je jednobrazna sa osnovnom dimenzijom ukoliko je dobijena isključivanjem nekih kolona iz osnovne dimenzije koje u konkretnom Data Mart-u nisu potrebne. I u ovom slučaju mora da bude ispoštovano pravilo da se zajedničke kolone isto imenuju, imaju iste opise i sadržaje.

Zbog velike važnosti, poželjno je da identifikacija usaglašениh dimenzija, njihovo implementiranje, održavanje i davanja prava pristupa različitim razvojnim timovima bude osnovni zadatak posebne radne grupe. Održavanje usaglašениh dimenzija zahteva da se svaka promena na jednoj dimenziji sprovede na ostalim usaglašениm dimenzijama i to kroz sve Data Mart podsisteme koji koriste kako osnovnu tako i njene usaglašene dimenzije. Smatra se da je jednobraznost dimenzija ključna u povezivanju različitih poslovnih procesa i deljenju podataka među njima.

Jednobraznost činjenica je princip koji podrazumeva da se u celom sistemu za skladištenje istim imenom nazivaju isključivo činjenice koje imaju isto značenje, vodeći pri tom računa o pravilima imenovanja atributa. Takođe, različitim poslovnim pojmovima, tj. njihovim interpretacijama u tabelama činjenica, obavezno se daju različita imena. Na ovaj način se izbegava mogućnost da se neuporedive činjenice kombinuju u računskim operacijama. Jedan od uzroka nekompatibilnosti činjenica može da bude i prikaz mera u tabelama činjenica u različitim mernim jedinicama ili u različitim valutama. Prihvatljivo rešenje je čuvanje istih mera, prikazanih u više mernih jedinica, u tabelama činjenica. Slično tome, u tabelama činjenica je moguće čuvati i iste mere prikazane u više valuta. Moguće rešenje je i uvođenje koeficijenta konverzije između mernih jedinica ili između posmatranih valuta. U tom slučaju korisnik, kroz računске operacije koje sprovodi u samom korisničkom alatu, podatke koje posmatra uporedo svodi na iste merne jedinice ili iste valute. Ovo je manje pouzdano rešenje i traži veću veštinu rukovanja korisničkim alatima.

4.3 Dimenziono modeliranje

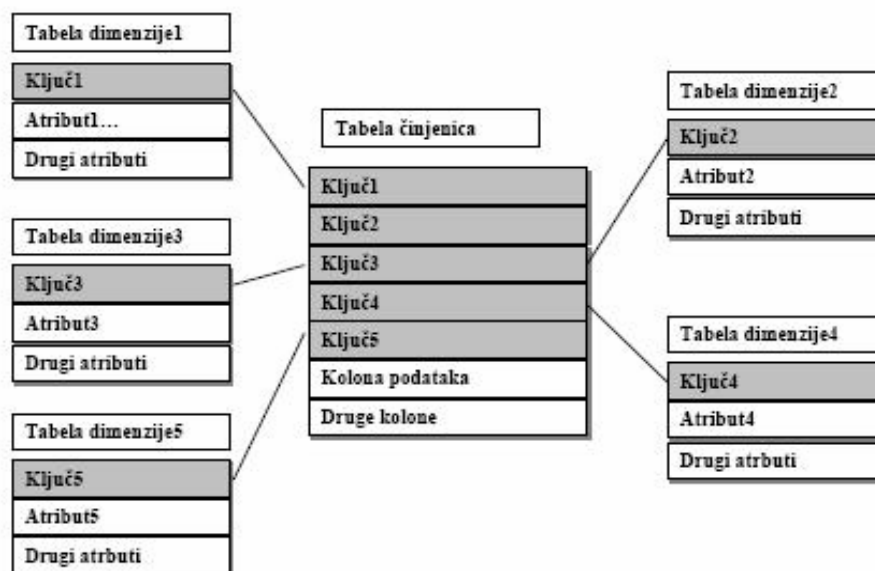
Sledeći korak u projektovanju sistema za skladištenje podataka je uspostavljanje veza između identifikovanih dimenzija i tabela činjenica, modeliranje vremenske dimenzije, kao i definisanje eventualnih hijerarhija na dimenzijama. Kao izlaz iz ove faze razvoja projekta dobija se tzv. dimenzioni model skladišta podataka.

4.3.1 Zvezdasto dimenziono modeliranje

Standardna tehnika koja se koristi u dimenzionom modeliranju sistema za skladištenje podataka je modeliranje korišćenjem zvezdastih (eng. *star*) šema, ili nekih njihovih modifikacija. Jedna od modifikacija je tzv. model pahuljice (eng. *snowflake*).

Zvezdastu strukturu čine jedna tabela činjenica koja je povezana sa više dimenzionih tabela. Dimenzione table su svojim primarnim ključem povezane sa delom složenog ključa table činjenica. Zvezdaste šeme su, po pravilu, denormalizovane. Denormalizacija i redundansa podataka u ovim modelima se uvodi u cilju poboljšanja performansi izvršavanja upita nad podacima iz skladišta.

Primer zvezdaste šeme opšteg tipa dat je na slici 10.



Slika 10. Zvezdasta šema u dimenzionom modeliranju

Osnovni princip dimenzionog modeliranja je zvezdasto modeliranje zato što zvezdaste šeme imaju sledeće karakteristike:

- Podržavaju multidimenzionalne analize.

- Poboljšavaju performanse i optimizovanost upita.
- Obezbeđuju intuitivan model sa stanovišta krajnjih korisnika. To je model koji odgovara njihovom načinu razmišljanja i korišćenja podataka.
- Obezbeđuju laku proširivost modela.

4.3.1.1 Tabele činjenica

Centralna, glavna tabela u zvezdastom dimenzionom modelu je tabela činjenica.

Osnovne karakteristike tabela činjenica su:

- Sadrže numeričke, kvantitativne podatke o poslovanju, tzv. poslovnu metriku. To su informacije koje su neophodne krajnjim korisnicima.
- U odnosu na dimenzione tabele sadrže veći broj slogova.
- Brzo se uvećavaju, uzimajući u obzir količinu podataka koja se upisuje pri svakom novom ciklusu osvežavanja baze skladišta.
- Podaci u njima mogu biti osnovni podaci najveće detaljnosti (atomske, nedeljivi na sitnije delove), ali i izvedeni i sumarni podaci. Izvedeni podaci su podaci dobijeni preračunavanjem iz osnovnih podataka koji mogu da potiču iz jednog ili više izvora. Izvedene vrednosti sačuvane u bazi podataka poboljšavaju efikasnost pristupa podacima. Izvedeni podaci formiraju se u toku ETL procesa izdvajanja i transformacije podataka iz činjenica koje su identifikovane u fazi analize poslovnog modela. Podaci se u ovom (izvedenom) obliku pojavljuju tek u dimenzionom modelu.
- Podaci u tabelama činjenica su, uglavnom, aditivni. To znači da nad njima mogu da se primenjuju agregatne funkcije sumiranja, izvlačenja maksimuma, minimuma, prosečne vrednosti i sl.
- Tabele činjenica su povezane stranim ključem sa primarnim ključem svake od povezanih dimenzionih tabela.
- Mogu da sadrže tzv. degenerativne ključeve (takav je npr. broj transakcije, ili broj porudžbenice). Ovo nisu poslovne mere, već atributi koji „igraju ulogu“ dimenzionih atributa u tabelama činjenica.
- Mogu da sadrže oznake, tzv. indikatore događaja. To su atributi koji uzimaju vrednosti iz domena sa malim brojem vrednosti (npr. prisutnost=1, konačni_podaci=1 i sl.). Ovo su polja koja se koriste za filtriranje podataka u tabelama činjenica.

Postoje, prema [2], tri osnovna tipa tabela činjenica:

- Transakcione tabele činjenica. Ove tabele sadrže podatke najnižeg, transakcionog nivoa, tj. svaku pojedinačnu transakciju poslovnog procesa. Ovi podaci omogućavaju detaljne analize, ali ne mogu da odgovore na čitav niz poslovnih pitanja za koje se zahtevaju strukture podataka višeg nivoa agregacije.
- Periodične presečne (eng. *periodic snapshot*) tabele činjenica. Ovo su tabele koje se osvežavaju u predviđenim vremenskim intervalima. Kod ove vrste tabela činjenica podaci se ne čuvaju na nivou pojedinačne transakcije. Na kraju utvrđenog vremenskog perioda (dana, nedelje, meseca i sl.) uzima se slika, presek stanja poslovnog procesa i upisuje u tabelu činjenica. Sledeće osvežavanje baze podataka je po isteku sledećeg perioda itd. U pojedinim slučajevima ovako dobijeni podaci predstavljaju agregirane transakcione podatke, koji se, tako agregirani, upisuju u tabelu činjenica. Struktura ovih tabela činjenica vrlo je slična strukturi transakcionih tabela. U nekim slučajevima, međutim, stanja poslovnih procesa na kraju vremenskog perioda ne mogu da se izvedu prostim agreiranjem dnevnih transakcija, već zahtevaju složenije operacije nad izvornim podacima. Takvi su, na primer, podaci o stanju sredstava na računu klijenta. Stanje na računu je podatak koji se na kraju dana ili meseca ne dobija prostim sabiranjem dnevnih transakcija, već sadrži u sebi složenu ekonomsku logiku. Logika izvođenja ovakvih podataka ugrađuje se u ETL procedure.
- Kumulativne presečne (eng. *accumulating snapshot*) tabele činjenica. Ovaj tip tabela činjenica se ne upotrebljava onoliko često koliko prethodna dva tipa, ali se sreće u praksi. Ovo su tabele činjenica koje se koriste onda kada je cilj da se predstave vremenska razdoblja u životnom ciklusu poslovnog subjekta (npr. proizvoda). U svakom slogu jedne ovakve tabele prati se proizvod tokom celog proizvodnog ciklusa, od trenutka njegove izrade, preko porudžbine, transporta, skladištenja, isporuke i naplate. Ove tabele činjenica sadrže, po pravilu, veći broj datumskih polja, koja se ažuriraju pri okončanju svake pojedinačne faze. U slučaju proizvoda u tabeli činjenica morali bi da postoje: datum proizvodnje, datum porudžbine, datum transporta, datum skladištenja itd. Svako datumsko polje u tabeli činjenica vezano je svojim spoljnim ključem sa fizički, ili logički odvojenom datumskom dimenzijom (ima ih onoliko koliko ima datumskih polja). Svaki od ovih referentnih datuma za poslovni subjekt se menja po okončanju jedne faze u životnom ciklusu tog subjekta. Dakle, za razliku od prethodnog tipa presečnih tabela, ovde je osim upisa novih slogova neophodno i ažuriranje postojećih slogova. Transakcione tabele činjenica podrazumevale su upis novog sloga pri svakoj transakciji. Periodične presečne tabele činjenica podrazumevale su jedan slog za svaku vremensku oderednicu. Kod ovog tipa tabela činjenica, međutim, jednom proizvodu odgovara jedan slog za ceo njegov životni ciklus. Ovakve tabele činjenica se smatraju težim za održavanje, pa su kao takve retke u modelima sistema za skladištenje podataka.

Da bi se dobila kompletna slika poslovnog procesa i obezbedili različiti aspekti posmatranja poslovnih činjenica od strane različitih grupa krajnjih korisnika, sistem za skladištenje podataka treba da sadrži kako transakcione tako i periodične presečne tabele činjenica. Redundansa u podacima postoji, ali je opravdana.

Sa aspekta mogućnosti agregiranja, podatke u tabelama činjenica delimo na:

- **Aditivne činjenice.** Ovo su poslovne mere koje su isključivo numeričke i mogu se agregirati duž svih dimenzija povezanih sa tom tabelom činjenica. Agregiranje se vrši prilikom izvršavanja upita. Ovo je osnovni tip podataka u tabelama činjenica. Ovi podaci su i osnovne komponente većine upita koji se postavljaju nad bazom skladišta. Aditivna mera je, na primer, ukupan iznos prodaje ili broj prodatih komada nekog proizvoda.
- **Poluaditivne (semiaditivne) činjenice.** Ovo su mere, koje se agregiraju samo duž pojedinih dimenzija povezanih sa tabelom činjenica. U ovom slučaju postoje dimenzione tabele povezane sa tabelom činjenica duž kojih se ove mere ne mogu agregirati. Takva mera je, na primer, stanje na računu klijenta koje se formira na osnovu dnevnih transakcija i to na kraju svakog radnog dana. Ovaj saldo je aditivan sa aspekta dimenzije klijenta ili dimenzije računa. Agregiranjem duž dimenzije klijenta može se dobiti zbirno stanje na računu svih poslovnih klijenata. Agregiranjem duž dimenzije računa dobija se zbirno stanje na svim računima jednog klijenta. Stanje na računu je mera koja nije aditivna duž vremenske dimenzije. Stanje na računu na kraju perioda (meseca, kvartala, godine i sl.) ne može se dobiti sabiranjem dnevnog stanja na tom računu, već se izvodi iz svih transakcija ostvarenih u celom periodu. Stanje na računu klijenta je podatak koji se isključivo posmatra na dnevnom nivou.
- **Neaditivne činjenice.** Ovo su poslovne mere koje nemaju svojstvo aditivnosti ni po jednoj dimenziji povezanoj sa tabelom činjenica. One se, najčešće, kombinuju (u matematičkim operacijama) sa nekim drugim činjeničnim podacima da bi se dobile aditivne ili poluaditivne poslovne mere koje se dalje analiziraju. Takva mera je npr. cena proizvoda po komadu (neaditivna mera) koja se, pomnožena brojem prodatih komada na dan (aditivna mera), dalje koristi u analizama kao ukupan iznos prodaje tog proizvoda na dan (aditivna mera).

4.3.1.2 Tabele dimenzija

Drugi tip tabela u zvezdastom dimenzionom modelu su tabele dimenzija. Osnovne karakteristike dimenzionih tabela su:

- Sadrže relativno statičke podatke koji se redje menjaju u odnosu na podatke u tabelama činjenica.

- Uglavnom sadrže opisne (deskriptivne) podatke. Ovi podaci se u upitima koriste za postavljanje uslova pri pretraživanju tabela činjenica i predstavljaju zaglavlja kolona u pregledima.
- Ove tabele su po obimu podataka koje sadrže manje od tabela činjenica.
- Primarni ključ u dimenzionim tabelama je surogat, veštački ključ (celobrojnog tipa). To je ključ koji je sistemski generisan i nezavisan od prirodnog ključa. U dimenzionoj tabeli prirodni ključ može da postoji kao atribut. Surogat ključevi su, uglavnom, kraći od prirodnih ključeva, značajno smanjuju veličinu tabela činjenica i poboljšavaju performanse upita. To su ključevi koji se ne menjaju čak ni kada dodje do promene prirodnog ključa u izvornom transakcionom sistemu. Primer prirodnog ključa je bar kod proizvoda u sistemu prodaje.
- Dimenzione tabele su povezane stranim ključem sa tabelama činjenica.
- Jedna od dimenzionih tabela uvek je i vremenska dimenzija.
- Podaci u dimenzionim tabelama uglavnom se retko menjaju, pa se kaže da su dimenzije sporomenjajuće (eng. *slowly changing*). Dinamika promena vrednosti dimenzionih atributa se meri mesecima i godinama, za razliku od tabela činjenica koje se dopunjavaju podacima sa svakim ciklusom osvežavanja baze skladišta.
- U zvezdastoj šemi tabele dimenzija su denormalizovane, tj. vrednosti atributa su upisane redundantno. Time se u tabelama dimenzija postižu:
 - Olakšano pretraživanje jer su svi povezani podaci na jednom mestu.
 - Poboljšane performanse upita.
 - Jednostavnost izrade modela.
- Smatra se da se uspešnost izrade dimenzionog modela i uspešnost njegove dalje implementacije ogleda, upravo, u bogatom opisu i potpunosti dimenzionih atributa. Zato se od dimenzionih tabela očekuje da sadrže što više atributa.

Prema [2], razlikujemo tri načina ažuriranja dimenzionih tabela:

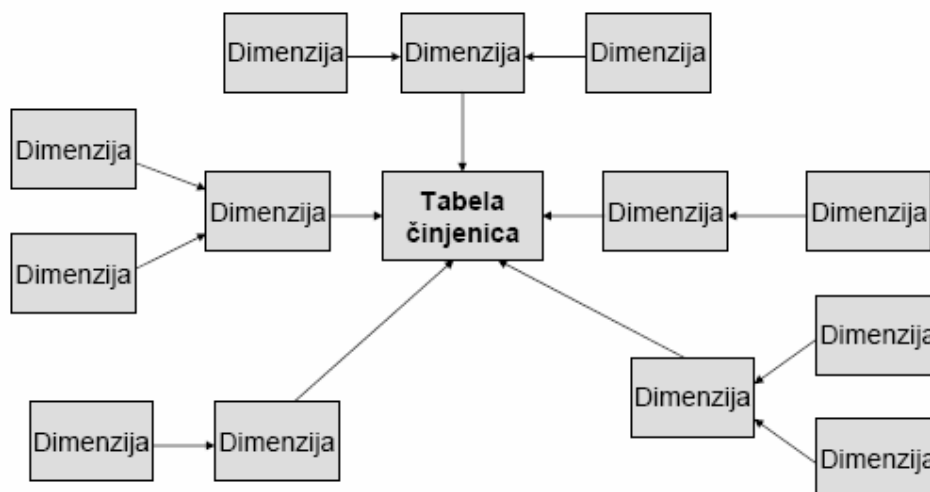
- Pri svakoj promeni vrednosti atributa nova vrednost se prepisuje preko prethodne. Na taj način se ne čuva istorija promena vrednosti tog atributa i svi činjenični podaci se gledaju isključivo u odnosu na poslednju, aktuelnu vrednost. Svi atributi u dimenzionim tabelama, osim primarnog ključa, mogu da se ažuriraju promenom sadržaja.

- Pri svakoj promeni vrednosti dimenzionog atributa upisuje se novi slog u tabeli dimenzije. Na ovaj način se čuva istorija promena vrednosti i obezbeđuju istorijske analize podataka na račun većeg obima podataka u bazi. Ova tehnika ažuriranja dimenzionih atributa zahteva evidentiranje perioda važenja svakog sloga u dimenzionoj tabeli.
- Čuvanjem, uporedo, tekuće i prethodne vrednosti dimenzionog atributa. Fizički, u tabeli postoje dva atributa, jedan sa tekućom vrednošću i drugi sa prethodnom vrednošću atributa čiji se istorijat prati. Dimenzione tabele obično sadrže i vremenski atribut u kome se čuva datum nastanka promene. Ova tehnika povećava složenost održavanja dimenzione tabele, ali nalazi svoju primenu u praksi. Ona omogućava tzv. „dvostruki“ pogled na poslovne procese: jedan sa aspekta tekuće vrednosti dimenzionog atributa i drugi sa aspekta prethodne vrednosti istog atributa (tj. šta bi bilo da do promene nije došlo). Ovaj pristup se koristi kada u samom poslovnom sistemu, ili od strane korisnika, postoje jasni zahtevi da se poslovne mere istovremeno posmatraju na oba načina.

Specifičan tip dimenzija su one koje u sistemu samo „glume“ dimenzije (eng. *role-playing*). To su dimenzije koje su fizički implementirane kao jedna osnovna tabela. Koriste se u slučajevima kada je potrebno da se na jednu tabelu činjenica spusti više stranih ključeva iz iste dimenzije. Ove dimenzije se formiraju kao pogledi nad istom baznom tabelom. Korisnički alat ih vidi kao odvojene dimenzije povezane (svaka svojim ključem) sa jednim datumskim poljem u tabeli činjenica. Pogledi se formiraju uobičajenom CREATE VIEW SQL naredbom i cilj im je da „igraju ulogu“ pravih dimenzija u sistemu.

4.3.2 Model pahuljice

Normalizacijom zvezdaste šeme dobija se tzv. model pahuljice koji je, mada u manjoj meri, takodje prisutan u projektovanju sistema za skladištenje podataka. Dimenziona tabela se normalizuje tako što se polja niske kardinalnosti izvlače u sopstvene tabele i povezuju spoljnim ključem sa glavnom dimenzionom tabelom. Uopšteni prikaz šeme pahuljice dat je na slici 11.



Slika 11. Uopšteni prikaz šeme pahuljice

Na ovaj način postiže se visok nivo normalizacije modela i smanjuje se veličina dimenzione tabele. Kako su dimenzione tabele znatno manjeg obima od tabela činjenica, to se ovim postupkom ne dobija velika ušteda prostora. Sa druge strane, model baze skladišta postaje složeniji i narušavaju se performanse upita. Zbog toga se ova tehnika retko koristi u sistemima za skladištenje podataka.

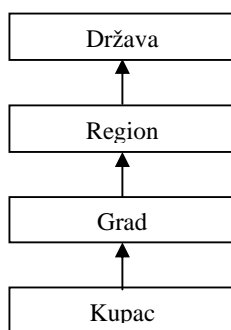
Izvesna normalizacija dimenzija opravdana je u slučaju „velikih“ dimenzija, koje zahtevaju poseban tretman. U pitanju su dimenzije koje sadrže milione slogova što značajno otežava njihovo pretraživanje. Ovakve su, na primer, dimenzije kupaca u prodajnim sistemima, ili dimenzija klijenata u telekomunikacionim ili bankarskim sistemima. Kako se u ovim dimenzionim tabelama čuvaju matični podaci o svakom pojedinačnom klijentu to u njima može da bude i više miliona stavki. Posebno je otežano pretraživanje po obeležjima kao što su: pol, starost, broj dece, godišnji prihodi, nivo obrazovanja i sl. To su, upravo, obeležja po kojima se najčešće postavljaju upiti i koja su glavni filter za izvlačenje podataka iz tabela činjenica. Smatra se da bi performanse upita bile značajno narušene kada bi se ove dimenzije uobličile u normalizovani, pahuljasti model. Kada je tretman velikih dimenzija u pitanju, najpogodnijom se smatra tehnika formiranja minidimenzija koje bi svojim jedinstvenim identifikatorom bile povezane sa dimenzionom tabelom klijenata, ili direktno sa tabelom činjenica. Preporuka je da se atributi, koji se najčešće koriste u pretraživanju ovih dimenzija, izvuku iz dimenzije u posebnu minidimenziju. U minidimenziji, u tom slučaju, nalaze se sve kombinacije vrednosti ovih atributa koje mogu da se pojave u skupu svih fizičkih lica glavne dimenzije. Svaka kombinacija ima svoj jedinstveni identifikator. Taj identifikator se, kao spoljni ključ, spušta u dimenzionu tabelu klijenata ili direktno u tabelu činjenica. Broj slogova minidimenzije je značajno manji nego u glavnoj dimenziji i pretraživanje vrednosti ovako izolovanih atributa je neuporedivo brže. Sledeće o čemu, u ovakvim sličajevima, treba razmišljati je da se atributi koji sadrže kontinuirane vrednosti, kao što su godine starosti ili prihod, čuvaju

grupisani u intervalima. Tako se veličina minidimenzije smanjuje, a samim tim i performanse upita značajno poboljšavaju.

4.3.3 Hijerarhije. Tehnike „bušenja“ podataka

Hijerarhije poslovnog sistema opisuju organizacionu strukturu i logičke veze tipa „roditelj-dete“ unutar podataka. Hijerarhijski podaci se nalaze među dimenzionim podacima. Svaka dimenzija može da sadrži jednu ili više hijerarhija.

Primer hijerarhije u sistemu prodaje prikazan je na slici 12.



Slika 12. Hijerarhija kupaca u sistemu prodaje

Vremenska dimenzija, po pravilu, sadrži višestruke hijerarhije. Jednu hijerarhiju čine: kalendarski dan, nedelja, mesec, kvartal, godina. Druga hijerarhija je: fiskalna nedelja, fiskalni mesec, fiskalni kvartal, fiskalna godina za potrebe računovodstva.

Postojanje hijerarhija nad podacima u dimenzijama omogućava jednu od ključnih tehnika koja se primenjuje u analizama podataka u skladištima, a to je *bušenje* (eng. *drill*) podataka. *Bušenje* omogućava dobijanje podataka na višem, odnosno nižem nivou detaljnosti:

- *Bušenje naniže* (eng. *drill down*) podrazumeva prolaz do podataka veće granularnosti, odnosno podataka višeg nivoa detaljnosti.
- *Bušenje naviše* (eng. *drill up*) je obrnuti proces koji omogućava prolaz ka podacima nižeg nivoa detaljnosti, odnosno podacima manje granularnosti.

Bušenje se može vršiti na nekoliko načina:

- Duž jedne hijerarhije.
- Prelaskom sa hijerarhije na nehijerarhijski atribut.
- Prelaskom sa jedne na drugu hijerarhiju, preko nehijerarhijskog atributa.

Poprečno bušenje (eng. *drill accross*) je postupak kombinovanja podataka iz više tabela činjenica. Nad svakom tabelom činjenica se, pri tome, formiraju nezavisni upiti. Rezultati tih upita se, zatim, uparuju spoljnim spajanjem nad zajedničkim dimenzijama tabela činjenica. Nivo detaljnosti podataka koje ovakav postupak može da vrati određen je najmanjom zajedničkom dimenzionalnošću dve tabele činjenica. Zajednička dimenzionalnost, uglavnom, podrazumeva agregaciju ovih tabela činjenica, ili bar jedne od njih. Ukoliko se ovakvi upiti često postavljaju, preporuka je da se formira agregirana tabela činjenica koja se čuva u portfoliju svih agregata sistema za skladištenje.

5 AGREGATI U SKLADIŠTIMA PODATAKA

Sumarne tabele činjenica (agregati) su tabele činjenica koje sadrže agregirane slogove osnovnih tabela činjenica, tj. onih tabela u kojima se nalaze najdetaljniji podaci. Agregirani podaci se, najčešće, dobijaju kao rezultat primene neke od agregatnih funkcija (npr. sumiranje, traženje proseka i sl.) na detaljnijim podacima. Dakle, jedina razlika koja postoji između osnovnih i sumarnih tabela činjenica je u nivou detaljnosti podataka. Dok osnovna tabela činjenica može da sadrži podatke o svakoj transakciji prodaje nekog proizvoda, to njene sumarne tabele sadrže samo delimično agregirane (tzv. lako agregirane) ili potpuno agregirane podatke. Tako se iz podataka koji su u osnovnoj tabeli sačuvani na dnevnom nivou, u tabeli agregata delimičnim agregiranjem dobijaju podaci na nedeljnom nivou (npr. nedeljna prodaja nekog proizvoda). Daljim sumiranjem dobijaju se podaci sve višeg nivoa agregiranosti (npr. mesečna prodaja proizvoda i sl.). Implementacija sumarnih tabela činjenica može da zahteva agregiranje nekih od osnovnih dimenzija sistema, ili potpuno izostavljanje nekih dimenzija iz modela. Iz osnovne datumske dimenzije može da se izvede agregirana vremenska dimenzija u kojoj se nalaze podaci o nedeljama, mesecima ili godinama. Iz dimenzije kupaca može da se izvede dimenzija gradova, okruga ili država. Prvo i osnovno pravilo skladištenja podataka, koje mora da bude ispoštovano, je da se vremenska dimenzija nikada ne izostavlja. Ona mora da postoji kako u osnovnim zvezdastim šemama, tako i u agregatnim zvezdastim šemama. Drugo pravilo je da dimenzije izvedene delimičnim sumiranjem osnovne dimenzije moraju da budu jednobrazne⁴.

Osnovni cilj agregiranja tabela činjenica je poboljšanje performansi upita, ali i planiranje kapaciteta, jer što su podaci detaljniji oni zahtevaju sve veći prostor u bazi. Zahtev za prostorom postaje posebno značajan nakon izvesnog vremenskog perioda, npr. posle pet do deset godina punjenja skladišta podataka.

Osnovne prednosti agregata u sistemima za skladištenje, ujedno i razlozi za njihovo implementiranje, su:

- Obezbeđuju kraće vreme odziva prilikom analitičkog izveštavanja, što je sa korisničkog aspekta jedna od najvažnijih performansi sistema.
- Obezbeđuju optimizovano i poboljšano korišćenje računarskih resursa, memorije, disk prostora i sl.
- Olakšavaju i ubrzavaju analitičke tehnike bušenja podataka sa višeg na niži nivo detaljnosti i obrnuto.

Da bi se implementirani agregati najoptimalnije iskoristili potrebne su softverske komponente (tzv. navigatori, eng. *Aggregate Navigator, Summary Navigator, Summary Advisor*) namenjene radu sa agregatima. Ove komponente omogućavaju da postavljeni korisnički upit prepozna postojanje

⁴ Principi jednobraznosti dimenzija razmatrani su u poglavlju 4.2.5.

agregata u sistemu, odnosno one transformišu upit krajnjeg korisnika tako da on pronadje i uključi najadekvatniji implementirani agregat. Najadekvatniji agregat je, uglavnom, najmanja osnovna tabela koja može da odgovori postavljenom upitu.

Navigatori imaju i svoju metabazu, u kojoj se nalaze opisi implementiranih sumarnih tabela. Metabaza može da beleži statistiku postavljanja upita. Na osnovu toga se dolazi do zaključka koji se agregati u sistemu koriste i koliko često, koje bi agregate trebalo implementirati da bi se ubrzali spori upiti i sl.

Softverske komponente namenjene radu sa agregatima mogu biti implementirane na dva načina:

- U sklopu baze podataka. Ovo je bolji pristup, jer prednosti softverskih komponenti mogu da iskoriste sve aplikacije koje rade nad tom bazom podataka. Takva je npr. Oracle 9i baza. Metabaza je zajednička i centralizovana.
- Kao deo korisničkih alata. Ovo je manje pogodan ali češći slučaj. Jedan od razloga leži u tome što puno baza podataka ne sadrži integrisane komponente za rad sa agregatima. Svaki korisnički alat ima svoje specifičnosti u radu s agregatima i održava svoju zasebnu metabazu, što je jedan od osnovnih nedostataka ovog pristupa.

U slučaju kada se ne raspolaze ni bazom podataka, a ni alatima koji omogućavaju dinamički rad sa agregatima, moguće je raditi i bez tih pogodnosti. Ovo je manje fleksibilan, ali u praksi rasprostranjen pristup koji ne omogućava skladištu podataka da iskoristi sve prednosti rada sa agregatima i ograničava rad sa njima. U ovom slučaju nema softverske komponente koja u fazi izvršavanja upita iz portfolia svih agregata izvlači najpogodniji (najoptimalniji, najmanji) agregat koji zadovoljava uslove upita i upit preusmerava sa osnovne tabele činjenica na njega. Zato neko drugi, a to je IT lice, mora da, za svaki postavljeni upit, unapred odredi osnovnu ili sumarnu tabelu nad kojom se upit izvršava. To znači da moraju unapred da se sagledaju svi predefinisani upiti koji će se postavljati nad skladištem podataka. Svaki nepredvidjeni, ad-hok upit, koji ne odgovara nijednom predefinisanom upitu izvršavaće se nad osnovnom tabelom činjenica.

Pokušaj da se u samom upitu referencira neki od agregata ima svojih nedostataka:

- Svaka izmena ili brisanje agregata upit čini nevažećim.
- Dodavanje novog, optimalnijeg agregata ne poboljšava performanse upita jer on nema mogućnost dinamičkog preusmeravanja izvršavanja.
- Osvežavanje agregata mora da se odvija uporedo sa osvežavanjem osnovnih tabela. Ukoliko to nije moguće, celo skladište podataka se stavlja van upotrebe dok se ne završi kompletno osvežavanje baze. Da bi mogli da se izvršavaju upiti, agregat na koji se eksplicitno referencira mora uvek da bude validan i dostupan, jer ne postoji

moгуćnost njegovog dinamićkog iskljućivanja i preusmeravanja upita na neku drugu tabelu skladišta.

Punjenje i osveŹavanje agregata radi se slićno punjenju i osveŹavanju osnovnih tabela. Preporuka pri punjenju je da se sprovodi kroz odvojene procese, tj. posebnom procedurom za svaku sumarnu tabelu. Ovo se preporućuje, pre svega, zbog lakšeg odrŹavanja i maksimalnog izbegavanja redundanse u procesima punjenja. OsveŹavanja agregata se mogu obavljati inkrementalno i potpuno (kompletno osveŹavanje). Inkrementalno osveŹavanje podrazumeva da se pri svakom osveŹavanju agregata zahvataju samo novi i, retko, izmenjeni slogovi osnovne tabele. Ovaj pristup zahteva da se u slogovima osnovne tabele ćuva vreme poslednje izmene sloga kako bi se, pri novom osveŹavanju agregata, identifikovao skup slogova koji će biti zahvaćeni procedurama za osveŹavanje. Potpuno osveŹavanje agregata podrazumeva brisanje celokupnog sadrŹaja sumarne tabele i punjenje pri svakom ciklusu osveŹavanja. Potpuno osveŹavanje je jednostavniji pristup, ali svaki put zahteva obradu celog skupa podataka i ne ćuva kljućeve u tabelama. Inkrementalno osveŹavanje je sloŹenije, ali se radi sa manjom kolićinom podataka i ćuvaju se kljućevi u tabelama. Agregati se osveŹavaju nakon osveŹavanja osnovnih tabela ili uporedo sa njima u istom procesu, i to sledećim redosledom: osnovne tabele dimenzija, osnovne tabele ćinjenica, agregirane dimenzije ukoliko postoje, agregirane tabele ćinjenica. Dinamika osveŹavanja agregata uvek je usaglašena sa dinamikom osveŹavanja osnovnih tabela i odvija se u istim vremenskim intervalima.

Komercijalni proizvodi na trŹištu nude svoja rešenja za implementiranje agregata u skladištima podataka. Takva rešenja su Oracle materijalizovani pogledi (eng. *Materialized Views*) i DB2 materijalizovane tabele (eng. *Materialized Query Tables*). Oba ova rešenja se koriste iskljućivo za implementaciju sumarnih tabela ćinjenica, ali ne i sumarnih dimenzija. Oracle materijalizovani pogledi su sumarne tabele koje se formiraju SQL naredbom `CREATE MATERIALIZED VIEW ime_pogleda AS SELECT...` Razlika u odnosu na uobićajene, relacione poglede je u tome što se rezultati ovih upita ćuvaju (materijalizuju) u bazi podataka u vidu tabele. Kada krajnji korisnik postavi upit koji ukljućuje tabele, kolone i veze identićne onima koje su u definiciji pogleda, optimizator ugradjen u bazi podataka prevodi upit i prebacuje izvršavanje na materijalizovani pogled. Vraćanje rezultata upita se znaćajno ubrzava. Materijalizovani pogledi ne zahtevaju procese punjenja i osveŹavanja kao ostale tabele u bazi skladišta podataka.

Dakle, dimenzije povezane sa sumarnim tabelama ćinjenica mogu biti osnovne dimenzije sistema ili delimićno agregirane sumarne dimenzije. Situacije u kojima agregatna zvezdasta šema sadrŹi i agregiranu dimenziju su, uglavnom, teško razrešive na nivou baze podataka. Sistem za upravljanje bazom podataka „zna“ da se iz jedne osnovne tabele izvodi jedan ili više agregata. Ono što baza podataka „ne zna“ to je da postoji veza izmedju agregirane dimenzije i agregirane tabele ćinjenica. Zato proizvodjaći kao što su Oracle i IBM nude moгуćnost delimićnog agregiranja dimenzija na logićkom nivou, bez fizićke implementacije agregirane dimenzije u bazi. Agregirane dimenzije se ne implementiraju, već se definišu preko hijerarhija u osnovnim dimenzijama sistema. Svaki nivo hijerarhije, pored ostalih

atributa, mora da sadrži i jedinstveni identifikator. Taj jedinstveni identifikator se, umesto stranog ključa iz agregirane dimenzije, spušta u sumarnu tabelu činjenica, tj. u materijalizovani pogled ili materijalizovanu tabelu. Najoptimalnije rešenje je da ključ svakog hijerarhijskog nivoa bude prost, tj. sastavljen od jednog atributa. Ukoliko to nije slučaj, najbolje je da se formira surogat ključ koji bi jedinstveno odredjivao svaki nivo hijerarhije. Surogat ključ ostaje nedostupan za krajnje korisnike i služi samo da bi se postigla jedinstvenost u nivoima hijerarhije.

Zaključak je da postojanje sumarnih tabela u sistemu za skladištenje značajno poboljšava njegove performanse. Najbolje je da se već u fazi analize i definisanja modela skladišta podataka uoče i dokumentuju svi potencijalni agregati, čak i oni koji kasnije neće biti implementirani. Vrednost agragata, sa aspekta performansi sistema, najjednostavnije se određuje na osnovu broja slogova osnovne tabele koji su sumirani u jednom agregiranom slogu. Ako je taj odnos bar 20:1 smatra se da je to značajno poboljšanje performansi sistema i ta sumarna tabela treba da bude implementirana.

6 EKSTERNI PODACI U SKLADIŠTIMA PODATAKA

Sa stanovišta jednog preduzeća eksterni su svi oni podaci koji ne postoje u sistemima tog preduzeća, već u njega dolaze iz spoljnog okruženja i to, najčešće, u nestruktuiranom, nepredvidivom formatu. Baza skladišta podataka se smatra pogodnom za smeštanje ovakvih podataka. Eksterni podaci dolaze iz časopisa, sa Veb stranica, putem elektronske pošte i sl. Poslovni analitičari ove podatke koriste u svojim izveštajima, analizama i pregledima. Oni ove podatke, najčešće, koriste tako što ih pročitaju na VEB stranici ili u nekom časopisu i upišu u svoj izveštaj. Na taj način se trajno gubi veza eksternog podatka sa njegovim izvorom. Nakon izvesnog vremena postaje nemoguće da se ti podaci ponovo izvuku iz istog izvora i iskoriste. Većina eksternih podataka (npr. oni koji dolaze sa Interneta) je nestruktuirana i samim tim nekorisna u izvornom obliku i zahteva ozbiljne rekonstrukcije pre upotrebe.

Neki od problema na koje se nailazi u radu sa eksternim podacima su:

- U većini slučajeva se ne zna unapred kada će biti dostupni novi podaci. Teško je predvideti vremenski trenutak ili događaj koji uslovljava dostupnost novih podataka. Često je potrebno (kao u slučaju VEB podataka) osmisliti programe za praćenje koji bi signalizirali pristizanje novih podataka. Eksterni podaci su nepredvidivi kako po pitanju izvora, tako i po pitanju vremena. To znači da mogu doći iz bilo kog izvora u bilo kom trenutku.
- Eksterni podaci su neprilagodjeni i nekompatibilni sa internim podacima koji već postoje u sistemu za skladištenje. Potrebno ih je prestrukturirati, preformatirati i usaglasiti sa internim podacima po pitanju ključeva.
- Strukture ovih podataka su zahtevne, kako po pitanju prostora, tako i po pitanju njihove obrade. Ovde se, pre svega, misli na podatke kao što su slike i zvučni zapisi. Dok se slike čuvaju u predviđenom formatu, zvučni podaci se čuvaju digitalizovano i pri pozivu se prevode u zvučni format. Tehnologije namenjene obradi i čuvanju ovakvih podatke su relativno nove. Čuvanje ovih podataka zahteva veliki fizički prostor pa se, uglavnom, čuvaju na nekom sekundarnom medijumu. U cilju efikasnog pristupa ovim podacima i poboljšanja performansi pribegava se naprednim tehnikama indeksiranja. Indeksiranje omogućava da se upiti, sasvim ili delimično, izvršavaju preko indeksa, bez direktnog pristupa eksternim podacima. Na osnovu indeksa može da se vidi koji su to eksterni podaci potrebni u izvršenju upita. Na taj način se, preko indeksa, eksterni podaci povezuju sa bazom skladišta i u skladište se prevlače samo oni podaci za kojima postoji potreba.

Zahtevni eksterni podaci mogu da se čuvaju u samom skladištu podataka, ukoliko je to pogodno i nije previše skupo. Osim toga, mogu da se smeštaju i izvan baze skladišta, a da se u metapodacima čuva informacija o tome gde se

oni nalaze. Kada se eksterni podaci povežu sa bazom skladišta oni postaju dostupni svim organizacionim delovima preduzeća. Time se izbegava situacija u kojoj svaki organizacioni deo te podatke dobija, prepisuje, čuva i održava za sebe, što se pokazalo kao neracionalno i nepouzdana rešenje koje uslovljava greške, ali i dupliranje vremena i poslova u organizacionim delovima jednog preduzeća.

Što se modeliranja struktura eksternih podataka tiče, teško je, čak i pogrešno, na ove nestruktuirane podatke primenjivati model postojećih, internih podataka. Eventualno, može da se identifikuje podskup ovih podataka koji je kompatibilan sa postojećim internim podacima.

Svaki podatak, bio on interni ili eksterni, ima svoj životni vek, nakon čega se pristupa nekom načinu njegovog arhiviranja. Eksterni podaci mogu, nakon izvesnog vremena, da budu uklonjeni iz baze skladišta. Oni se tada sklanjaju na neko mesto koje je jeftinije za održavanje, pri čemu se u metapodacima upisuje referenca na novu arhivsku lokaciju tih podataka.

7 FIZIČKI MODEL SKLADIŠTA PODATAKA

Poslednji korak pre puštanja sistema za skladištenje podataka u produkcionu rad je fizička implementacija formiranog dimenzionog modela. Na nivou fizičke implementacije dimenzioni model se dopunjava specifičnostima konkretne baze podataka. Dakle, dok je dimenzioni model nezavisan od platforme i tehnologije, fizički model skladišta podataka zavisi od osobina baze podataka i servera na kome će biti implementiran. Smatra se da je uspešnost fizičke implementacije jedan od najbitnijih preduslova za uspeh celog projekta uvođenja sistema za skladištenje.

Kada se govori o performansama sistema za skladištenje podataka misli se, pre svega, na optimizaciju aktivnosti:

- Punjenja podataka.
- Pristupa podacima.
- Čuvanja podataka.
- Arhiviranja podataka.
- Brisanja podataka iz sistema.
- Praćenja podataka.

Pre izrade fizičkog modela koriste se projektantske tehnike koje imaju za cilj poboljšanje performansi sistema, od kojih je većina već pomenuta u ovom radu:

- Pažljivo određivanje nivoa granularnosti podataka.
- Spajanje u jednu tabelu činjenica koje se koriste u istom upitu.
- Uključivanje redundantnih podataka u dimenzioni model.
- Uključivanje izvedenih podataka u dimenzioni model.
- Implementiranje sumarnih tabela.
- Pažljivo imenovanje objekata uz poštovanje standarda imenovanja.
- Definisane ključeva, formata podataka, null vrednosti i veza između podataka u sistemu.

Uspešnost fizičke implementacije se ogleda, pre svega, u postignutoj brzini upisa podataka u hardverski prostor i brzini vraćanja podataka iz njega pri izvršavanju upita. Tehnike koje se na fizičkom nivou primenjuju u cilju dodatnog poboljšanja performansi sistema za skladištenje su particionisanje (podela), indeksiranje i paralelizam.

7.1 Particionisanje

U sistemima za skladištenje podataka se ne postavlja pitanje da li je potrebno particionisanje, već kako ga izvesti. Mnogi analitičari smatraju da su upravo granularnost i pravilno particionisanje ključni za uspešnost projektovanja i implementacije skladišta podataka. Particionisanje podataka predstavlja razdvajanje podatka u zasebne, manje fizičke jedinice kojima se nezavisno pristupa. Pristup podacima je olakšan, što je posebno važno u

sistemima koje odlikuje masovnost podataka. Osnovne metode partitionisanje su:

- Horizontalno partitionisanje. Izbor kriterijuma za partitionisanje je odluka IT projektanta, ali je u sistemima za skladištenje podataka skoro obavezno da se partitionisanje vrši po datumskoj i eventualno po još nekoj dimenziji. Tako u sistemima osiguravajućih agencija, particije mogu biti određene godinom i linijom osiguravajućih polisa. Grupisanje u particije bi išlo po sistemu: 2001/zdravstvene polise, 2002/zdravstvene polise, 1999/životno osiguranje, 2000/životno osiguranje itd.
- Partitionisanje intervala. Slično horizontalnom, zasniva se na grupisanju podataka prema intervalima vrednosti određenih atributa. U istoj particiji se nalaze samo slogovi u kojima su vrednosti atributa u definisanom skupu vrednosti.
- Vertikalno, po kolonama tabele. Primenjuje se u sistemima u kojima se često pristupa pojedinačnim kolonama tabela. Ukoliko se sadržaji nekih kolona tabela retko menjaju takve kolone se grupišu u posebne particije i na taj način razdvajaju od kolona čiji se sadržaji često menjaju. Preporuka je i da se dugački opisni, tekstualni dimenzioni atributi odvoje u posebne particije.
- Heš partitionisanje. Ovaj način omogućava partitionisanje uz minimalne konfiguracione zahteve. Sistem za upravljanje bazom podataka automatizovano, korišćenjem heš funkcije raspoređuje podatke po particijama.

7.2 Indeksiranje

Pravilan izbor indeksa je od posebne važnosti u sistemima za skladištenje podataka zbog velike količine podataka koja se u njima nalazi. Broj indeksa u sistemima za skladištenje je znatno veći nego u transakcionim sistemima. Indeksi se uvode da bi povećali brzinu izvršavanja upita. Oni treba da podrže najrazličitije kombinacije upita koji se mogu postaviti u konkretnom sistemu za skladištenje.

Strategija indeksiranja zasnovana je na sledećim pravilima:

- Indeksi se postavljaju nad kolonama koje se često pojavljuju u WHERE klauzuli upita. One nisu obavezno deo primarnog ili stranog ključa.
- Neki transakcioni sistemi ograničavaju broj indeksa. Sistemi za skladištenje podataka su po tom pitanju mnogo fleksibilniji.
- Sporo izvršavanje upita je znak neadekvatnog ili nepostojećeg indeksiranja. U tom slučaju radi se o loše organizovanim indeksima ili nedovoljnom broju indeksa nad tabelama. Pri tome treba imati u vidu

da previše postavljenih indeksa zahteva veliki hardverski prostor i usporava proces upisivanja podataka u skladište.

- Preporuka je da se indeksi ne koriste nad tabelama sa malim brojem slogova. U ovakvim slučajevima prednosti indeksa su zanemarljive u odnosu na cenu njihovog održavanja.

U sistemima za skladištenje podataka koriste se: B-drvoidni indeksi, bitmapirani indeksi, ali i specifični tipovi indeksa kao što su sekundarni, privremeni, dinamički i rang indeksi.

7.3 Paralelizam

Paralelizam omogućava da se složeni zahtevi, u cilju povećanja brzine, izvršavaju istovremeno na jednom ili više procesora. Većina sistema za upravljanje bazama podataka danas ima mogućnost paralelnog izvršavanja upita. Na taj način se jedan zahtev za podacima deli između više serverskih procesa, tj. sama baza podataka vrši deljenje upita na procese koji se istovremeno izvršavaju. Paralelizam na nivou baze podataka može značajno da ubrza proces punjenja podataka u skladište. Paralelni režim rada omogućava i: paralelno formiranje tabela i indeksa, paralelno sortiranje, paralelno uzimanje sigurnosnih kopija (eng. *backup*) i oporavak (eng. *recovery*) sistema i sl.

8 SKLADIŠTA PODATAKA I TEHNOLOGIJE

Zbog svih navedenih razlika koje postoje između sistema za skladištenje podataka i standardnih transakcionih sistema, sistemi za skladištenje podataka imaju specifične tehnološke zahteve. Pod tehnološkim zahtevima podrazumevaju se podešavanja, parametrizacija i sistemski rešenja koja se sprovode u cilju poboljšanja performansi baze podataka i uspešne implementacije sistema za skladištenje.

U ovom poglavlju biće sumirani osnovni zahtevi koji se postavljaju pred tehnologijama koje treba da obezbede uspešno funkcionisanje sistema za skladištenje.

Osnovni tehnološki zahtevi, prema [1], skladišta podataka su sledeći:

- Uspešan rad sa velikim količinama podataka, koje se u sistemima za skladištenje mere gigabajtima i terabajtima. U transakcionim sistemima količina podataka se izražava u megabajtima i gigabajtima. Količina podataka je velika zato što skladišta podataka mogu da sadrže detaljne, ali i istorijske i sumarne podatke.
- U direktnoj vezi sa količinom podataka je i rad sa različitim medijima za smeštanje tih podataka. To mogu biti: glavna memorija, proširena memorija, keš memorija, disk, magnetna traka i sl. Postoje razlike u brzini i ceni pristupa različitim medijumima. Zbog toga, kao i zbog količine podataka u skladištu, preporuka je da sistem za skladištenje bude smešten na više različitih medijuma. Dakle, podaci mogu da se čuvaju na više fizičkih uređaja, pa je neophodno da se omogući i rad sa podacima u paralelnom režimu.
- Jedan od osnovnih tehnoloških zahteva sistema za skladištenje je brz i lak pristup podacima. Tu se, pre svega, misli na mogućnost indeksiranja korišćenjem pomenutih specifičnih tipova indeksa kao što su sekundarni, privremeni, dinamički i rang indeksi.
- Mogućnost praćenja aktivnosti (eng. *monitoring*) u skladištu podataka. Na taj način se otkrivaju eventualne potrebe za reorganizovanjem podataka ili promenom načina indeksiranja. Prati se brzina i efikasnost pristupa podacima, meri se preostali prostor za smeštanje podataka i planiraju se kapaciteti.
- Neophodan je odgovarajući jezički interfejs (uglavnom zasnovan na SQL-u), koji je efikasan i omogućava korišćenje jednog ili više specifičnih indeksa u upitima. To mogu biti: jezici za statističke analize, jezici za istraživanje podataka, jezici za pristup podacima u bazi, jezici za rad sa predefinisanim upitima i sl.
- Efikasno punjenje podataka u bazu skladišta (slog po slog ili skup slogova odjednom). Upis podataka u bazu skladišta može da se, kako

je pomenuto, odvija paralelno i to kroz nekoliko procesa koji se nezavisno i istovremeno izvršavaju.

- Obezbediti kompaktnost podataka koja podrazumeva smeštanje podataka na najmanjem mogućem fizičkom prostoru. Kompaktno čuvanje podataka je posebno važno u sistemima za skladištenje podataka, jer su to relativno stabilni sistemi u kojima se retko rade operacije izmena postojećih podataka. Kada se podaci kompaktno čuvaju procesi upisa podataka u bazu i vraćanja podataka iz baze su efikasniji.
- Neophodno je da se obezbedi korišćenje složenih ključeva. Složeni ključevi su uvek prisutni u tabelama činjenica.
- Koncept zaključavanja slogova, koji u transakcionim sistemima onemogućava da više korisnika menja sadržaj istog sloga tabele u isto vreme, u sistemima za skladištenje nema veliki značaj. Softver koji obezbeđuje ovu funkcionalnost (eng. *lock manager*) zauzima značajne resurse čak i kada se ne radi izmena slogova u tabeli. Zato je jedan od zahteva sistema za skladištenje da se omogući njegovo uključivanje i isključivanje, po potrebi.
- Važan tehnološki zahtev je što brži oporavak sistema u slučaju havarije. Oporavak se vrši pomoću podataka sa sekundarnih medija (magnetna traka, disk i sl.). Kako su baze skladišta velike predviđa se delimični oporavak, dok se opcija potpunog oporavka baze smatra skoro nemogućom. Od koristi može da bude i korišćenje dijagnostičkih alata koji omogućavaju da sam sistem za upravljanje bazom podataka otkrije grešku i locira oštećene podatke.

Kako skladišta podataka ne predviđaju izmene i brisanje slogova, to još neke (veoma zahtevne) funkcionalnosti sistema za upravljanje bazama podataka ovde gube na značaju. To su: tačke preseka stanja sistema (eng. *checkpoint*), mrtve petlje i sl.

Osim toga, baze parametrizovane za rad sa transakcionim sistemima u trenutku upisa i izmene sloga rezervišu prostor koji može da zauzima i do 50% praznog prostora u bloku. Sistemi za skladištenje nemaju potrebu za zauzimanjem slobodnog prostora pri upisu sloga, jer se ne očekuje izmena tih podataka.

Jedno od pitanja koje se postavlja je da li postoji mogućnost promene tehnologije u trenutku kada je sistem za skladištenje u upotrebi. Do potrebe za promenom tehnologije može da dodje zbog porasta količine podataka, dostupnosti novih tehnologija na tržištu, neadekvatnosti postojeće tehnologije ili povećanog broja korisnika skladišta podataka. Pre donošenja konačne odluke preporuka je da se odgovori na sledeća pitanja: da li će promena tehnologije zadovoljiti postavljene zahteve, može li da se izvede i na koji način konverzija starog sistema za upravljanje bazom podataka u novi, mogu li i kako da se konvertuju softverski programi? U opštem slučaju,

promena tehnologije je moguća i kod sistema za skladištenje podataka koji su u produkciji.

9 METAPODACI

Važna komponenta sistema za skladištenje podataka su metapodaci (tj. „podaci o podacima“). Metapodaci su oduvek bili deo informacionih sistema, ali su oni u svetu skladištenja podataka posebno značajni. Metapodaci obezbeđuju efikasnije korišćenje skladišta podataka. Oni omogućavaju krajnjim korisnicima da sagledaju mogućnosti skladišta, da saznaju odakle da počnu svoje analize, da brzo dodju do saznanja da li postoje i, ako postoje, gde se nalaze traženi podaci. Metapodaci skladišta sadrže detaljne opise lokacija, struktura i značenja podataka, kao i opise ključeva i indeksa nad podacima. U metapodacima su dokumentovana pravila preslikavanja, algoritmi i poslovna pravila koja su korišćena u fazi prevodjenja i sumiranja podataka.

Dok su u standardnim transakcionim sistemima korisnici metapodataka bili isključivo IT stručnjaci, to su u sistemima za skladištenje podataka i krajnji korisnici ti koji pristupaju metabazi i koriste metapodatke. Merilo uspešnosti implementacije skladišta podataka je, između ostalog, dostupnost i lakoća snalaženja krajnjih korisnika u metabazi.

Dakle, osnovna pitanja na koja metapodaci treba da odgovore krajnjim korisnicima su:

- Koje informacije i o kojim poslovnim subjektima su sadržane u skladištu?
- Od kada se podaci o nekom subjektu sakupljaju i skladište?
- Kako se pristupa odredjenim podacima, tj. koji upiti se nad njima mogu postavljati?
- Koja poslovna pravila i pretpostavke su ugrađene u informacije iz skladišta? To znači, iz kojih izvornih podataka i primenom kojih algoritama je izveden podatak u skladištu. Zahteva se detaljan opis svakog podatka i to poslovno razumljivim rečnikom. Korisnik se nikada ne sme dovesti u situaciju da nagadja kako je i iz kog izvornog podatka neki sumarni podatak izveden.
- Gde se u skladištu nalaze traženi podaci i koja je njihova starost i obim?
- Šta znače konkretne vrednosti podataka?
- Ko su vlasnici odredjenih podataka? Odnosno, ko je kontakt osoba koja može da objasni nejasnoće vezane za sadržaje podataka?

Ono što je IT stručnjacima dodatno važno to su metapodaci koji se odnose na procese izdvajanja, prevodjenja u punjenja podataka u skladište, tzv. ETL metapodaci. ETL metabaza sadrži podatke o izvorima podataka (nazivima

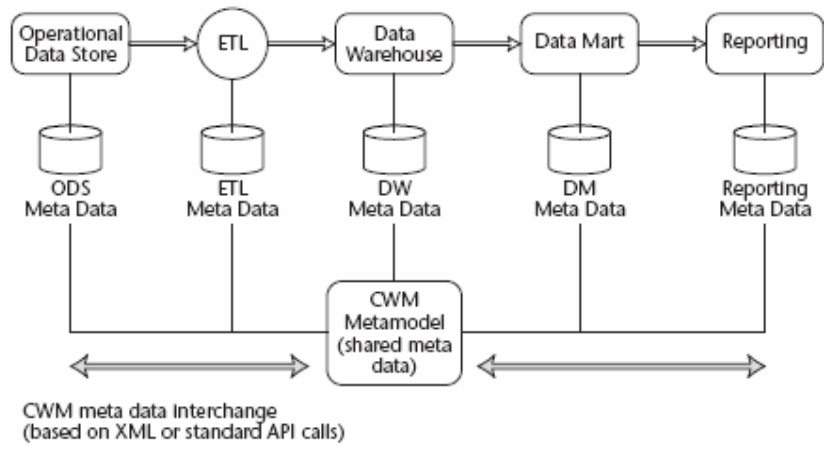
izvornih tabela, atributa i ključeva) i o tome na koji način su izvorni podaci prevedeni u sadržaje i formate koji se nalaze u skladištu. Tu su podaci o primenjenim pravilima izdvajanja, sumiranja, prečišćavanja, prenosa i punjenja podataka. U metapodacima se nalaze i vremenske odrednice formiranja i izmene svakog objekta metabaze.

Još jedan aspekt sistema za skladištenje podataka daje poseban značaj metapodacima. To je istorijski aspekt, odnosno postojanje istorijskih podataka i promena koje su se u skladištu odigravale u toku njegovog životnog ciklusa. Podaci u skladištu su stari, u proseku, pet do deset godina. U toku tog vremenskog perioda normalno je da dolazi do promena strukture skladišta. Zadatak metabaze je upravo čuvanje informacija o strukturnim promenama i vremenu njihovog dešavanja. U metapodacima smeštene su informacije o tome kako su se i kada menjale definicije podataka, pravila izvodjenja i primenjeni algoritmi. Bez ovih informacija gotovo je nemoguće uporedno pratiti vremenske serije podataka i razumeti sadržaj istorijskih podataka.

Svaki proces i proizvod koji je uključen i razvoj i implementaciju skladišta podataka ima svoje metapodatke. To su metabaze koje zavise od proizvoda i tehnologije koji su korišćeni prilikom formiranja metapodataka. Alati za modeliranje imaju svoje metabaze, ETL alati svoje, većina specijalizovanih korisničkih alata opet ima svoje metabaze. Najzad, sam sistem za upravljanje bazom podataka ima svoj rečnik podataka (eng. *Data Dictionary*). Svi ovi metapodaci ne moraju da budu smešteni u istoj bazi podataka u kojoj je i skladište podataka. Ono što je bitno u slučaju postojanja više odvojenih metanivoa, to je da se utvrdi da li postoji komunikacija, tj. razmena informacija između njih.

Kada je reč o metapodacima u sistemima za skladištenje, danas se smatra da je neformalni standard tzv. zajednički metamodel skladišta (eng. *Common Warehouse Metamodel, CWM*) promovisan od strane OMG grupe (*Object Management Group*). Ovaj standard uvodi jedinstveni jezik i mehanizme razmene metapodataka (xml razmena ili pozivi API procedura) između aplikacija, alata i proizvoda različitih proizvođača prisutnih na tržištu. To su komercijalni transakcioni sistemi, izveštajni alati, alati za analitičke upite, alati za istraživanje podataka i sl.

Grafički prikaz razmene metapodataka primenom CWM standarda dat je na slici 13.

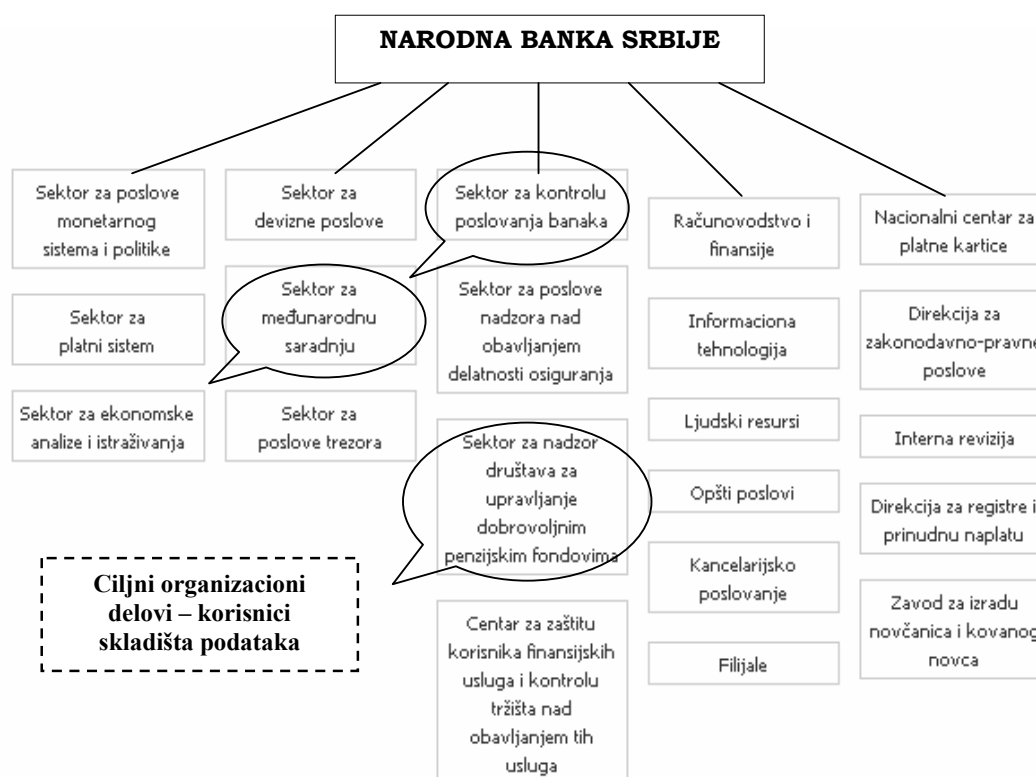


Slika 13. Prikaz integracije metapodataka korišćenjem zajedničkog metamodela

***II PRIMENE KONCEPATA SKLADIŠTENJA
PODATAKA U NARODNOJ BANCI SRBIJE***

1 UVODNE NAPOMENE

U ovom delu rada opisani su sistemi za skladištenje podataka implementirani u tri sektora Narodne banke Srbije: Sektoru za međunarodnu saradnju, Sektoru za nadzor društava za upravljanje dobrovoljnim penzijskim fondovima i Sektoru za kontrolu poslovanja banaka. Mesto ovih sektora u organizacionoj strukturi Narodne banke Srbije prikazano je na slici 1.



Slika 1. Osnovna organizaciona šema Narodne banke Srbije

Ova skladišta podataka su razvijana u Narodnoj banci Srbije u periodu od 2000. do 2009. godine, kroz tri odvojena projekta koja su realizovana u različitim vremenskim intervalima i na kojima su bili angažovani različiti projektni timovi. Sva tri skladišta se i danas aktivno koriste i predstavljaju osnovu analitičkog rada i izveštavanja u sektorima u kojima su implementirana. Činjenica je da u Narodnoj banci Srbije ne postoji jedinstveno skladište podataka koje bi obuhvatilo sve poslovne funkcije Centralne banke, već su ovi moduli razvijani manje ili više nezavisno, prema potrebama poslovanja i prema zahtevu krajnjih korisnika, a pre svega prema zahtevu rukovodstva u ciljnim sektorima. Pored ovde opisanih skladišta podataka, u Narodnoj banci je trenutno u upotrebi još nekoliko izveštajnih sistema (jedan veći i dva manja) koji su razvijeni primenom konceptata skladištenja podataka. U razvoju, kao i danas u održavanju ovih skladišta podataka, učestvuju relativno mali timovi sastavljeni od zaposlenih u IT

sektoru i ključnih poslovnih korisnika. Rezultati do kojih se došlo su proistekli iz zajedničkog višemesečnog rada zaposlenih u Informacionoj tehnologiji, korisnika u ciljnim sektorima i, naravno, rukovodstva koje je svaki od ovih projekata iniciralo, odobrilo i promovisalo kod svojih zaposlenih.

U ovom poglavlju navedene su opšte karakteristike ovih sistema za skladištenje: opis poslovnog sistema koji je obuhvaćen skladištem podataka, opis zatečenog stanja, zahtevi, motivi i ciljevi sa kojima se ušlo u implementaciju ovih skladišta, ali i rezultati dobijeni na kraju, uz sve pozitivne i negativne efekte. Na kraju ovog poglavlja ukratko su opisani dalji planovi vezani za unapredjenje postojećih i dalji razvoj skladišta podataka u Narodnoj banci.

Detaljniji prikaz tehničke implementacije ova tri modela dat je u prilogima ovog rada. U prilogima je za sva tri modela priložen opis dimenzija i tabela činjenica. Na kraju su prikazani primeri izveštaja koje su krajnji korisnici samostalno uradili u alatima predvidjnim za rad sa skladištima podataka.

Napominje se da je automatizacija poslovnih procesa u tri pomenuta sektora kroz implementaciju skladišta podataka prikazana, kako u ovom poglavlju tako i u prilogima, do nivoa detaljnosti koji ne ugrožava tajnosti informacija.

2 OPIS POSLOVNIH SISTEMA. OBUHVATNOST SKLADIŠTA

2.1 Penzijski fondovi

Prvi sistem za skladištenje podataka koji je opisan u ovom radu koristi se u Sektoru za nadzor društava za upravljanje dobrovoljnim penzijskim fondovima. Sektor je u Narodnoj banci počeo sa radom 2006. godine od kada je posebno aktivirana nadzorna funkcija Narodne banke u ovom segmentu finansijskog delovanja. Naime, jedna od funkcija Narodne banke Srbije je nadzor nad društvima za upravljanje dobrovoljnim penzijskim fondovima. Ova nadzorna funkcija je, prema [9], deo opšte funkcije uspostavljanja pouzdanog finansijskog sistema. Narodna banka Srbije propisuje normativne osnove, obavlja kontrolu i reaguje preventivno u cilju zaštite uložениh sredstava članova fondova, tj. u cilju adekvatnog upravljanja tim sredstvima. Automatizacija poslovnih procesa ovog sektora Narodne banke Srbije kroz implementaciju skladišta podataka radjena je u periodu od avgusta 2008. godine do februara 2009. godine. Tim koji je učestvovao u razvoju ovog skladišta brojao je troje zaposlenih iz IT sektora i dvoje ili troje poslovnih korisnika, zavisno od faze razvoja. Danas je skladište podataka osnova za analize i izveštavanje u sektoru. Ovo skladište podataka je razvijano medju poslednjima u Narodnoj banci, u vreme kada je nekoliko sličnih projekata već bilo završeno. To je omogućilo da se na ovom skladištu primene sva pozitivna iskustva od ranije i da se izbegnu neke tipične početničke greške koje su do tada pravljene. Ovo skladište je možda i najsvēobuhvatnije od pomenuta tri, jer su njime obuhvaćene skoro sve poslovne informacije kojima ciljni sektor raspolaže i koje analizira. Dakle, to je jedan celovit analitički i izveštajni sistem na kome se zasniva poslovanje sektora.

Najznačajnije poslovne kategorije koje su obuhvaćene ovim sistemom su: hartije od vrednosti, nepokretnosti, gotovina, depoziti, potraživanja, obaveze, penzijski planovi, transakcioni troškovi, naknade i podaci o članovima i individualnim računima članova fondova. Zahtev korisnika je bio da sistem obuhvati i podatke Beogradske berze, koji se na dnevnom nivou preuzimaju sa Veb stranice berze. U prvim razgovorima sa korisnicima je kao osnovni cilj budućeg skladišta navedeno da ono mora da omogući analizu kompletne neto i bruto imovine svakog pojedinačnog fonda, kao i celog sistema penzijskih fondova u zemlji. Imovina fonda je kategorija koja se dobija zbirnom analizom više grupa podataka po propisanoj metodologiji sektora. Zadatak je bio da se sistem postavi tako da se iz osnovnih komponenti lako formira ukupna imovina fonda. Obrnuto, polazeći od ukupne imovine trebalo je da se omogući prolaz korisnicima do svake pojedinačne, najdetaljnije komponente tog agregata.

Analiza imovine fondova je sigurno najznačajnija poslovna funkcija koja je u ovom sistemu uspešno implementirana. Detaljniji opis svih poslovnih funkcija koje su obuhvaćene skladištem podataka dat je u prilogu A.

Ovaj primer je u radu najdetaljnije opisan, posebno u poglavlju sa priložima. Ostala dva primera su manje detaljna, tj. opisane su razlike i specifičnosti svakog skladišta bez ponaljanja koncepata i principa koji su zajednički u svim primerima.

2.2 Kreditni odnosi sa inostranstvom

U drugom primeru u radu opisano je skladište podataka koje je, ujedno, najstariji sistem za skladištenje podataka koji je razvijen u Narodnoj banci Srbije. Ovo skladište se već deset godina, tačnije od sredine 2000. godine koristi u Sektoru za međunarodnu saradnju. Za proteklih deset godina ovaj sistem se menjao, doradjivao i proširivao shodno promenama poslovnih odluka i propisa i prilagođavao novim poslovnim potrebama i zahtevima.

Ukratko, u Narodnoj banci Srbije obavljaju se poslovi registrovanja i evidentiranja kredita koji se uzimaju od stranih lica, kao i kredita koji se odobravaju stranim licima u skladu sa zakonom o kreditnim poslovima sa inostranstvom, odlukom o načinu, rokovima i obrascima za registrovanje i evidentiranje kreditnih poslova sa inostranstvom i uputstvom Narodne banke Srbije za popunjavanje obrazaca za registrovanje i evidentiranje kreditnih poslova sa inostranstvom. Banke, preduzeća i druga pravna lica dužna su da registruju kreditne poslove sa inostranstvom u Narodnoj banci Srbije. Ovaj sistem za skladištenje obuhvata poslove vezane za kreditne odnose sa inostranstvom i to većim delom poslovne funkcije praćenja kreditnog zaduženja domaćih privrednih i finansijskih subjekata (banaka, preduzeća, državnih organa, Republike, Centralne banke i sl.) u inostranstvu. Drugi deo sistema koji je opisan u ovom radu obuhvata kreditne poslove u kojima su domaći privredni i finansijski subjekti kreditori, dakle poslove u kojima domaća lica kreditiraju inostrane firme. Ovaj aspekt kreditnih poslova sa inostranstvom je dugo bio u drugom planu, ali je činjenica da u poslednje vreme sve više dobija na značaju. Stoga se i analize podataka vezane upravo za kreditna odobrenja inostranstvu sve više aktiviraju. Poslovi kreditnog zaduženja i kreditnog odobrenja predstavljaju samo jedan deo poslova koje sektor obavlja u Narodnoj banci Srbije. Kako se kroz prijave kreditnog zaduženja i kreditnog odobrenja vidi kompletna kreditna aktivnost u zemlji koja se prati u Centralnoj banci, to je upravo ovaj deo poslovnih funkcija bio najpogodniji za razvoj skladišta podataka.

2.3 Bilansni podaci poslovnih banaka

Treći sistem za skladištenje podataka napravljen je za potrebe i po zahtevu Sektora za kontrolu poslovanja banaka Narodne banke Srbije. Skladište koje je opisano u ovom radu predstavlja deo opšteg informacionog sistema koji je implementiran u sektoru. Model obuhvata poslove vezane za praćenje i analizu podataka o prometu i stanju na knjigovodstvenim računima poslovnih banaka u Srbiji. U sklopu razvoja ovog skladišta podataka implementirana je metodologija za formiranje bilansa stanja i bilansa uspeha poslovnih banaka na osnovu dostavljenih podataka o prometu i stanju na knjigovodstvenim računima. U skladište podataka uključene su i poslovne

kategorije na osnovu kojih se vrši praćenje i analiza kreditne i depozitne aktivnosti banaka po različitim aspektima. Poslovi praćenja i analize prometa i salda na knjigovodstvenim računima, bilansa stanja i uspeha kao i kreditne i depozitne aktivnosti poslovnih banaka predstavljaju samo jedan deo poslova koje sektor obavlja u Narodnoj banci. Činjenica je da su upravo ove poslovne funkcije od posebnog značaja za kontrolu bankarskog sektora u zemlji. Takođe, analize koje se rade na osnovu podataka sa knjigovodstvenih računa spadaju medju najsloženije analize kojima se bavi relativno mali broj analitičara. Zahtevi za ad-hok pregledima su brojni i raznovrsni. Od analitičara se očekuje da reaguju brzo na sve zahteve rukovodstva, posebno poslednjih godina kada je supervizija bankarskog sektora dobila poseban značaj. Zbog svih ovih razloga, a u dogovoru sa korisnicima, 2007. godine je implementirano skladište podataka sa ciljem da pokrije najkritičnije poslovne funkcije analize kako osnovnih knjigovodstvenih podataka tako i svih povezanih kategorija izvedenih iz osnovnih podataka o prometu i stanju na knjigovodstvenim računima poslovnih banaka. U roku od četiri meseca uradjen je model skladišta, inicijalno su napunjeni istorijski podaci od 31.12.2003 godine i počelo se sa redovnim mesečnim osvežavanjem baze skladišta. U julu 2008. godine stupilo je na snagu novo, izmenjeno uputstvo o evidentiranju podataka na knjigovodstvenim računima, koje je značajno izmenilo prethodni koncept praćenja i analize ovih podataka. Uvedene su neke potpuno nove poslovne kategorije (sektorska i granska podela, valutna struktura računa, ročnost itd.). Zato je u periodu od maja do avgusta 2008. godine uradjena revizija, dopuna i proširivanje postojećeg modela prema novim zahtevima. Dogovor sa korisnicima je bio da se na prezentacionom nivou odvoje u posebnu zonu tzv. stari podaci koji važe zaključno sa 30.06.2008. godine od tzv. novih podataka koji se formiraju prema trenutno važećem uputstvu i to od 31.07.2008. godine. Uspostavljene su veze izmedju starih i novih podataka i omogućeno je uporedno praćenje onih poslovnih mera koje su zajedničke za sve kontne okvire. Primer tehničke implementacije ovog skladišta koji je dat u prilogu C zasniva se na modelu koji je postavljen sa poslednjom izmenom kontnog okvira i to je trenutno važeći prošireni model skladišta. Specifično je da se jedan deo podataka iz ovog skladišta koristi i u drugim sektorima Narodne banke, pre svega u Monetarnom sektoru. Projektni tim koji je učestvovao u svim fazama, razvoja, izmene i održavanja ovog skladišta je od 2007. godine do danas više puta menjan. U proseku, na razvoju i održavanju ovog skladišta uvek je bilo dvoje ljudi iz IT sektora i dvoje aktivno uključenih krajnjih korisnika.

3 OKRUŽENJE. ULAZNI PODACI. NEDOSTACI POSTOJEĆIH SISTEMA

3.1 Penzijski fondovi

Podatke koji su izvorni za ovo skladište elektronski dostavljaju kastodi banke⁵, društva za upravljanje dobrovoljnim penzijskim fondovima i Beogradska berza. Podaci se dobijaju u datotekama u xml formatu i to u formi koja je propisana uputstvom o elektronskoj razmeni podataka sa kastodi bankama, odnosno društvima za nadzor dobrovoljnih penzijskih fondova, kao i potpisanim protokolom o razmeni podataka sa Beogradskom berzom. Svi podaci se smeštaju u transakcionu, prihvatnu bazu podataka. Ovaj transakcioni sistem je izvor podataka za skladište. On se dnevno puni novim podacima koji dolaze elektronskom razmenom. Izvorni sistem je specifičan u tom smislu što ne postoji pristup krajnjih korisnika tom sistemu. Sve operacije punjenja, izmena i brisanja transakcionih podataka obavljaju se isključivo automatizovano. Ovako struktuiran transakcioni sistem omogućava brz upis u bazu bez ikakvih transformacija izvornih podataka. U istim tabelama nalaze se inicijalno poslani podaci i sve njihove naknadno dospele korekcije, što posebno opterećuje bazu i otežava njeno pretraživanje. Pre implementacije skladišta podataka nije postojao informacioni sistem posebno projektovan u cilju brzog, efikasnog izveštavanja i analiza podataka za potrebe sektora, već su nad skupom prihvatnih relacionih tabela radjeni predefinisani, statički izveštaji.

Ovako organizovan relacioni sistem nije prilagodjen analitičkom izveštavanju. Glavni nedostaci su:

- Podaci su u izvornom obliku i potrebno ih je agregirati i transformisati da bi predstavljali kvalitetnu poslovnu informaciju.
- Sistem nije subjekt orijentisan, već se podaci o istom poslovnim pojmu nalaze na različitim mestima u sistemu. Potrebno ih je regrupisati i objediniti prema osnovnim poslovnim subjektima.
- Prisutni su konflikti u nazivima, formatima i značenjima poslovnih pojmova u različitim delovima sistema.
- Svaka izmena u opisu xml sadržaja uslovljava izmene postojećih ili pravljenje novih prihvatnih tabela. Rešenje zavisi prvenstveno od toga šta je sa stanovišta brzog prijema podataka jednostavnije i lakše za implementaciju. I jedno i drugo rešenje zahtevaju modifikaciju izveštaja koji rade nad konkretnim tabelama.
- Uglavnom ne postoje uspostavljene veze i kontinuitet podataka u toku mnogobrojnih izmena uputstava o razmeni podataka. Brzina prijema i

⁵ Kastodi banka - banka koja obavlja poslove vođenja računa hartija od vrednosti za račun klijenta i postupa po nalogu klijenta[10]

upisa podataka je prioritet i pri svakoj promeni uputstva o razmeni pribegava se najbržem i najefikasnijem rešenju.

- Izveštavanje koje zahteva kombinovanje svih raspoloživih podataka i praćenje vremenskih serija i istorijata promena pokazalo se teško izvodljivo iz ovakvih struktura.
- Relaciona baza podataka, koja inače nije optimizovana za brza izveštavanja, postajala je sve više opterećena novim podacima i odziv je bivao sve sporiji.
- Izrada statičkih izveštaja zahtevala je ozbiljnu IT podršku, ljudske resurse i vreme.

Osim ovih podataka sektor, po prirodi svog posla, koristi i podatke koji dolaze iz spoljnih izvora. Takvi podaci prikupljaju se neautomatizovano i to putem Interneta, elektronske pošte, sa zvaničnih sajtova ili iz časopisa. Ovi podaci nisu deo skladišta podataka.

3.2 Kreditni odnosi sa inostranstvom

Poslovi praćenja kreditnog zaduženja i kreditnog odobrenja u zemlji koje obavlja Narodna banka obuhvataju aktivnosti:

- Evidentiranja i registrovanja kredita koji se uzimaju od stranih lica, kao i kredita koji se odobravaju stranim licima.
- Evidentiranje i praćenje svih promena koje se odnose na anuitetne planove otplate glavnice i kamate.
- Evidentiranje i praćenje izvršenih plaćanja obaveza po glavnici i kamati, kao i drugih plaćanja po osnovu kreditnog posla.
- Praćenje svih promena koje se odnose na potpisane kreditne ugovore.

Sve navedene aktivnosti evidentiraju se u transakcionoj bazi podataka koja je projektovana pre skladišta i u kojoj se nalaze podaci još iz 70-ih godina prošlog veka. Proces rada je u početku nalagao ručni unos podataka sa propisanih obrazaca u transakcionu bazu. Poslednje dve godine zaživeo je sistem elektronskog dostavljanja podataka Narodnoj banci od strane ovlašćenih banaka kod kojih se registruje kreditni posao. Trenutna situacija je takva da se transakciona baza puni većim delom iz dospelih xml datoteka, manjim delom automatizovano iz drugih aplikacija razvijenih u Narodnoj banci i najmanjim delom ručnim korisničkim unosom. Trenutno se u transakcionoj bazi nalaze podaci o 11000 aktivnih prijavi kreditnog zaduženja u inostranstvu i oko 1000 aktivnih prijavi kreditnog odobrenja, sa planovima otplate i stawkama izvršenih plaćanja na najdetaljnijem dnevnom nivou.

Ova transakciona baza je izvor podataka za skladište.

Činjenica je da je transakciona baza kreditnih poslova sa inostranstvom jedna od najsloženijih, ali i najpažljivije projektovanih i najzaštićenijih baza u Narodnoj banci Srbije. Podaci u njoj su verovatno najstariji od svih podataka koji postoje u bazama Narodne banke Srbije. Skup ugrađenih logičkih aplikativnih i baznih kontrola je medju najsloženijima. Nad ovom bazom podataka postoje razvijene aplikacije u okviru kojih je urađen veliki broj statičkih izveštaja nad dnevnim, ažurnim podacima. Izveštaji se razvijaju u IT sektoru na zahtev korisnika i koriste se u kontrolne svrhe ili za praćenje dnevnih promena, čak i posle uvođenja skladišta podataka.

Razlozi zbog kojih se i pored veoma razvijenog i aplikativno pokrivenog transakcionog sistema počelo sa razvojem skladišta podataka su:

- Sve veća potreba za brzom i efikasnom analizom vremenskih serija podataka o stanju duga, planovima otplate i izvršenim plaćanjima glavnice i kamate, pod različitim uslovima i sa različitim aspektata. Potreba za brzim odgovorom na postavljene zahteve direktno je uslovljena dešavanjima u spoljnom okruženju u kome je aspekt kreditnog zaduživanja dobijao sve više na značaju.
- Brzina formiranja izveštaja i pregleda za potrebe rukovodstva Narodne banke, drugih organizacionih delova Narodne banke, ali i spoljnog okruženja kao što su organi Vlade i Ministarstva zahtevala je sve veće osamostaljivanje poslovnih analitičara u formiranju izveštaja i sve veće izuzimanje svakodnevne IT podrške.
- Organizacija podataka u transakcionoj bazi u kojoj je za svaki kredit sačuvana svaka najsitnija stavka plana i plaćanja pokazala se neadekvatnom za analitičko izveštavanje kakvo je trebalo obezbediti. Analitičko izveštavanje zahtevalo je rad sa podacima na višem nivou detaljnosti. Trebalo je iz transakcionih podataka izvesti čitav niz poslovnih kategorija vezanih za praćenje stanja duga po različitim aspektima. To su kategorije koje se prate uglavnom na mesečnom nivou i isključivo agregirane.
- Analiza podataka i principi izveštavanja u sektoru zahtevali su postojanje tzv. zamrznutih podataka, odnosno presečnih skupova transakcionih podataka koji bi se formirali na tačno odredjeni dan i nadalje ostajali nepromenjeni. Dok su podaci u transakcionoj bazi ažurni u realnom vremenu, podaci u skladištu se formiraju periodičnim presekom transakcionih podataka. Ono što je bio zahtev korisnika to je da uvek imaju tačno, zamrznuto stanje duga i planova otplate na odredjeni datum. Priroda posla nalaže da se često analize rade unazad, za prošle datume ili za vremenski period, pa je istorijat stanja duga i planova otplate ovde od posebnog značaja. Postojeća transakciona baza i sam sistem evidentiranja kreditnih poslova ne omogućavaju „vraćanje“ stanja unazad. Čak i da se uspe u tome, složenost takvog zahvata i mogućnost greške su veoma veliki.

Iz navedenih razloga, u trenutku kada je transakcioni sistem već bio stabilan i sveobuhvatan, počelo se sa projektovanjem skladišta podataka.

3.3 Bilansni podaci poslovnih banaka

Okruženje u Sektoru za kontrolu poslovanja banaka koje je zatečeno pre uvođenja skladišta podataka je skoro identično okruženju koje je prethodilo uvođenju skladišta podataka koje je opisano u poglavlju 3.1. U ovom sektoru i tada je bio, a i danas je aktuelan sistem razmene podataka sa spoljnim okruženjem u vidu xml datoteka. Sve aktivne poslovne banke u Srbiji su u obavezi da najkasnije do 15. dana u mesecu dostave podatke o prometu i stanju na svojim knjigovodstvenim računima i to na poslednji dan prethodnog meseca. Podaci se dostavljaju ovlašćenoj službi Narodne banke u skladu sa uputstvom Narodne banke o obavezi i načinu prikupljanja, obrade i dostavljanja podataka o stanju i strukturi plasmana, potraživanja i obaveza banaka. Tako dobijeni mesečni podaci se, kao što je napred pomenuto, smeštaju u transakcione tabele koje po strukturi odgovaraju propisanim xml datotekama i struktuirane su prvenstveno da obezbede brz upis primljenih podataka.

Podaci poslovnih banaka dobijeni redovnom mesečnom razmenom predstavljaju izvorne podatke za skladište.

Pre uvođenja skladišta u sektoru nije postojao izveštajni sistem koji obuhvata praćenje i analize podataka o bilansu stanja i bilansu uspeha banaka po svim aspektima, kao i analize kreditne i depozitne aktivnosti banaka. Bilans stanja i bilans uspeha su kategorije koje se izvode iz osnovnih knjigovodstvenih podataka po propisanoj metodologiji Narodne banke. Krediti i depoziti su kategorije koje se takodje izvode iz osnovnih knjigovodstvenih podataka i predstavljaju deo redovnog izveštavanja sektora. Pre uvođenja skladišta podataka pravljani su predefinisani, statički izveštaji prema zahtevima koji su dolazili iz sektora. Ovi izveštaji izvršavali su se direktno nad prihvatnim, transakcionim tabelama koje odlikuje velika količina podatka i neoptimizovanost za brzo, analitičko izveštavanje. Zahtevi za ad-hok izveštavanjem su poslednjih godina postali učestali i tražio se sve brži odgovor zaposlenih u IT sektoru.

Poslovne potrebe i okruženje u zemlji koje je fokusirano na nadzor bankarkog sektora i kontrolu njihovog poslovanja inicirali su da se počne sa uvođenjem novog izveštajnog sistema u sektoru.

Odluka da se taj izveštajni sistem implementira kroz skladište podataka bila je logična, jer:

- Već su postojala značajna iskustva sa razvojem skladišta podataka u banci, na primerima Sektora za međunarodnu saradnju i Monetarnog sektora.
- Poslovne funkcije o kojima je reč su pogodne za skladište podataka, jer se radi o isključivo analitičkim funkcijama. Činjenica je da je skoro

100-procentna aktivnost ovog sektora zasnovana na izveštavanju iz dobijenih podataka, a u cilju supervizije banaka. Dakle, nema unosa, izmene i brisanja podataka. Svi ulazni podaci se dostavljaju elektronski, ali i sve njihove korekcije stižu elektronski. Trebalo je da se sagledaju svi aspekti u kontekstu kojih se analiziraju knjigovodstveni podaci, da se identifikuju poslovni subjekti i poslovne mere i da se na osnovu toga formira model skladišta podataka.

- Posebni zahtevi iz sektora su bili da se postepeno celokupne izveštajne aktivnosti prebace sa zaposlenih u IT sektoru na poslovne analitičare. Osamostaljivanje je bilo važno da bi se dobilo na efikasnosti, brzini i kvalitetu analiza, zbog čega se i krenulo u projekat implementacije skladišta.
- Trebalo je obezbediti istorijski aspekt podataka. Uputstvo o praćenju i evidentiranju plasmana, obaveza i potraživanja poslovnih banaka (tj. kontni okvir) je od 2003. godine promenjeno 4 puta. Sa promenom kontnog okvira uglavnom se menja, formalno ili suštinski, metodologija izrade bilansa stanja i bilansa uspeha, pa i metodologija izvodjenja kreditnih i depozitnih kategorija. Promene kontnog okvira su u nekoliko navrata bile takve da je najteži zadatak u implementaciji skladišta bio uspostavljanje kontinuiteta podataka i obezbedjivanje vremenskih serija za praćenje odredjenih poslovnih mera. Dogovor sa korisnicima je bio da se u skladište podataka uključe podaci od 31.12.2003 godine i da se kontinuitet poslovnih mera ispoštuje koliko god je to moguće.

4 ZAJEDNIČKI CILJEVI I MOTIVI ZA UVODJENJE SKLADIŠTA PODATAKA

Skladišta podataka koja se razvijaju u Narodnoj banci predstavljaju sisteme koji su odvojeni od transakcionih sistema, strukturirani, i logički i fizički optimizovani za potrebe analitičkog izveštavanja. Ono što je krajnjim korisnicima uvek bio imperativ je samostalnost u izradi izveštaja, brzina izrade izveštaja, kombinovanje podataka po potrebi, jednostavno uključivanje u sistem novih grupa podataka.

Osnovna dva razloga zbog kojih se krenulo u projekte razvoja skladišta podataka bila su:

- Olakšavanje i unapredjenje izveštajnih, analitičkih aktivnosti na kojima se posebno insistira u Centralnoj banci.
- Rasterećenje i poboljšanje performansi transakcionih sistema prebacivanjem glomaznog izveštajnog aparata na skladišta podataka, razdvajanjem transakcionog i novog izveštajnog sistema. Transakcioni podaci sada mogu da se periodično arhiviraju (što do sada nije praktikovano zbog istorijskih podataka) u cilju bržeg rada transakcione baze, a da se pri tom ne naruši izveštajni sistem u kome se nalazi sva istorija promena podataka.

Skladišta podataka ni u jednom slučaju ne predstavljaju prostu kopiju operativnih podataka već su rezultat reorganizacije i restrukturiranja podataka. Cilj je uvek da skladište podataka ne bude narušeno ni u slučaju da postojeći transakcioni sistemi bude zamenjeni nekim boljim, naprednijim rešenjem.

Kada se govori o ciljevima i motivima sa kojima se ulazi u svaki od projekata razvoja skladišta podataka, to su:

- Brži izveštaji. Svako skladište je projektovano u svrhu brzog i efikasnog izveštavanja. Performanse baze podataka na kojoj su sistemi fizički implementirani omogućavaju brze upite, jer je to baza podataka parametrizovana isključivo za skladištenje podataka.
- Lakši razvoj izveštaja. Podaci su organizovani i strukturirani tako da omoguće lako pravljenje izveštaja. Alati koji se koriste u ovu svrhu su napredni i prilagodjeni upravo brzom analitičkom izveštavanju.
- Ad-hok upiti. Korisnici imaju mogućnost direktnog pristupa podacima u realnom vremenu, što im omogućava da, ukoliko dobro poznaju strukturu baze, sami prave upite i za kratko vreme dobiju podatke kao rezultate tih upita. Stari sistemi standardnih izveštaja to nisu omogućavali, što je bila jedna od najvećih zamerki krajnjih korisnika.

- Objedinjenost podataka. Iz različitih izvora podaci su objedinjeni i grupisani tako da mogu da se uporedo analiziraju, onako kako odgovara poslovnim procesima u ciljnom sektoru. Cilj je dobijanje što sveobuhvatnije slike o poslovnom domenu koji sektor prati i analizira.
- Intuitivno radno okruženje i tehnološki napredni korisnički alati. Sa aspekta krajnjih korisnika prezentacioni sloj je uvek intuitivan, terminološki i poslovno prilagodjen samostalnom formiranju izveštaja. Nazivi i značenja poslovnih kategorija potpuno odgovaraju poslovnom rečniku krajnjih korisnika. Korisnicima se nude odgovarajući korisnički alati za rad sa ovim tipom baza podataka. To su alati namenjeni pretraživanju skladišta podataka, sa optimizacijom izvršavanja upita. Alati omogućavaju „bušenja“ i „sečenja“ podataka duž hijerarhija dimenzija, automatsko agregiranje činjenica, kao i napredne matematičke funkcije. Zato svaki projekat razvoja skladišta podataka predstavlja za nas tehnološki napredak.
- Sistem prilagodjen operativcima, poslovnim analitičarima i upravljačkoj strukturi. U razgovorima i intervjuima sa zaposlenima koji će koristiti sisteme za skladištenje profilisale su se ove tri grupe krajnjih korisnika. Osnovni zadatak uvek je da se isprojektuje i implementira sistem koji može da zadovolji potrebe svih grupa korisnika. O korisnicima skladišta podataka biće više reči u poglavlju 6.

5 PROJEKTOVANJE SKLADIŠTA PODATAKA

5.1 Fazni razvoj sistema

Sistemi za skladištenje podataka prikazani u ovom radu razvijani su iterativno, u više faza. Cilj ovog pristupa je da svaka faza uključi u sistem nove grupe podataka da bi se na kraju dobio sveobuhvatni izveštajni sistem jednog sektora. Iterativni razvoj se u praksi pokazao kao najbolji izbor, jer se na taj način podaci uvode grupisani u logičke celine. Time se vreme potrebno za implementaciju svake faze skraćuje tako da se novi sistemi brže uvode u produkciju. Kako su, uglavnom, u pitanju krajnji korisnici koji se do sada nisu susretali sa sličnim sistemima i nisu imali iskustva u pretraživanju ovakvih podataka vremenom se kao najbolje rešenje pokazalo postepeno upoznavanje sa mogućnostima sistema. Razvoj sistema po fazama omogućava projektnom timu praćenje ponašanja korisnika i evidentiranje njihovih potreba. Iskustvo je pokazalo da, kako razvoj sistema napreduje, pojavljuju se (i najčešće budu prihvaćene) primedbe korisnika u odnosu na predloženi model. Fazni razvoj omogućava projektnom timu da ispravi greške nad manjim skupom podataka. Edukacija korisnika za korišćenje alata namenjenih analitičkom izveštavanju takodje se, po pravilu, lakše odvijala onda kada se sistem postepeno proširivao.

Definisanje faza u razvoju i implementaciji skladišta uvek je bilo rezultat razgovora sa krajnjim korisnicima o prioritetima u implementaciji, ali i rezultat iskustva projekatana koji donose konačne odluke kada je tehnička implementacija u pitanju.

U razvoju skladišta podataka koje obuhvata poslove iz domena analize i praćenja poslovanja penzijskih fondova posebno je došao do izražaja koncept faznog iterativnog razvoja. Ovaj sistem je razvijan u vreme kada je, na osnovu iskustava od ranije, definitivno usvojen iterativni pristup kao standard u projektovanju i razvoju skladišta podataka. Detaljniji opis svake od pet identifikovanih faza razvoja dat je u prilogu A ovog rada, dok su ovde ukratko navedene:

- Uvodjenje grupa podataka koje učestvuju u formiranju ukupne i neto imovine fonda: hartija od vrednosti, nepokretnosti, depozita, gotovine, potraživanja i obaveza. Na ovaj način, već nakon prve faze razvoja sistema, korisniku je data mogućnost da detaljno pretražuje svaku od navedenih kategorija podataka po relevantnim dimenzijama, ali i da dobije uvid u ukupnu i neto imovinu fonda koja se izvodi iz ovih kategorija. Takodje, korisnik je već sada mogao da raščlani imovinu na svoje najdetaljnije komponente i da izračuna procenat učešća svake komponente u imovini fonda. Ova faza je, po broju poslovnih kategorija koje se prate, najobuhvatnija i zato je projektantski izbor bio da to bude prva faza u razvoju sistema.

- Uvodjenje podataka o uplatama, isplatama, naknadama, transakcionim troškovima i članovima fondova.
- Uvodjenje podataka o penzijskim planovima, obveznicima uplata, pojedinačnim ugovorima i transferima.
- Uvodjenje podataka o trgovanju na Beogradskoj berzi. U ovoj fazi u sistem su uključeni i pokazatelji trgovanja koji se čuvaju na dnevnom nivou.
- Uvodjenje visoko-agregiranih podataka o individualnim računima članova fondova. Ovo je, ujedno, i najstroženiji podmodel celog modela, pre svega zbog složene metodologije izvodjenja svake poslovne mere. Poslovne potrebe kao i postojanje poverljivih podataka u izvornim podacima o individualnim računima su razlozi zbog kojih oni nisu uključeni u skladište na detaljnijem nivou, već preračunati u poslovne mere koje mogu biti publikovane.

Skladište podataka koje obuhvata kreditne odnose sa inostranstvom implementirano je kroz dva odvojena Data Mart podsistema: „Kreditna zaduženja“ i „Kreditna odobrenja“. Ova dva podsistema su razvijana nezavisno jedan od drugog. Najpre je u produkciju uveden podsistem „Kreditna zaduženja“, a nekoliko meseci kasnije i podsistem „Kreditna odobrenja“. Pri projektovanju skladišta pošlo se od podataka koji su najobimniji i od analiza koje su korisnicima najznačajnije, pa su zahtevale i najveće angažovanje IT zaposlenih.

Redosled uvodjenja podsistema bio je uslovljen sledećim razlozima:

- Praćenje i analiza podataka vezanih za zaduženost domaćih firmi, banaka, državnih organa i institucija zemlje bili su uvek od najveće važnosti. Praćenje podataka o kreditima koje domaća lica odobravaju stranim firmama dobija više na značaju u poslednje vreme kada se taj aspekt kreditnih odnosa sve više aktivira. Ranije su kreditna odobrenja bila mnogo manje interesantna analitičarima kreditnih odnosa sa inostranstvom.
- Obim podataka je u domenu kreditnih zaduženja mnogo veći. Takva je situacija i danas, a u vreme projektovanja skladišta podataka to je bilo posebno izraženo.

Kod svakog od ova dva Data Mart podsistema, uočljive su dve faze kroz koje su podsistemi prošli u svom razvoju, a to su: uvodjenje visoko-agregiranih podataka o stanju duga na dan i uvodjenje detaljnih podataka o planovima otplata obaveza, odnosno planovima naplata potraživanja. Grupe podataka o planovima otplata i naplata su u podsisteme uključene čak nekoliko godina nakon početka korišćenja skladišta, onda kada su se uvidele sve prednosti ovog sistema i kada se posebno aktiviralo praćenje planova otplata.

Iako su sa poslovne tačke gledišta kreditna zaduženja i kreditna odobrenja dva potpuno odvojena sistema, tehnička implementacija ova dva Data Mart-a

ima više sličnosti nego razlika. Analogija postoji u opisu i strukturi dimenzija, u opisu tabela činjenica, dinamici i principima osvežavanja tabela dimenzija i tabela činjenica. U prilogu B ovog rada su detaljnije opisane dimenzije i činjenice koje su uključene u Data Mart kreditnih zaduženja. Pomenuta je analogija sa kreditnim odobrenjima, ali i neke razlike koje postoje između ovih podсистema.

Skladište podataka koje obuhvata knjigovodstvene i bilansne podatke poslovnih banaka nema tako izražene faze u razvoju kao što je to slučaj u prethodna dva primera. Redosled uvođenja poslovnih kategorija je bio utvrđen, ali je cela implementacija trajala relativno kratko tako da se u ovom slučaju ne može govoriti o striktno odvojenim fazama razvoja koje su sukcesivno testirane i puštane u produkciju. Rokovi su bili kratki, a podaci koje skladište obuhvata su takvi da je sistem pušten u produkciju onda kada su sve predviđene poslovne kategorije bile uključene. Ipak, projektni tim je imao uvid u neke od standardnih pregleda koji se u ovom sektoru rade prema zahtevu najvišeg rukovodstva. Obavljeni su razgovori sa analitičarima o tome koje se vrste ad-hok pregleda najčešće prave, kakvi se zahtevi uglavnom postavljaju i koje su analize najznačajnije i najsloženije. Kako nije postojalo celovito transakciono aplikativno rešenje pre uvođenja skladišta, intervjui sa korisnicima su najviše pomogli da se utvrdi šta se to od zaposlenih u IT sektoru do sada tražilo i šta je bilo najteže dobiti. Na osnovu ovoga definisan je (više u okviru samog projektnog tima) redosled uvođenja poslovnih kategorija u sistem, i to: najpre podaci o prometu i stanju na knjigovodstvenim računima, zatim bilans stanja i bilans uspeha i na kraju podaci o kreditnoj i depozitnoj aktivnosti banaka.

Ono što je zajedničko za sva tri skladišta podataka je to da uključuju jedan deo najdetaljnijih podataka, uz mogućnost njihovog agregiranja u samim upitima, ali i drugi deo podataka koji je unapred preračunat i agregiran do nivoa granularnosti potrebne za spovodjenje analiza. Zato su u sisteme najpre uvođeni najdetaljniji podaci, pa tek u sledećoj fazi agregati izvedeni iz osnovnih podataka.

5.2 Pripremna zona

Pripremna zona kao deo arhitekture iskorišćena je u primerima penzijskih fondova i kontrole poslovanja banaka. U implementaciji ovih sistema za skladištenje podataka, pripremna zona postoji u prostoru između izvorne transakcione baze i samog skladišta. Sve karakteristike i sama namena pripreme zone zajednički su za oba primera.

Pripremna zona je implementirana kao skup normalizovanih relacionih tabela i fizički je razdvojena od baze skladišta. Tabele iz pripreme zone su nedostupne krajnjim korisnicima, tj. nad njima se ne radi izveštavanje. One služe za prihvatanje izvornih podataka i njihovu pripremu za dalju obradu.

Pripremna zona je uvedena iz dva razloga:

- 1 Podaci su u transakcionim šemama razdvojeni u puno tabela. Svaka tabela je kopija strukture datoteke sa xml sadržajem preko koje se ti podaci razmenjuju. Kako se menjao opis datoteka, shodno promenama uputstva o razmeni podataka, tako su i transakcione tabele doradjivane. Često je sa promenom opisa xml sadržaja u datotekama pravljena nova transakciona tabela, tako da su podaci o istom subjektu fizički razdvojeni. Za dalju obradu, transformaciju i usaglašavanje podataka bilo je najbolje rešenje da se najpre svi podaci objedine u pripremnoj zoni skladišta.
- 2 Kako je baza skladišta sasvim razdvojena od transakcione baze, razmena podataka izmedju izvornih sistema i baze skladišta se odvija komunikacijom izmedju dva servera što značajno usporava prenos i obradu podataka. Pošto su se nad izvornim podacima očekivale velike transformacije (primenom prostih matematičkih operacija, ali i složenijih ekonomskih formula) uvedena je priprema zona da bi se ubrzao ETL proces. Podaci se u pripremu zonu upisuju iz izvornih sistema, uz minimalna usaglašavanja podataka po značenju i imenovanju. To je proces koji se relativno brzo završava čime se transakciona baza oslobadja i omogućava se da ona nesmetano radi. Sve dalje složene ETL operacije odvijaju se u procesima izmedju pripreme zone i baze skladišta. Kako su u oba slučaja poslovne potrebe uslovile da se osvežavanje baze skladišta obavi bar jednom u radno vreme, osnovni cilj je bio da se rad transakcione baze ne uspori. Pri tome se imao u vidu veliki broj aplikacija koje neprekidno rade nad transakcionom bazom, kao i veliki broj aktivnih korisnika tih aplikacija.

Prilikom osvežavanja baze i jednog i drugog skladišta najpre se puni priprema zona, potom se osvežavaju dimenzije pa tabele činjenica u skladištu.

5.3 Ažuriranje skladišta podataka. ETL strategije

Zajednička karakteristika sva tri skladišta je da su ETL procedure pisane tako da se svaka tabela činjenica puni posebnom procedurom u zasebnom procesu. Kod nekih tabela činjenica redosled pozivanja procedura je bitan i unapred definisan. Takve su, na primer, tabele izračunatih indeksa kao i tabela sa podacima o individualnim računima članova fondova u primeru penzijskih fondova. Punjenje ovih tabela činjenica uvek se obavlja nakon punjenja tabela činjenica koje se konsultuju pri ovim obračunima. Takodje, obračun ukupne vrednosti imovine i procenta učešća svake stavke u imovini fonda uvek se obavlja posle svih ETL procedura koje pune podatke koji ulaze u sastav imovine. U sistemima koji imaju pripremu zonu ETL procedure koje pune tabele činjenica u skladištu kao izvorne podatke imaju isključivo podatke u pripremu zoni i nikada ne konsultuju transakcionu bazu. Zato se, kao što je rečeno, procedure za preuzimanje izvornih podataka i njihovo prebacivanje u pripremu zonu pozivaju pre bilo koje ETL procedure koja puni skladište. Transakciona baza se konsultuje pri punjenju pripreme zone, ali i pri osvežavanju dimenzionih tabela. Dimenzione tabele se

osvežavaju nakon punjenja pripremljene zone, a pre osvežavanja tabela činjenica.

Kada je reč o osvežavanju podataka u svakom od skladišta, dva osnovna zahteva su bila:

- Identifikacija novih slogova u izvorima podataka zbog osvežavanja tabela činjenica.
- Izbor strategije osvežavanja dimenzionih tabela.

U svakom od primera primenjena su različita rešenja u osvežavanju skladišta, zavisno od dogovora, mogućnosti i potreba posla.

5.3.1 Penzijski fondovi

Punjenje skladišta podataka obavlja se kroz inicijalno punjenje baze istorijskim podacima i kroz periodična osvežavanja. Inicijalnim punjenjem u bazu su ubačeni istorijski podaci kojima sektor raspolaže od početka automatizovane razmene podataka sa spoljnim okruženjem. Najstariji podaci u skladištu su dnevni kumulativni podaci kastodi banaka i oni postoje od 15.11.2006. godine. Ostale grupe podataka su mlađe jer je njihovo dostavljanje naknadno definisano. Najmladji su podaci o individualnim računima članova fondova koji se elektronski razmenjuju od 01.01.2009. godine. Pre toga, podaci su stizali putem elektronske pošte, najčešće u vidu Excel tabela. Svi ti podaci preuzeti su iz internih arhiva sektora, prečišćeni, preformatirani, usaglašeni i inicijalno ubačeni u bazu skladišta.

Periodično osvežavanje baze odvija se u dogovorenim terminima i to automatizovanim procedurama. Glavno osvežavanje baze obavlja se u noćnom režimu kada je rasterećeno skladište, a i transakcioni sistemi koji služe kao izvori podataka. Zahtev korisnika i potrebe posla uslovile su da se još jedno osvežavanje baze definiše i u toku dana posle 13 časova, tj. nakon prijema svih podataka putem elektronske razmene. ETL procedure koje pune bazu podacima o računima članova fondova najduže traju i zauzimaju najviše resursa. Zato se ove procedure aktiviraju isključivo u noćnom režimu.

Identifikacija novih slogova za punjenje tabela činjenica obavlja se na osnovu datuma formiranja sloga. Datum formiranja sloga je uključen u svaku tabelu činjenica i uvek se puni sistemskim datumom. Iz izvornih tabela izdvajaju se slogovi koji su upisani nakon poslednjeg datuma formiranja sloga u tabeli činjenica. Taj skup slogova biće upisan u skladište pri sledećem osvežavanju. Polje „Datum formiranja sloga“ važno je i kod osvežavanja dimenzionih tabela. Kod dimenzionih tabela se, pored datuma upisa sloga, upoređivanje sa izvornim podacima vrši i po prirodnom ključu. Dakle, iz transakcione tabele se najpre izdvajaju slogovi koji su upisani nakon poslednjeg osvežavanja odgovarajuće dimenzione tabele. U svakom slogu tako izdvojenog skupa ispituje se vrednost prirodnog ključa (matičnog broja banke, broja penzijskog fonda, datuma, ISIN broja hartije od vrednosti itd.).

Ukoliko izvorni slog sadrži novu vrednost prirodnog ključa on će biti upisan u dimenzionu tabelu pri sledećem osvežavanju. Svaki slog u izvornim podacima koji sadrži vrednost prirodnog ključa koja postoji u dimenzionoj tabeli upoređuje se atribut po atribut sa odgovarajućim atributima u dimenzionoj tabeli.

Ukoliko se pronadje razlika u vrednosti nekog od dimenzionih atributa sprovodi se jedna od sledeće dve akcije:

- 1 Zatvara se tekući slog u dimenzionoj tabeli i otvara novi slog sa istom vrednošću prirodnog ključa a novom vrednošću surogat primarnog ključa. Na ovaj način čuva se istorijat promena u dimenziji poslovne banke i dimenziji fondova. Svaki slog u ovim dimenzijama ima dva datumska polja koja označavaju period važenja podataka u slogu. Kod zatvaranja sloga popunjava se „datum do“ i otvara novi slog sa istim prirodnim ključem i novim surogat ključem. „Datum od“ novog sloga postavlja se na sledeći datum u odnosu na datum zatvaranja prethodnog sloga. U slučaju prestanka rada banke ili fonda slog se zatvara i ne otvaraju se novi slogovi.
- 2 Vršiti se zamena vrednosti atributa u dimenzionim tabelama novim vrednostima. Na ovaj način osvežavaju se dimenzije hartija od vrednosti, valuta i okruga. Ovo su dimenzije kod kojih se istorijat promena ne čuva i analize nad podacima skladišta se uvek sprovode u odnosu na tekuću vrednost dimenzionih atributa.

Da bi datum formiranja sloga mogao da se iskoristi pri osvežavanju skladišta, neophodno je da i izvorne transakcione tabele imaju predviđen datum upisa podataka. U slučaju ovog skladišta identifikacija novih slogova za upis ne može da se vrši upoređivanjem datuma podataka, već isključivo upoređivanjem datuma upisa podataka u bazu. Poslovni sistem je koncipiran tako da se i za „stare“ datume podataka dostavljaju korekcije koje ETL procedure treba da „prepoznaju“ i obrade. Izuzetak je skup podataka sa berze koji se uvek odnosi na tekući dan i nema korekcija prethodnih stanja. U ovom slučaju dovoljno je izdvojiti izvorne podatke koji imaju datum stanja veći od najvećeg datuma stanja u tabeli činjenica. Obrada podataka kastodi banaka i društava za nadzor morala je da uzme u obzir i eventualne korekcije. Iz tog razloga, najpre je uradjeno inicijalno punjenje skladišta postojećim podacima iz izvornih tabela, a zatim je u svakoj izvornoj tabeli formirano sistemsko datumsko polje čime su omogućena dalja osvežavanja skladišta.

5.3.2 Kreditni odnosi sa inostranstvom

Punjenje ovog skladišta podacima počelo je 2000. godine. Sve do januara 2001. godine presek podataka iz transakcione baze nije se obavljao redovno, već po potrebi, odnosno na zahtev korisnika. Dakle, u periodu od januara do decembra 2000. godine u skladištu postoje podaci tek za poneki mesec. Od januara 2001. godine obavlja se redovan mesečni presek. Podaci koji se odnose na planove otplate uključeni su u izveštajni sistem 2003. godine i od

tada se redovno skladište. Poslovne potrebe kao i priroda podataka i tipovi analiza koje se nad njima rade odredile su dinamiku ažuriranja skladišta podataka. Osvežavanje skladišta podataka obavlja se jednom mesečno. Sve do 2008. godine procedura za osvežavanje skladišta pokretala se na zahtev korisnika i to je imalo svoje opravdanje. Naime, da bi se dobili kvalitetni i potpuni mesečni podaci o kreditnoj aktivnosti neophodno je da se u transakcionoj bazi nadju sve mesečne promene po kreditima koji se prate. Kako datum do koga ovlašćena banka dostavlja izveštaje Narodnoj banci nije preciziran to je stav krajnjih korisnika da treba sačekati sve izveštaje koji se očekuju i tek onda raditi mesečni presek. Činjenica je da su korisnici jedini koji znaju kada se „zatvara“ mesec za izveštavanje. Problem je bio što se nikada nije znalo tačno vreme kada će se pokretati ETL procedure za presek podataka i što je zahtev korisnika za presekom podataka uvek dolazio u radno vreme kada je ceo sistem najopterećeniji. 2008. godine počelo se sa osvežavanjem skladišta podataka 20. dana u mesecu za prethodni mesec i ETL procedure se automatizovano pokreću u 17 časova, dakle nakon radnog vremena. Dok se podaci o stanju duga i prometu osvežavaju svakog meseca podaci o planovima otplata se osvežavaju kvartalno-četiri puta godišnje. Jedan od razloga leži u velikoj količini podataka o planovima otplata koja se sa svakim osvežavanjem upisuje u bazu skladišta. Taj broj podataka se trenutno kreće oko 600000 slogova za jednu tabelu činjenica. Drugi razlog je poslovne prirode. Naime, planovi otplata su relativno sporomenjajuća kategorija, tako da je po mišljenju korisnika kvartalno praćenje promena optimalno i zadovoljavajuće. Osvežavanje skladišta se odvija u tačno određenom redosledu i to najpre osvežavanje svih dimenzionih tabela, zatim osvežavanje podataka o stanju duga i prometu za kreditna zaduženja, osvežavanje planova otplata po kreditnim zaduženjima i na kraju osvežavanje podataka o kreditnim odobrenjima. Nakon završenog mesečnog preseka podataka izveštajni mesec se zatvara. Svi podaci koji ulaze u transakcioni sistem nakon toga deo su narednog izveštajnog meseca i upisuju se u skladište u sledećem ciklusu osvežavanja.

Identifikacija novih slogova za punjenje tabela činjenica obavlja se na osnovu datuma stanja. Tabele činjenica ne sadrže datum upis sloga. U njima se, isključivo, radi upis novih slogova dok se brisanja i izmene ne sprovede. Osvežavanje svih dimenzija vrši se analogno dimenziji hartija od vrednosti u primeru penzijskih fondova. Dakle, za novu vrednost prirodnog ključa u izvornoj tabeli, upisuje se slog u odgovarajuću dimenziju. Za postojeće vrednosti prirodnog ključa radi se izmena ostalih atributa u dimenziji, bez ispitivanja razlika.

5.3.3 Bilansni podaci poslovnih banaka

Inicijalno punjenje skladišta podacima je uradjeno 2007. godine. Inicijalnim punjenjem obuhvaćeni su istorijski podaci od 31.12.2003. Od avgusta 2007. godine obavlja se redovno osvežavanje baze. Procedure za osvežavanje se pokreću u noćnom režimu i to svaki dan. Predlozi da se osvežavanje baze radi u toku radnog vremena kako bi se dobilo na ažurnosti i blagovremenosti su vrlo brzo odbačeni. Razlog je bio taj što su ove procedure zahtevne i značajno su ometale rad ostalih aplikacija. Osim toga, vreme trajanja

preseka podataka nije zanemarljivo i danas se kreće od 45 do 60 minuta. Iako se radi o mesečnim podacima, ETL procedure se automatizovano pokreću svake noći i ispituju da li u izvornim transakcionim tabelama ima novih podataka. Novi podaci se identifikuju na osnovu datuma upisa podataka, a ne na osnovu datuma stanja. Datum upisa je jedini relevantan pokazatelj zato što je uputstvom o evidentiranju i dostavljanju podataka predviđena mogućnost da se, u slučaju greške ili ispravke, naknadno dostave ispravljeni podaci za neki od prethodnih meseci. Dakle, osvežavanje baze se ne radi svaku noć već kad ima novih podataka u izvoru, ali se svaku noć radi provera izvornih tabela. ETL procedurama je predviđen upis novih slogova u bazu skladišta i brisanje i ponovni upis ispravljenih podataka. Izmene slogova se u slučaju ovog skladišta nikada ne rade. Kako je koncept tabele činjenica namenjene za uporedne analize i tabele činjenica koja sadrži podatke neophodne za formiranje pokazatelja profitabilnosti takav da se pored podataka koji su važeći na datum stanja prate i podaci na neke prošle vremenski određene datume, to svako brisanje podataka pokreće dalje brisanje slogova u kojima se izmenjeni podaci referišu. Što se obima podataka tiče, najviše novih slogova (oko 175000) sa presekom se upiše u tabelu činjenica namenjenu za uporedne analize. Osnovna tabela, u kojoj su podaci o prometu na računima, se sa svakim presekom dopuni za oko 75000 novih slogova. U agregirane tabele u kojima su podaci sačuvani na nivou datuma i banke sa presekom se upiše onoliko slogova koliko ima aktivnih banaka u Srbiji na datum stanja (trenutno aktivnih banaka ima 34). Izuzetak je tabela u kojoj su podaci o kreditima i depozitima u koju se sa presekom upiše dva puta više slogova jer su podaci prikazani u hiljadama i milionima. Sve ovo važi za trenutna osvežavanja baze i ta statistika se menjala sa promenama kontnog okvira, uputstva i metodologije formiranja bilansa. U ovom primeru postoji definisan redosled osvežavanja tabela činjenica tako što se prvo osvežava osnovna tabela prometa i stanja na knjigovodstvenim računima, zatim izvedena tabela za uporedne analize i na kraju agregati. Opis tabela činjenica, veze medju njima i redosled osvežavanja detaljnije su opisani u prilogu C.

Od dimenzija u ovom sistemu osvežavaju se samo vremenska dimenzija i dimenzija poslovnih banaka. Ono što je zajedničko u ovom primeru i primeru penzijskih fondova to je da ovi sistemi „dele“ datumsku dimenziju i dimenziju poslovnih banaka. Ove dve dimenzije su projektovane 2006. godine sa ciljem da budu centralne dimenzije za sva skladišta podataka koja se nadalje budu razvijala u Narodnoj banci. Zato se ove dve dimenzije ažuriraju centralizovano, pre pokretanja bilo koje ETL procedure u skladištima koja ove dimenzije uključuju. Sva skladišta podataka projektovana od 2006. godine sadrže zajedničku datumsku dimenziju kao osnovnu dimenziju. Zajednička dimenzija poslovnih banaka postoji u ne svim, ali u većini tih skladišta, obzirom da je u Narodnoj banci najveći broj analiza zasnovan upravo na podacima o bankarskom sektoru u zemlji. Poslovna banka je jedan od osnovnih subjekata u svim poslovnim procesima u Narodnoj banci Srbije i element većine analiza koje se rade.

5.4 Fizička implementacija. Statistika rasta baza

Za fizičku implementaciju skladišta podataka opisanih u primerima penzijskih fondova i kontrole poslovanja banaka odabrana je posebna baza podataka fizički razdvojena od transakcionih operativnih baza. Ova baza podataka je parametrizovana i optimizovana za skladištenje podataka, sa podignutim OLAP opcijama. U vreme kada su ova skladišta puštena u produkcionu rad ona su fizički bila implementirana na istoj instanci baze podataka na kojoj se nalaze i transakcioni sistemi. Rezultat je bio sporije izvršenje upita nad podacima skladišta i neiskorišćenost svih prednosti koje imaju baze podataka optimizovane za rad sa skladištima podataka. Osim toga, bio je evidentan otežan rad i narušavanje performansi transakcionih sistema. Svaka složenija ETL aktivnost, kao i samo pretraživanje velikih baza skladišta drastično je narušavalo rad transakcionih sistema. Fizičko razdvajanje transakcionih baza i baze skladišta bilo je uslovljeno i samom veličinom skladišta i trendom njegovog rasta. Zato je u periodu do kraja 2008. godine postavljena i parametrizovana posebna instanca baze podataka na koju su prebačene baze podataka za većinu skladišta koja su u upotrebi u Narodnoj banci. Činjenica je da još uvek postoje skladišta podataka, od kojih je najobimnije skladište kreditnih odnosa, koja su fizički implementirana na istoj instanci kao transakcione baze. Ovo su veoma stara skladišta, radjena u starim tehnologijama i njihovo prebacivanje na novu bazu podataka zahteva i tehnološko unapređenje. Ovo je ozbiljniji zahvat za koji treba više vremena i razradjen plan migracije i testiranja. Ipak, planirano je da se u najskorije vreme i skladište podataka za kreditne odnose tehnološki unapredi i prebaci na zajedničku bazu namenjenu i optimizovanu za skladištenje podataka.

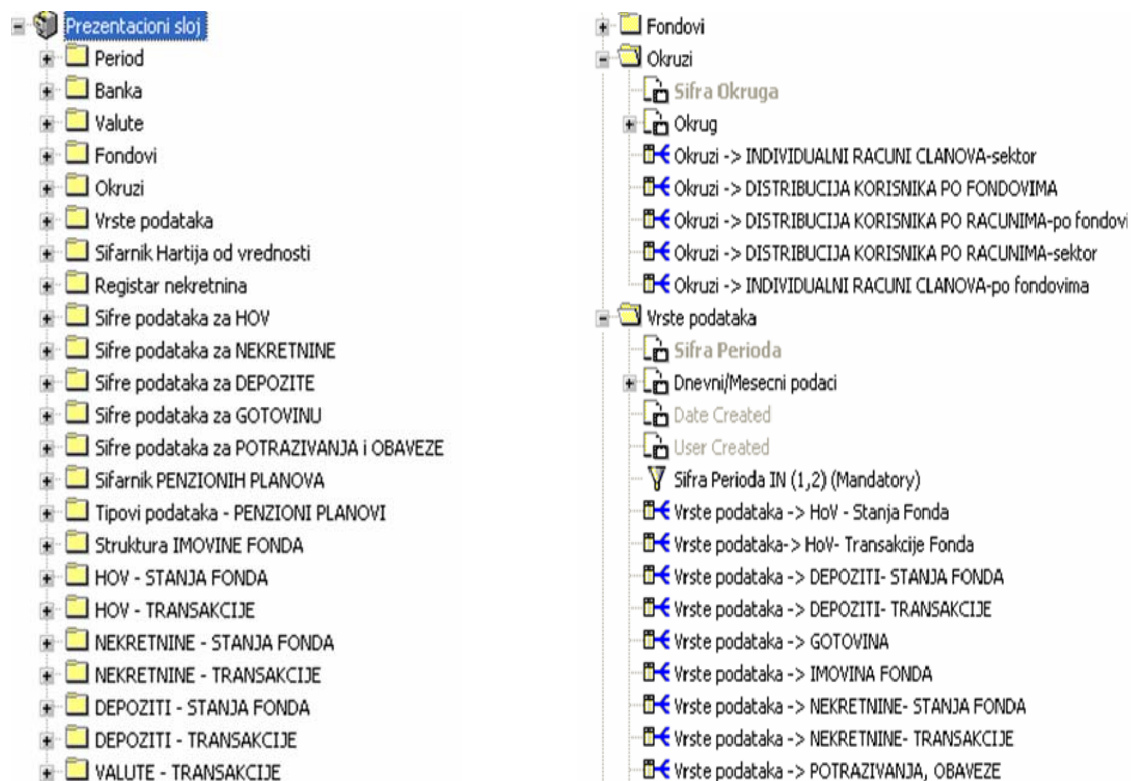
Na kraju ovog poglavlja dat je uporedni pregled porasta baze skladišta penzijskih fondova praćeno u polugodišnjem periodu, baze skladišta kreditnih odnosa sa inostranstvom praćeno u dvogodišnjem periodu i baze skladišta bilansnih podataka poslovnih banaka praćeno u periodu od godinu dana:

<u>Baza skladišta penzijskih fondova</u>		<u>Baza skladišta kreditnih odnosa</u>		<u>Baza skladišta bilansnih podataka poslovnih banaka</u>	
<i>Datum evidencije</i>	<i>Veličina baze skladišta u GB</i>	<i>Datum evidencije</i>	<i>Veličina baze skladišta u GB</i>	<i>Datum evidencije</i>	<i>Veličina baze skladišta u GB</i>
01.03.2009	1.40917969	07.05.2008	7.0181234	27.02.2009	5.37
15.03.2009	1.83544922	01.01.2009	8.1464844	05.01.2010	5.48
01.04.2009	1.90698242	01.04.2009	9.1230469	01.02.2010	5.75
15.04.2009	2.19799805	15.06.2009	9.5454102	01.03.2010	5.92
01.05.2009	2.20434572	15.08.2009	10.2587891	01.04.2010	6.23
15.05.2009	2.51708984	01.12.2009	11.0419922	01.05.2010	6.40
01.06.2009	2.52465821	01.03.2010	11.8161621	15.05.2010	6.45
15.06.2009	2.92089844				
01.07.2009	2.92089844				
15.07.2009	3.21533203				
01.08.2009	3.22753906				

5.5 Prezentacioni sloj

Korak nakon implementacije skladišta podataka bio je izrada prezentacionog nivoa skladišta, tj. „slike“ baze skladišta prilagodjene pomenutim grupama krajnjih korisnika. Nezavisno od proizvođača aktuelnih BI alata na tržištu i nezavisno od samog izabranog alata, praćena su osnovna pravila prikazivanja podataka iz skladišta poslovnim korisnicima. Ono što je moglo da se zaključi iz razgovora sa korisnicima (posebno sa anličarima koji su najviše uključeni u samu izradu pregleda) je da je njima najbitnije da ponudjeni alati i cela organizacija prezentacionog sloja bude intuitivna i fleksibilna za rad i da u potpunosti prati njihove poslovne procese. Srodne grupe podataka, koje se često posmatraju zajedno, su grupisane i u samom alatu. Većina alata danas prisutnih na tržištu omogućava grupisanje podataka iz skladišta u poslovne zone shodno poslovnim procesima koje oslikavaju. Dimenzije su naznačene i vizuelno odvojene od poslovnih mera. Uz aktivno učestvovanje korisnika napravljene su hijerarhije nad dimenzijama i imenovani su svi objekti na prezentacionom nivou. Imenovanja su uradjena isključivo u skladu sa poslovnim terminologijom, a ne sa terminologijom baza podataka.

Prikaz dela prezentacionog sloja (opšte i detaljnije) uradjenog nad bazom skladišta penzijskih fondova dat je na slici 2.



Slika 2. Opšti i detaljniji prikaz dela prezentacione slike skladišta podataka o penzijskim fondovima

Veze između podataka su većim delom implementirane u samoj bazi podataka, dok je jedan deo veza urađjen na nivou samog alata za administriranje korisničkog interfejsa. Sva pomoćna polja (sekvence za sortiranje, pomoćni podaci za izvodjenja, jedinstveni ID identifikatori, polja koja povezuju mere sa dimenzijama i sl.) su sakrivena od korisnika. U primeru penzijskih fondova je nad jednom tabelom podignuto više tzv. korisničkih tabela sa filterima koji su takodje sakriveni od korisnika i „vide“ se samo u generisanom SQL upitu. U alatu za administriranje su napravljene hijerarhije i definisana sortiranja podataka. Većina izvedenih polja je formirana na nivou baze podataka, ali je jedan manji broj formiran i u samom alatu za administriranje. Ono što je bitno je da su sva „dešavanja“ u bazi podataka, prilikom formiranja i izvršavanja upita, potpuno nevidljiva za krajnje korisnike.

U cilju sigurnosti informacija iz skladišta i zaštite podataka koji se prezentuju pristup je dozvoljen samo autorizovanim korisnicima. Sistem ograničenog pristupa podacima iz skladišta, definisanjem privilegija, primenjen je kako na nivou baze podataka tako i na nivou samog korisničkog alata.

5.6 Završne napomene

Trenutno u Narodnoj banci ne postoji jedinstveno, sveobuhvatno skladište podataka koje bi uključivalo sve poslovne subjekte koji se analiziraju i koje bi moglo da podrži sve poslovne procese koji se u banci odvijaju. Princip koji je prihvaćen kao standardan je izrada više odvojenih Data Mart podsistema. Svaki podsistem obuhvata poslovne procese jednog sektora u banci, ili deo poslova sektora koji predstavljaju jednu celinu i koje obavlja određena grupa ljudi. Ono na čemu se poslednjih godina insistira to je postojanje i centralizovano održavanje zajedničkog skupa dimenzija za sve Data Mart-ove, po uzoru na Kimballov princip magistrale u projektovanju skladišta. Sva „stara“ skladišta o kojima je bilo reči biće, nakon tehnološkog unapredjenja, takodje povezana sa skupom zajedničkih dimenzija.

Vremenom su se definisale zajedničke aktivnosti kroz koje se prolazilo pri projektovanju i implementaciji svakog pojedinačnog skladišta podataka. To su:

- Izrada modela skladišta podataka koje se zasniva na zvezdastim šemama uz eventualne modifikacije. Model se iterativno doradjuje i proširuje kako se prolazi kroz faze razvoja skladišta.
- Formiranje jedinstvene baze podataka sa svim grupama podataka kojima ciljni sektor raspolaže, a koji su zahtevom korisnika i IT analizom označeni kao kandidati za uključivanje u jedinstveno skladište podataka.
- Inicijalno punjenje baze podataka istorijskim podacima iz izvornog sistema. Obim i „starost“ podataka zavise od konkretnih poslovnih potreba i definišu se u dogovoru sa korisnicima.

- Osvežavanje baze podataka novim serijama podataka, kao i korekcijama prethodno zahvaćenih podataka, ukoliko postoje. Dinamika osvežavanja baze određena je prema potrebama posla, ali i prema zahtevima krajnjih korisnika.
- Formiranje standardnog skupa izveštaja od strane programera, koji bi poslužili kao polazna osnova za samostalno korišćenje predviđenih korisničkih alata.
- Formiranje radnog okruženja u predviđenim korisničkim alatima. Radno okruženje se formira iz fizičke implementacije skladišta i ono je intuitivno, logički organizovano i poslovno razumljivo sa aspekta krajnjeg korisnika.

Kada je reč o projektnim timovima koji su učestvovali u razvoju i koji danas rade na održavanju skladišta podataka, to su, u svakom slučaju, veoma mali timovi. Raspoloživost ljudskih resursa i veliki broj projekata na kojima se radi utiču na to da u fazi razvoja bude uključeno dvoje do troje zaposlenih u IT sektoru. Na održavanju svakog skladišta podataka nikada nije angažovano više od dvoje zaposlenih. Tim od dvoje ljudi pokazao se kao dovoljan, jer je evidentno da su zahtevi krajnjih korisnika koji dolaze ka IT sektoru znatno manjeg obima i da je samostalnost korisnika u korišćenju skladišta vremenom sve veća.

6 KORISNICI SKLADIŠTA PODATAKA

Kao što je napred rečeno, jedan od osnovnih ciljeva i motiva za uvođenje novih izveštajnih sistema zasnovanih na skladištima podataka bio je da se autorima izveštaja omogući da brzo prave i distribuiraju izveštaje koji zadovoljavaju potrebe širokog kruga korisnika. Vremenom su se izdvojile tri grupe korisnika koje su od početka produkcionog života skladišta podataka koristile, svaka prema svojim potrebama, informacije iz skladišta. To su:

- Analitičari poslovanja, koji aktivno učestvuju u izradi izveštaja. Ovo je grupa korisnika koji su obučeni za rad u BI alatima. Oni u potpunosti poznaju strukturu i veze koje postoje medju podacima u skladištu. Organizacija prezentacionog sloja radjena je uz aktivno učestvovanje upravo ovih korisnika, jer su oni ti koji prave standardne izveštaje, ali rade i ad-hok analize. Oni služe kao podrška ostalim korisnicima skladišta. Od njih se očekuje da prikupljaju i analiziraju informacije, utvrđuju trendove i pojavu anomalija u podacima. Oni imaju potrebu da informacije u skladištu razumeju do najdubljeg nivoa. Grubi statistički podaci kažu da u ovu grupu korisnika uglavnom spada oko 15%-20% svih korisnika skladišta. Izuzetak su korisnici u Sektoru za međunarodnu saradnju. Ukupan broj korisnika ovog skladišta je mali, ali su svi obučeni analitičari poslovanja specijalizovani za praćenje kreditnih odnosa sa inostranstvom koji aktivno učestvuju u izradi izveštaja i analizama podataka iz skladišta. Oni najduže rade sa BI alatima, dobro poznaju organizaciju skladišta i potpuno su samostalni u radu. Podrška koja se zahteva od IT lica je u delu analiza i izveštavanja svedena na minimum. Korisnici operativci zaposleni u ovom sektoru su uglavnom orijentisani na dnevne podatke iz transakcionog sistema.
- Menadžeri, odnosno upravljačka struktura. Ovo su korisnici delimično obučeni za rad sa BI alatima, jer se oni retko bave analizama, već koriste informacije i rezultate analiza u cilju donošenja odluka. Oni su uglavnom orijentisani na dinamičke izveštaje (uradjene po zahtevu) u kojima se prate ključni indikatori poslovanja. Oni retko samostalno izradjuju izveštaje već, uglavnom, minimalno koriguju već napravljene najznačajnije preglede. Veze izmedju objekata skladišta i sama organizacija prezentacionog sloja su im relativno poznati, onoliko koliko je potrebno za osnovno snalaženje u skladištu podataka. Iskustvo je pokazalo da je ove korisnike najbolje snabdeti skupom predefinisanih izveštaja iz bitnih oblasti poslovanja.
- Ostali korisnici informacija-operativci, izvršiooci. U ovu grupu spadaju zaposleni koji informacije iz skladišta koriste dnevno i to da bi efikasno obavljali svoj posao. Oni koriste izveštaje uradjene nad onim podskupom podataka iz skladišta koji je u domenu njihovog opisa posla. Nikada nemaju potrebu da sagledaju celokupno poslovanje. U svom ograničenom domenu mogu da koriste i tehnike bušenja i filtriranja podataka. Karakteristično za ove korisnike je da

svakodnevno koriste isključivo jedan deo informacija iz skladišta, tj. „parče“ iz svog domena rada. Oni su takoreći neobučeni za rad u BI alatima. Delimično poznaju veze između objekata u skladištu i to samo u onom delu modela koji je njima interesantan. Zato retko samostalno prave izveštaje, već uglavnom koriste predefinisane preglede. U slučaju skladišta podataka prikazanih u primerima penzijskih fondova i kontrole poslovanja banaka broj ovakvih korisnika je bio čak oko 75%. Ovde je većina operativaca orijentisana na korišćenje gotovih pregleda uz minimalne izmene koje se najčešće svode na izmene filtera u upitima.

Svi korisnici skladišta, oni koji prave preglede kao i oni koji ih samo koriste, su prošli obuku za korišćenje alata koji su predviđeni za rad sa podacima iz skladišta. Vremenom se iz velike grupe polaznika izdvojila grupa koja vrlo uspešno i aktivno koristi podatke iz skladišta, dok su ostali pasivni korisnici gotovih pregleda. Generalno, podrška zaposlenih u IT sektoru pri izradi pregleda je sve manja. Dogovoreni princip je da na početku IT projektni tim inicijalno uradi veći broj standardnih pregleda u koje ima uvid i koji se zahtevaju sa korisničke strane. Nakon obavljene obuke napredniji korisnici, uz pomoć gotovih pregleda, počinju da rade samostalno. Dalji napredak u korišćenju BI alata je individualan i zavisi od složenosti poslovnih analiza, složenosti modela samog skladišta, ali i ličnih afiniteta korisnika. Činjenica je da su, na primer, analize koje se rade nad knjigovodstvenim i bilansnim podacima poslovnih banaka veoma složene, da je poslovna problematika poznata malom broju analitičara i da je sam model skladišta složeniji od ostalih modela koji su radjeni. Zato i osamostaljivanje u korišćenju ide sporije nego u drugim slučajevima implementacije.

Što se tiče rada u korisničkim alatima, zadatak zaposlenih u IT sektoru uvek je bio da se korisniku olakša da shvati veze između podataka i ovlada tehnikom izvlačenja, kombinovanja, ukrštanja informacija i prikazivanja u željenom formatu (tabela, graf i sl.). Napredniji korisnici skladišta podataka relativno lako su uspeli da ovladaju sledećim tehnikama:

- Pravljenje ad-hok upita.
- Izbor željenih stavki koje će se uključiti u upite.
- Menjanje upita pivotiranjem, sortiranjem i dodavanjem ili ukljanjanjem polja.
- Formatiranje prikaza dobijenih rezultata.
- Deljenje pregleda sa drugim korisnicima.
- Pronalaženje i formatiranje izuzetaka u podacima, kao i odstupanja od zadate vrednosti.
- Korišćenje predefinisanih uslova.
- Definisavanje novih uslova.
- Definisavanje parametara.
- Korišćenje parametara pri definisanju uslova.
- Uključivanje suma, medjusuma i procenata.
- Pravljenje izvedenih polja.
- Korišćenje parametara u izračunavanjima.
- Pravljenje i izmene grafova kao formata za prikaz podataka.

- Svi tipovi bušenja podataka duž hijerarhija na dimenzijama: Bušenje na gore ili na dole, bušenje ka povezanom polju, bušenje ka detalju.
- Zadavanje vremena izvršavanja izveštaja.
- Izvoz rezultata izveštaja u druge aplikacije.
- Slanje izveštaja elektronskom poštom.

7 DALJI PLANOVI

Dalji planovi vezani za unapredjenje skladišta podataka u Narodnoj banci Srbije obuhvataju nekoliko aktivnosti:

- Izmeštanje baze skladišta kreditnih poslova sa inostranstvom sa transakcione baze podataka na posebnu bazu. Kao što je rečeno, to je baza optimizovana za rad sa podacima iz skladišta i na njoj se nalaze ostala skladišta podataka. Izmeštanje podrazumeva i tehnološko unapredjenje skladišta koje je projektovano pre 10 godina na tada aktuelnim platformama. Alati koje korisnici upotrebljavaju su, u današnjem trenutku, zastareli i treba ih zameniti nekim od BI alata koji su danas prisutni na tržištu. U sektoru postoji više desetina urađenih pregleda, izveštaja i analiza nad podacima iz skladišta. Ono što će biti jedan od najtežih zadataka je da se migracija sprovede što „bezbolnije“ po IT i korisnike, tj. da se metabaza korisničkog alata u kojoj su sačuvane definicije, opisi i veze objekata sa prezentacionog sloja, kao i opisi svih pregleda koje su korisnici do sada uradili migrira bez većih gubitaka.
- Ozbiljnije bavljenje problemom fragmentiranja, odnosno logičkog i fizičkog particionisanja podataka. Činjenica je da particionisanje tabela na fizičkom nivou, kao jedna od ključnih stavki za poboljšanje performansi izvršavanja upita, nije bilo adekvatno iskorišćeno u do sada razvijenim skladištima podataka. Nova znanja i stečeno iskustvo doveli su do toga da se o problemu particionisanja počne ozbiljnije da razmišlja. Procena je da bi najpre trebalo da se reorganizuju podaci o planovima otplata u sistemu kreditnih odnosa, jer je ovih podataka najviše i upiti koji rade nad njima su najsporiji. Kada je o ovom sistemu reč, vremenom se pokazalo da se podaci stariji od godinu dana vrlo retko posmatraju, tako da se datum važenja podataka odmah nameće kao jedan od ključeva za particionisanje podataka. Knjigovodstveni i bilansni podaci poslovnih banaka su, takodje, kandidati za reorganizaciju. Razlog je sličan: podataka je mnogo, a iskustvo je pokazalo da se podaci stariji od 2 godine znatno redje koriste i mogu se izmestiti u posebne particije.
- Arhiviranje istorijskih podataka. Pitanje arhiviranja podataka se posebno potencira u poslednje vreme i to je jedna od trenutno aktuelnih tema u IT krugovima. IT sa svoje informatičke tačke gledišta pokreće pitanje arhiviranja kako transakcionih baza, tako i baza skladišta podataka. Neophodnost arhiviranja starih podataka je posebno izražena u slučaju skladišta podataka o kreditnim odnosima sa inostranstvom. Ovo skladište je postalo preopterećeno istorijskim podacima koji su prestigili starost od 10 godina. Zahtev za arhiviranjem grupa kredita dolazi sa obe strane: od strane projektanata i administratora baze u IT sektoru, ali i od strane krajnjih korisnika. Korisnici žele da sa svoje poslovne strane arhiviraju sve „stare“, neaktivne kredite, u koje spadaju krediti bivših jugoslovenskih

republika, ali i svi domaći krediti zaključeni pre 2000. godine. Ono na čemu korisnici insistiraju je da im se u slučaju potrebe, arhivirani podaci u najkraćem roku stave na uvid. Ono na čemu se trenutno radi u IT sektoru je definisanje načina i metoda, tj. neke vrste strategije arhiviranja i dearhiviranja podataka iz svih skladišta koja se koriste u banci. Skladište podataka o kreditnim odnosima sa inostranstvom će biti među prvima na kojima će se arhiviranje sprovesti, jer su podaci u njemu najstariji.

8 ZAKLJUČAK

Poslovna inteligencija je oblast u okviru informacionih tehnologija čiji je cilj da informacione potencijale jednog preduzeća stavi u funkciju donošenja najkvalitetnijih odluka i ostvarivanja strateških ciljeva preduzeća. Poslovna inteligencija uključuje razne vrste tehnologija i alata koji predstavljaju nadogradnju skladištenju podataka. Zato se kaže da je skladištenje podataka osnova, tj. arhitektura na kojoj se zasnivaju sistemi poslovne inteligencije. Skladištenje podataka je danas našlo primenu u skoro svim oblastima rada i odlučivanja, svuda gde je akcenat na informaciji kao osnovnom resursu, a to su: proizvodnja, telekomunikacije, osiguranje, medicina, bankarstvo, uslužne delatnosti i sl. Cilj je efikasnost, pravovremenost i pre svega tačnost dobijenih informacija na osnovu kojih se prate ključni indikatori poslovanja, trendovi, uočavaju problemi i brzo reaguje.

Ono što je u ovom radu novo je to što su opisane primene osnovnih koncepata i metoda skladištenja podataka u domenu poslovanja Centralne banke. Poslovanje Centralne banke je specifično i bitno se razlikuje od poslovanja ostalog bankarskog sektora u zemlji. Većina poslovnih funkcija proizilazi iz same uloge Narodne banke koja podrazumeva sveobuhvatno sagledavanje, praćenje i analizu kompletnog finansijskog sektora u zemlji da bi se delovalo preventivno, stimulatивно ili kazneno.

Ono što se u IT sektoru pokušava poslednjih godina je da se velika količina podataka koja neprestano dolazi u Centralnu banku iz različitih izvora grupiše prema poslovnim subjektima, sistematizuje, transformiše u kvalitetnu informaciju i ponudi poslovnim korisnicima kroz BI alate. U Narodnoj banci akcenat je na analizi podataka, efikasnom i pouzdanom izveštavanju da bi se brzo donosile odluke. Višegodišnje iskustvo je pokazalo da je, ipak, BI koncept u oblasti analiza i izveštavanja za ovakvo poslovno okruženje najprihvatljiviji. Ne treba zanemariti poteškoće kojih je bilo i još uvek ih ima kad god se uvodi nova tehnologija. Obuka zaposlenih u IT sektoru za rad na ovakvim projektima, vreme od nekoliko meseci potrebno da se neko skladište uvede, otpor jednog dela poslovnih korisnika prema novim tehnologijama i obuka korisnika za korišćenje podataka iz skladišta su neki od problema koji i dalje postoje. Ono što je postignuto je efikasniji, samostalniji i kvalitetniji analitički rad korisnika. Kad se sve sumira, dobija se pozitivan rezultat uporedne analize cene i postignutih rezultata.

Zaključak je da su svi članovi timova koji su učestvovali u projektima razvoja skladišta sa strane IT sektora a, pre svega, sa strane poslovnih korisnika zadovoljni onim što je dobijeno sa uvođenjem skladišta podataka. Dalji planovi su da se intezivira rad na projektima ovog tipa. S tim u vezi planira se dalje usavršavanje zaposlenih u IT sektoru u oblasti skladištenja podataka i inteligentnog poslovanja uopšte. Neizbežno je i praćenje tehnoloških trendove na tržištu da bi se efikasno odgovorilo potrebama savremenog poslovnog okruženja.

III PRILOZI: PRIMERI IMPLEMENTACIJE

A. Penzijski fondovi

Implementacija skladišta podataka u Sektoru za nadzor
društava za upravljanje dobrovoljnim penzijskim fondovima
Narodne banke Srbije

A.1. Izvori podataka

Podaci koji su uključeni u ovaj sistem za skladištenje podataka dolaze iz tri izvora, i to isključivo elektronski preko ovlašćene službe Narodne banke. Izvori podataka su:

- Kastodi banke. Za sada to je trinaest poslovnih banaka u Srbiji koje imaju dozvole za obavljanje kastodi delatnosti. Iz ovog izvora podaci dolaze dnevno i uvek se odnose na prethodni radni dan. Podaci koje dostavljaju kastodi banke svrstani su u nekoliko grupa:
 - Podaci o stanju dužničkih hartija od vrednosti i akcija u portfoliu fonda.
 - Podaci o dnevnom trgovanju dužničkim hartijama od vrednosti i akcijama.
 - Podaci o stanju nepokretnosti u portfoliu fonda.
 - Podaci o dnevnom trgovanju nepokretnostima.
 - Podaci o položenim i povučenim depozitima.
 - Podaci o gotovini kojom fond raspolaže.
 - Podaci o stanju potraživanja i obavezama fonda.

Kastodi banka ove podatke dostavlja za svaki pojedinačni fond koji svoje aktivnosti na tržištu obavlja preko te banke.

- Društva za upravljanje dobrovoljnim penzijskim fondovima. Iz ovog izvora podaci dolaze do desetog dana u mesecu i uvek se odnose na prethodni kalendarski mesec. Veći deo podataka predstavlja dnevne promene i to za svaki radni dan u prethodnom mesecu, dok manji broj podataka, kao što su podaci o transakcionim troškovima, predstavlja stanje na poslednji radni dan u prethodnom mesecu. Podaci koje dostavljaju društva za nadzor svrstani su, takodje, u nekoliko grupa:
 - Podaci o stanju dužničkih hartija od vrednosti i akcija u portfoliu fonda.
 - Podaci o trgovanju dužničkim hartijama od vrednosti i akcijama.
 - Podaci o stanju nepokretnosti u portfoliu fonda.
 - Podaci o trgovanju nepokretnostima.
 - Podaci o depozitima.
 - Podaci o gotovini.
 - Podaci o stanju potraživanja i obavezama fonda.
 - Podaci o transakcionim troškovima i naknadama.
 - Podaci o penzijskim planovima.
 - Podaci o članovima fonda.
 - Podaci o individualnim računima članova fonda

Za sve navedene grupe podataka dozvoljene su korekcije podataka. Korekcije se dostavljaju isključivo elektronski i to ponovljenim slanjem podataka. Korekcije dolaze periodično, po potrebi, ili prema posebnom zahtevu ukoliko se uoči greška u primljenim podacima.

- Beogradska berza. Podaci Beogradske berze se ne razmenjuju na isti način kao podaci iz prethodna dva izvora. To su podaci koji se automatizovano, jednom dnevno prevlače sa Veb stanice Beogradske berze i smeštaju u prihvatne tabele, strukturirane tako da podrže direktno učitavanje iz preuzetih datoteka. Podaci su dostupni svakog radnog dana za tekući dan. Za ove podatke nisu predviđene korekcije nakon preuzimanja.

Sistem prihvatanja elektronski dostavljenih podataka zaživeo je tri godine pre početka projektovanja ovog skladišta podataka. Svi prihvaćeni istorijski podaci zahvaćeni su pri inicijalnom punjenju baze skladišta.

Novi izveštajni sistem ima osnovnu namenu da objedini podatke kastodi banka, društava za nadzor i Beogradske berze, da ih organizuje tako da budu dostupni i laki za korišćenje i da korisnicima omogući da na brz i jednostavan način izradjuju sopstvene izveštaje i formiraju ad-hok upite.

A.2. Opis poslovnih funkcija

Na osnovu identifikovanih izvora podataka i analizom poslovnih procesa u sektoru definisan je skup poslovnih funkcija koje treba da budu podržane od strane novog izveštajnog sistema.

Poslovne funkcije označene kao kandidati za implementaciju u novom sistemu su:

- Analize podataka koji se odnose na stanje dužničkih hartija od vrednosti i akcija u portfoliju fonda. Analize se obavljaju zbirno za sve hartije u portfoliju fondova, ali i pojedinačno i to po svakom obeležju hartije od vrednosti (ISIN broju⁶, simbolu⁷, valuti, izdavaocu, itd.).
- Detaljne i zbirne analize podataka o transakcijama dužničkim hartijama od vrednosti i akcijama na tržištu hartija od vrednosti.
- Uporedne analize podataka o stanjima hartija od vrednosti u portfoliju fonda i podataka o trgovanju tim hartijama. Uporedne analize izvode se po zajedničkim dimenzijama ovih kategorija podataka.
- Analize podataka koji se odnose na stanje nekretnina u portfoliju fonda. Analize treba da budu zbirne za sve nepokretnosti u registru, kao i pojedinačne i to po svakoj registrovanoj nekretnini.
- Analize podataka koji se odnose na trgovinu nepokretnostima na tržištu, po svim relevantnim merama (cena, prihod, troškovi i sl.).
- Uporedne analize podataka o tome koje nepokretnosti fond poseduje i koje nepokretnosti kupuje i prodaje u dnevnim transakcijama.

⁶ ISIN (eng. *International Securities Identification Number*) broj – jedinstveni identifikacioni broj hartije od vrednosti^[10]

⁷ Trgovački simbol hartije od vrednosti na Beogradskoj berzi

- Analize podataka koji se odnose na položene i povučene depozite fonda, kao i stanje gotovine. Ovo su podaci koji se, prema uputstvu o razmeni podataka, dostavljaju razvrstani po kategorijama koje su značajne sektoru za analize i izveštavanje. Sistem treba da obezbedi zbirno i analitičko izveštavanje o depozitima i gotovini do nivoa svake pojedinačne kategorije. Depoziti i gotovina se prate i uporedno.
- Analize podataka koji se odnose na potraživanja i obaveze fonda.
- Formiranje i analize neto i ukupne imovine pojedinačnih fondova i celokupnog sistema fondova u zemlji. Ove mere se formiraju na osnovu prethodno navedenih kategorija, prema propisanoj metodologiji. Sistem mora da obezbedi uvid u iznos imovine, ali i hijerarhijski pregled komponenti iz kojih je ona izvedena. Takodje, prati se i procenat učešća svake komponente imovine u ukupnom iznosu, kako na nivou pojedinačnog fonda, tako i na nivou sistema svih fondova.
- Analize podataka koji se odnose na transakcione troškove i naknade. Ovo su kategorije koje se prate uporedno i pojedinačno. Sistem mora da obezbedi detaljan uvid u sve elemente iz kojih su ove poslovne mere sastavljene.
- Analize podataka koje društva za nadzor dostavljaju o članovima svojih fondova, izvršenim uplatama i isplatama sredstava, ali i računima članova fondova.
- Uporedne analize podataka kastodi banaka i društava za nadzor koji su za iste poslovne subjekte dostavljeni iz ova dva izvora. Pri projektovanju skladišta podataka izvršilo se usaglašavanje zajedničkih podataka kako po imenovanju tako i po formatu, sadržaju i značenju. Napravljena je unija mera o istom poslovnom subjektu koje potiču iz ova dva izvora.
- Analize podataka o dnevnim transakcijama dužničkim hartijama od vrednosti i akcijama na Beogradskoj berzi. U sklopu ovih analiza vrši se i uporedna analiza podataka sa berze i podataka koje društvo za nadzor ili kastodi banka dostavljaju iz svojih evidencija.
- Formiranje pokazatelja poslovanja fondova, prema propisanoj metodologiji.

A.3. Fazni razvoj sistema

Ideja je bila da se u prvoj iteraciji razvoja skladišta sagledaju grupe podataka koje učestvuju u formiranju imovine fonda i da se taj podmodel prvi implementira. Na taj način korisnicima su najpre omogućene poslovno najznačajnije analize, kao npr. analize imovine fondova po svim aspektima. Naredne iteracije uvodile su jednu po jednu poslovnu kategoriju koje ne utiču na imovinu fonda i odvojeno se prate u sektoru. Poslednja iteracija bila

je uvođenje složenog podmodela individualnih računa članova fondova, čime su svi predviđeni podaci bili uključeni u skladište.

I faza:

U ovoj fazi identifikovane su skoro sve poslovne dimenzije koje će se koristiti kako se sistem bude proširivao. To su: vremenska dimenzija, dimenzija poslovnih banaka, dimenzija penzijskih fondova, šifarnik hartija od vrednosti, registar nepokretnosti i dimenzija valuta. Kako se svaka od navedenih šest grupa podataka, koje ulaze u imovinu fonda, razvrstava na podgrupe, podgrupe dalje na svoje podgrupe itd. po definisanim šiframa razvrstavanja, uvedena je dimenzija šifara podataka. Dimenzija šifara podataka je hijerarhijski organizovana kako bi se omogućio prolaz do najnižeg nivo podataka. Na kraju ove faze u sistem je uključena i dimenzija u kojoj će se čuvati samo indikator koji razdvaja dnevne podatke kastodi banaka od mesečnih podataka društava za nadzor.

II faza:

Uvođenje podataka o uplatama, isplatama, naknadama, transakcionim troškovima i članovima fondova. To su podaci koji se prate na nivou datuma i fonda i ne raščlanjuju se detaljnije. Zahtev je bio da sistem omogućí uporedno praćenje ovih podataka. Podaci o članovima fondova su jedini iz ove grupe podataka koji se dostavljaju sa stanjem na poslednji radni dan u mesecu. Stoga je njihovo upoređivanje sa ostalim podacima moguće na mesečnom a ne na dnevnom nivou. U ovoj fazi uvedene su samo nove tabele činjenica. Potrebne dimenzije su identifikovane u prvoj fazi.

III faza:

Uvođenje podataka o penzijskim planovima, obveznicima uplata, pojedinačnim ugovorima i transferima. Ova faza podrazumevala je uključivanje nove dimenzije šifarnika penzijskih planova i dimenzije u kojoj se čuvaju podaci o tipovima penzijskih planova. Kako su i ovim podacima predstavljena stanja na poslednji radni dan u mesecu, uporedno praćenje sa podacima uvedenim u drugoj fazi je moguće najdetaljnije na mesečnom nivou.

IV faza:

Uvođenje podataka o trgovanju na Beogradskoj berzi. Podaci su na dnevnom nivou razvrstani po svakoj pojedinačnoj hartiji kojom se trguje, tako da su vremenska dimenzija i šifarnik hartija od vrednosti jedine dve dimenzije ove tabele činjenica. Ovo su dimenzije koje su identifikovane i imlementirane još u prvoj fazi razvoja sistema. Uvođenjem podataka o trgovanju na berzi u sistem su uključeni i pokazatelji trgovanja koji se čuvaju na dnevnom nivou. Na dnevnom nivou obračunavaju se i pokazatelji poslovanja fondova. Kako se ovi pokazatelji prate uporedo i kako su podaci iste granularnosti svi su smešteni u jednoj tabeli činjenica povezanoj samo sa datumskom dimenzijom.

V faza:

Uvodjenje podataka o individualnim računima članova fondova. Dimenzije koje je iskorišćene iz prethodnih faza su vremenska dimenzija i dimenzija fondova. Uvedena je nova dimenzija okruga jer je zahtev bio da se svaka poslovna mera iz ovih tabela prati po regionalnoj strukturi, tj. po okruzima. U ovoj fazi razvoja sistema implementirane su dve tabele činjenica, sa preko pedeset poslovnih mera. Svaka poslovna mera izvedena je iz izvornih podataka složenom ekonomskom logikom. Ovaj podmodel ne sadrži nijedan izvorni podatak, već samo izvedene poslovne mere. Zbog neaditivnosti poslovnih mera, podaci na nivou celog sistema ne mogu da se izvedu iz podataka koji su sačuvani na nivou svakog pojedinačnog fonda. Zato su podaci po fondovima i podaci na nivou sistema posebno računati i smešteni u razdvojene tabele činjenica.

A.4. Opis poslovnih dimenzija

Iz svih podataka koje sektor prati i analizira, izvedene su osnovne poslovne kategorije, tj. poslovne dimenzije i poslovne mere.

Skup dimenzija sistema je definisan na osnovu nivoa granularnosti podataka u tabelama činjenica. Nivo granularnosti u tabelama činjenica određen je prema potrebama sve tri grupe korisnika koji će koristiti ovaj sistem. Poslovne potrebe, a i zahtevi korisnika koji su inicirali izgradnju skladišta podataka, uslovili su da se određene grupe podataka čuvaju na najdetaljnijem nivou, uz mogućnost njihovog agregiranja. Ukoliko bi se neka od potrebnih dimenzija isključila iz sistema dobili bi se podaci manje detaljnosti. Iz razgovora sa krajnjim korisnicima i samim uvidom u njihove potrebe, došlo se do zaključka da bi isključivanjem najdetaljnijih podataka iz skladišta poslovne analize bile osiromašene. Kako korisnici nemaju alternativni informacioni sistem koji daje uvid u detaljne, transakcione podatke oni će svoje analitičke aktivnosti sprovoditi isključivo nad skladištem podataka. Zato je dogovoreno da se u skladište podataka uključe i neki najgranularniji podaci. Kasnije bi se, po potrebi, identifikovali i implementirali agregati, u cilju poboljšanja performansi upita.

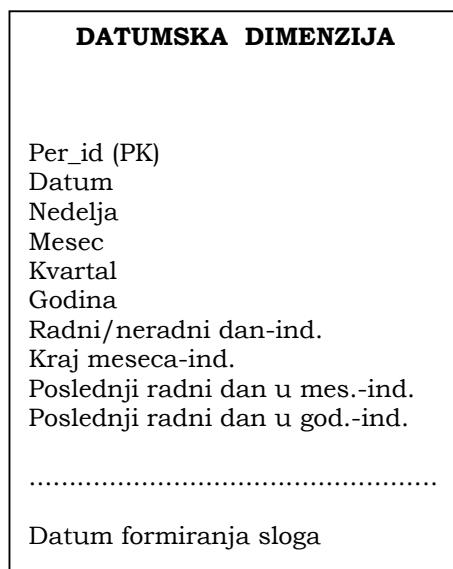
A.4.1. Vremenska dimenzija

Vremenska ili datumska dimenzija je dimenzija koja postoji u svim sistemima za skladištenje podataka, odnosno u svim nezavisno razvijanim Data Mart podsistemima. Svi sistemi za skladištenje podataka, koji su kroz primere analizirani u ovom radu, dele istu datumsku dimenziju, koja je i fizički implementirana kao jedna tabela. Na ovoj dimenziji zasnivaju se još neki sistemi za skladištenje podataka koji nisu opisani u ovom radu.

Datumska dimenzija čuva podatke do nivoa dana i to za svaki kalendarski radni i neradni dan, od 01.01.2000. godine. Početni datum određuje datum od kada postoje podaci u ma kom implementiranom sistemu koji se oslanja na tu datumsku dimenziju. Datumska dimenzija se puni dnevno, podacima

o tekućem danu i to automatizovano, procedurom koja se pokreće pre svih ETL procedura za osvežavanje baza skladišta podataka. Tabela datuma se isključivo puni novim podacima, dok operacije brisanja i izmene podataka nad njom nisu dozvoljene. Osvežavanje ove tabele je nezavisno od dinamike ažuriranja sistema za skladištenje podataka. Tabela datuma sadrži polja: datum, oznaka kalendarske nedelje, oznaka meseca u godini, oznaka kvartala, oznaka godine, identifikator da li je dan radni ili neradni (zbog potreba filtriranja podataka), identifikator da li je dan kalendarski kraj meseca ili ne (takodje zbog potreba filtriranja podataka). Osim ovih, dimezija sadrži i par specifičnih polja. Ona se koriste u nekim sistemima koji zahtevaju nestadardnu podelu meseca i godine koja odstupa od kalendarske podele (slično fiskalnom mesecu, tj. fiskalnoj godini). Na ovaj način napravljena je uniformna datumska dimezija koja zadovoljava potrebe do sada implementiranih sistema za skladištenje podataka.

Prikaz datumske dimezije dat je na slici A.1.



Slika A.1. Datumska dimezija u sistemu skladišta podataka

Prirodni ključ ove dimezije (polje Datum) zamenjen je surogat ključem (polje PER_ID) koji je kao deo ključa spušten u svaku tabelu činjenica koja je implementirana u sistemu. Nad podacima iz datumske dimezije definisana je hijerarhija:

```

Godina
  Kvartal
    Mesec
      Nedelja
        Dan,
  
```

koja služi za „bušenje“ podataka naniže i naviše duž definisanih atributa.

A.4.2. Dimenzija poslovnih banaka

Slično datumskoj dimenziji, dimenzija poslovnih banaka je zajednička većini do sada implementiranih sistema za skladištenje podataka u Narodnoj banci Srbije, jer se svi podaci koji se analiziraju u ovim sistemima, uglavnom posmatraju u kontekstu poslovnih banaka.

Prikaz dimenzije poslovnih banaka dat je na slici A.2.

POSLOVNE BANKE
Ban_id (PK)
Matični broj
Naziv banke
Skraćeni naziv banke
Adresa
Grad
Republika-šifra
Republika-naziv
Status banke
Datum otvaranja
Datum zatvaranja
.....
Datum formiranja sloga

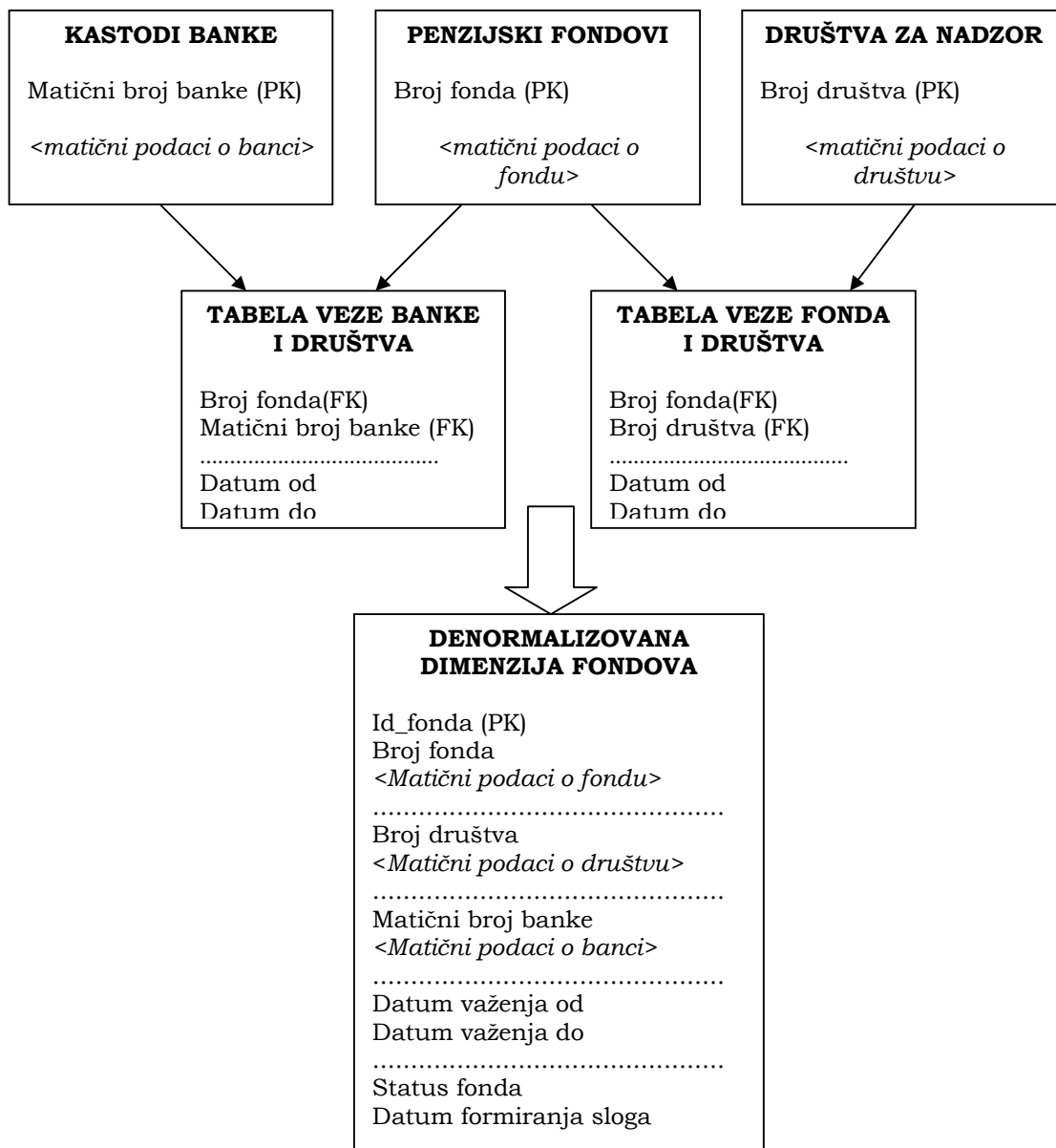
Slika A.2. Dimenzija poslovne banke

Prirodni ključ ove dimenzije (polje `Matični_broj`) zamenjen je surogat ključem (polje `Ban_id`) koji je, kao deo ključa, spušten u svaku tabelu činjenica koja je povezana sa ovom dimenzijom. Osim matičnog broja dimenzija banaka sadrži i: naziv banke, skraćeni naziv banke (zbog preglednijeg izveštavanja), ime grada, adresu, šifru i naziv republike itd. Status banke sadrži podatke o tome da li je banka aktivna, zatvorena ili pripojena nekoj drugoj poslovnoj banci. Pored datuma formiranja sloga (koji je uvek sistemski datum), ova dimenzija ima i dva datumska polja u kojima se čuva period važenja podataka. Naime, u slučaju zatvaranja banke ili njenog pripajanja drugoj poslovnoj banci, zatvara se period važenja podataka. Pri svakom evidentiranju nove poslovne banke u tabelu se upisuje slog koji dobija novu vrednost `Ban_id` ključa. Svaka promena naziva banke takodje inicira zatvaranje sloga i otvaranje novog sa istim matičnim brojem i novim nazivom. Praćenje istorijata promena naziva banaka je od posebne važnosti u skoro svim izveštavanjima koja se obavljaju nad podacima iz skladišta. Dimenzija poslovnih banaka ažurira se dnevno na osnovu podataka o poslovnim bankama koji postoje u transakcionim sistemima. U transakcionim sistemima nalaze se tekući podaci o bankama koje su aktivne, a svi prethodno važeći podaci smeštaju se u arhivske tabele. ETL procedure za ažuriranje dimenzije poslovnih banaka identifikuju promene u tekućim transakcionim podacima, konsultuju arhivu podataka i ažuriraju dimenziju banaka. Ažuriranje dimenzije podrazumeva upis novih slogova i

eventualno popunjavanje datuma važenja sloga prilikom deaktiviranja banke. Osim datuma važenja, nijedno drugo polje u tabeli se ne ažurira. Brisanje podataka iz tabele se nikada ne sprovodi.

A.4.3. Dimenzija penzijskih fondova

Jedna od ključnih dimenzija ovog sistema za skladištenje podataka je dimenzija penzijskih fondova. To je dimenzija na koju se naslanja većina implementiranih tabela činjenica. Dimenzija penzijskih fondova formirana je i ažurira se na osnovu podataka o fondovima, društvima za upravljanje i kastodi bankama iz transakcionog sistema. U transakcionom sistemu odvojeno se održavaju podaci o fondovima, kastodi bankama i društvima za upravljanje. Posebno se održavaju i veze koje postoje medju tim entitetima. Normalizovana šema, prikazana na slici A.3, zamenjena je denormalizovanom dimenzijom fondova koja pored osnovnih podataka o penzijskom fondu sadrži i matični broj i naziv kastodi banke, kao i matični broj i naziv društva za upravljanje u čijoj je nadležnosti.



Slika A.3. Zamena normalizovane šeme denormalizovanom dimenzijom fondova

Ova dimenzija, kao i prethodno opisane, sadrži dva datumska polja koja označavaju period važenja sloga. Pri promeni bilo kog obeležja fonda u transakcionom sistemu slog u tabeli dimenzije se zatvara i otvara novi sa novom vrednošću surogat ključa (atribut Id_fonda). Slog se zatvara i pri promeni naziva kastodi banke ili naziva društva za upravljanje, kao i u slučaju pripajanja fonda nekoj drugoj kastodi banci ili drugom društvu za upravljanje. Slog se, konačno, zatvara u slučaju prestanka rada fonda.

Nad dimenzijom penzijskih fondova definisane su hijerarhije:

Naziv društva za upravljanje
Naziv penzijskog fonda i

Naziv kastodi banke
Naziv penzijskog fonda,

koje omogućavaju dobijanje sumarnih podataka na nivou kastodi banke i društva za nadzor i „bušenje“ naniže do nivoa pojedinačnog fonda. Kako su nazivi, a ne matični brojevi, obeležja koja su krajnjim korisnicima značajna i sva izveštavanja se obavljaju po njima, to su obe hijerarhije bazirane isključivo na nazivima. Uvodjenje novih hijerarhija nad dimenzijom se jednostavno sprovodi i ne utiče na postojeći izveštajni sistem.

A.4.4. Dimenzija hartija od vrednosti

Šifarnik hartija od vrednosti uveden je kao dimenzija za sebe zbog potrebe korisnika za detaljnom analizom podataka o stanju dužničkih hartija od vrednosti i akcija u portfoliju fonda, kao i podataka o dnevnim transakcijama trgovanja ovim hartijama na tržištu. Podaci o hartijama od vrednosti se najdetaljnije posmatraju upravo na nivou svake pojedinačne hartije (tj. ISIN broja kao njenog prirodnog ključa). Eventualno izostavljanje ove dimenzije iz sistema dalo bi mogućnost samo kumulativnog praćenja stanja svih hartija u portfoliju fonda na dan, što bi značajno osiromašilo analitičko izveštavanje iz novog sistema.

Dimenzija hartija od vrednosti sadrži sve attribute značajne za opis i praćenje jedne hartije. To su: ISIN broj, simbol, CFI kod⁸, matični broj i naziv izdavaoca itd. U dimenziju je dodat izvedeni identifikator koji, na osnovu CFI koda, razdvaja tip hartije na dužničku hartiju i akciju, čime se poboljšavaju performanse pretraživanja podataka. Još jedan izvedeni identifikator uveden je takodje iz razloga boljih performansi upita. To je identifikator koji, na osnovu valute, razdvaja poreklo hartije na domaću i stranu.

Podaci o hartijama od vrednosti dobijaju se putem razmene datoteka sa propisanim xml sadržajem. Podaci dolaze od društava za nadzor dobrovoljnih penzijskih fondova i, delom, sa sajta Beogradske berze. ETL procedura koja ažurira ovaj šifarnik prati oba izvora i upoređuje podatke iz njih sa postojećim podacima u dimenziji. Nove grupe podataka upisuju se u tabelu u bazi. Ukoliko se identifikuje promena u izvornom sistemu nekog od atributa već postojeće hartije u dimenziji, ti atributi se programski ažuriraju.

Prikaz dimenzije hartija od vrednosti dat je na slici A.4.

⁸ CFI (eng. *Classification of Financial Instruments*) kod - niz slovnih znakova kojima se identifikuje vrsta hartija od vrednosti, pojedinačna prava iz te hartije od vrednosti i druga bitna obeležja koja se određuju s obzirom na vrstu te hartije od vrednosti[10]

ŠIFARNIK HARTIJA OD VREDNOSTI
Id_hartije (PK)
ISIN broj
CFI kod
Datum emitovanja
Datum dospeća
Matični broj izdavaoca
Naziv izdavaoca
Valuta hartije
Simbol hartije
Naziv hartije
Akcija/dužnička-ind.
Domaća/strana-ind.
.....
Datum formiranja sloga

Slika A.4. Dimenzija hartija od vrednosti

A.4.5. Dimenzija hijerarhijskog šifriranja podataka

Sektor je definisao metodologiju klasifikacije podataka o: hartijama od vrednosti, nepokretnostima, gotovini, depozitima, potraživanjima i obavezama. Ta klasifikacija predviđa hijerarhijsko (drvoidno) raščlanjivanje podataka svake od ovih grupa do atomičkog nivoa. Svaki podatak koji stigne putem elektronske razmene za ove grupe podataka šifriran je nekom od propisanih šifara ali isključivo sa atomičkog nivoa hijerarhije.

Primeri nekih definisanih podela po kategorijama podataka dati su sledećom šemom. Brojevi navedeni u zagradama predstavljaju jedinstvene propisane šifre svake kategorije podataka, odnosno svakog nivoa hijerarhije.

AKCIJE (1)

- Akcije domaćih lica (11)
 - Akcije domaćih lica - listing A (111)
 - Akcije domaćih lica - listing B (112)
- Akcije stranih lica (12)

NEPOKRETNOSTI (4)

- Zgrade (40)
 - Stambene zgrade (401)
 - Stanovi (404)
 - Poslovni prostor (405)
- Zemljište (41)
- Ostalo (42)

POTRAŽIVANJA (7)

- Potraživanja od društava za upravljanje (70)
 - Potraživanja od društava za transakcione troškove (700)
 - Ostala potraživanja od društava (709)

Ostala potraživanja (71)

Potraživanje po osnovu prodaje HOV (710)

Potraživanje po osnovu dividendi (712)

Potraživanje po osnovu prodaje nekretnina (713)

Potraživanje po osnovu uplate nove emisije HOV (714)

Potraživanje po osnovu dospele nenaplaćene kamate (715)

Ostala potraživanja (719)

Sve identifikovane hijerarhije šifriranja podataka grupisane su i objedinjene u jednoj dimenzionoj tabeli prikazanoj na slici A.5.

ŠIFRE_PODATAKA
Id_šifre (PK)
Dnevni/mesečni-ind.
Vrsta podatka-sort_id
Vrsta podatka
Nivo1-sort_id
Nivo1-šifra
Nivo1-naziv
Nivo2-sort_id
Nivo2-šifra
Nivo2-naziv
Nivo3-sort_id
Nivo3-šifra
Nivo3-naziv
Datum od
Datum do
.....
Datum formiranja sloga

Slika A.5. Dimenzija šifara podataka

Jedan slog u tabeli predstavlja jednu hijerarhijsku granu od najvišeg do atomičkog nivoa. Hijerarhije su posebno definisane kako za dnevne podatke kastodi banaka tako i za mesečne podatke društava za nadzor. Zato je u dimenziju uključeno polje identifikatora koje razdvaja te dve grupe podataka. Analizom je utvrđeno da je za svaku od šest grupa podataka, koliko je obuhvaćeno šifriranjem, dubina hijerarhije različita i da je maksimalna dubina do trećeg nivoa. Svaki nivo sadrži svoju šifru i pun propisan naziv. Kako je redosled pojavljivanja podataka iz šifarnika unapred definisan i bitan sa poslovnog aspekta, za svaki nivo u hijerarhiji uključeno je polje sort_id koje će poslužiti kao sekvenca za sortiranje podataka na prezentacionom sloju. Datumska polja označavaju period važenja konkretne hijerarhijske grane. Vremenska odrednica bila je neophodna jer se struktura i značenje ovih šifara menja prema potrebama sektora, a praćenje istorijata tih promena je postavljeno kao obavezan zadatak. Tabela ima primarni surogat ključ koji jedinstveno identifikuje hijerarhijsku granu i koji je spušten u tabele činjenica. Ova dimenzija, kao i dimenzija hartija od vrednosti, biće korišćena samo u prvoj fazi razvoja sistema dok se budu uvodile grupe podataka koje formiraju ukupnu i neto imovinu fonda (hartije od vrednosti, nepokretnosti, depoziti, gotovina, potraživanja i obaveze).

Tabela je inicijalno napunjena prema propisanoj metodologiji i dopunjena ranije važećim šiframa zbog historijata praćenja podataka. Ova dimenzija se retko (tj. „sporo“) menja i njene izmene direktno su vezane za izmene zvaničnog uputstva o razmeni podataka. Stoga se ona neće automatski ažurirati, već će je IT tim održavati shodno promeni regulative.

Nad podacima u dimenziji definisana je hijerarhija:

Vrsta podataka

Nivo1_naziv

Nivo2_naziv

Nivo3_naziv.

Hijerarhija je napravljena po nazivima, jer se u analizama i izveštavanjima šifra podatka retko koristi.

A.4.6. Ostale dimenzije

Dimenzije koje se ovde analiziraju mnogo su manje po obimu podataka i jednostavnije po strukturi. To su:

- Registar nepokretnosti. Dimenzija se ažurira na osnovu podataka dobijenih razmenom datoteka sa xml sadržajem sa društvima za nadzor. Podaci se odnose na nekretnine koje su u vlasništvu penzijskih fondova, kao i na nekretnine kojima se u toku dana trgovalo na tržištu. Izvorni podaci se analiziraju i upoređuju sa postojećim podacima u registru nekretnina. Matični podaci o svakoj novoj nekretnini upisuju se u registar. Registar nekretnina ima svoj surogat ključ koji je spušten u tabelu činjenica. Trenutno u evidenciji postoje samo dve nekretnine, tako da je ova dimenzija po obimu minorna. Bez obzira na obim uključena je u sistem od početka razvoja jer bi eventualno kasnije aktiviranje trgovanja nekretninama uzrokovalo potrebu za dodavanjem nove dimenzije u sistem. Kako bi to poremetilo definisanu granularnost podataka najverovatnije bi svi podaci o nekretninama morali da budu obrisani i ponovo napunjeni podacima veće detaljnosti. Nad ovom dimenzijom radi se isključivo upis novih podataka.

Prikaz atributa iz registra nekretnina dat je na slici A.6.

REGISTAR NEKRETNINA
Id_nekretnine (PK)
Adresa
Broj katastra
Šifra mesta
Naziv mesta
Šifra opštine
Naziv opštine
Površina
Cena
.....
Datum formiranja sloga

Slika A.6. Matični podaci o nekretninama

- Šifarnik valuta. Šifarnik valuta održava se na osnovu podataka iz transakcionog sistema u kome se evidentiraju nove valute, promene naziva postojećih valuta, kao i promene šifara postojećih valuta. Dimenzija sadrži: šifru valute, naziv valute i jedinstveni sirogat ključ. U transakcionom sistemu šifra valute je jedinstveni ključ tabele. Procedura za ažuriranje dimenzije identifikuje nove šifre valuta i upisuje ih u tabelu dimenzija. Procedura zatim evidentira promene naziva postojećih valuta i ažurira naziv u dimenziji valuta. Dakle, atribut „naziv valute“ uvek sadrži tekuću vrednost. Praćenje istorijata promene naziva nije važno sa aspekta poslovnog sistema, a i sam transakcioni sistem je koncipiran tako da se arhiva promena ne prati. Izveštavanje se gotovo uvek radi po troznačnoj šifri valute (EUR,USD,RSD..), retko po nazivu valute. Iz svih ovih razloga odustalo se od praćenja istorije promene naziva valute, jer bi implementacija bila teška a koristi za krajnjeg korisnika neznatne.

Šifarnik valuta je, kao dimenzija, uključen u model u prvoj fazi razvoja projekta da bi omogućio praćenje imovine fondova po valutnoj strukturi. Ovaj šifarnik ne služi za kursiranje podataka iz valute u valutu, već je prikazivanje podataka u više valuta razrešeno u samim tabelama činjenica.

- Šifarnik okruga. Ovaj šifarnik identifikovan je kao potencijalna dimenzija, iako se on koristi tek u poslednjoj fazi projektovanja sistema kada budemo uvodili složen podmodel individualnih računa članova fondova. Dimenzija je malog obima i ima onoliko slogova koliko ima okruga u Srbiji, svega deset. Sadrži šifru i pun naziv okruga. Napunjena je inicijalno i nije predviđeno njeno automatsko ažuriranje, jer se menja izuzetno retko, skoro nikada.
- Minidimenzija tipova podataka. Ova dimenzija nije deo poslovnog sistema već projektantsko rešenje problema uporednog praćenja zajedničkih poslovnih mera dobijenih iz dva različita izvora: od kastodi banaka i od društava za nadzor. Podaci kastodi banaka čuvaće se u

skladištu podataka kao „dnevni podaci“, a podaci društava za nadzor kao „mesečni podaci“. Jedan od najtežih zadataka u projektovanju sistema bio je da se identifikuju zajedničke poslovne mere koje se dostavljaju iz oba izvora i da se razreše konflikti u nazivima i značenju tih mera. Nakon toga, pronađene su razlike ova dva skupa podataka i na kraju napravljena njihova unija. Zahtev i poslovne potrebe korisnika uslovile su da se u sistemu čuvaju podaci iz oba izvora. Omogućeno je njihovo uporedno praćenje u cilju otkrivanja grešaka u podacima ili nepravilnosti u poslovanju. Ova dimenzija sadrži samo dva sloga: Dnevni podaci (id=1) i Mesečni podaci (id=2). Početna ideja bila je da se ovaj prosti identifikator spusti u svaku tabelu činjenica u kojoj se nalaze podaci iz oba izvora. Nedostatak tog rešenja bilo bi otežano kombinovanje podataka iz više tabela činjenica. Ovim rešenjem sve tabele činjenica u kojima se nalaze podaci iz oba izvora imaju spušten jedinstveni ključ iz iste dimenzione tabele i po tom ključu se povezuju. Dimenzija je inicijalno napunjena i ne ažurira se.

A.5. Opis tabela činjenica

Nakon identifikovanja nivoa granularnosti i osnovnih dimenzionih tabela sistema identifikovane su poslovne mere koje će biti grupisane u tabele činjenica. Poslovne mere identifikovane su prema poslovnim subjektima koje sistem obuhvata. Uvodjenje poslovnih mera razdvojeno je po fazama razvoja sistema za skladištenje.

A.5.1. I faza

U prvoj fazi razvoja sistem je obuhvatio pet velikih grupa podataka: hartije od vrednosti (dužničke hartije i akcije), nepokretnosti, depozite, gotovinu, potraživanja i obaveze. Za svaku od ovih grupa podataka prati se dnevno stanje u portfoliju fonda, a za neke od njih kao što su hartije od vrednosti, nekretnine i depoziti, prate se i dnevne transakcije trgovanja. Osnovna podela koja se nameće je na podatke o stanju na dan i podatke o transakcijama trgovanja na dan. Podaci o stanju i transakcijama se prate uglavnom nezavisno, ali je zahtev da se sistem koncipira tako da i uporedne analize budu moguće (npr. stanje u portfoliju hartija od vrednosti posmatrati uporedo sa dnevnim transakcijama). Takođe, podaci se, uglavnom, prate nezavisno po grupama i vrlo retko se ukrštaju medjusobno. Postoji potreba da se povremeno kombinuju podaci o depozitima i gotovini na nivou zajedničke dimenzionalnosti. Sistem je, dakle, koncipiran tako da su na prezentacionom nivou grupe podataka razdvojene kako bi se korisniku napravilo intuitivno okruženje lako za korišćenje. U isto vreme, fizička implementacija je takva da svi podaci koji su obuhvaćeni prvom fazom razvoja mogu da se kombinuju i posmatraju uporedo, tamo gde to ima smisla i na nivou koji predstavlja zajedničku dimenzionalnost.

Kada se grupe podataka posmatraju uporedo, to je uvek na nekom agregiranom nivou, ali kada se svaka grupa podataka analizira za sebe

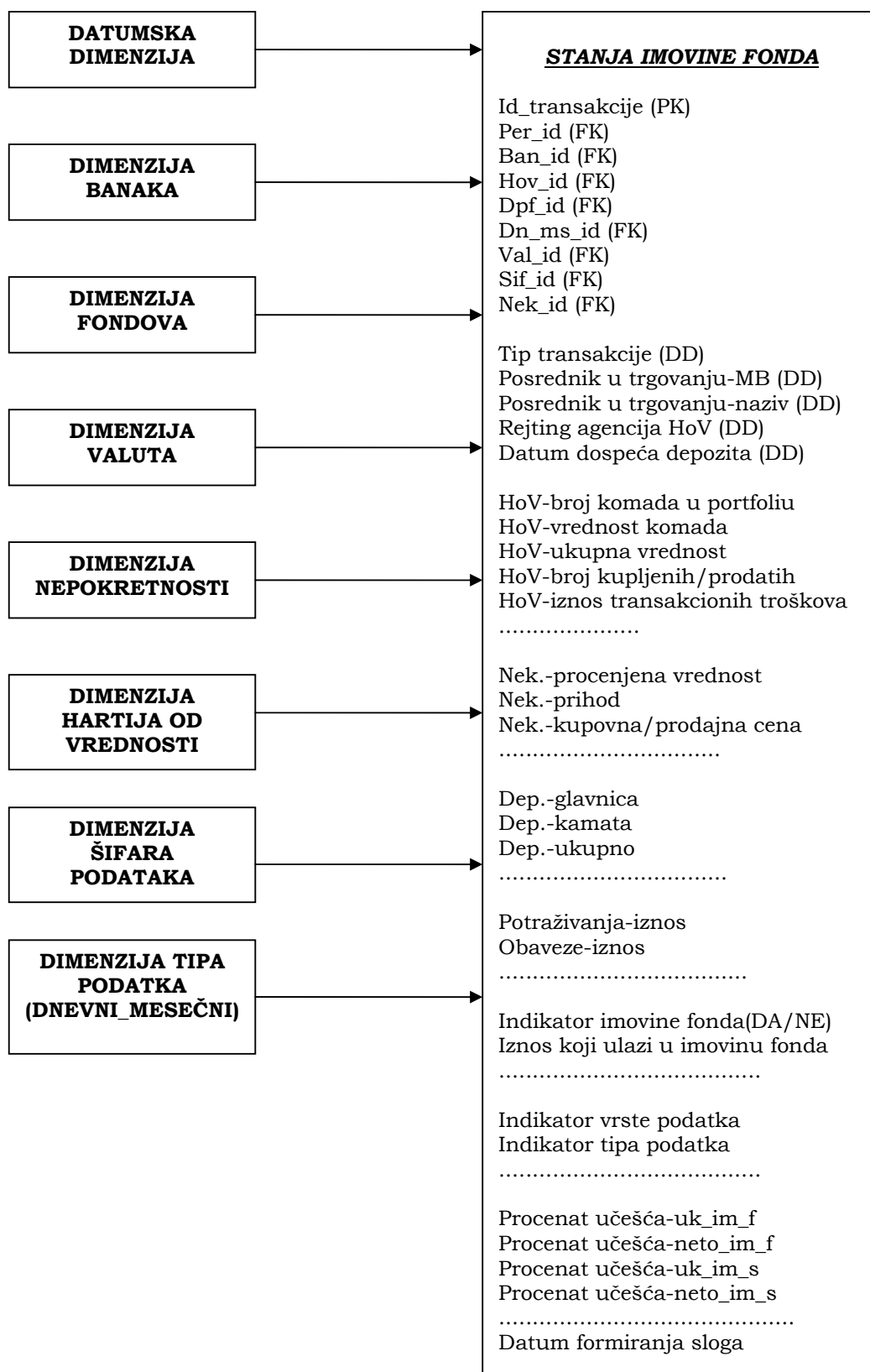
neophodno je da se obezbedi najdetaljnija moguća analiza po odgovarajućim dimenzijama. Zato se u tabeli činjenica nalaze podaci najveće detaljnosti i ona je povezana sa svim dimenzijama koje su relevantne za ove grupe podataka.

Identifikovane poslovne mere grupišemo prema poslovnim subjektima koje pratimo. Broj poslovnih mera koje se analiziraju po ovim kategorijama iznosi nekoliko desetina. Ovo su neki primeri poslovnih mera koje su uključene u sistem za svaku navedenu grupu podataka:

- Hartije od vrednosti – broj komada konkretne hartije u portfoliu fonda, vrednost hartije po komadu, ukupna vrednost svake hartije u portfoliu fonda (ovo je izvedena mera dobijena kao proizvod broja hartije i vrednosti hartije po komadu), broj prodatih/kupljenih komada konkretne hartije na tržištu na dan, ukupna vrednost prodatih/kupljenih hartija na dan, iznos transakcionih troškova, posrednik (broker) preko koga je trgovanje obavljeno i sl.
- Nepokretnosti – vrednost pojedinačne nekretnine, prihod ostvaren nad konkretnom nekretninom, kupovna/prodajna cena u trgovanju na dan i sl.
- Depoziti – iznos glavnice depozita, iznos kamate, ukupan iznos (izveden kao zbir glavnice i kamate), kamatna stopa, iznos povučenih depozita na dan, iznos položenih depozita na dan i sl.

Sve činjenice identifikovane u prvoj fazi razvoja biće implementirane kroz jednu osnovnu tabelu i povezane sa svim dimenzijama. Ideja je da se odvojene tabele činjenica povezane sa samo sebi svojstvenim dimenzijama logički formiraju na prezentacionom nivou, a da fizička implementacija bude zajednička. U tu svrhu zajednička tabela činjenica dopunjena je identifikatorima grupa podataka koji će poslužiti za filtriranje podataka pri logičkom razdvajanju podskupova činjenica. Tabele činjenica koje korisnik bude video biće konformni podskupovi zajedničke tabele činjenica.

Delimičan prikaz objedinjene tabele činjenica koja sadrži kompletne podatke o stanju imovine fonda i njegovim dnevnim aktivnostima na tržištu dat je na slici A.7.



Slika A.7. Implementacija tabele činjenica u prvoj fazi razvoja sistema za skladištenje

Ova tabela činjenica sadrži osnovne, atomske podatke ali i izvedene podatke dobijene primenom matematičkih operacija (kao što je ukupna vrednost hartija), ili primenom složenije ekonomske logike. Takvi su podaci o procentu učešća svake pojedinačne stavke u ukupnoj ili neto imovini fonda, kao i procenat učešća stavke u ukupnoj i neto imovini celog sistema fondova. Ovi izvedeni podaci su preračunati i čuvaju se u tabeli činjenica jer ih je teško dobiti pri samom izvršenju upita, a i performanse takvih upita drastično opadaju. Pomenuti procenti učešća su bitni pokazatelji stanja imovine fonda i često se koriste u analizama.

Indikator vrste podatka uveden je da razdvoji šest posmatranih grupa podataka i omogući logičku podelu tabele činjenica na podskupove. Indikator tipa podatka ima namenu da svaku takvu grupu podeli na podatke o stanjima u portfoliju i podatke o dnevnim transakcijama koji će, takodje, logički biti razdvojeni. Logičke celine su podskupovi kako kolona tako i vrsta zajedničke bazne tabele. Ovako izvedene konformne činjenice su povezane samo sa odgovarajućim dimenzijama iz celog skupa identifikovanih dimenzija (hartije od vrednosti se neće povezivati sa registrom nepokretnosti, depoziti i gotovina neće se povezivati ni sa registrom nepokretnosti ni sa šifarnikom hartija od vrednosti i sl.). Takodje, svaki podskup tabele činjenica povezuje se sa odgovarajućim podskupom skupa svih definisanih šifara. Skup svih šifara u dimenziji šifara podataka filtrira se pomoću polja „Vrsta_podatka“, koje jedinstveno određuje grupu podataka na koju se konkretna šifra odnosi.

Najvažnije poslovne mere koje treba da se dobiju ovom implementacijom su iznosi ukupne i neto imovine fonda. U tabeli činjenica identifikator imovine fonda omogućava filtriranje onih podataka koji ulaze u obračun te imovine (npr. podaci o transakcijama trgovanja ne ulaze u obračun imovine fonda već samo podaci o stanjima u portfoliju fonda). Iznos koji ulazi u obračun imovine fonda izračunava se u ETL procedurama pre punjenja tabele činjenica i to iz atomičkih podataka. Tako izračunat čuva se na najdetaljnijem nivou, tj. na nivou sloga u tabeli činjenica. Postignuto je da se podaci o imovini fonda logički razdvoje od ostalih podataka iz tabele činjenica, a da ostanu povezani sa istim skupom dimenzija. Podskup tabele činjenica koji će omogućiti uvid samo u imovinu fonda biće povezan sa svim prikazanim dimenzijama kako bi se omogućile detaljne analize imovine po svim aspektima praćenja. Ta tabela činjenica sadržaće podskupove kako vrsta tako i kolona bazne tabele.

Na slici A.7. se vidi da su u tabelu činjenica uključene i neke veštačke „degenerativne“ (u oznaci DD) dimenzije. To su atributi kao što su: broj transakcije, posrednik u trgovanju, rejting agencija koja je izvršila kategorizaciju hartije od vrednosti i sl. Ovo nisu poslovne mere i uvode se umesto klasičnih dimenzija. Cilj njihovog uvođenja je da se podaci svedu na najdetaljniji nivo (kao broj transakcije) ili da se detaljnije opišu podaci u slogu (kao što su matični broj i naziv posrednika u trgovanju hartijom od vrednosti). Ovo su podaci po kojima se vrše analize poslovnih mera, analogno podacima u dimenzijama.

Zahtev korisnika je bio da se svaki od podataka u sistemu, za koji to ima smisla i koji se poslovno tako prati, prikaže u originalnoj valuti u kojoj je

podatak dostavljen iz izvornog okruženja, ali i u dinarima i evrima. Projektantski izbor je da se, zbog lakšeg snalaženja u ionako velikoj količini podataka, svaka poslovna mera u ovoj tabeli činjenica, koju ima smisla pratiti po valutama, čuva tri puta, i to: u originalnoj valuti, kursirana u evre na dan podataka i kursirana u dinare na dan podataka.

Podatak o stanju hartija u portfoliju fonda tako je predstavljen u tri valute kroz posebne poslovne mere u tabeli činjenica:

HoV_stanje_u_portfoliju_VAL,
HoV_stanje_u_portfoliju_EUR,
HoV_stanje_u_portfoliju_RSD itd.

Većina mera u ovoj tabeli činjenica je aditivna, dakle može se agregirati duž dimenzija i tako agregirana posmatrati na manje detaljnom nivou. Neke poslovne mere kao što je broj hartija i kamatna stopa nisu sasvim, ili nisu uopšte aditivne. Broj hartija se može agregirati po predviđenim dimenzijama, ali samo u kontekstu jedne konkretne hartije. Tako je sa aspekta dimenzije hartija od vrednosti to neaditivna mera, a sa aspekta ostalih dimenzija aditivna. Dakle, broj hartija je semiaditivna mera. Proizvod broja hartija i vrednosti jednog komada konkretne hartije daje ukupnu vrednost hartije i to je mera koja je aditivna sa aspekta svih dimenzija modela. Kamatna stopa kod depozita ili gotovine je potpuno neaditivna mera jer se ne agregira ni po jednoj dimenziji. To je mera koja se posmatra samo na najnižem nivou detaljnosti. Da bi se omogućio uvid u vrednost kamatne stope na agregiranom nivou obračunava se prosečna kamatna stopa prema prihvaćenoj metodi ponderisanja. U tu svrhu primenjuje se matematička formula nad atomičkim podacima, i tako izvodi mera koja se kasnije ponaša kao aditivna u sistemu (tj. iznos kamatne stope pomnožen iznosom glavnice je aditivna mera koja, na bilo kom agregatnom nivou, podeljena ukupnom glavnicom istog nivoa predstavlja traženu prosečnu kamatnu stopu).

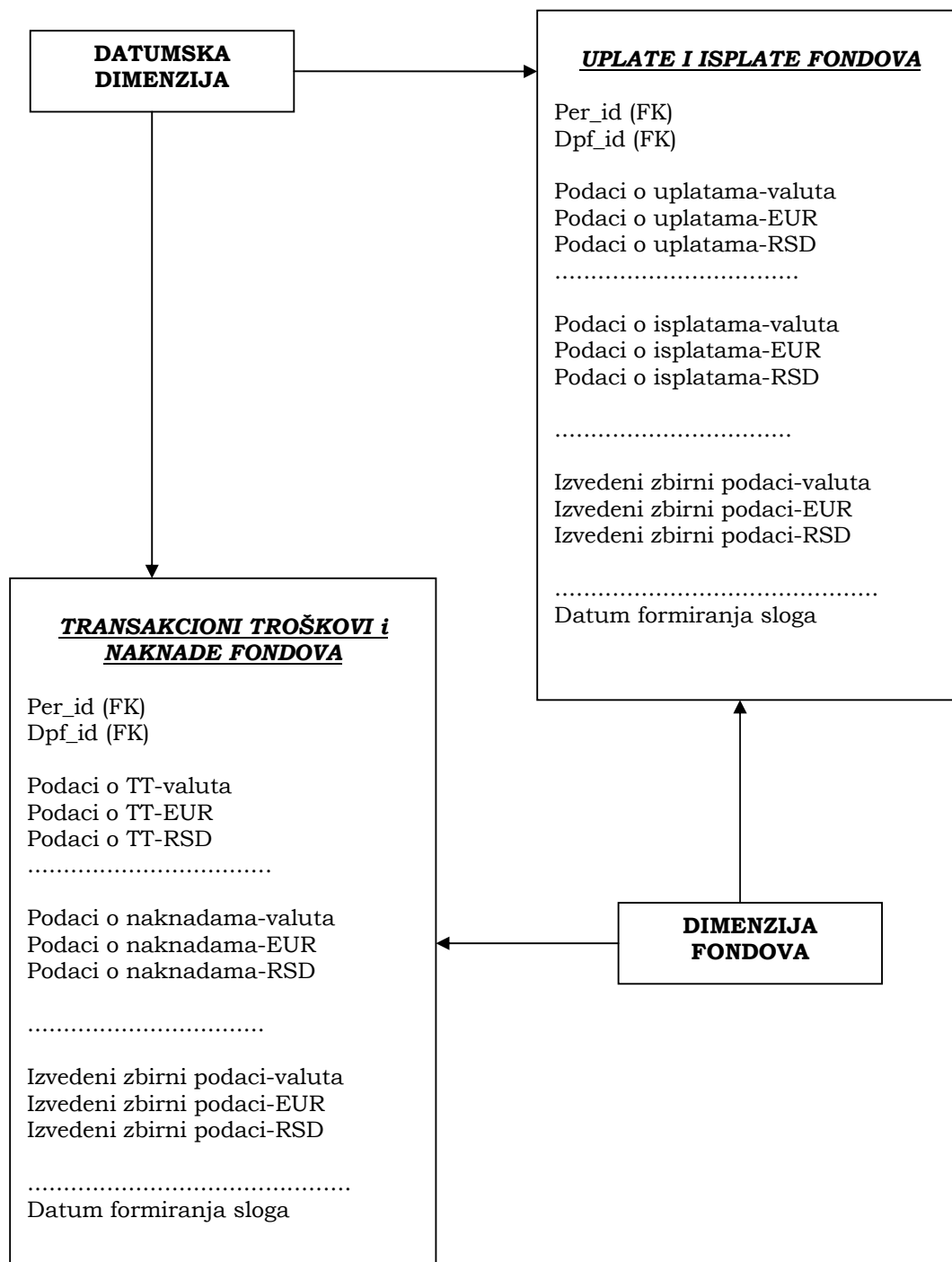
A.5.2. II faza

U sledećoj fazi razvoja sistema u skladište podataka su uključeni podaci o uplatama, isplatama, naknadama i transakcionim troškovima fondova. Ovo su podaci koji se najdetaljnije čuvaju na nivou datuma i penzijskog fonda, tako da su implementirane samo nove tabele činjenica. Potrebne dimenzije su uvedene u prvoj fazi.

Poslovne mere su grupisane u dve tabele činjenica i to: podaci o uplatama i isplatama u jednoj, a podaci o transakcionim troškovima i naknadama u drugoj tabeli. Zajedno su grupisani podaci koji se, najčešće, uporedno prate. U praksi se pokazalo da se vrlo retko rade međusobna kombinovanja podataka iz dve tabele činjenica. Projektantska odluka je bila da se oni fizički implementiraju kroz dve razdvojene tabele činjenica umesto kroz jednu objedinjenu tabelu. U slučajevima kada se bude zahtevalo kombinovanje činjenica iz ove dve tabele, u samom upitu će se izvršiti spoljno spajanje po zajedničkim dimenzijama datuma i fonda. Sve mere u tabelama činjenica su aditivne sa aspekta obe dimenzije. Sve mere su prikazane u originalnoj valuti, dinarima i evrima. Podaci o uplatama, isplatama, transakcionim

troškovima i naknadama dobijaju se razmenom podataka sa društvima za nadzor penzijskih fondova.

Uprošćen prikaz podmodela razvijenog u drugoj fazi dat je na slici A.8.

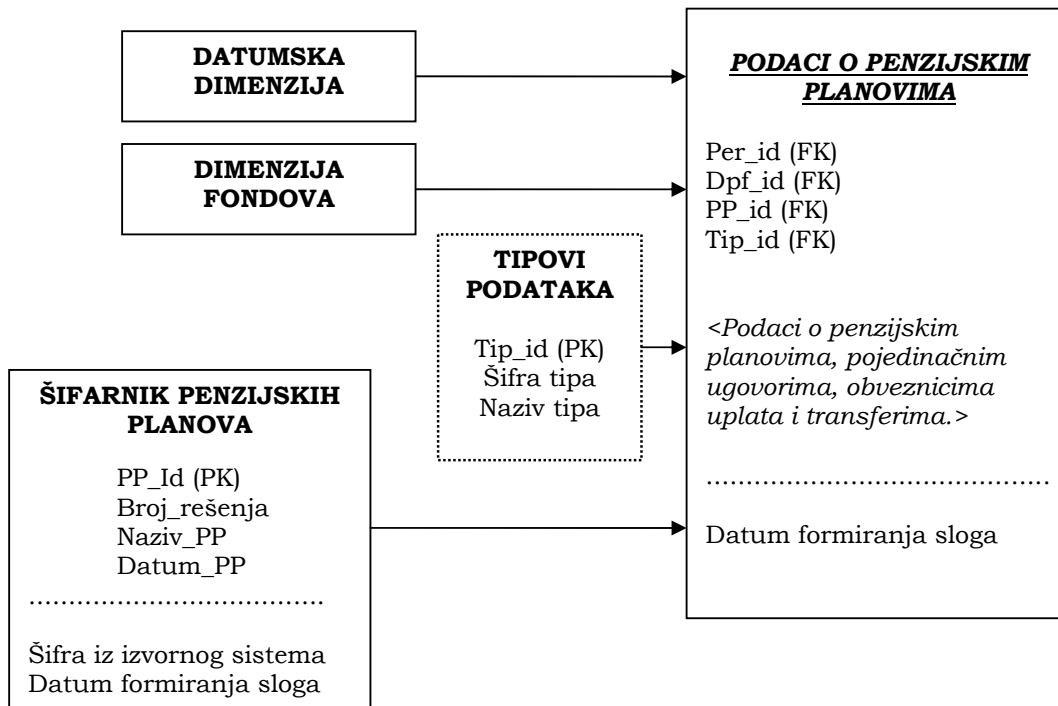


Slika A.8. Implementacija tabela činjenica u drugoj fazi razvoja sistema za skladištenje

A.5.3. III faza

U trećoj fazi razvoja skladišta u sistem su uvedene dve nove dimenzije i jedna tabela činjenica. Proširljivost zvezdastog modela dobijenog u prve dve faze razvoja sistema omogućila je da se skup dimenzija dopuni bez narušavanja postojećeg modela. U ovoj fazi u sistem su uključeni podaci o penzijskim planovima, obveznicima uplata, pojedinačnim ugovorima i transferima. Ove kategorije se, uglavnom, posmatraju nezavisno od ostalih podataka. Podaci se razmenjuju sa društvima za nadzor jednom mesečno i predstavljaju stanje fondova po ovim poslovnim kategorijama na poslednji radni dan u prethodnom mesecu. Datumska dimenzija i dimenzija penzijskih fondova su zajedničke na nivou celog skladišta. Nova dimenzija koja je uvedena u ovoj fazi je šifarnik penzijskih planova koji se osvežava pre svakog osvežavanja tabele činjenica. Šifarnik penzijskih planova sadrži matična obeležja jednog penzijskog plana, kao što su: broj rešenja, naziv i datum plana. Dimenzija se isključivo puni novim podacima pri svakom osvežavanju. Postojeći podaci se ne brišu i ne menjaju. Tabela sadrži surogat primarni ključ, ali je zadržana i šifra plana koja predstavlja prirodni ključ u izvornom sistemu. Šifra je veza sistema za skladištenje i izvornog transakcionog sistema u koji se podaci smeštaju po prijemu datoteka sa xml sadržajem. Na osnovu te šifre pri svakom ažuriranju dimenzije mogu da se identifikuju novi slogovi (nove šifre plana) u izvornoj tabeli. Samo ti slogovi se zahvataju ETL procedurom kojom se puni ova dimenziju. Druga dimenzija koju je uvedena u ovoj fazi je minidimenzija koja sadrži samo identifikator vrste podatka. Identifikator određuje da li se upisani slog odnosi na penzijski plan, obveznika uplate, pojedinačni ugovor ili stavku transfera. Na ovaj način se grupišu podaci o penzijskim planovima po metodologiji sektora i društva za nadzor poštuju ovu klasifikaciju pri dostavljanju podataka.

Uprošćen prikaz podmodela razvijenog u ovoj fazi dat je na slici A.9.



Slika A.9. Dimenzije i tabela činjenica uvedeni u trećoj fazi razvoja sistema za skladištenje

Šifarnik tipova podataka sadrži četiri sloga i menja se samo sa promenom uputstva o razmeni podataka. Zato je ovaj šifarnik inicijalno napunjen i ne ažurira se automatski. Pored jedinstvenog surogat ključa, uvedena su polja šifra i naziv tipa. Šifra će poslužiti kao sekvenca za sortiranje podataka na prezentacionom nivou. Ova dimenzija je posebno označena na slici jer je, na neki način, njeno postojanje u sistemu neobavezno. Isti efekat postigao bi se i da su umesto nje uvedene dve DD dimenzije u tabeli činjenica. DD dimenzije (šifra i naziv tipa) bi se punile pri upisu sloga u tabelu činjenica. Ako bi se u toku života sistema za skladištenje prešlo sa tabele dimenzije na DD dimenzije izmene u samoj tabeli činjenica bi bile minimalne. Za istorijske podatke polja bi se popunila na osnovu jedinstvenog identifikatora koji je kao strani ključ spušten iz dimenzije u tabelu činjenica. ETL procedura za punjenje tabele činjenica bi, takodje, pretrpela manje izmene. Dakle, ovakva promena bi mogla da se sprovede relativno brzo i sistem ne bi bio narušen.

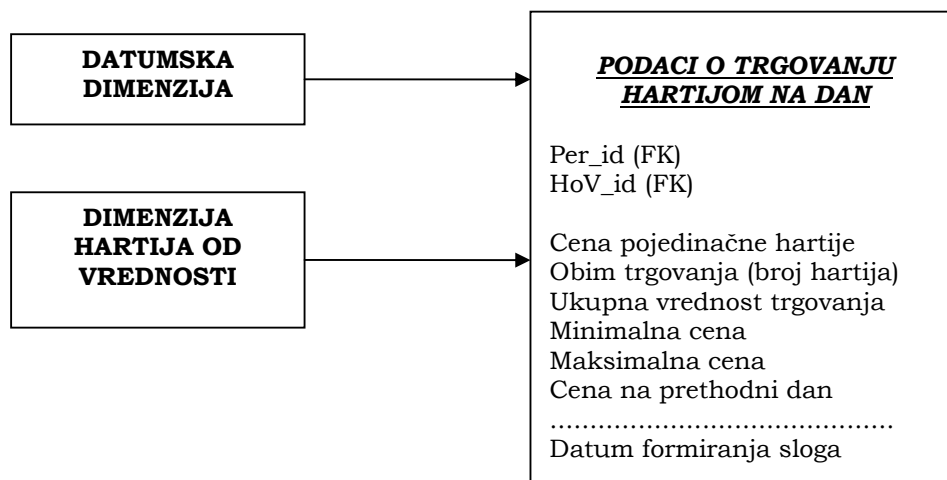
A.5.4. IV faza

U ovoj fazi projekta analizirani su i uključeni u skladište podaci iz trećeg izvora podataka - Beogradske berze. Podaci se, prema definisanom protokolu, preuzimaju sa Veb strane berze jednom dnevno i to nakon završetka rada berze. Podaci su na nivou dana i postoje za svaki radni dan berze. Ovo su mere koje se ne koriguju, već se tabele činjenica isključivo pune novim dnevnim podacima.

Uvodjenje ove grupe podataka nije zahtevalo proširivanje modela novim dimenzijama. Obe tabele činjenica, kako je u ovoj fazi implementirano, oslanjaju se na postojeću datumsku dimenziju i postojeći šifarnik hartija od

vrednosti. Prva tabela činjenica puni se podacima o dnevnom trgovanju hartija od vrednosti na berzi, pa je granularnost ovih podataka na nivou datuma i hartije od vrednosti.

Uprošćen prikaz podmodela razvijenog u ovoj fazi dat je na slici A.10.



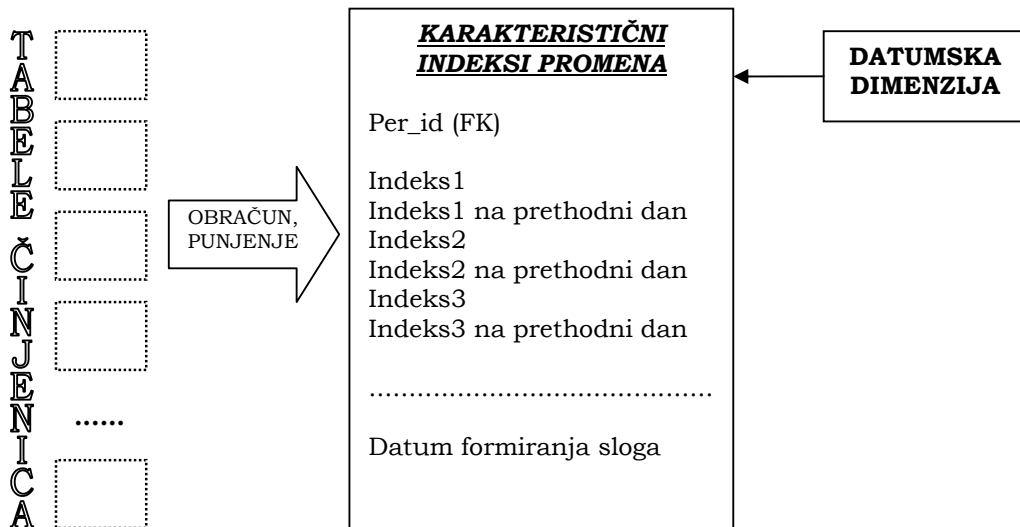
Slika A.10. Tabela činjenica sa podacima o dnevnom trgovanju hartijama na berzi

Mere koje su prikazane na slici A.10. uglavnom su neaditivne sa aspekta vremenske dimenzije i dimenzije hartija od vrednosti. Cena pojedinačne hartije, minimalna cena trgovanja, maksimalna cena trgovanja i broj komada jedne hartije su podaci koji se uvek analiziraju u kontekstu konkretne hartije od vrednosti i isključivo na dan. Ove mere se ne agregiraju ni po jednoj od dve dimenzije. Ukupna vrednost trgovanja je jedina aditivna mera u ovoj tabeli činjenica. Podaci iz ove tabele činjenica često se uporedno posmatraju sa podacima o trgovanju hartijama od vrednosti koji potiču iz drugih izvora i uključeni su u sistem u prvoj fazi. Uporedna analiza je omogućena, ali samo na nivou zajedničke dimenzionalnosti, tj. posmatrano na dan i u odnosu na konkretnu hartiju.

Polje „Cena na prethodni dan“ je izvedeno izračunavanjem u ETL proceduri pri punjenju tabele činjenica. Standardni korisnički zahtev je bio da se računa promena cene kao razlika tekuće i prethodne vrednosti za konkretnu hartiju. Zbog što jednostavnijeg korišćenja korisničkih alata bez složenih kalkulisanja, izbor je bio da se u svakom slogu uz tekuću čuva i prethodna vrednost ove poslovne mere.

Na sličan način prate se promene vrednosti atributa i u drugoj tabeli činjenica. U ovoj tabeli činjenica čuvaju se karakteristični izvedeni indeksi, tj. pokazatelji dnevnog stanja. To su podaci koji su preračunati na dnevnom nivou i zahtevaju samo datumsku dimenziju.

Uprošćen prikaz tabele indeksa dat je na slici A.11.



Slika A.11. Implementacija tabele činjenica sa izvedenim indeksima obračunatim na osnovu podataka o neto imovini fondova i investicionim jedinicama

Indeks1, koji je i najinteresantniji za korisnike, je izveden složenom ekonomskom logikom iz podataka koji već postoje u sistemu, ali su sačuvani na detaljnijem nivou, tj. na nivou datuma i fonda. Primenom formule u ETL proceduri izvodi se ovaj dnevni pokazatelj. Dakle, da bi bio moguć obračun indeksa i punjenje tabele činjenica koja je prikazana na slici A.11. preduslov je da budu napunjene tabela činjenica u kojima se nalaze podaci koji ulaze u obračun, a koje su opisane u prvoj fazi. Obračun indeksa isključivo konsultuje prethodno napunjene tabele u skladištu a ne izvorne podatke. Procena je bila da bi indeks definisan na ovaj način bilo nemoguće izvesti korišćenjem korisničkih alata. Čak i da se uspe u tome performanse takvog upita bi, zbog obima i složenosti operacija, bile značajno narušene. Zato se ovi indeksi čuvaju preračunati kako na tekući tako i na prethodni dan. Ono što je bitno za ovu tabelu činjenica je da se ona uvek puni posle tabela u kojima se nalaze mere koje se koriste u obračunu. Takodje, izmene podataka u tabeli indeksa moraju da budu sinhronizovane sa izmenama u tabelama činjenica koje „čita“ ETL procedura kojom se indeksi izvode. U ETL proceduri se ispituje postojanje promene u osnovnim podacima. U slučaju promena brišu se slogovi u tabeli indeksa za konkretan datum, zatim se ponavlja obračun i punjenje podataka.

A.5.5. V faza

Najsloženiji zahtev koji se postavljao pred sistemom za skladištenje podataka bila je implementacija podmodela članova fondova i njihovih računa. Ova faza je najduže trajala i zahtevala je posebno veliko angažovanje zaposlenih u sektoru koji rade na ovim poslovima.

Složenost implementacije ovog podmodela uslovlili su sledeći razlozi:

- Velika količina ulaznih podataka koje treba obraditi, obzirom da na dnevnom nivou ulazni podaci sadrže oko 200000 slogova. Masovnost podataka proizilazi iz činjenice da se podaci dostavljaju za svako pojedinačno fizičko lice (preko jedinstvenog matičnog broja) o kome postoji evidencija u nekom od penzijskih fondova u zemlji. Poznato je da su u svim poslovnim sistemima (bankarskim, telekomunikacionim i sl.) najobimniji upravo podaci o klijentima, fizičkim licima.
- Poverljivost podataka koji mogu da budu publikovani samo u propisanoj formi. Prilikom implementacije moralo je da se vodi računa da podaci koji su poverljivi ostanu sakriveni kako u pripremljenoj zoni, tako i u zoni samog skladišta, a da opet budu iskoristljivi za formiranje poslovnih mera. Princip je bio da se u samoj bazi podataka, kao i na prezentacionom nivou poverljivi podaci ne prikazuju u čitljivom obliku.
- Složenost i raznovrsnost analiza koje se nad ovim podacima rade. Uvidom u preglede, izveštaje i ad-hok analize koje se traže sagledani su svi aspekti praćenja poslovnih mera i identifikovane dimenzije koje će biti uključene u sistem.
- Veliki broj poslovnih mera koje se prate. Tabele činjenica koje su u ovoj fazi implementirane broje preko 60 poslovnih mera. Pored velikog broja mera, složenost je uslovila i neaditivnost većine mera. Stoga je posebno tretirana implementacija tabele činjenica po pojedinačnim fondovima, a posebno za ceo sistem fondova. U oba slučaja poslovne mere su morale da se izvode iz osnovnih podataka primenom propisane metodologije.
- Složena predefinisana metodologija dobijanja svake poslovne mere iz izvornih podataka.
- Postojanje podataka u internim evidencijama sektora koji su pre elektronske razmene dolazili putem elektronske pošte i koje je trebalo preformatirati, prečistiti i inicijalno napuniti u bazu skladišta. Podaci su preuzeti iz Excel tabele, usaglašeni po sadržaju, formatu i šifarski prilagodjeni. Tek tako prečišćeni podaci bili su uporedivi sa podacima koji su od uvođenja skladišta podataka počeli da dolaze elektronskim putem u propisanoj formi i sa ugrađenim šifarskim kontrolama.

Podaci o računima članova fondova su na mesečnom nivou i dostavljaju ih društva za nadzor fondova. Ovo su podaci o računima članova sa stanjem na poslednji kalendarski dan u mesecu. Struktura izvornih podataka je vrlo jednostavna. Ulazni slog sadrži: datum važenja podataka, broj fonda, jedinstveni broj fizičkog lica (u nečitljivom obliku), broj računa, poštanski broj, stanje na računu, bruto iznos uplate i datum poslednje uplate na račun. Posebnu važnost u ovoj fazi je imala pripremljena zona. Pošto su konačne tabele činjenica zahtevale složene proračune bilo je neprihvatljivo da ETL procedure za punjenje skladišta rade nad izvornom transakcionom bazom. Ove ETL procedure su veoma zahtevne, dugo traju i zauzimaju značajne resurse. Prihvatljivije rešenje je bilo da se u pripremljenu zonu što efikasnije prebace izvorni podaci i već tu sprovedu osnovne transformacije i

priprema za ciljno punjenje skladišta. Nadalje ETL procedure za punjenje skladišta rade isključivo nad tabelom u pripremljenoj zoni. Kako su u ovoj oblasti standardne analize koje se rade prema polu, starosnoj dobi i okruzima to su procedurama koje pune pripremljenu zonu iz matičnog broja izvedeni podaci o polu, datumu rođenja, godinama starosti i datumu kada će lice steći uslov za povlačenje uloženi sredstava (napunjene 52 godine). Iz dobijenog poštanskog broja izveden je podatak o šifri okruga (prema zvaničnom šifriranju okruga u zemlji). Šifrirani podatak o jedinstvenom matičnom broju lica je zadržan jer je jedino na osnovu njega bilo moguće izbrojati članove po raznim kriterijumima, što će biti neophodno pri punjenju skladišta. Nikakva izvodjenja složenih poslovnih mera u pripremljenoj zoni nisu radjena, već uglavnom priprema dimenzionih podataka

Izgled tabele u pripremljenoj zoni dat je na slici A.12.

FAZA V – PRIPREMNA ZONA
Datum važenja podataka
Broj fonda iz transakcionog sistema
Jedinstveni broj lica (šifriran)
Pol
Datum rođenja
Godine starosti
Datum sticanja uslova (52 godine)
Domaće/Strano lice identifikator
Broj računa
Poštanski broj
Šifra okruga
Iznos uplata
Stanje na račun
.....
Datum formiranja sloga

Slika A.12. Računi i članovi fondova – izgled tabele u pripremljenoj zoni

U ovoj fazi izgradnje skladišta podataka u sistem je uključena prethodno pomenuta dimenzija okruga, formirana na osnovu zvaničnog šifarnika okruga u zemlji. Pored ove dimenzije i standardnih dimenzija datuma i fondova pojavila se potreba za formiranjem još jedne dimenzije koja bi obuhvatila podatke o polu, godinama i vrsti fizičkog lica. Naime, poslovne mere se nikada ne prate u kontekstu pojedinačnog lica, već isključivo kumulativno prema polu, godinama života, vrsti lica i vremenu sticanja uslova za povlačenje uplaćenih sredstava. Prema pravilima modeliranja velikih dimenzija (posebno u slučajevima sa fizičkim licima) u posebnu minidimenziju su izdvojene sve moguće kombinacije polova i godina koje su moguće za svaku vrstu lica (fizičko i strano lice). Dogovoreno je da se godine grupišu u intervale koji su unapred definisani, a da se umesto datuma sticanja uslova za povlačenje sredstava prati mesec i godina sticanja uslova. Tako je broj mogućih kombinacija smanjen i značajno ubrzano pretraživanje po ovim obeležjima. Inače, obeležja izdvojena u ovoj dimenziji su ona po kojima se najčešće rade pretraživanja podataka o računima članova.

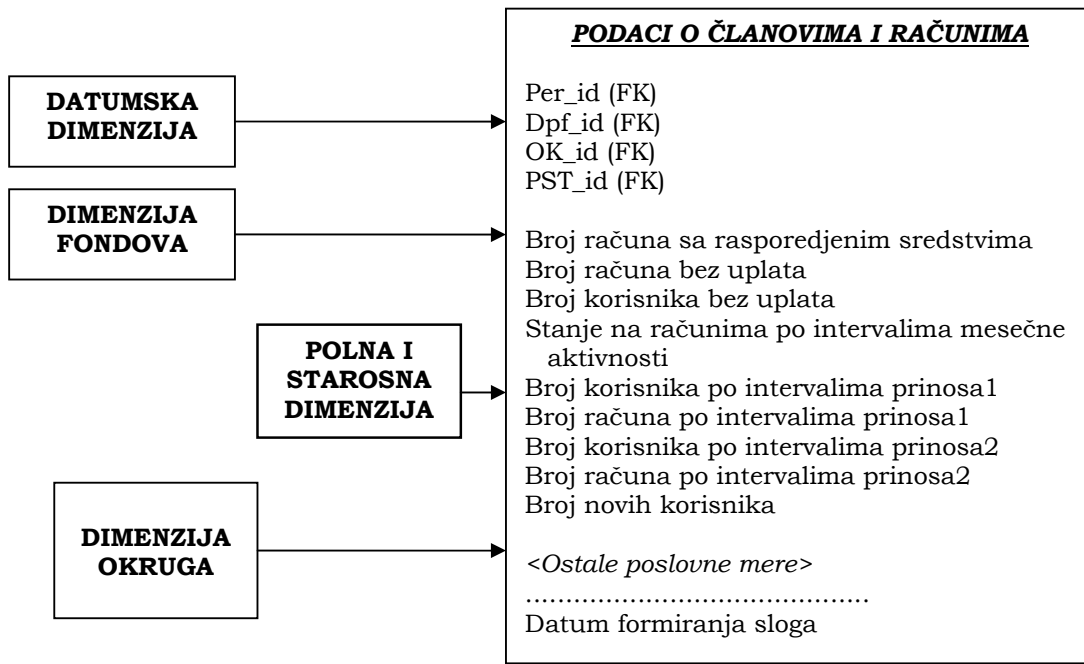
Dobijena dimenzija je prikazana na slici A.13.

POLNA I STAROSNA MINIDIMENZIJA
Jedinstveni ID sloga (PK)
Domaće/Strano lice identifikator
Pol (u oznaci „Muški“, „Ženski“)
Starosni intervali (npr. 0-18, 19-35..)
Mesec/Godina sticanja uslova za isplatu

Slika A.13. Minidimenzija sa svim mogućim kombinacijama polne i starosne strukture članova

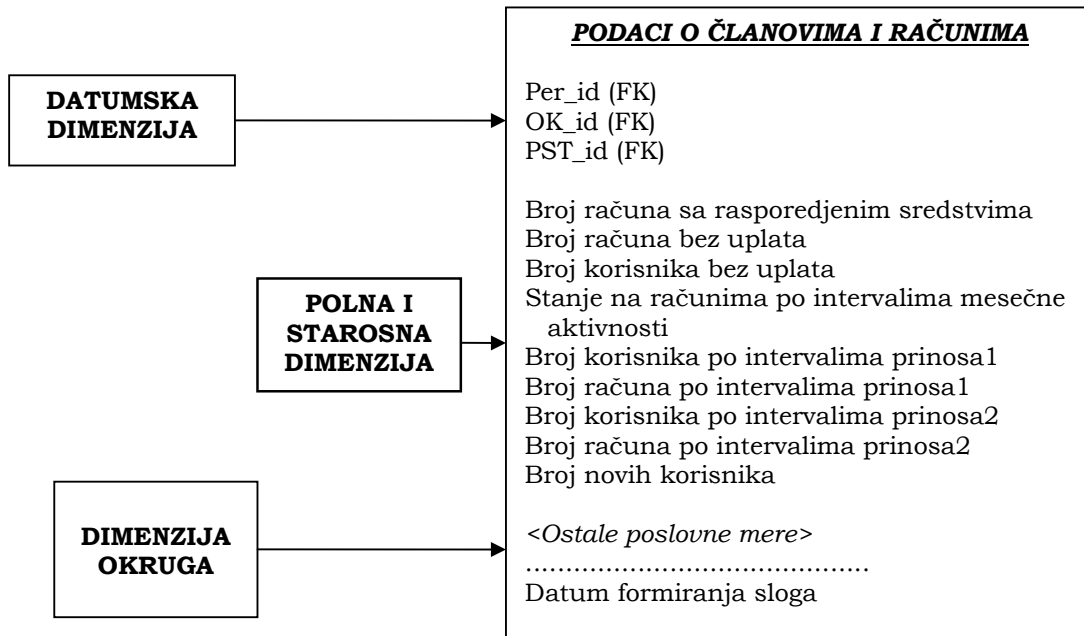
Tabele činjenica implementirane u ovoj fazi ne sadrže nijedan izvorni podatak. Svaka poslovna mera izvodi se u ETL procedurama složenom ekonomskom logikom. Korisnici ovde nemaju pristup osnovnim, detaljnim podacima već samo izvedenim, pa je zato je bilo neophodno da se sagledaju sve poslovne mere koje će eventualno biti potrebne u analizama. Pomenuto je da je u ovom slučaju većina mera neaditivna sa aspekta penzijskih fondova, odnosno podaci za ceo sistem fondova u zemlji se ne mogu dobiti agregiranjem podataka po fondovima. Zato su implementirane dve tabele činjenica sa istim opisom poslovnih mera, i to: jedna po fondovima i druga za ceo sistem. Svaka mera u tabeli činjenica za sistem izvedena je iz osnovnih podataka primenom iste logike, samo na višem nivou grupisanja. Osnovni razlog neagregatnosti mera je „preklapanje“ korisnika po fondovima. Tj. kako su ista fizička lica članovi više fondova istovremeno, to bi sabiranje podataka o aktivnosti korisnika davalo netačne rezultate, uvećane upravo za iznos „preklapanja“.

Uprošćeni izgled tabele činjenica na nivou fondova dat je na slici A.14.



Slika A.14. Računi članova fondova - tabela činjenica po fondovima

Skoro identična ovoj tabeli činjenica je i tabela činjenica prikazana na slici A.15. sa merama izračunatim na nivou sistema fondova.



Slika A.15. Računi članova fondova - tabela činjenica za sistem fondova

Informacije koje su deo standardnih analiza a nisu obuhvaćene ovim tabelama činjenica odnose se na tzv. distribuciju računa po korisnicima, kao i distribuciju korisnika po fondovima. Distribucija korisnika po fondovima

dobija se kao odgovor na pitanje: „Koliko je korisnika u sistemu član dva tri, četiri ili više fondova istovremeno?“ Distribucije računa po korisnicima dobijaju se kao odgovori na pitanja: „Koliko je korisnika u fondu koji poseduju dva, tri, četiri ili više računa?“ i „Koliko je korisnika u celom sistemu fondova koji poseduju dva, tri, četiri ili više računa?“ Sve tri grupe podataka su pokušane da se dobiju preko standardnih pogleda (dobijenih CREATE VIEW naredbom) u kojima se najpre za svakog korisnika izračunava broj računa, odnosno broj fondova u kojima ima račune. Zatim se, grupisano po tako dobijenim brojevima računa i fondova, broje korisnici koji su članovi u toliko fondova, odnosno korisnici koji poseduju toliko računa. Pogledi se implementiraju kroz jedan bazni upit, ali se to rešenje pokazalo kao sporo nad postojećom količinom podataka. Zato su za ove tri grupe podataka napravljeni materijalizovani pogledi. Tri materijalizovana pogleda su dobijena korišćenjem slične SQL naredbe kao i standardni pogledi, ali su fizički implementirani kao tabele, tako da je izvršavanje upita nad materijalizovanim pogledima takoreći trenutno. Naravno, pošto se i podaci o distribucijama posmatraju u kontekstu pola, okruga, starosnih intervala i vrste fizičkog lica ovi inače dimenzioni podaci su spuštteni u same poglede. Pogledi se osvežavaju istom dinamikom kao i tabele činjenica, dakle u noćnom režimu nakon izvršavanja svih ETL procedura.

A.6. Primeri izveštavanja

Na narednim slikama dati su primeri izveštaja koji su radjeni nad podacima o hartijama od vrednosti, imovini fondova i računima članova fondova. Prikazani su pregledi koji su radjeni samostalno od strane korisnika analitičara, a za potrebe šireg kruga zaposlenih u sektoru. Na prikazanim izveštajima primenjene su tehnike bušenja, filtriranja, grupisanja, sortiranja podataka, kao i definisanje i formatiranje izuzetaka i odstupanja od traženih vrednosti. Korišćeni su tabelarni i grafoidni prikazi podataka.

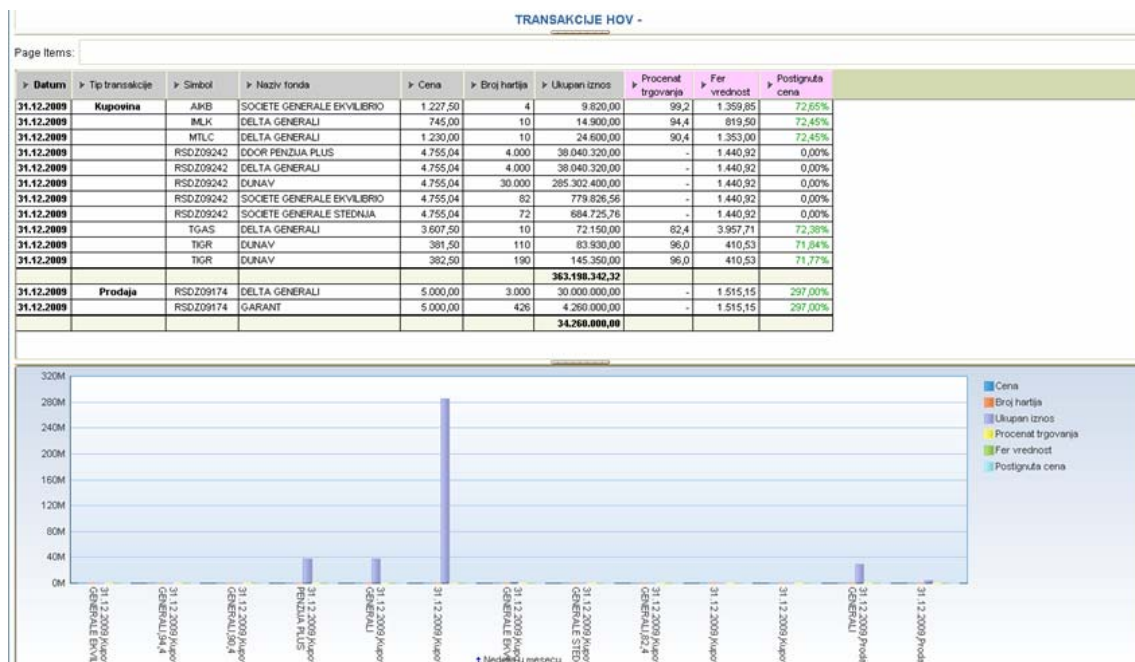
Napominjem da su u sva tri primera prikazane strukture i formati pregleda, dok su podaci sakriveni u skladu sa politikom poverljivosti informacija i zabrane njihovog publikovanja.

DOMAĆI/STRANI - Vrsta lica : Domaće lica

Page Items: Vrsta lica: Domaće lica

	Strano lica <All>	U fazi akumulacije						U fazi poslačenja					
		Procentualno učešće	Broj novih korisnika	Procentualno učešće	Broj korisnika sa sredstvima	Udeo sa sredstvima	Iznos sredstava	Procentualno učešće	Prosečan iznos sredstava (za sve korisnike)	Prosečan iznos sredstava (za ove 80 imaju)	Ukupan broj korisnika	Procentualno učešće	Broj novih korisnika
2009	4.218.597,14		1.510.740,00		3.486.827,21	85,76%	157.686.172.709,20		55.099,32	46.990,71	179,26		
Januar	337.039,86		71.280,00		278.575,97	87,71%	10.722.215.117,28		46.894,72	39.112,21	17,75		
DOOR PENZIJA PLUS	50.933,58	100,00%	5.775,00	99,06%	42.098,50	82,63%	1.999.275.629,90	100,00%	57.861,43	51.221,64	-	-	-
DELTA GENERALI	77.118,31	100,00%	6.600,00	98,36%	63.741,15	97,18%	2.496.954.918,91	99,78%	47.728,07	35.927,66	-	-	-
DUNAV	145.461,64	99,99%	26.510,00	98,37%	120.229,45	95,17%	4.606.537.095,85	99,99%	46.681,71	35.880,95	12,42	100,00%	-
GARANT	26.754,46	99,99%	17.545,00	99,69%	22.113,56	42,49%	477.912.303,48	100,00%	26.331,32	45.331,11	-	-	-
HYPO	2.016,26	99,91%	990,00	94,74%	1.666,51	97,98%	49.398.701,95	99,99%	36.115,21	26.964,21	-	-	-
NLB NOVA PENZIJA	7.172,27	99,80%	1.320,00	75,00%	5.928,15	88,84%	175.958.111,58	99,61%	36.163,74	29.843,58	-	-	-
RAIFFESEN FUTURE	19.058,61	99,96%	4.345,00	96,34%	15.752,65	97,17%	860.368.427,40	99,93%	66.544,78	50.095,75	1,77	100,00%	-
SOCIETE GENERALE BKVL	223,63	99,21%	2.915,00	96,36%	184,84	73,81%	1.098.473,66	100,00%	7.240,56	7.175,86	-	-	-
SOCIETE GENERALE STECK	331,90	100,00%	4.895,00	98,89%	274,33	62,03%	821.115,92	100,00%	3.646,83	4.300,45	-	-	-
TRIGLAV	7.969,19	100,00%	385,00	100,00%	6.586,83	19,42%	53.890.138,62	100,00%	9.968,18	37.545,69	3,55	100,00%	-
Februar	337.503,11		71.940,00		278.958,85	87,76%	10.983.695.394,22		48.016,07	40.020,33	17,75		
DOOR PENZIJA PLUS	50.917,61	100,00%	15.180,00	100,00%	42.085,30	82,08%	2.000.310.023,04	100,00%	57.909,52	51.611,62	-	-	-
DELTA GENERALI	77.020,70	100,00%	7.205,00	100,00%	63.660,47	97,85%	2.594.111.623,43	99,78%	49.648,02	37.114,64	-	-	-
DUNAV	145.376,45	99,99%	21.725,00	100,00%	120.159,04	95,16%	4.695.448.898,53	99,99%	47.610,61	36.597,90	12,42	100,00%	-
GARANT	27.105,89	99,99%	13.090,00	100,00%	22.404,02	42,94%	508.871.085,25	100,00%	27.873,55	47.141,53	-	-	-
HYPO	2.053,53	99,91%	1.155,00	100,00%	1.697,32	97,67%	58.080.605,76	99,99%	41.691,81	31.226,26	-	-	-
NLB NOVA PENZIJA	7.193,57	99,80%	2.585,00	100,00%	5.945,75	88,08%	183.479.338,83	99,61%	37.597,89	31.223,83	-	-	-
RAIFFESEN FUTURE	19.099,43	99,96%	4.400,00	100,00%	15.786,39	97,30%	892.744.430,79	99,93%	68.901,31	51.802,15	1,77	100,00%	-
SOCIETE GENERALE BKVL	276,88	99,36%	1.650,00	100,00%	228,85	74,36%	1.465.688,87	100,00%	7.803,14	7.676,27	-	-	-
SOCIETE GENERALE STECK	475,67	100,00%	4.455,00	100,00%	393,16	57,84%	1.579.601,13	100,00%	4.895,15	6.191,32	-	-	-
TRIGLAV	7.983,39	100,00%	495,00	100,00%	6.598,57	20,83%	57.604.100,50	100,00%	10.636,21	37.349,18	3,55	100,00%	-
Mart	339.380,92		121.880,00		280.510,94	87,38%	11.381.651.547,88		49.435,46	41.384,00	15,97		
DOOR PENZIJA PLUS	50.847,83	100,00%	4.180,00	100,00%	41.862,31	81,93%	2.046.086.551,61	100,00%	59.550,29	53.169,84	-	-	-
DELTA GENERALI	76.946,76	100,00%	9.185,00	100,00%	63.516,70	97,68%	2.698.678.108,81	99,77%	51.786,19	38.764,88	-	-	-
DUNAV	146.283,41	99,99%	58.190,00	100,00%	120.908,67	94,67%	4.842.211.244,92	99,99%	46.794,33	37.701,18	12,42	100,00%	-
GARANT	27.515,89	99,99%	14.830,00	100,00%	22.742,90	43,62%	536.186.489,67	100,00%	28.723,47	48.171,88	-	-	-
HYPO	2.660,54	99,73%	18.810,00	99,13%	2.199,03	84,46%	67.640.620,06	99,97%	37.476,60	32.459,55	-	-	-

Slika A.16. Računi članova fondova-pregled po godinama, mesecima i fondovima, sa obačunom procenta učešća



Slika A.17. Transakcije hartijama od vrednosti-kombinacija podataka iz dve tabele činjenica po zajedničkoj dimenzionalnosti

SEKTORSKA STRUKTURA IMOVINE

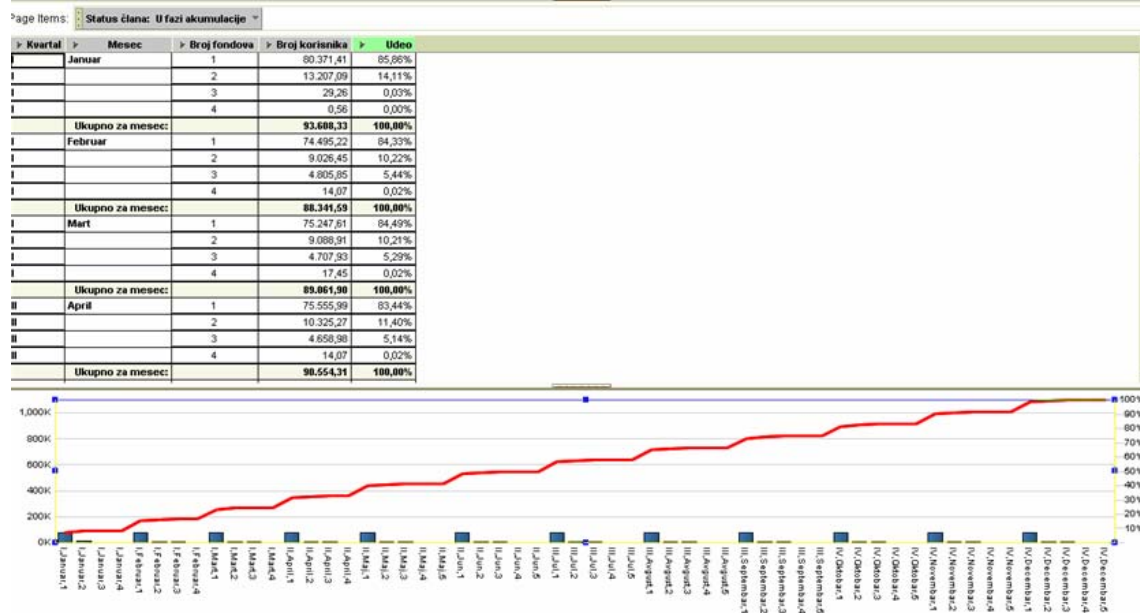
	Iznos imovine	Iznos imovine (u EUR)	% učešća u ukupnoj imovini svih fondova
2008	80.249.786.262,21	3.352.675,99	36,376,27%
▶ Hartije od vrednosti	40.510.433.824,79	1.692.444,90	19,051,33%
▶ Akcije	14.552.315.036,99	607.966,62	6,530,41%
▶ -	14.552.315.036,99	607.966,62	6,530,41%
▶ Dužničke hartije od vrednosti	25.958.118.787,79	1.084.478,29	12,520,92%
▶ -	25.958.118.787,79	1.084.478,29	12,520,92%
▶ Nepokretnosti	946.220.822,33	39.531,21	298,84%
▶ Depoziti	3.448.437.756,71	144.068,83	2,025,56%
▶ Gotovina	35.390.874.953,62	1.478.559,97	15,064,69%
▶ Potraživanja	320.881.270,48	13.405,78	159,59%
▶ Potraživanja od društava za upravljanje	3.251.512,85	135,84	2,14%
▶ Ostala potraživanja	317.629.757,63	13.269,94	157,44%
▶ Obaveze	-367.052.385,23	-15.334,71	-223,71%
2009	129.916.248.136,70	5.427.641,01	36,352,35%
▶ Hartije od vrednosti	85.842.847.002,92	3.586.342,46	23,167,44%
▶ Akcije	8.420.980.481,69	351.811,72	1,877,80%
▶ Bankarski sektor u zemlji	517.705.080,18	21.628,68	107,78%
▶ Domaće banke	517.705.080,18	21.628,68	107,78%
▶ Drugi pružatelji	4.178.720.351,25	174.578,58	842,36%
▶ Vajanje ruda i kamena i prerada čelika industrija	2.042.356.067,91	85.325,55	411,24%
▶ Aktivnosti u vezi s nekretnima, iznajmljivanje i poslovne aktivnosti, ostale komunalne, društvene i lične uslužne aktivnosti	2.136.364.263,35	69.253,03	431,12%
▶ -	3.724.555.050,25	155.604,46	927,65%
▶ Dužničke hartije od vrednosti	77.421.866.521,24	3.234.530,74	21,289,65%
▶ Javni sektor	43.012.368.252,24	1.796.970,72	10,844,13%
▶ Republički organi i organizacije	43.012.368.252,24	1.796.970,72	10,844,13%
▶ -	34.609.498.269,00	1.437.560,02	10,445,52%
▶ Nepokretnosti	1.663.407.186,73	69.493,83	371,67%
▶ Depoziti	6.804.720.252,27	284.287,60	1,896,88%
▶ Gotovina	35.474.718.369,24	1.482.062,78	10,877,67%
▶ Potraživanja	482.297.460,37	20.149,42	186,32%
▶ Potraživanja od društava za upravljanje	3.908.294,72	163,26	1,64%
▶ Ostala potraživanja	478.389.165,65	19.986,14	184,68%
▶ Obaveze	-351.742.135,83	-14.695,06	-147,64%
UKUPNO:	210.166.844.398,41	8.780.317,00	72,728,62%

Slika A.18. Struktura imovine fondova. Raščlanjenje imovine na komponente. Prikaz bušenja podataka duž strukture imovine. Obračun procenta učešća svake komponente u ukupnoj imovini na svim nivoima hijerarhije

UPOREDNI PREGLED DOSTAVLJENE I IZRAČUNATE VREDNOSTI NETO IMOVINE										
-u dinarima-										
Od : '01.01.2009' , Do : '31.12.2009'										
Page Items: Godina: 2009 Mesec: Maj										
	Ukupna vrednost HoV u dinarima	Vrednost nekretnina u dinarima	Gotovina	Depoziti	Mesečni podaci			NETO IMOVINA izračunata	NETO IMOVINA (dostavljen podatak)	RAZLIKA
					Neto imovina u dinarima	Potraživanja	Obaveze			
01.01.2009	7.435.497.766,20	65.210.882,63	985.598.477,78	207.461.681,49	1.813.597.682,63	5.411.470,26	7.564.632,57	1.638.044.707,77	2.463.931,67	-818.339.068,10
▶ DOOR PENZIJA PLUS	1.737.249.805,18	-	79.278.693,43	50.110.200,00	321.808.333,26	388.796,03	1.111.378,58	285.454.706,14	965.434.654,13	-144.991.242,41
▶ DELTA GENERALI	1.578.929.964,25	-	296.632.372,39	69.665.400,00	433.260.829,41	546.591,44	2.193.324,67	389.672.819,48	1.299.795.406,18	-195.235.734,12
▶ DUNAV	3.194.988.188,42	65.210.882,63	382.327.223,02	122.220.000,00	770.120.875,08	3.334.688,28	3.253.413,16	691.796.091,25	2.318.385.729,51	-347.878.526,12
▶ GARANT	293.208.110,59	-	62.064.367,14	14.866.400,00	85.725.344,50	766.557,74	244.442,02	76.865.254,59	257.178.695,30	-38.865.233,11
▶ HYPO	30.691.786,13	-	12.060.229,08	-	11.865.135,86	36.643,83	12.362,25	10.665.432,60	35.595.763,54	-5.353.177,11
▶ NLB NOVA PENZIJA	145.810.377,35	-	13.704.853,46	5.377.680,00	31.485.273,16	28.938,30	336.073,14	28.316.364,65	94.468.764,85	-14.189.221,11
▶ RAIFFESEN FUTURE	390.362.960,89	-	138.776.646,54	23.515.126,00	148.441.388,16	301.040,34	121.548,81	133.484.691,08	445.328.817,75	-66.913.387,12
▶ SOCIETE GENERALE EKVILIBRIO	2.043.296,32	-	133.270,32	49.644,99	400.374,27	870,27	1.032,72	358.998,72	1.201.134,82	-188.526,10
▶ SOCIETE GENERALE STEDNJA	4.569.870,88	-	171.524,01	133.926,50	821.302,34	2.774,34	19.183,17	737.587,39	2.463.931,67	-371.182,12
▶ TRIGLAV	57.643.402,19	-	1.049.298,37	1.723.302,00	9.669.806,58	4.569,70	271.874,04	8.692.189,87	29.808.709,80	-4.388.842,11
01.01.2009	7.606.662.159,09	65.077.047,16	947.633.665,69	291.128.281,49	1.817.847.331,77	5.748.117,64	7.040.523,01	1.633.761.378,82	2.463.931,67	-828.359.522,12
▶ DOOR PENZIJA PLUS	1.844.878.590,49	-	53.576.766,44	50.110.200,00	322.007.210,15	234.533,64	2.473.152,22	289.648.486,62	966.031.290,75	-145.065.828,13
▶ DELTA GENERALI	1.640.032.798,96	-	282.673.230,75	69.665.400,00	434.456.254,88	653.474,09	2.303.264,24	398.786.225,77	1.383.381.738,46	-195.814.178,41
▶ DUNAV	3.210.437.714,89	65.077.047,16	381.486.261,37	122.220.000,00	772.525.132,73	3.642.876,60	1.881.655,28	693.888.637,30	2.317.598.574,16	-348.831.704,14
▶ GARANT	348.501.919,34	-	48.206.904,08	14.866.400,00	85.757.471,40	795.196,90	137.243,23	76.886.883,76	257.274.986,94	-38.886.976,17
▶ HYPO	40.332.699,51	-	9.662.116,83	-	11.873.198,63	38.983,22	15.011,94	10.671.252,02	35.619.926,82	-5.357.742,12
▶ NLB NOVA PENZIJA	146.141.376,93	-	13.553.931,98	5.377.680,00	31.537.322,24	33.118,53	36.508,83	28.370.663,52	94.612.912,86	-14.285.189,11
▶ RAIFFESEN FUTURE	312.071.839,27	-	157.233.610,32	27.181.726,00	148.786.794,35	344.820,78	183.482,03	133.777.648,25	446.368.846,71	-67.086.732,11
▶ SOCIETE GENERALE EKVILIBRIO	2.046.018,01	-	144.270,12	49.644,99	406.942,27	481,66	1.693,65	366.824,65	1.228.839,83	-183.353,12
▶ SOCIETE GENERALE STEDNJA	4.568.079,19	-	179.881,28	133.926,50	824.309,11	573,64	19.548,70	741.089,46	2.472.952,85	-371.740,11
▶ TRIGLAV	57.670.822,50	-	916.892,52	1.723.302,00	9.672.696,01	5.058,57	8.962,89	8.703.398,48	29.818.378,22	-4.354.884,12
01.01.2009	7.636.154.354,87	65.186.605,17	953.274.961,23	285.017.281,49	1.824.056.940,38	5.602.961,16	6.895.969,64	1.638.411.687,24	2.463.931,67	-822.088.949,13
▶ DOOR PENZIJA PLUS	1.850.827.863,54	-	55.900.787,47	43.999.300,00	322.663.754,51	256.730,81	2.407.811,11	298.234.284,28	968.088.843,64	-145.366.656,11
▶ DELTA GENERALI	1.647.579.188,93	-	282.396.178,07	69.665.400,00	435.656.196,56	745.126,07	1.082.403,18	391.969.286,87	1.387.581.665,49	-196.443.138,11
▶ DUNAV	3.224.721.411,05	65.186.605,17	383.138.380,85	122.220.000,00	775.263.297,22	3.318.708,28	2.530.154,58	696.451.879,87	2.325.878.148,80	-358.165.884,11
▶ GARANT	349.527.169,42	-	49.139.199,46	14.866.400,00	86.458.911,79	809.566,51	212.206,91	77.510.669,39	258.378.329,15	-39.218.145,17
▶ HYPO	40.350.995,49	-	9.668.243,38	-	11.880.593,47	41.324,79	17.547,59	10.676.973,00	35.642.136,82	-5.362.884,41
▶ NLB NOVA PENZIJA	146.523.781,99	-	13.695.052,04	5.377.680,00	31.678.362,53	37.366,35	50.376,72	28.495.627,60	95.838.837,94	-14.278.632,12
▶ RAIFFESEN FUTURE	312.245.893,78	-	157.480.603,45	27.181.726,00	148.979.813,07	387.396,39	226.328,18	133.934.471,63	446.943.988,65	-67.198.488,12
▶ SOCIETE GENERALE EKVILIBRIO	2.050.712,08	-	149.223,77	49.644,99	410.492,92	521,31	2.062,49	369.194,67	1.231.491,87	-184.976,11
▶ SOCIETE GENERALE STEDNJA	4.574.207,82	-	182.962,41	133.926,50	827.018,83	635,45	19.896,92	743.495,13	2.481.881,29	-372.992,12
▶ TRIGLAV	57.753.130,77	-	1.524.330,34	1.723.302,00	10.038.499,49	5.585,20	47.183,96	9.831.285,58	30.115.799,64	-4.528.837,14

Slika A.19. Uporedna analiza istog podatka dobijenog izračunavanjem i dostavljenog iz spoljnog izvora. Praćenje odstupanja u cilju kontrole ispravnosti dostavljenih podataka i verodostojnosti izveštavanja o poslovanju fondova

DISTRIBUCIJA KORISNIKA PO BROJU FONDOVA U KOJIMA SU ČLANOVI - Status člana : U fazi akumulacije



Slika A.20. Pregledi nad podacima o članovima fondova i njihovim računima su sastavni deo redovnog statističkog izveštavanja. Pregled je radjen nad materijalizovanim pogledom

B. Kreditni odnosi sa inostranstvom

Implementacija skladišta podataka u Sektoru za međunarodnu saradnju Narodne banke Srbije

B.1. Projektovanje skladišta podataka

B.1.1. Dimenzije u podsystemima kreditnih zaduženja i kreditnih odobrenja

Skup dimenzija podsystema je definisan nakon dogovora sa korisnicima o tome kog nivoa detaljnosti će biti poslovne mere u tabelama činjenica i koji su aspekti posmatranja tih poslovnih mera interesantni sa analitičkog stanovišta. Kada se određivao nivo detaljnosti podataka u skladištu imalo se u vidu da je korisnicima na raspolaganju i transakcioni sistem sa velikim brojem statičkih izveštaja u kojima su prikazani svi najdetaljniji podaci. Zato je odlučeno da skladište podataka sadrži isključivo agregirane podatke o stanju duga na dan, a da se podaci o planovima korišćanja sredstava i podaci o planovima otplate glavnice i kamate čuvaju na najgranularnijem nivou. Korisnici su posebno insistirali na tome da se planovi otplate čuvaju na detaljnom nivou jer te podatke zamrznute na dan ne mogu da dobiju iz transakcionog sistema, a svako agregiranje ovih podataka umanjuje kvalitet analiza koje rade. Nakon definisanja nivoa granularnosti definisani su i aspekti posmatranja poslovnih mera vezanih za praćenje planova otplate, prometa i stanja duga. Rezultat je bio skup dimenzija koje su uključene u podsysteme kreditnih zaduženja i kreditnih odobrenja. To su:

- **Vremenska dimenzija.** Kao i vremenska dimenzija opisana u prvom primeru ova dimenzija sadrži podatke do nivoa datuma. Kako su sa poslovnog aspekta značajni samo poslednji dani u mesecu, to se u vremenskoj dimenziji nalaze samo poslednji dani u mesecu i to od januara 2000. godine do danas. Dimenzija se osvežava pre punjenja tabela činjenica i pre punjenja ostalih dimenzija. U nju se samo vrši upis podataka, dok se izmene i brisanje podataka ne sprovode. Ovo je zajednička dimenzija za podsysteme kreditnih zaduženja i kreditnih odobrenja.
- **Dimenzija prijave kreditnog zaduženja.** Ovo je ključna dimenzija celog podsystema kreditnih zaduženja, jer se atributi iz ove dimenzije konsultuju u svakom pregledu koji korisnici rade. Svaka analiza, bez izuzetka, radi se u kontekstu nekog od dimenzionih atributa ove dimenzije. Dimenzija sadrži više desetina atributa (preko 50 atributa) koji opisuju jednu prijavu kreditnog zaduženja u inostranstvu. Svi atributi izvode se iz transakcionog sistema u kome su svi podaci o zaključenom kreditu. Osobine kreditne prijave sastavni su deo obrasca preko koga se registruje kreditni posao u Narodnoj banci Srbije. Obrazac za registraciju kreditnog posla sadrži preko dvadeset grupa obeležja i to je jedan od najobimnijih i najdetaljnijih dokumenta u sistemu Narodne banke. Većina obeležja vezana je za propisani kodeks šifara. Obeležja kreditnog posla su u toku života kredita podložna promeni. Svaka promena evidentira se takodje preko propisanog obrasca za izmenu kreditne prijave. Denormalizovana dimenzija sadrži sva osnovna polja sa kreditne prijave, ali i izvedena polja i polja iz propisanih šifarnika. Poslovne potrebe odredile su da se u ovoj

dimenziji uvek nalazi samo poslednja, aktuelna vrednost obeležja sa kreditne prijave. Dakle, istorijat promena obeležja se ne čuva, već se pri svakom osvežavanju dimenzije stare vrednosti izmenjenih atributa prepisuju novim vrednostima. Vremenom se prijava kreditnog posla menjala shodno promenama odluke o evidentiranju kreditnih poslova sa inostranstvom. Mnoga obeležja kredita više nisu aktuelna, dok su u poslednje dve godine uključena brojna nova obeležja. Atributi u dimenziji kredita predstavljaju uniju svih obeležja jednog kreditnog posla koja su od 2000. godine do danas bila aktuelna. Dimenzija trenutno ima nešto više od 14000 slogova.

Delimični prikaz dimenzije kreditne prijave dat je na slici B.1.

DIMENZIJA KREDITNE PRIJAVE
Id_kredita (PK)
Broj prijave (prirodni ključ)
Broj okvirne prijave
Datum zaključenja ugovora
Ročnost kredita
Sektorska pripadnost
Republička pripadnost
Valuta kredita
Iznos kredita
Broj korišćenja kredita
Broj rata glavnice
Broja rata kamate
Vrsta kamatne stope
Promenljiva kamatna stopa
Matični broj garanta
Naziv garanta
Matični broj supegaranta
Naziv supergaranta
Oznaka šeme kredita
Matični broj osiguravajuće agencije
Naziv osiguravajuće agencije
Status kredita
.....

Slika B.1. Dimenzija prijave kreditnog posla zaključenog sa inokreditorom

- **Dimenzija stranih lica.** Dimenzija sadrži podatke o svim inokreditorima koji su učestvovali u kreditiranju domaćih lica od početka automatizovanog praćenja kreditnih zaduženja do danas. Podaci o stranim licima se preuzimaju iz transakcionog šifarnika koji održavaju ovlašćena lica u sektoru. Dimenzija sadrži samo aktuelne vrednosti atributa i, prema zahtevu korisnika, ne prati istorijat promena. Kako se, za razliku od dimenzije kredita, obeležja u ovoj dimenziji jako retko menjaju, to se u ovu dimenziju najčešće radi upis novih slogova i vrlo retko izmena postojećih podataka.

Prikaz dimenzije stranih lica dat je na slici B.2.

STRANA LICA
Id_stranog_lica (PK)
Šifra stranog lica (prirodni ključ)
Naziv stranog lica
Držva stranog lica
Sedište osnivača
Vrsta stranog lica
Grupa stranog lica
.....

Slika B.2. Dimenzija stranih lica-inokreditora u poslu kreditnog zaduženja i dužnika u poslu kreditnog odobrenja

Ovo je zajednička dimenzija u podsistemima kreditnih zaduženja i kreditnih odobrenja. U domenu kreditnih zaduženja strana lica predstavljaju kreditore, dok su u kreditnim odobrenjima strana lica dužnici po odobrenom kreditnom poslu. Da bi se na prezentacionom nivou razdvojili podaci o stranim licima koji se odnose na zaduženja od podataka o stranim licima dužnicima u poslovima odobrenja kredita u bazi skladišta su napravljena dva pogleda nad istom tabelom. Svaki od ova dva pogleda je uključen kao dimenzija u odogovrajući Data Mart i logički je povezan ključem sa odgovarajućim tabelama činjenica. Ova dimenzija danas sadrži podatke o oko 4000 stranih lica.

- **Dimenzija namena kreditnog zaduženja.** Ovo je vrlo jednostavna minidimenzija koja sadrži oko 40 slogova sa propisanim šiframa i nazivima namena po osnovu kojih se obavljaju kreditni poslovi zaduženja sa inostranstvom. Šifarnik namena je sastavni deo zvaničnog kodeksa šifara i kao takav se vrlo retko ažurira. Uvodjenje novih šifara u sistem ili izmena postojećih naziva podrazumeva izmenu uputstva o evidentiranju kreditnih poslova sa inostranstvom što zahteva zvaničnu pravnu proceduru. Održavanje ovog transakcionog šifarnika je u nadležnosti sektora. Procedura za ažuriranje dimenzije u skladištu identifikuje nove slogove ili izmene postojećih slogova u šifarniku i, prema tome, radi upis slogova ili izmenu postojećih slogova u dimenziji. Jedinstveni identifikator sloga je primarni ključ u ovoj dimenziji.
- **Dimenzije domaćih lica-obveznika po kreditnim poslovima.** U okviru ove zajedničke dimenzije razmatra se više poslovnih subjekata uključenih u skladište podataka kroz poddimenzije sa specifičnom fizičkom implementacijom. Naime, svi podaci o dužnicima kod kreditnog zaduženja, korisnicima kredita, poslovnim bankama koje predstavljaju ovlašćene banke preko kojih se obavlja kreditni posao, podnosiocima prijava, ali i poveriocima i davaocima kredita kada su kreditna odobrenja u pitanju fizički se nalaze u zajedničkoj tabeli u bazi skladišta podataka. Ta zajednička tabela obveznika po kreditnim poslovima zaduženja i odobrenja osvežava se iz transakcionog

šifarnika koji se održava u sektoru. Dimenzija, kao i prethodne, sadrži poslednje, aktuelne vrednosti atributa. U nju se vrši upis novih slogova i izmene atributa u postojećim slogovima, shodno promenama u šifarniku. Prikaz zajedničke dimenzije obveznika dat je na slici B.3.

DOMAĆA LICA-OBVEZNICI
Id_obveznika (PK)
Matični broj (prirodni ključ)
Naziv domaćeg lica
Mesto
Republika
Vrsta obveznika
Delatnost domaćeg lica
.....

Slika B.3. Dimenzija domaćih lica-obveznika po kreditnom poslu

Nezavisno od zajedničke fizičke implementacije poslovne kategorije dužnika, korisnika i ovlašćene poslovne banke u oblasti kreditnih zaduženja su zasebne kategorije. Stoga je bilo neophodno da se i u modelu, bar na logičkom, prezentacionom nivou ove dimenzije vide „svaka za sebe“. Potrebe korisnika da rade analize po nekoj od ovih kategorija ili po više njih istovremeno, uslovlila je da se formiraju tzv. dimenzije koje „igraju ulogu“ (*eng. Role-playing*). Dimenzije su implementirane kao pogledi formirani uobičajenom CREATE VIEW SQL naredbom iz osnovne dimenzije. Svaka od njih sadrži jedinstveni identifikator preko koga je na prezentacionom nivou povezana sa tabelom činjenica. Tako je, preko pogleda, u tabelu činjenica logički spušteno više primarnih ključeva iz iste dimenzije. Na ovaj način su iz dimenzije obveznika formirane dimenzije:

- Dužnika, korisnika, ovlašćene banke u Data Mart-u kreditnih zaduženja.
- Kreditora i poverioca u Data Mart-u kreditnih odobrenja.
- **Dimenzija prijave kreditnog odobrenja.** Pominjem je ukratko zato što je uradjena u potpunoj analogiji sa dimenzijom prijave kreditnog zaduženja koja je napred detaljnije opisana. Dve dimenzije kreditnih prijava su i fizički i logički odvojene. Pone se u odvojenim procesima iz različitih izvornih transakcionih tabela. Ono što je slično to je svrha i opis ovih dimenzija. U ovoj dimenziji se nalaze obeležja koja figurišu na propisanoj prijavi kreditnog odobrenja i to sa svojim tekućim vrednostima. Ta obeležja su većim delom već prikazana u dimenziji prijave kreditnog zaduženja. Treba imati na umu da su poslovi kreditnog zaduženja i kreditnog odobrenja sa inostranstvom skoro identični samo suprotno usmereni poslovni procesi. Zato su i obeležja jednog kredita većim delom zajednička, bilo da se radi o odobrenom ili uzetom kreditu. Dimenzija trenutno ima oko 1100 slogova.

- **Dimenzija izvora sredstava kod kreditnog odobrenja.** Ovo je dimenzija koja figuriše samo kod kreditnih odobrenja i to za stare, istorijske podatke. Prema najnovijem uputstvu o evidentiranju kreditnih poslova izvori sredstava za kreditiranje se ne prijavljuju Narodnoj banci Srbije. Dimenzija i dalje postoji u modelu upravo zbog analize kredita zaključenih do aprila 2008. godine. Dimenzija je jednostavne strukture i malog obima. Ona sadrži desetak slogova sa šiframa i nazivima izvora sredstava koji su u periodu do aprila 2008. godine bili propisani zvaničnim kodeksom šifara. Jedinstveni identifikator sloga je primarni ključ u ovoj dimenziji. Procedura za ažuriranje dimenzije radila je upis novih ili izmene postojećih slogova, mada su se i nad ovim šifarnikom vrlo retko sprovodile izmene. Danas se ovaj transakcioni šifarnik ne ažurira. Shodno tome, procedura za osvežavanje dimenzije je isključena iz redovne procedure osvežavanja skladišta podataka.

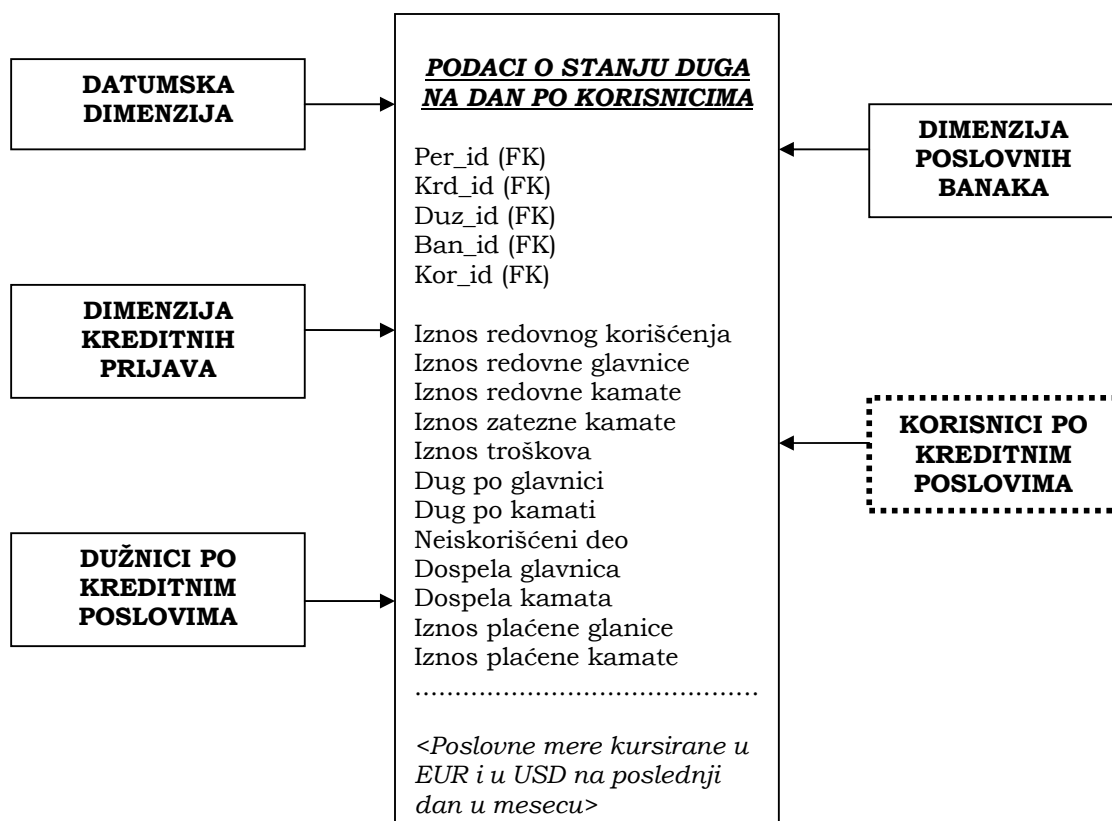
B.1.2. Poslovne mere u podsistemu kreditnih zaduženja

Poslovne mere u oba Data Mart-a grupisane su u dve grupe podataka. Prva grupa su poslovne mere koje se koriste u okviru praćenja, analiza i izveštavanja o prometu i stanju duga na dan po različitim aspektima kreditnih poslova. Druga grupa su podaci o planovima otplate glavnice i kamate, kao i podaci o planovima korišćenja preostalih sredstava iz kredita. Planovi otplata su posebno važni jer se na osnovu njih dobijaju informacije o ukupnim kreditnim obavezama u zemlji u budućem periodu i tako omogućava planiranje sredstava koja se moraju obezbediti za servisiranje duga prema inostranstvu.

Poslovne mere vezane za praćenje stanja duga se izvode u ETL procedurama. Izvorni podaci se nalaze u više transakcionih tabela koje detaljno na nivou svake pojedinačne transakcije čuvaju podatke o korišćenim sredstvima, planiranim otplatama i izvršenim plaćanjima. U skladištu podataka podaci o stanju duga su izvedeni i transformisani iz izvornih podataka. U ovim tabelama nema najgranularnijih podataka. Pravila izvodjenja i transformacije su definisana u saradnji sa korisnicima i ugradjena u ETL procedure. Struktura i organizacija transakcionog sistema je takva da bi izvodjenje ovih poslovnih mera u sklopu nekog statičkog izveštaja bilo skoro nemoguće. Zato je, za praćenje stanja duga, skladište podataka jedini izveštajni sistem koji se koristi u sektoru.

Podaci o stanju duga nalaze se u četiri tabele činjenica i to: po korisnicima kredita, po inokreditorima, po namenama kredita i zbirno. Svi podaci koji se prate u sklopu kreditnih zaduženja sa inostranstvom prate se sa nekog od ova četiri aspekta. Zato su implementirane najpre tri tabele činjenica u kojima su podaci po korisnicima, inokreditorima i namenama, a zatim je u cilju boljih performansi izvršavanja upita u sistem uključena i četvrta, agregirana tabela.

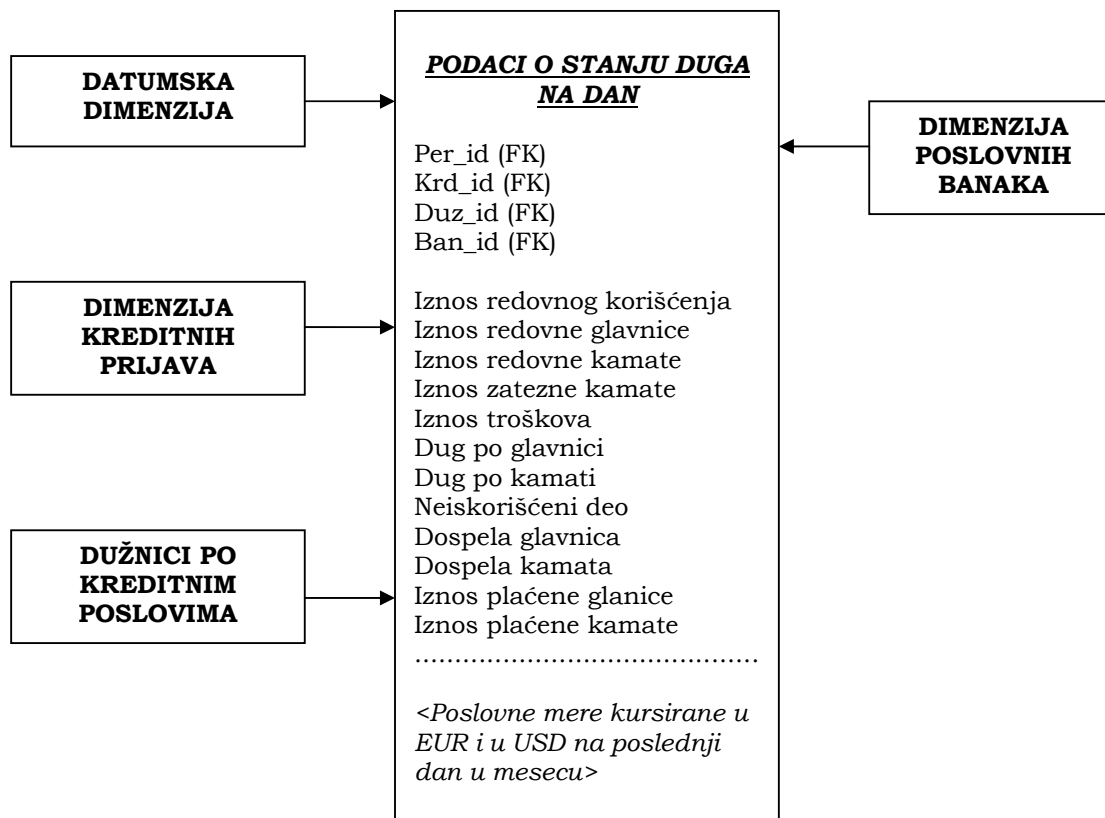
Prikaz jedne zvezdaste strukture koja predstavlja deo modela za praćenje stanja duga dat je na slici B.4.



Slika B.4. Zvezdasta struktura za praćenje stanja duga po korisnicima kredita

Slično ovoj zvezdastoj šemi implementirane su još dve zvezdaste šeme. U prvoj je umesto dimenzije korisnika kredita (posebno označena na prethodnoj slici) uključena dimenzija namena. U drugoj je umesto korisnika kredita uključena dimenzija stranih lica-inokreditora. Agregirana zvezdasta šema je implementirana na kraju. U njoj su isključene dimenzije korisnika, namena i inokreditora i dobijeni su zbirni podaci na nivou datuma, kredita, dužnika i poslovne banke.

Prikaz agregirane zvezdaste strukture dat je na slici B.5.

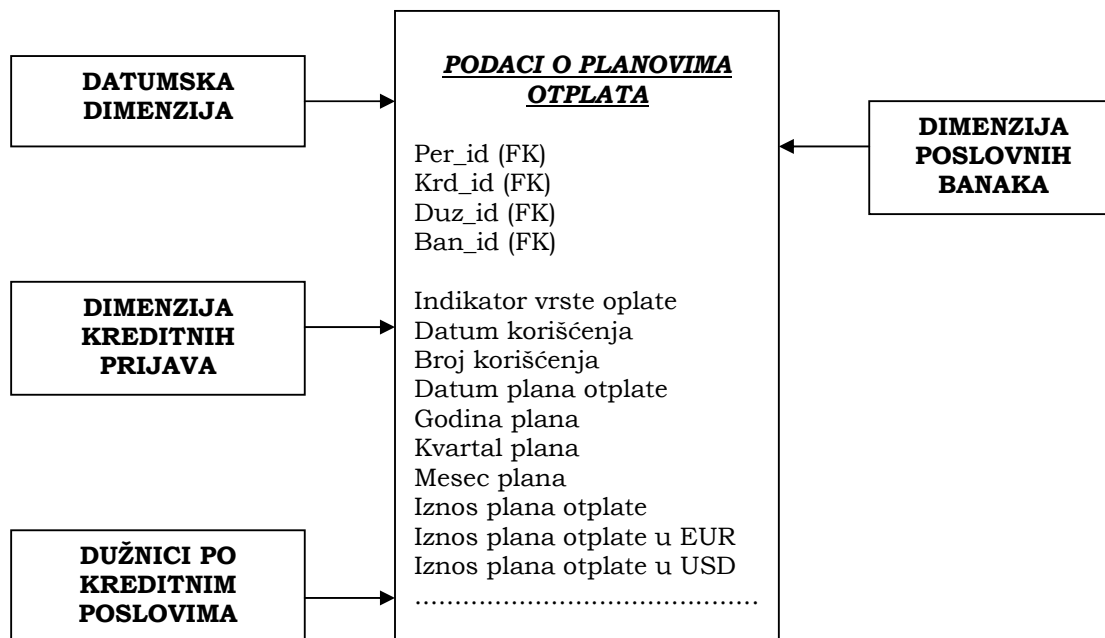


Slika B.5. Zvezdasta struktura za zbirno praćenje stanja duga

U svakoj tabeli činjenica u ovom modelu nalaze se kursirani iznosi svih poslovnih mera, i to u evrima i dolarima na poslednji kalendarski dan u mesecu za koji se presek podataka radi. Kursiranje je primenjeno u tabelama koje se odnose na stanje duga, kao i u tabelama koje se odnose na planove korišćenja i otplate. Ovo važi za oba podsistema: kreditna zaduženja i kreditna odobrenja.

Kada je reč o planovima korišćenja i otplate glavnice i kamate zahtev korisnika je bio da se u skladište podataka uključe najgranularniji podaci. Pokušaji da se, zbog odbima podataka, postigne dogovor oko nekog nivoa sumarnosti ovih podataka nisu dali rezultate, jer je stav analitičara da bi svako agregiranje najdetaljnijih podataka umanjilo kvalitet njihovih analiza. Neprihvatljivo je bilo i rešenje da se iz već postojećeg transakcionog sistema izvode detaljni pregledi, a da u skladište udju bar delimično sumirani podaci. Razlog je taj što ovako koncipiran transakcioni sistem nije u mogućnosti da u potpunosti „vrati stanje unazad“. Priroda posla nalaže da se vrlo često traže podaci sa stanjem na dan od pre tri, šest meseci, ili godinu dana. Zato je korisnicima uvek bilo bitno da imaju zamrznuto, nepromenljivo stanje svih planova otplate na određeni datum. Analize planova korišćenja i otplate danas se isključivo rade nad podacima iz skladišta. Planovi korišćenja implementirani su kroz četiri zvezdaste šeme (po korisnicima, po namenama, po inokreditorima i zbirno). Planovi otplate glavnice i kamate su zbog slične

strukture i čestog uporednog analiziranja, implementirani kroz zajedničke zvezdaste šeme. U svakoj od četiri tabele činjenica nalazi se indikator vrste otplate (sa vrednostima „Glavnica“ i „Kamata“), koji služi za filtriranje podataka. Na slici B.6. prikazana je jedna zvezdasta šema kroz koju su implementirani zbirni podaci o planovima otplata glavnice i kamate.



Slika B.6. Zvezdasta struktura za zbirno praćenje planova otplate glavnice i kamate

B.1.3. Poslovne mere u podsistemu kreditnih odobrenja

Podsistem kreditnih odobrenja uveden je nekoliko meseci nakon kreditnih zaduženja. U to vreme već su se videli prvi rezultati, prednosti i mane modela koji je implementiran. Činjenica je da je podsistem kreditnih zaduženja bio neka vrsta pilot projekta, prvi pokušaj da se u nekom segmentu poslovanja Narodne banke implementira skladište podataka. Zato su svi naredni sistemi, medju koje spadaju kreditna odobrenja, penzijski fondovi, sistem za analizu bilansnih podataka poslovnih banaka i još neki sistemi koji se danas koriste, radjeni na bazi iskustava koje smo imali sa kreditnim zaduženjima. Koristeći pozitivna iskustva u vreme kada se sistem kreditnih zaduženja već ustalio uradjen je model kreditnih odobrenja u potpunoj analogiji sa prethodnim sistemom. Dimenzije koje su povezane sa svakom tabelom činjenica su:

- Vremenska dimenzija, koja je zajednička za oba Data Mart-a.
- Dimenzija prijava kreditnog odobrenja, koja je i fizički i logički odvojena od napred pomenute dimenzije prijava kreditnog zaduženja.

Ona je po strukturi, opisu atributa i načinu osvežavanja vrlo slična dimenziji koja je prikazana na slici B.1.

- Dimenzija dužnika, koja je u ovom slučaju podskup opisane zajedničke dimenzije stranih lica, implementirana je kroz bazni pogled.
- Dimenzija poverilaca, koja je podskup opisane zajedničke dimenzije domaćih lica-obveznika, takodje je implementirana kroz bazni pogled. Ova dimenzija ima više istorijski značaj, jer se koristi u analizama kreditnih poslova zaključenih do aprila 2008. godine. Novi krediti se ne posmatraju sa aspekta poverilaca, ali je dimenzija i dalje deo modela podataka.

Tabele činjenica su, kao u prethodnom primeru, grupisane u dve grupe i to: podaci o prometu i stanju duga na dan i podaci o planovima korišćenja i planovima naplata potraživanja. Svaka grupa podataka se, analogno prethodnom, ovde posmatra na tri načina: po kreditorima (domaćim licima-davaocima kredita), po izvorima sredstava za kreditiranje i zbirno. Zvezdaste šeme u kojima su podaci po kreditorima imaju, osim pomenutih obaveznih dimenzija, uključenu i dimenziju kreditora, tj. davaoca kredita odobrenog inostranstvu. Ova dimenzija je još jedna dimenzija implementirana kao pogled nad napred pomenutom zajedničkom dimenzijom domaćih lica-obveznika. Zvezdaste šeme u kojima su podaci grupisani prema izvorima sredstava za kreditiranje imaju, osim pomenutih obaveznih dimenzija, uključenu i dimenziju izvora sredstava za kreditiranje. Napred je rečeno da je ova dimenzija zadržana isključivo zbog istorijskih podataka, jer se podaci o izvorima sredstava ne prijavljuju Narodnoj banci od aprila 2008. godine. Zbirne zvezdaste šeme ne uključuju nijednu od ove dve dimenzije i sadrže podatke agregirane do nivoa obaveznih dimenzija u sistemu. Sve poslovne mere se, u ETL procedurama, kursiraju u dolare i evre na poslednji dan u mesecu za koji se radi presek transakcionih podataka. Tako kursirane poslovne mere deo su tabela činjenica. Podaci o prometu i stanju duga su izvedeni, transformisani i agregirani do nivoa pomenutih dimenzija. Podaci o planovima korišćenja i naplata su uključeni u najgranularnijem obliku.

B.1.4. Tabela medjuvalutnih odnosa

Tabela medjuvalutnih odnosa se ne svrstava ni medju dimenzije ni medju tabele činjenica već se razmatra nezavisno kao posebna pomoćna tabela u sistemu. Ova tabela je u skladište podataka uključena naknadno kada su i potrebe za ad-hok kursiranjem podataka bile sve veće. Napred je rečeno da se u svakoj tabeli činjenica u oba Data Mart-a pored svake poslovne mere prikazane u originalnoj valuti sa kreditne prijave nalaze i njoj odgovarajuće mere kursirane u dolare i evre. Ove mere kursirane su na poslednji dan izveštajnog meseca za koji se presek podataka radi. Vremenom se pojavila potreba za kursiranjem podataka iz originalne valute u neku drugu valutu i to na izabrani datum. Ograničenje koje je uvedeno je da taj izabrani datum bude uvek poslednji dan u mesecu i da valuta u koju se kursira bude dinar, evro ili dolar. Implementirana tabela medjuvalutnih odnosa sadrži medjuvalutne odnose izmedju svih originalnih valuta u kojima su zaključeni

kreditni poslovi i tri dogovorene valute i to na poslednji dan svakog kalendarskog meseca. Tabela je uključena u skladište 2007. godine i napunjena je retroaktivno, tako da se u njoj nalaze podaci od januara 2001. godine do danas. Uvodjenje ove tabele omogućilo je korisnicima da direktno na pregledima kursiraju podatke u željenu valutu na željeni dan. Najčešće korišćeni kursirani podaci i dalje su poslovne mere prekursorane u ETL procedurama i smeštene u tabele činjenica. Za sporadična, ad-hok kursiranja po potrebi korisnika tabela medjuvalutnih odnosa se pokazala kao dobro rešenje. Grafički prikaz tabele medjuvalutnih odnosa dat je na slici B.7. U narednom poglavlju se medju prikazanim pregledima koji su uradjeni nad podacima iz ovog skladišta nalazi i primer u kome je za potrebe kursiranja iskorišćena tabela medjuvalutnih odnosa.

<u>TABELA MEDJUVALUTNIH ODNOSA</u>
Datum kursa
Valuta sa kreditne prijave
Odnos valute prema dinaru
Odnos valute prema evru
Odnos valute prema dolaru

Slika B.7. Tabela medjuvalutnih odnosa

B.2. Primeri izveštavanja

Na prezentacionom nivou formirane su dve poslovne oblasti: jedna pod nazivom „Kreditna zaduženja“ i druga pod nazivom „Kreditna odobrenja“. U sektoru postoji interna podela na grupu analitičara koji prate jednu ili drugu poslovnu oblast. Generalno, podaci iz ove dve oblasti se nikada ne kombinuju, već se svaka posmatra i analizira za sebe. Do danas postoji uradjeno više desetina standardnih i ad-hok pregleda nad podacima iz ovog skladišta. U ovom poglavlju je prikazano nekoliko pregleda koji su samostalno uradjeni od strane korisnika u sektoru.

PLANOWI OTPLATE GLAVNICE I KAMATE PO SEKTORIMA, ROČNOSTI I PRIVREDNIM GRANAMA													
Page Items:		Datum preseka: 31.12.2008		Sektorska pripadnost: Negarantovani privatni									
Ročnost kredita	Delatnost dužnika	Godina	2008		2009		2010		2011		2012		
			Tip	Glavnica	Kamata	Glavnica	Kamata	Glavnica	Kamata	Glavnica	Kamata	Glavnica	Kamata
A. Kratkoročni krediti	Dr. komunalne, društvene i licne usluge		0	11,042	2,450,000	60,852	-	-	-	-	-	-	
	Finanjsko posredovanje		-	42,416	10,125,000	467,330	-	-	-	-	-	-	
	Građevinarstvo		285,714	50,344	11,114,286	505,534	-	-	-	-	-	-	
	Hoteli i restorani		-	-	300,000	24,333	-	-	-	-	-	-	
	Poljoprivreda, sumarstvo i vodoprivreda		3,215,873	134,447	11,239,841	346,748	-	-	-	-	-	-	
	Poslovi s nekretninama, iznajmljivanje		3,720,595	95,620	722,407,679.00	8,735,501	-	-	-	-	-	-	
	Preradivačka industrija		17,499,885	1,763,554	727,637,234.30	4,211,265	-	-	-	-	-	-	
	Saobracaj, skladištenje i veze		63,619	4,209	13,910,420	433,410	-	-	-	-	-	-	
	Trgovina na veliko i malo, opravka		10,789,408	781,894	71,177,293.32	8,578,674	400,000	4,799	-	-	-	-	
	Vadjenje ruda i kamena		46,855,002	587,716	277,507,093.15	3,083,162	-	-	-	-	-	-	
	Dr. komunalne, društvene i licne usluge		5,707,709	1,402,204	20,251,334	3,631,160	13,207,953	3,285,507	6,370,443	1,620,019	4,130,497	1,138,571	
B. Srednjoročni Dugoročni krediti	Finanjsko posredovanje		735,687,997.63	25,307,420	46,992,3290.38	762,652,446.79	308,215,721.66	743,999,071.07	509,052,041.63	72,109,035.97	569,204,076.42	705,933,397.76	477,599
	Građevinarstvo		24,592,053	4,325,264	73,007,955.38	40,754,995	759,899,902.50	39,004,388	67,583,295	24,216,565	47,117,104	20,349,993	38.0
	Hoteli i restorani		2,522,514	347,282	5,487,197	3,142,000	12,547,881	2,222,878	6,167,809	1,712,139	5,082,812	1,414,309	4.1
	Neraspoređeno		1,256,636	914,777	76,488	7,320	84,049	19,886	0	0	0	0	0
	Obrazovanje		66,667	2,116	23,529	597,334	805,043	579,705	805,043	512,497	793,278	446,544	7
	Poljoprivreda, sumarstvo i vodoprivreda		12,850,347	3,211,487	23,922,404	6,360,087	20,927,438	4,218,100	13,206,599	3,109,549	13,747,232	2,281,516	8.4
	Poslovi s nekretninama, iznajmljivanje		778,187,58.85	22,979,240	67,164,059.99	72,390,279.83	422,234,068.05	98,462,637	362,827,925.49	61,737,596	274,833,848.07	52,792,891	73,008
	Preradivačka industrija		75,668,48.55	27,535,814	626,327,678.86	98,869,012	3,597,347,515	67,005,776	257,183,537.82	52,162,945	205,964,958.15	35,705,070	75,810
	Proizv. el. energije, gasa i vode		2,505	55,914	579,799	317,121	535,036	286,621	1,830,896	236,427	424,348	97,333	4
	Ribarstvo		15,625	139,250	1,486,594	617,748	1,939,707	483,038	2,040,400	346,554	1,588,113	213,519	1.2
	Saobracaj, skladištenje i veze		12,593,205	6,288,172	36,586,548.18	71,072,627	750,657,44.68	55,115,255	746,559,108.49	47,768,677	257,762,468.85	33,739,167	72.6
	Trgovina na veliko i malo, opravka		7,063,705.94.57	20,627,847	5,076,826.72.33	7,075,963,67.82	398,468,882.96	71,424,240	278,300,884.84	58,985,804	226,306,667.63	40,455,882	75,528
	Vadjenje ruda i kamena		1,609,708	875,883	70,486,869	3,276,357	20,920,327	1,051,502	18,064,791	455,118	430,476	32,981	
	Zdravstveni i socijalni rad		125,000	41,106	705,710	97,247	425,754	59,339	319,935	34,780	187,991	18,103	
	UKUPNO:		666,700,343	115,595,814	3,521,678,651	649,241,408	1,878,168,383	486,422,743	1,694,212,510	373,997,165	1,531,878,258	294,625,878	989.7

Slika B.8. Primer pregleda planova otplate grupisanih po sektorima, ročnosti kredita i delatnosti dužnika

PREGLED KREDITA KOJE REGISTRUJE NBS									
Page Items:		Datum preseka: 31.12.2009		Vrsta kamatne stope: Fiksna kamatna stopa					
Grupa kredita	Zajam	Dužnik	Valuta	Broj okvirne prijave	Broj prijave	Datum ugovora	Ukupan iznos kredita	Dug po glavnici	Dospela nenaplaćena kamata
Pariski klub - konsolidovani kratkoročni krediti	PK 04 Japan METI_JPY 900,9mIn	Republika Srbija	JPY	700833	760834	31-JUL-2001	0	671,103,389	83,887,946
					780835	31-JUL-2001	0	229,494,807	28,685,609
					UKUPNO ZA OKVIRNU PRIJAVU:		0	900,598,196	112,573,554
	PK 04_Belgija EUR 4,76mIn (2001)	Dzavna zajednica SCO	EUR	700649	760649	31-JUL-2001	0	4,655,969	581,996
					780649	31-JUL-2001	0	99,454	12,432
					UKUPNO ZA OKVIRNU PRIJAVU:		0	4,755,422	594,428
	PK 04_Francuska COFACE EUR 4,15 mln (2001)	Dzavna zajednica SCO	EUR	700807	760809	31-JUL-2001	0	4,104,278	513,035
					780810	31-JUL-2001	0	46,828	5,854
					UKUPNO ZA OKVIRNU PRIJAVU:		0	4,151,106	518,888
	PK 04_Holandija EUR 0,72mIn (2001)	Dzavna zajednica SCO	EUR	700663	760663	31-JUL-2001	0	562,415	70,302
					780663	31-JUL-2001	0	155,689	19,461
				UKUPNO ZA OKVIRNU PRIJAVU:		0	718,103	89,763	
PK 04_Spanija EUR 0,05mIn (2001)	Dzavna zajednica SCO	EUR	-	760656	31-JUL-2001	0	47,948	5,993	
				UKUPNO ZA OKVIRNU PRIJAVU:		0	47,948	5,993	
PK 04_Srbija_Nemacka HERMES EUR 19,7mIn (2001)	Dzavna zajednica SCO	EUR	700687	760689	31-JUL-2001	0	16,828,161	2,103,520	
				780690	31-JUL-2001	0	1,677,827	209,728	
				790691	31-JUL-2001	0	1,152,771	144,096	
				UKUPNO ZA OKVIRNU PRIJAVU:		0	19,658,760	2,457,345	
Pariski klub - ODA krediti	PK 03 Italija Artigiancase 1,2mInUSD	Republika Srbija	USD	-	760873	28-DEC-2001	0	1,171,184	1,171,184
				UKUPNO ZA OKVIRNU PRIJAVU:		0	1,171,184	1,171,184	
PK 03 Italija Artigiancase 17,6mInEUR	Republika Srbija	EUR	700866	760868	28-DEC-2001	0	14,324,854	14,324,854	
				780869	28-DEC-2001	0	1,373,501	1,373,501	
				790870	28-DEC-2001	0	1,561,404	1,561,404	
				UKUPNO ZA OKVIRNU PRIJAVU:		0	17,259,759	17,259,759	
PK 03 Japan JIBC_JPY 1,97mIn	Republika Srbija	JPY	700849	760851	28-DEC-2001	0	1,855,123,366	1,855,123,366	
				790852	28-DEC-2001	0	115,356,386	115,356,386	
				UKUPNO ZA OKVIRNU PRIJAVU:		0	1,970,479,752	1,970,479,752	
PK 03_Amerika USAID USD 5,14 mln (2001)	Dzavna zajednica SCO	USD	700742	760744	28-DEC-2001	0	4,840,501	4,840,501	
				790745	28-DEC-2001	0	300,995	300,995	
				UKUPNO ZA OKVIRNU PRIJAVU:		0	5,141,496	5,141,496	
PK 03_Nemacka DEG EUR 0,59mIn (2001)	Dzavna zajednica SCO	EUR	700751	760752	28-DEC-2001	0	249,058	249,058	
				780753	28-DEC-2001	0	139,227	139,227	
				790754	28-DEC-2001	0	190,388	190,388	
				UKUPNO ZA OKVIRNU PRIJAVU:		0	578,673	578,673	
PK 03_Srbija_Nemacka KfW EUR 202,3mIn (2001)	Dzavna zajednica SCO	EUR	700683	760685	28-DEC-2001	0	190,430,327	165,430,327	
				790686	28-DEC-2001	0	11,841,452	11,841,452	
				UKUPNO ZA OKVIRNU PRIJAVU:		0	202,271,779	177,271,779	
Pariski klub - amba dužnik	PK 01 Japan METI IPV	Republika Srbija	IPV	700849	760849	28-DEC-2001	0	7,481,877,119	6,040,074,568

Slika B.9. Primer pregleda registrovanih kredita grupisanih po grupama kredita, šiframa zajmova, dužniku, valuti i okvirnoj kreditnoj prijavi

PREGLED STANJA DUGA PO BANAKARSKIM GRUPAMA						
Page Items:		Datum: 31.12.2009				
BANKARSKA GRUPA	NOVI NAZIV INOKREDITORA	NAZIV VESTE DUZNIKA	Broj prijave	Dug po glavnici u EUR	Dug po kamati u EUR	Iznos troskova u EUR
EFC	EFC Eurobank S.A.	Poslovna banka		1,285,717.22	4,714,621.36	0.00
		Preduzece		5,142,868.90	1,565,447.78	0.00
	EFC New Europe Funding B.V.	Preduzece		84,842,340.47	67,591,590.29	145,005.01
	EFC private bank (Luxemburg) s.a.	Preduzece		7,769,325.49	11,086,729.20	1,027,068.15
	Eurobank EFC Cyprus Limited, Kipar	Preduzece		15,858,890.58	26,214,878.76	80,630.24
	Eurobank EFC Factors S.A., Athens	Preduzece		953,829.43	633,386.18	0.00
	UKUPNO:			115,842,972.09	111,806,653.57	1,252,703.40
Hypo	HYP0 ALPE-ADRIA-BANK AD Podgorica	Preduzece		1,800,004.11	3,997,376.60	0.00
	Hypo Alpe-Adria-Bank, Klagenfurt	Preduzece		85,208,616.65	161,182,169.61	1,013,828.36
	Hypo Leasing doo Ljubljana	Preduzece		68,963.08	71,204.80	2,855.80
	Hypo group Netherlands finance B.V.	Poslovna banka		75,018,295.35	212,039,229.49	0.00
		Preduzece		219,463,211.58	262,996,461.48	345,418.80
	Hypo-alpe-adria-bank d.d. Zagreb	Preduzece		0.00	322,351.50	4,663.80
	UKUPNO:			381,549,090.78	640,608,793.48	1,366,766.76
Raiffeisen	Central Eastern European Finance Agency, Amsterdam	Preduzece		23,518,850.90	44,055,839.63	876,768.11
	RI Eastern European Finance B.V.	Preduzece		157,523,178.43	180,531,895.06	531,055.92
	Raiffeisen Bank Rt. Budimpesta	Preduzece		5,329,882.34	11,623,492.48	527,727.38
	Raiffeisen banka d.d. Maribor (stari naziv Raiffeisen Krehova banka d.d. Maribor)	Preduzece		1,987,284.34	1,758,450.61	925.84
	Raiffeisenbank Austria dd Podruznica Vinkovci	Preduzece		257,143.44	214,745.39	0.00
	Raiffeisenbank s.s. Prag	Preduzece		5,420,472.77	12,540,895.13	535,734.49
	Raiffeisen Zentralbank Osterreich, Wien	Preduzece	> 613976	3,620,479.45	4,413,851.97	49,074.56
			> 614354	1,170,002.67	617,234.81	381,479.43
			> 615495	282,637.79	340,425.64	0.00
			> 615546	570,229.87	698,252.81	0.00
			> 616198	36,765.02	69,016.82	0.00
Raiffeisenbank Austria dd, Zagreb	Preduzece		0.00	233,040.30	0.00	
	UKUPNO:			199,716,927.02	262,183,581.72	2,951,109.49
Alpha	ABC Factors S.A., Atina	Preduzece		103,958.31	94,071.23	0.00
	ALPHA Bank London Limited	Preduzece		12,028,479.03	10,388,003.24	76,937.86
	Alpha Bank A.E. Athens	Poslovna banka		30,342,926.50	31,355,604.30	0.00
	Alpha bank Cyprus Ltd, Nicosia	Preduzece		29,180,933.81	33,246,845.42	346,774.02
	Alpha bank s.s., Filijala u Londonu	Preduzece		61,418,821.88	58,109,518.93	1,143,498.54
		UKUPNO:			133,075,119.53	133,194,043.12
NLB	NLB Leasing" d.o.o. Ljubljana	Preduzece		23,801,197.26	29,846,964.23	0.00
	LJB Internationals Handelsbank A.G. Frankfurt/M	Preduzece		9,361,030.99	14,364,792.59	847,296.78

Slika B.10. Primer pregleda stanja duga na dan 31.12.2009. grupisano po bankarskim grupama, inokreditoru i dužniku

Dug Republike Srbije						
Page Items:		Datum preseka: 31.12.2009		Datum kursa: 31.12.2008		
Sektorska pripadnost	Grupa zajma	Sema kredita	Rocnost kredita	Dug - kursirano u EUR na datum preseka	Dug - kursirano u EUR na datum kursa	
Drzavni sektor	EIB_EUR 56,36mln_Reprogram 10 st	EIB - regulisani stari dug	Dugorocni krediti	13,582,655	13,582,655	
	PK_Rusija ostatak od USD 180 mln (Vlade - Stalne članice Pariskog kluba	Dugorocni krediti	3,277,167	3,343,237	
	SSSR-kredit od USD 750 mil.	Kliring	Dugorocni krediti	14,147,881	14,433,111	
	-	Centralne banke-kreditni zakljuceni pre 20.1	Dugorocni krediti	70,928,627	72,358,590	
	-	Kliring	Dugorocni krediti	5,570,262	5,682,562	
	-	Ostali strani kreditori_kreditni zakljuceni pre	Dugorocni krediti	688,230	702,105	
		UKUPNO ZA SEKTOR:			108,194,822	110,102,259
Drzavno garantovani	SSSR-kredit od USD 750 mil.	Kliring	Dugorocni krediti	1,264,858	1,290,358	
	-	EUROFIMA	Dugorocni krediti	16,416,955	16,372,929	
	-	IFC - neregulirani stari dug	Dugorocni krediti	98,421	100,406	
	-	Mesovite banke-kreditni zakljuceni pre 20.1	Dugorocni krediti	247,412	247,413	
	UKUPNO ZA SEKTOR:			18,027,646	18,011,106	
Negarantovani privatni	SSSR-kredit od USD 750 mil.	Kliring	Dugorocni krediti	778,000	793,685	
	-	Centralne banke-kreditni zakljuceni pre 20.1	Dugorocni krediti	282,064	282,064	
	-	IFC - neregulirani stari dug	Dugorocni krediti	3,501,924	3,506,333	
	-	Kliring	Dugorocni krediti	46,901,582	47,847,146	
	-	Komercijalne banke-kreditni zakljuceni pre	Dugorocni krediti	63,864,188	65,049,859	
	-		Kratkorocni krediti	4,783,488	4,830,167	
	-	Mesovite banke-kreditni zakljuceni pre 20.1	Dugorocni krediti	33,024,631	33,380,718	
	-		Kratkorocni krediti	12,959,323	13,101,684	
	-	Ostali strani kreditori_kreditni zakljuceni pre	Dugorocni krediti	127,997,613	129,857,692	
	-		Kratkorocni krediti	14,345,032	14,376,523	
	UKUPNO ZA SEKTOR:			308,437,846	313,025,872	

Slika B.11. Primer kursiranja podataka korišćenjem pomoćne tabele međuvalutnih odnosa

C. Bilansni podaci poslovnih banaka

Implementacija skladišta podataka u Sektoru za kontrolu banaka
Narodne banke Srbije

C.1. Projektovanje skladišta podataka

C.1.1. Dimenzije u skladištu podataka

Neke od dimenzija u ovom modelu opisane su detaljnije zato što su ili posebno značajne za koncept celog skladišta ili su specifične sa stanovišta same tehničke implementacije. Ukratko su pomenute one dimenzije koje su ili zajedničke sa sistemom za skladištenje koji je opisan u prvom primeru ili su veoma jednostavne strukture i lake za implementaciju. Neke od ovih dimenzija su uključene u skladište još 2007. godine i koriste se za analizu podataka iz svih sistemom obuhvaćenih kontnih okvira. Neke od dimenzija su uključene sa poslednjom izmenom uputstva i primenljive su samo na podatke koji se dostavljaju od 31.07.2008. godine. Osnovne dimenzije su:

- **Vremenska dimenzija.** Ovo je vremenska dimenzija koja je opisana u prvom primeru. Osim ova dva sistema za skladištenje u Narodnoj banci su razvijena još dva skladišta podataka koja dele istu dimenziju. U dimenziji se nalaze podaci o svakom kalendarskom danu od 01.01.2000. godine, ali su za ovo skladište podataka značajni samo poslednji dani u mesecu. Zato je u dimenziju uključen indikator, tj. atribut koji označava da li je dan poslednji u mesecu ili ne. On se koristi pri filtriranju podataka u svim onim sistemima u kojima podaci predstavljaju stanja poslovnih kategorija na poslednji dan u mesecu.
- **Dimenzija poslovne banke.** Ovo je takodje zajednička dimenzija za ovo skladište i skladište opisano u prvom primeru, ali i za još jedno skladište podataka implementirano u Monetarnom sektoru Narodne banke. Kako se skladište opisano u ovom primeru zasniva isključivo na podacima poslovnih banaka, to je ova dimenzija ključna u celom sistemu. Skoro sve analize koje se rade, rade se u kontekstu podataka iz ove dimenzije. U većini pregleda figurišu ili podaci iz ove dimenzije ili filteri postavljeni nad podacima iz ove dimenzije.

Vremenska dimenzija i dimenzija poslovnih banaka osvežavaju se u režimu koji je opisan u prvom primeru. Dimenzije koje se navode u nastavku su specifične samo za ovo skladište podataka. Ono što je karakteristično za sve te dimenzije je da su podaci u njima propisani uputstvom o evidentiranju i dostavljanju podataka o plasmanima, obavezama i potraživanjima banaka. Dakle, izmene nad podacima u ovim dimenzijama moguće su samo u slučaju promene kontnog okvira ili neke druge zvanične regulative. Zato su ove dimenzije inicijalno napunjene podacima za svaki kontni okvir i ne ažuriraju se dnevno. U ovom skladištu podataka osvežavaju se samo vremenska dimenzija, dimenzija poslovnih banaka i tabele činjenica.

Dimenzije karakteristične za ovo skladište su:

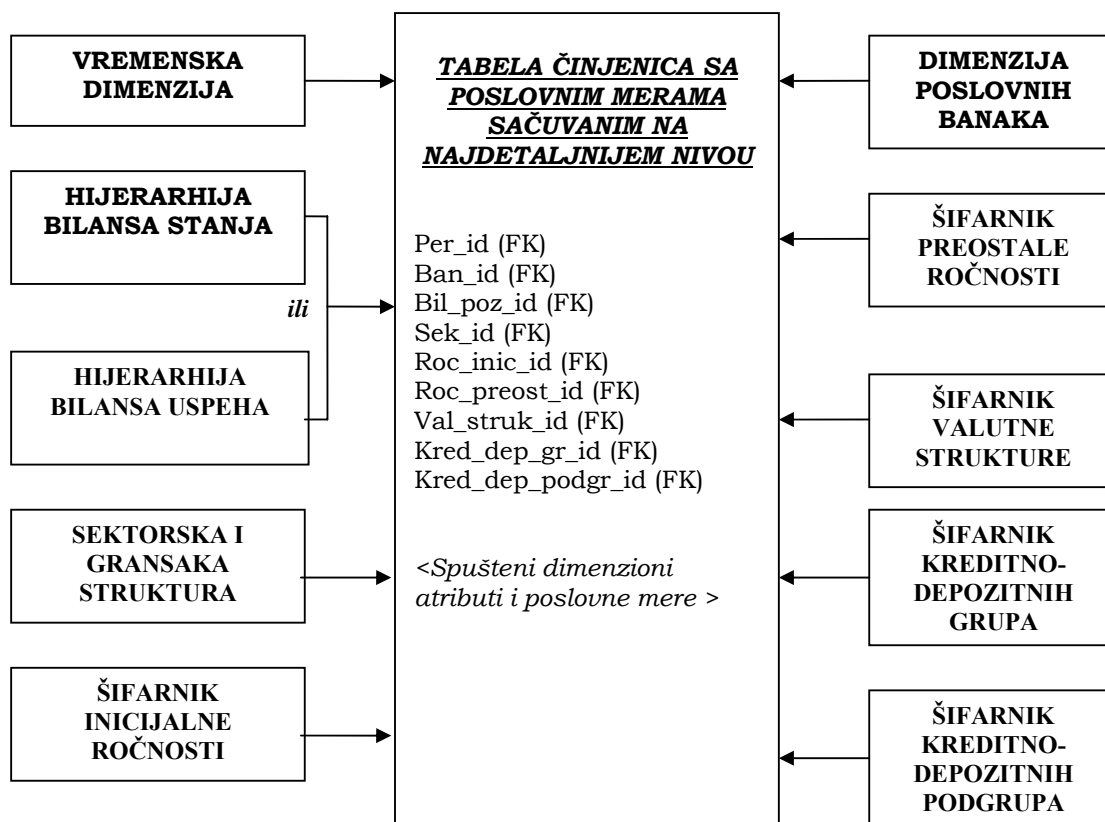
- **Hijerarhijski strukturiran bilans stanja.** Hijerarhijska struktura bilansa sadrži sve nivoe bilansa stanja od osnovnog, kao što su AKTIVA, PASIVA, VANBILANS itd. do najnižeg. Time se korisnicima omogućava bušenje podataka duž cele bilansne hijerarhije i to po svim granama bilansa. Metodologija hijerarhijskog formiranja bilansa stanja propisana je u Narodnoj banci tako da se svaka bilansna pozicija formira iz jednog ili više knjigovodstvenih računa. Svaka bilansna pozicija ima svoj jedinstveni identifikator, šifru i naziv. Svaki nivo na jednoj grani u bilansnoj hijerarhiji sadrži šifru i naziv bilansne pozicije, ali i jedinstveni identifikator pozicije koji jedinstveno određuje to mesto u hijerarhiji. Jedinstveni identifikatori pozicija koje se nalaze na najnižim nivoima hijerarhije spuštene su u vidu spoljnog ključa u tabele činjenica. Podaci na višim hijerarhijskim nivoima dobijaju se agregiranjem osnovnih podataka u samom upitu. Dimenzija sadrži onoliko baznih slogova koliko je moguće dobiti hijerarhijskih grana iz definicije bilansa. Hijerarhije nisu iste dubine. Jedna bilansna pozicija može da se nadje u jednoj i samo jednoj hijerarhijskoj grani. Tako se u svaki slog tabela činjenica koje su povezane sa ovom dimenzijom spušta samo jedan jedinstveni identifikator bilansne pozicije koja je poslednja na nekoj grani hijerarhije. Dimenzija sadrži podatke o strukturi bilansa za četiri kontna okvira koliko je obuhvaćeno skladištem podataka. Zato su u ovu dimenziju, kao i u dimenziju bilansa uspeha uključena dva datumska polja koja označavaju period važenja sloga. Pri punjenju tabela činjenica se, u ETL procedurama, ispituje datum važenja i zahvataju se samo bilansne pozicije koje su važeće na datum stanja podataka.
- **Hijerarhijski strukturiran bilans uspeha.** Analogno bilansu stanja uradjena je dimenzija bilansa uspeha. Ove dve dimenzije su sa stanovišta implementacije slične, ali su poslovno dve sasvim odvojene kategorije. Ne postoji nikakvo preklapanje između ova dva bilansa: pozicije su sasvim razdvojene i ne postoji knjigovodstveni račun koji se pojavljuje u oba bilansa. Analize podataka iz ova dva bilansa se, uglavnom, rade odvojeno. Za analize u kojima se ovi podaci kombinuju sa još nekim podacima iz sistema pod određenim uslovima, napravljene su posebne tabele činjenica sa izvedenim, visoko agregiranim poslovnim merama. Vremenska dimenzija, dimenzija poslovnih banaka, kao i obe pomenute hijerarhijske dimenzije uključene su u skladište podataka na samom početku implementacije i koriste se u analizama podataka iz svih kontnih okvira.
- **Šifarnici kreditno-depozitnih grupa i njihovih podgrupa.** Šifarnik kreditno-depozitnih grupa je jednostavne strukture. U njemu se nalaze šifre i nazivi poslovnih kategorija kao što su: Krediti, Depoziti, Hartije od vrednosti, Kamate i naknade itd. Svaka od ovih grupa formirana je na osnovu jednog ili više knjigovodstvenih računa. Preklapanje grupa ne postoji. U svaku tabelu činjenica sa kojom je povezan ovaj šifarnik spušten je jedinstveni identifikator sloga. Naknadno su u sistem uključene podgrupe ovih osnovnih grupa. Svaka od podgrupa obuhvata podskup knjigovodstvenih računa koji su definisani u okviru osnovne grupe. Kada se pojavila potreba da se grupe razvrstaju u

podgrupe i da se rade detaljnije analize podataka, najjednostavnije rešenje je bilo da se uključi nova dimenzija podgrupa i da se u tabele činjenica preko spoljnog ključa spusti i identifikator sloga nove dimenzije. Zahvaljujući proširljivosti modela uključivanje nove dimenzije je relativno brzo uradjeno. Tabele činjenica su proširene novim poljem koje je retroaktivno ažurirano vrednošću odgovarajućeg primarnog ključa iz nove dimenzije. Uradjene su i manje izmene u procedurama za osvežavanje tabela činjenica.

- **Šifarnik sektorske i granske strukture.** Ovo je jedan od značajnijih šifarnika jer se veliki broj analiza radi upravo u kontekstu ovih podataka. Šifriranje sektorske i granske strukture plasmana, obaveza i potraživanja propisano je uputstvom Narodne banke. Podele na sektore i grane, numeričke oznake i nazivi su striktno definisani. Osnovna sektorska podela obuhvata podelu na: bankarski sektor u zemlji, javni sektor, javna preduzeća, ostala preduzeća, stanovništvo, strana lica itd. Grane predstavljaju podpodelu osnovnih sektorskih grupacija. Grane su, takodje, striktno definisane kako šifarski tako i po nazivima. Tako se npr. bankarski sektor u zemlji deli na: domaće banke, centralnu banku, osiguravajuća društva, društva za upravljanje investicionim i penzijskim fondovima itd. Preduzeća, javna i ostala, dele se po granama delatnosti na: građevinarstvo, poljoprivredu, lov i šumarstvo, vadjenje ruda, hotele, zdravstvene ustanove i sl. Dimenzija sadrži sve propisane podele, sa šiframa i nazivima i ima oko 60 slogova. Svaki slog ima svoj jedinstveni identifikator. U dimenziju je, na kraju, ubačen i fiktivan slog sa sadržajem „Bez podele“ i to za sve one slogove u tabelama činjenica koji ne sadrže podatak o sektorskoj strukturi. Ovaj princip je zadržan i u dimenzijama koje su opisane u nastavku. Dok nije ubačen ovaj dodatni slog u dimenzije, prilikom izvršavanja upita u kome se tabela činjenica povezuje sa dimenzijom eliminisani su slogovi u kojima je vrednost odgovarajućeg spoljnog ključa null. Ovim rešenjem je izbegnuta potreba za spoljnim spajanjem pri izvršavanju upita, što se pokazalo kao dobro pa je primenjeno i u ostalim dimenzijama.
- **Šifarnik valutne strukture.** Ovo nije šifarnik valuta koji je napred pomenut, već šifarnik valutne strukture plasmana, obaveza i potraživanja, sa šiframa i nazivima koje su propisane uputstvom. Kako se valutna struktura posebno definiše za devizne, a posebno za dinarske plasmane (kao osiguranje od rizika), dogovoreno je da ovaj šifarnik obuhvati objedinjene devizne i dinarske podšifarnike uz uvodjenje dodatnog atributa koji sadrži identifikaciju plasmana (dinari ili devize). Time je olakšano često uporedno posmatranje dinarskih i deviznih plasmana po valutnoj strukturi.
- **Šifarnici propisanih ročnosti.** U okviru šifarnika ročnosti pominju se dva šifarnika. To su:
 - Šifarnik sa propisanim šiframa i nazivima inicijalne ročnosti plasmana, obaveza i potraživanja (to je ročnost obračunata od datuma plasmana do datuma dospeća).

- Šifarnik sa propisanim šiframa i nazivima preostale ročnosti plasmana, obaveza i potraživanja (to je ročnost obračunata od datuma stanja podataka do datuma dospeća).

Na slici C.1. dat je uprošćeni prikaz zvezdaste šeme u koju su uključene sve pomenute dimenzije. Neke od navedenih dimenzija su povezane sa većinom tabela činjenica u modelu. Ovo skladište podataka obuhvata najdetaljnije podatke, ali i neke složene izvedene, agregirane poslovne kategorije. Najdetaljniji podaci u modelu čuvaju se na nivou propisanog trocifrenog knjigovodstvenog računa i svih pomenutih dimenzija. Podaci koji su visokoagregirani čuvaju se na nivou datuma i poslovne banke. Podaci za ceo bankarski sektor ne nalaze se fizički sačuvani u bazi već se izvode agregiranjem u samim upitima, tako što se iz upita izostavi dimenzija banaka. Agregiranja po vremenskoj dimenziji se nikada ne sprovode. Podaci kojima se ovde raspolaže predstavljaju stanja na poslednji dan u mesecu i sve zbirne analize koje bi se iz toga izvodile na nivou kvartala ili godine su besmislene.



Slika C.1. Uprošćeni prikaz osnovnih dimenzija u skladištu

C.1.2. Tabele činjenica u skladištu podataka

Kod definisanja poslovnih mera i same strukture tabela činjenica pošlo se od toga da skladište podataka treba da sadrži najdetaljnije podatke, ali i izvedene, agregirane podatke i to do nivoa datuma i banke. Najpre su

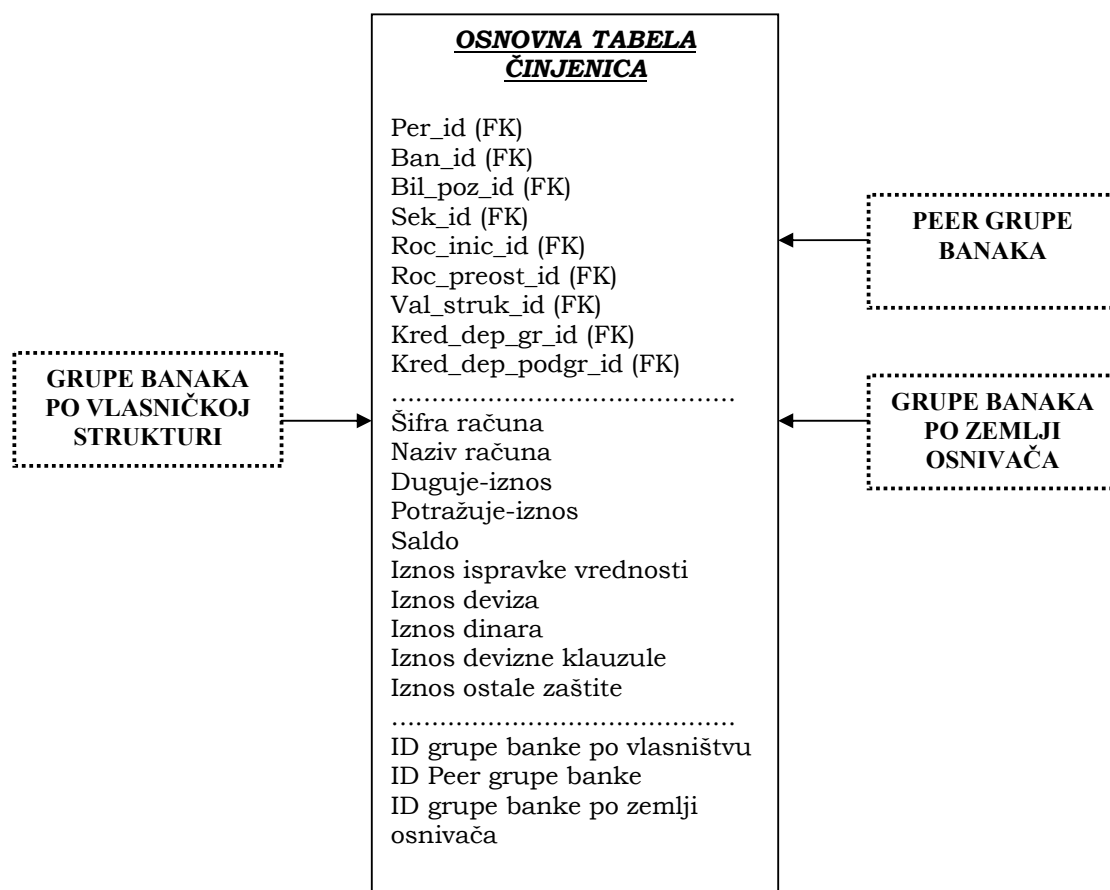
implementirane osnovne tabele činjenica sa detaljnim podacima, a tek u sledećoj fazi su iz njih izvedeni agregati.

Osnovna tabela činjenica u celom sistemu je tabela sa podacima o prometu i stanju na trocifrenim knjigovodstvenim računima banaka i to po kompletnoj analitici koja je propisana uputstvom Narodne banke. Šifarnik propisanih knjigovodstvenih računa u početku je uveden kao jedna od dimenzija sistema, ali je naknadno ova dimenzija izbačena a podaci o šifri i nazivu trocifrenog računa uključeni su u osnovnu tabelu činjenica kao izvedeni dimenzioni atributi. Analitiku trocifrenog računa određuju podaci o sektorskoj strukturi, inicijalnoj i preostaloj ročnosti i valutnoj strukturi što je obezbeđeno kroz dimenzije povezane sa tabelom činjenica. Jedno od pravila formiranja bilansa stanja i uspeha na osnovu knjigovodstvenih računa, i to u svim kontnim okvirima, je da se jedan račun može naći na jednoj i samo jednoj bilansnoj poziciji i to u samo jednom od ova dva bilansa. Tako može da se za trocifreni račun u svakom slogu tabele činjenica jednoznačno odredi jedinstveni identifikator odgovarajuće analitičke pozicije koja se nalazi na listu, tj. na poslednjem nivou neke bilansne grane. Taj jedinstveni identifikator bilansne pozicije je polje preko koga se slog u tabeli činjenica povezuje sa jednom od dve dimenzije bilansne hijerarhije. Zatim je definisana hijerarhija koja polazi od prvog bilansnog nivoa naniže i prelazi sa dimenzije na tabelu činjenica do trocifrenog računa. Ovim rešenjem je omogućeno bušenje duž bilansne hijerarhije i dalje spuštanje na nivo samog računa za najdetaljnije analize. Svaka analiza, detaljna ili opšta može da se obavi u kontekstu kompletne analitike preko povezanih dimenzija. Veza sa dimenzijama kreditno-depozitnih grupa i podgrupa implementirana je takodje na osnovu činjenice da se svaki trocifreni račun mapira za samo jednu od ovih kategorija. Ono što je postignuto je da sve grupe podataka (kreditni, depoziti, hartije od vrednosti, kamate, neknade i sl.) mogu da se raščlane na svoje podgrupe, da se po potrebi „spuste“ do osnovnog trocifrenog računa i analiziraju detaljno po sektorima, valutnoj strukturi, ročnosti itd.

Još jedna specifičnost koja se pojavila pri projektovanju ovog skladišta podataka bila je potreba da se uvedu tri tipa grupisanja banaka koji se ne koriste ni u jednom od ostalih skladišta podataka u Narodnoj banci. To su grupisanja: po vlasničkoj strukturi (na državne, privatne i strane banke), po tzv. peer grupama banaka (na grupu 1, grupu 2 i grupu 3) i grupisanje banaka po zemljama osnivača. Ovo su podele koje nisu od početka bile uključene u dimenziju banaka, a i nisu karakteristike od opšteg značaja za ostale sisteme. Zato je projektantska odluka bila da se u sve tabele činjenica uvedu tri nova dimenziona atributa u kojima bi se sa svakim presekom podataka za konkretnu banku u slogu čuvali njeni podaci o grupacijama. U početku su analize radjene preko ovih atributa u tabelama činjenica. Kasnije se, sa upotrebom skladišta, uvidelo da je bolje rešenje da se u sistem uključe tri nove minidimenzije u kojima se nalaze podaci o ovim grupama. Tako da su u skladištu koji je trenutno u upotrebi, osim navedenih dimenzija, prisutne i tri minidimenzije povezane spoljnim ključem sa tabelama činjenica. Uvodjenje dimenzija umesto atributa u tabelama činjenica bila je jednostavna akcija koja je završena u roku od dva dana, bez prekida ili ikakvog narušavanja rada skladišta.

Osnovna tabela činjenica ima oko 30 poslovnih mera. Neke od njih su zajedničke za sve kontne okvire, neke su važeće samo za podatke iz starih kontnih okvira, a neke su nove, ubačene sa poslednjom velikom izmenom 2008. godine. Jedine izvorne poslovne mere, dobijene direktno iz izvornih podataka, su podaci o dugovnom i potražnom prometu na računu. Sve ostale mere su izvedene jednostavnim aritmetičkim operacijama ili složenom ekonomskom logikom koja je ugrađena u ETL procedure. Inače, ETL procedure kojima se osvežava ova tabela činjenica, kao i sama metodologija formiranja bilansnih podataka, spadaju u najsloženije ETL procedure koje su implementirane u skladištima podataka u Narodnoj banci.

Na slici C.2. je prikazan deo osnovne tabele činjenica u kojoj su podaci o prometu i stanju na nivou trocifrenih računa. Posebno su označene tri naknadno uvedene dimenzije grupa banaka.



Slika C.2. Osnovna tabela činjenica sa najdetaljnijim podacima na nivou trocifrenog računa sa analitikom

Četri tabele činjenica koje su dalje opisane u ovom primeru su zasnovane na ovoj osnovnoj tabeli činjenica i iz nje se izvode. Samim tim, neophodan uslov za osvežavanje ovih tabela činjenica je da se osvežavaju u tačno propisanom

redosledu i to: najpre osnovna tabela činjenica, zatim tabela za uporedne analize i nakon njih ostale tabele činjenica bez propisanog reda.

Tabela za uporedne analize je karakteristična upravo za ovo skladište i to je bilo prvi put da je problem brzih uporednih datumskih analiza u skladištima podataka rešen na ovaj način. Analizom pregleda koji su deo standardnog izveštavanja u sektoru došlo se do zaključka da se u skoro svakoj analizi uporedno posmatraju i analiziraju podaci na: zadati datum, podaci na kraj prethodnog meseca, podaci na kraj prethodnog kvartala, podaci na kraj prethodne godine i podaci na kraj pret-prethodne godine u odnosu na zadati datum. Analize podrazumevaju međusobna upoređivanja, izračunavanja promena, porasta, indeksa i razlika među njima. Dakle, nije bilo dovoljno te podatke prikazati filtriranjem datumske dimenzije, već ih je trebalo tretirati kao izvedene poslovne mere koje mogu da se međusobno upoređuju. Ova tabela je povezana sa istim dimenzijama kao i osnovna tabela činjenica, što znači da sve uporedne analize mogu da se rade po svim aspektima propisane analitike. Tabela ne sadrži trocifreni račun i obuhvata samo odabrani deo poslovnih mera iz osnovne tabele koje se uporedno prate. To su: ukupan saldo, iznos dinara u ukupnom saldu, iznos deviza, iznos devizne klauzule i iznos ispravke vrednosti. U vreme kada je ovo skladište implementirano ovo je bilo najbolje rešenje za uporedne analize jer alati koji su bili na raspolaganju nisu omogućavali da se na drugi način dodje do vremenski distanciranih podataka. Čak i da su korišćeni takvi alati postavljalo se pitanje brzine izvršavanja upita. Činjenica je da su knjigovodstveni podaci jedni od najbrojnijih. Imajući u vidu detaljnost i složenost analiza koje se rade svako preračunavnaje u samom upitu verovatno bi bilo sporo. Ovako pripremljeni podaci sačuvani u bazi dali su kao rezultat brze upite i lako korišćenje.

Na slici C.3. dat je delimičan prikaz ove tabele činjenica.

<u>TABELA ZA UPOREDNE ANALIZE</u>
<Spušteni spoljni ključevi kao u osnovnoj tabeli činjenica>
.....
Bruto saldo na dan
Bruto saldo-prethodni mesec
Bruto saldo-prethodni kvartal
Bruto saldo-prethodna godina
Bruto saldo-pret-prethodna godina
Iznos deviza na dan
Iznos deviza-prethodni mesec
Iznos deviza-prethodni kvartal
Iznos deviza-prethodna godina
Iznos deviza-pret-prethodna godina
.....

Slika C.3. Delimični prikaz tabele činjenica za uporedne datumske analize po svim aspektima

Treća tabela činjenica koja je opisana u ovom primeru uvedena je da bi se omogućile analize podataka o kreditima i depozitima koje nisu mogle da se izvedu iz osnovne tabele činjenica. Ova tabela sadrži oko 70 poslovnih mera koje su izvedene iz osnovnih podataka o saldu na računima, sektorskoj i granskoj strukturi i iz pravila formiranja kreditnih i depozitnih grupa. Tabela se puni nakon osnovne tabele, jer se sve poslovne mere izvode iz podataka u osnovnoj tabeli činjenica a ne iz izvornih podataka. Tabela je agregirana do nivoa datuma i banke i sadrži podatke u hiljadama i milionima dinara. Pri punjenju tabele činjenica za svaku banku se radi upis sloga sa podacima u hiljadama i upis još jednog sloga sa podacima prikazanim u milionima. Ovo rešenje je trebalo da olakša izradu pregleda u kojima se veliki broj poslovnih mera prikazuje u jednoj ili drugoj mernoj jedinici. Poslovne mere iz ove tabele činjenica se, uglavnom, sabiraju, oduzimaju, dele, ili upoređuju korišćenjem drugih operacija, što je iz osnovne tabele činjenica bilo teško izvodljivo.

Na slici C.4. prikazan je samo deo poslovnih mera koje su uključene u ovu tabelu činjenica.

<u>KREDITI I DEPOZITI</u>
Per_id (FK)
Ban_id (FK)
Hiljade/Milioni indikator
.....
Ukupni krediti
Ukupni krediti-bankarski sektor
Ukupni krediti-.banke u zemlji
Ukupni krediti-.centralna banka
Ukupni krediti-strana lica
Ukupni depoziti u aktivni
Ukupni depoziti u pasivi
Ukupni depoziti-bankarski sektor
Ukupni depoziti-sva preduzeća
Kreditna zaduženja u inostranstvu
Kreditna zaduženja-strana fizička lica
Kreditna zaduženja-strane banke
.....
<Identifikatori grupa banaka >

Slika C.4. Deo tabele sa izvedenim podacima o kreditima i depozitima

Sledeća agregirana tabela činjenica sadrži podatke o nekoj vrsti statistike izvedene iz osnovnih podataka o bankama na dan. Podaci se takodje izvode iz osnovne tabele činjenica, pa je uslov da se i ova tabela činjenica osvežava nakon osnovne tabele. Struktura ove tabele činjenica se pokazala pogodnom u analizama u kojima je trebalo uporedno posmatrati podatke o bilansnoj sumi banke, ukupnim kreditima, vanbilansnoj aktivni, kreditima iz inostranstva ili depozitima. Ovo su poslovne kategorije izvedene i agregirane za potrebe konkretnih analiza. U tabeli se za svaku banku nalazi sačuvan neto iznos za svaku od ovih kategorija, rang banke po toj kategoriji na dan u

odnosu na ostale banke i iznos učešća banke u ukupnom iznosu bankarskog sektora za svaku kategoriju.

Na slici C.5. prikazan je deo tabele sa izvedenim podacima o učešću i rangiranju banaka po osnovu vrednosti određenih kategorija na dan.

<u>STATISTIKA BANKE NA DAN</u>
Per_id (FK)
Ban_id (FK)
.....
Iznos bilansne sume
Rang po osnovu bilansne sume
Ucesce banke u ukupnoj bilansnoj sumi
Iznos ukupnih kredita banke
Rang po osnovu iznosa kredita
Ucesce banke u ukupnim kreditima
Iznos vanbilansne aktive banke
Rang po osnovu vanbilansne aktive
Ucesce banke u ukupnoj vanbilansnoj aktivi
.....
<Identifikatori grupa banaka >

Slika C.5. Statistika po bankama-podaci o učešću i rangiranju banke u bankarskom sistemu na dan po osnovu vrednosti poslovnih kategorija

Kao poslednja tabela činjenica u ovom primeru opisana je tabela koja je poslednja implementirana. Ona je najobuhvatnija od svih tabela činjenica, što znači da su u njoj agregirani najraznovrsniji podaci sakupljeni iz osnovne tabele činjenica i tabele za uporedne analize. Tu su delom podaci iz bilansa stanja, a delom podaci iz bilansa uspeha. Deo podataka odnosi se na kredite i depozite. Pri izvodjenju nekih mera konsultuje se i sektor stanovništva. Neki od podataka posmatraju se samo na dan, a za neke su uključeni i podaci na poslednji dan prethodne godine da bi se pratile promene. Na osnovu podataka u ovoj tabeli rade se analize pokazatelja profitabilnosti po bankama i celog bankarskog sektora (uključujući i ROA i ROE pokazatelje). Neki od pokazatelja su izvedeni u ETL procedurama i nalaze se obračunati u bazi skladišta, dok se ostali izvode u samim upitima na osnovu pripremljenih pomoćnih podataka iz ove tabele činjenica. Tabela se osvežava na kraju ciklusa osvežavanja baze skladišta. Inicijalno, podaci u ovoj tabeli su bili agregirani do nivoa datuma. Kasnije se pojavila potreba da se ovi pokazatelji prate i na nivou svake pojedinačne banke pa je detaljnost podataka povećana i ponovljeno je punjenje kompletnog sadržaja tabele. Danas se podaci u ovoj tabeli čuvaju na nivou datuma i banke.

Na slici C.6. prikazan je deo ove tabele činjenica.

POKAZATELJI PROFITABILNOSTI

Per_id (FK)
Ban_id (FK)

.....
 Finansijski rezultat pre oporezivanja
 Neto prihodi od kamata
 Prihodi od kamata
 Rashodi od kamata
 Prihodi od naknada
 Prihodi od kursnih razlika
 Prihodi od kamate kod stanovništva
 Rashodi od kamate kod stanovništva
 Neto prihodi od kamata i naknada
 Iznos kamatonosne pasive
 Iznos akcijskog kapitala
 Prihodi od promene vrednosti
 Iznos obavezne rezerve
 Krediti dati stanovništvu
 Depoziti stanovništva

.....
 <Identifikatori grupa banaka >

Slika C.6. Tabela činjenica na osnovu koje se analiziraju pokazatelji profitabilnosti banaka i celog bankarskog sektora

C.2. Primeri izveštavanja

Na narednim slikama prikazano je nekoliko pregleda koji su uradjeni nad podacima iz ovog skladišta i koji su deo standardnih analiza u sektoru.

AKTIVA-BANKARSKI SEKTOR
-u milionima-

	31.12.2000					31.03.2001					31	
	IZNOS	Prošlog mesec	Godiš	Kraj godine	Godiš	IZNOS	Prošlog mesec	Godiš	Kraj godine	Godiš	IZNOS	Prošlog mesec
▶ GOTOVINA I GOTOVINSKI EKIVALENTI	88.009	770.925	199	0	0	73.563	1.092.636	119	222.246	101	85.227	1.075.483
▶ OPOZIVI DEPOZITI I KREDITI	91.134	1.970.785	81	0	0	101.452	1.375.104	129	230.135	135	145.081	2.155.133
▶ U dinarima	27.273	762.631	62	0	0	31.185	334.301	154	68.877	138	52.372	901.980
▶ U stranoj valuti	63.858	1.208.154	92	0	0	70.267	1.020.803	120	161.258	133	92.708	1.253.152
▶ DATI KREDITI I DEPOZITI	352.613	5.048.119	122	0	0	364.276	5.260.283	121	890.438	125	395.701	5.545.128
▶ HARTIJE OD VREDNOSTI	6.503	99.648	115	0	0	5.697	91.884	108	16.422	106	30.027	368.165
▶ UDELI (UČEŠĆA)	1.787	26.230	119	0	0	2.100	26.080	141	4.513	142	2.414	35.388
▶ OSTALI PLASMANI	12.408	216.900	100	0	0	13.268	196.516	118	31.334	129	13.894	198.204
▶ POTRAŽIVANJA ZA KAMATU, NAKNADU I PO OSNOVU PROMENE FER VREDNOSTI DERIVATA	3.217	51.687	109	0	0	4.345	63.771	119	8.123	163	4.454	75.696
▶ U dinarima	3.015	49.622	106	0	0	4.181	60.910	120	7.614	168	4.264	72.673
▶ Potraživanja za kamatu u dinarima	4.102	58.189	123	0	0	5.309	76.449	121	10.359	157	6.066	95.693
▶ Potraživanja za naknadu u dinarima	480	6.646	136	0	0	523	7.411	123	1.212	132	619	8.397
▶ Potraživanja po osnovu prodaje	34	419	144	0	0	39	599	115	87	139	41	726
▶ Potraživanja po osnovu promene fer vrednosti derivata	22	2.040	19	0	0	44	1.727	45	55	245	46	771
▶ Druga potraživanja	62	650	167	0	0	60	758	139	157	117	91	1.233
▶ Ispravka vrednosti potraživanja za kamatu i za naknadu, potraživanja po osn	-1.685	-18.322	161	0	0	-1.795	-24.035	120	-4.256	129	-2.600	-34.147
▶ U stranoj valuti	201	2.065	170	0	0	164	2.861	100	508	99	190	3.023
▶ NEMATERIJALNA ULAGANJA	2.886	28.411	128	0	0	2.154	38.642	123	5.267	125	2.352	31.225
▶ OSNOVNA SREDSTVA I INVESTICIONE NEKRETNINE	21.840	291.566	131	0	0	21.839	313.341	122	55.151	121	22.187	315.126
▶ STALNA SREDSTVA NAMENJENA PRODAJI I SREDSTVA POSLOVANJA KOJE SE OBUŠTAVLJA	71	2.144	58	0	0	71	977	128	179	122	155	2.149
▶ U dinarima	71	2.144	58	0	0	71	977	128	179	122	155	2.149
▶ ODLOŽENA PORESKA SREDSTVA	841	12.723	115	0	0	828	12.481	116	2.124	119	702	10.622
▶ OSTALA SREDSTVA	5.886	89.830	116	0	0	6.034	88.362	119	14.863	124	10.751	130.104
▶ U dinarima	4.994	77.390	113	0	0	5.499	74.160	129	12.612	133	7.633	98.516
▶ U stranoj valuti	892	11.461	136	0	0	535	14.202	66	2.252	73	3.118	40.588
▶ Potraživanja po osnovu avansa datih za obrtna sredstva u stranoj valuti	15	231	112	0	0	14	180	131	37	111	18	258
▶ Potraživanja po osnovu avansa datih za trajna ulaganja u stranoj valuti	4	55	118	0	0	4	58	121	9	131	4	53
▶ Potraživanja od zaposlenih u stranoj valuti	9	237	66	0	0	11	155	125	23	149	12	181

Slika C.7. Bilans stanja bankarskog sektora-aktiva, uporedno po datumima. Primer bušenja podataka duž bilansne hijerarhije

PREGLED UČEŠĆA GRUPA BANAKA PO VLASNIŠTVU
 Datum: 31.12.2008, 31.12.2009, 31.01.2010

	Bilansna smera			Vanbilansna aktivna				Ukupni krediti				Ukupni depoziti	
	31.01.2010	31.12.2009	31.12.2008	31.01.2010	31.12.2009	31.12.2008	31.01.2010	31.12.2009	31.12.2008	31.01.2010	31.12.2009		
Strane banke	73.80%	74.29%	75.34%	89.12%	88.22%	88.12%	75.05%	75.67%	75.23%	69.13%	70.36%		
Alpha Bank Srbija A.D. - Beograd	3.44%	3.41%	3.28%	5.05%	5.61%	4.81%	2.97%	2.76%	2.52%	3.89%	3.76%		
Banca Intesa A.D. - Beograd	14.03%	14.25%	14.08%	6.51%	6.98%	6.16%	15.70%	15.69%	15.00%	15.61%	16.01%		
Credit Agricole banka Srbija A.D. - Novi Sad	1.73%	1.85%	-	0.71%	0.86%	-	1.79%	1.72%	-	1.78%	1.81%		
Erste Bank A.D. - Novi Sad	2.33%	2.41%	2.77%	0.99%	1.15%	1.74%	2.57%	2.66%	3.28%	2.31%	2.35%		
Eurobank EFG A.D. - Beograd	5.90%	6.80%	6.98%	17.93%	13.52%	10.54%	6.11%	7.37%	7.84%	6.77%	8.16%		
Findomestic banka A.D. - Beograd	0.80%	0.80%	0.68%	0.41%	0.46%	0.41%	0.63%	0.61%	0.68%	0.41%	0.44%		
Hypo Alpe-Adria-Bank A.D. - Beograd	6.57%	6.44%	6.44%	5.55%	6.14%	5.77%	6.83%	6.69%	6.23%	4.85%	4.92%		
KBC banka A.D. - Beograd	0.80%	0.80%	0.92%	0.19%	0.20%	0.17%	0.89%	0.90%	1.02%	0.98%	0.97%		
Marfin Bank A.D. - Beograd	1.14%	1.10%	0.96%	0.46%	0.57%	0.74%	1.31%	1.26%	0.90%	0.78%	0.80%		
Meridian Bank - Credit Agricole Group A.D. -	-	-	1.99%	-	-	0.79%	-	-	2.02%	-	-		
Moskovska banka a.d. - Beograd	0.11%	0.11%	0.10%	0.05%	0.07%	0.04%	0.11%	0.10%	0.13%	0.09%	0.11%		
NLB Continental banka A.D. - Novi Sad	-	-	1.35%	-	-	0.41%	-	-	1.24%	-	-		
NLB LHB banka Beograd A.D. - Beograd	-	-	0.87%	-	-	0.56%	-	-	0.67%	-	-		
NLB banka Beograd A.D. - Beograd	2.06%	2.06%	-	0.98%	1.07%	-	1.81%	1.80%	-	2.44%	2.44%		
OTP banka Srbija A.D. - Novi Sad	2.11%	2.08%	2.72%	0.72%	0.79%	0.79%	2.12%	2.08%	2.77%	0.92%	0.93%		
Opportunity banka A.D. - Novi Sad	0.20%	0.20%	0.25%	0.01%	0.01%	0.01%	0.19%	0.19%	0.25%	0.08%	0.07%		
Parus Bank A.D. - Beograd	2.13%	2.08%	1.78%	0.81%	0.93%	2.02%	2.61%	2.55%	2.27%	1.83%	1.73%		
ProCredit Bank A.D. - Beograd	2.96%	3.01%	3.65%	0.53%	0.58%	0.55%	2.63%	2.66%	3.37%	2.90%	2.87%		
Raiffeisen Banka A.D. - Beograd	9.12%	8.96%	9.05%	20.41%	20.73%	21.11%	6.86%	6.88%	8.89%	7.67%	7.40%		
Societe Generale banka Srbija A.D. - Beograd	4.83%	4.62%	4.05%	4.74%	4.65%	4.54%	5.52%	5.38%	4.19%	3.48%	3.44%		
Unicredit Bank Srbija A.D. - Beograd	6.31%	6.28%	5.04%	7.40%	7.26%	6.73%	7.50%	7.52%	4.78%	6.50%	6.43%		
Vojvodanska banka A.D. - Novi Sad	3.95%	4.03%	4.88%	6.21%	6.51%	5.28%	3.64%	3.82%	4.07%	4.60%	4.58%		
Volskbank A.D. - Beograd	3.28%	3.20%	3.50%	9.47%	10.13%	9.95%	3.26%	3.23%	3.11%	1.24%	1.15%		
Drzavne banke	18.05%	17.48%	16.01%	6.45%	6.89%	6.79%	16.68%	16.10%	15.37%	22.00%	21.05%		
Privatne banke	8.36%	8.23%	8.68%	4.41%	4.89%	5.10%	8.27%	8.20%	9.42%	8.90%	8.69%		
Agroindustrijska komercijalna banka "AIK ba	5.16%	5.06%	4.70%	3.51%	3.89%	3.65%	4.91%	4.93%	5.45%	5.12%	4.90%		
JUBMES banka A.D. - Beograd	0.35%	0.35%	0.40%	0.29%	0.32%	0.43%	0.35%	0.36%	0.42%	0.18%	0.18%		
Kosovsko Metohijska banka A.D. - Zvečan	0.03%	0.03%	0.03%	0.00%	0.00%	0.00%	0.01%	0.01%	0.02%	0.03%	0.03%		
Metals banka A.D. - Novi Sad	-	-	1.05%	-	-	0.45%	-	-	0.80%	-	-		
Privredna banka Beograd A.D. - Beograd	1.32%	1.32%	1.07%	0.27%	0.33%	0.25%	1.31%	1.26%	1.04%	1.61%	1.61%		
Univerzal banka A.D. - Beograd	1.50%	1.47%	1.43%	0.34%	0.35%	0.32%	1.69%	1.64%	1.69%	1.95%	1.87%		

Slika C.8. Pregled učešća banaka u ukupnom iznosu za bankarski sektor po različitim kategorijama. Grupisanje banaka je uradjeno po vrsti vlasništva

KREDITI - SEKTORSKA STRUKTURA
 -u milionima-

age Items: Datum: 31.01.2010 : AKTIVA

	Opoziti krediti				Po repo transakcijama				Po transakcionim računima				Overnight		
	dinari	devize	d.klauz.	ost.zas.	dinari	devize	d.klauz.	ost.zas.	dinari	devize	d.klauz.	ost.zas.	dinari	devize	d.klauz.
Bankarski sektor u zemlji	2,682	0	0	0	152,967	0	0	0	84	0	0	0	1,833	16,771	0
Centralna banka	0	0	0	0	152,967	0	0	0	0	0	0	0	0	0	0
Domaće banke	2,682	0	0	0	0	0	0	0	33	0	0	0	1,833	16,771	0
Društva za osiguranje	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Društva za upravljanje investicionim fondovima i dobrovoljnim penzijskim fondovima	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Privredno društvo koje obavlja poslove finansijskog lizinga	0	0	0	0	0	0	0	0	36	0	0	0	0	0	0
Drugo pravno lice koje se pretežno bavi finansijskom delatnošću	0	0	0	0	0	0	0	0	14	0	0	0	0	0	0
Povezane finansijske organizacije koje ne ulaze u krug konsolidacije	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Povezane finansijske organizacije koje ulaze u krug konsolidacije	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Javna preduzeća	0	0	0	0	0	0	0	0	3,019	0	0	0	0	0	0
Druga preduzeća	0	0	0	0	0	0	0	0	16,633	0	0	0	-0	0	0
Poljoprivreda, lov, šumarstvo, vodoprivreda i ribarstvo	0	0	0	0	0	0	0	0	-11	0	0	0	0	0	0
Vadjenje ruda i kamena i preradivačka industrija	0	0	0	0	0	0	0	0	6,380	0	0	0	-0	0	0
Proizvodnja i snabdevanje električnom energijom, gasom i vodom	0	0	0	0	0	0	0	0	10	0	0	0	0	0	0
Gradjevinarstvo	0	0	0	0	0	0	0	0	1,703	0	0	0	0	0	0
Trgovina na veliko i trgovina na malo, opravka motornih vozila, motocikala i predmeta za ličnu upotrebu i domaćinstvo	0	0	0	0	0	0	0	0	6,421	0	0	0	-0	0	0
Hoteli i restorani, saobraćaj, skladištenje i veze	0	0	0	0	0	0	0	0	1,153	0	0	0	-0	0	0
Obrazovanje, zdravstvo i socijalni rad	0	0	0	0	0	0	0	0	49	0	0	0	0	0	0
Aktivnosti u vezi s nekretnima, iznajmljivanje i poslovne aktivnosti, ostale komunalne, društvene i lične uslužne aktivnosti	0	0	0	0	0	0	0	0	924	0	0	0	0	0	0
Povezana preduzeća koja ne ulaze u krug	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0

Slika C.9. Sektorska podela ukupnih kredita u aktivi za bankarski sektor na dan. Pregled je radjen po podgrupama kredita. Prate se iznosi dinara, deviza, devizne klauzule i ostale zaštite

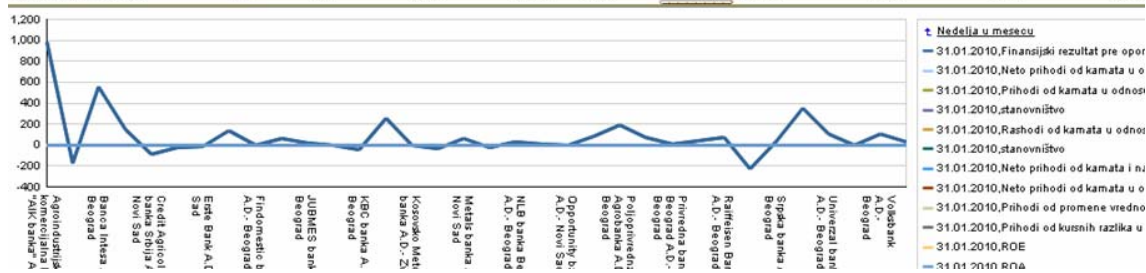
BILANS USPEHA - BANKARSKI SEKTOR
- u milionima -

	31.12.2008				31.12.2009				31.01.2010		
	Neto stanje	Kraj godine	Indeks /XII	Dinamika	Neto stanje	Kraj godine	Indeks /XII	Dinamika	Neto stanje	Kraj godine	Indeks /XII
PRIHODI	972,663	0	0	25,127	721,892	2,268,249	156	18,649	87,891	1,683,453	2
PRIHODI OD KAMATA	142,770	0	0	3,688	153,362	332,939	226	3,967	12,732	358,107	1
PRIHODI OD NAKNADA I PROVIZIJA	36,161	0	0	934	38,255	84,326	222	988	2,733	89,212	1
DOBICI PO OSNOVU PRODAJE HARTIJA OD VREDNOSTI PO FER VREDNOSTI KROZ BILANS USPEHA	88	0	0	2	187	205	446	5	15	435	1
DOBICI PO OSNOVU PRODAJE HARTIJA OD VREDNOSTI KOJE SU RASPOLOŽIVE ZA PRODAJU	832	0	0	21	193	1,941	49	5	14	451	1
DOBICI PO OSNOVU PRODAJE HARTIJA OD VREDNOSTI KOJE SE DRŽE DO DOSPEĆA	42	0	0	1	3	99	14	0	0	7	
DOBICI PO OSNOVU PRODAJE UDELA (UČEŠĆA)	120	0	0	3	67	280	117	2	0	156	
DOBICI PO OSNOVU PRODAJE OSTALIH PLASMANA	3	0	0	0	18	6	1,436	0	0	42	
PRIHODI OD KURSNIH RAZLIKA	527,621	0	0	13,630	272,054	1,230,413	108	7,028	41,422	634,420	3
PRIHODI OD DIVIDENDI I UČEŠĆA	896	0	0	23	55	2,089	13	1	0	128	
Ostali poslovni prihodi	3,324	0	0	86	2,879	7,730	182	74	129	6,713	
PRIHODI OD LIKIDANJA INDIKREKTNIH OTPISA PLASMANA I REZERVISANJA	70,743	0	0	1,828	95,465	164,973	283	2,466	4,825	222,623	1
PRIHODI OD PROMENE VREDNOSTI IMOVINE I OBAVEZA	190,064	0	0	4,910	159,153	443,220	176	4,112	26,022	371,149	3
RASHODI	941,082	0	0	24,311	703,687	2,194,603	157	18,179	85,175	1,640,999	2
RASHODI OD KAMATA	56,232	0	0	1,453	62,176	131,134	252	1,606	5,182	144,995	1
RASHODI NAKNADA I PROVIZIJA	8,306	0	0	215	9,334	19,371	241	246	765	22,233	1
GUBICI PO OSNOVU PRODAJE HARTIJA OD VREDNOSTI PO FER VREDNOSTI KROZ BILANS USPEHA	31	0	0	1	184	72	1,250	5	0	428	
GUBICI PO OSNOVU PRODAJE HARTIJA OD VREDNOSTI KOJE SU RASPOLOŽIVE ZA PRODAJU	213	0	0	6	4	497	4	0	0	10	
GUBICI PO OSNOVU PRODAJE HARTIJA OD VREDNOSTI KOJE SE DRŽE DO DOSPEĆA	0	0	0	0	0	0	0	0	0	0	
GUBICI PO OSNOVU PRODAJE UDELA (UČEŠĆA)	1	0	0	0	35	2	8,281	1	0	82	
GUBICI PO OSNOVU PRODAJE OSTALIH PLASMANA	-	-	-	-	53	0	0	1	0	124	
RASHODI PO OSNOVU KURSNIH RAZLIKA	575,985	0	0	14,880	309,453	1,343,198	113	7,994	60,312	721,644	4
RASHODI INDIKREKTNIH OTPISA PLASMANA I REZERVISANJA	94,454	0	0	2,440	128,292	220,267	285	3,314	6,590	299,177	1
Ostali poslovni rashodi	79,518	0	0	2,054	85,391	185,435	226	2,206	6,285	199,132	1
RASHODI OD PROMENE VREDNOSTI IMOVINE I OBAVEZA	126,341	0	0	3,264	108,566	294,628	180	2,805	5,942	253,175	1

Slika C.10. Bilans uspeha bankarskog sektora. Prikazan je prvi, osnovni nivo bilansne hijerarhije. Pregled je radjen nad tabelom činjenica za uporedne analize

Pokazatelji profitabilnosti bankarskog sektora
- po bankama -

	31.01.2010								
	Finansijski rezultat pre operacija	Neto prihodi od kamata u odnosu na bilansnu aktivu	Dobici od kamata u odnosu na stanovništvo	stanovništvo	Rashodi od kamata u odnosu na kamatno osiromašeno pasivo	stanovništvo	Neto prihodi od kamata u odnosu na operativne rashode	Neto prihodi od kamata u odnosu na operativne rashode	Dobici od promene vrednosti imovine i obaveza u uloznim prihodima
Agrarindustrijska komercijalna banka "AIK banka" A.D. - Niš	989	6.46%	4.70%	11.63%	2.39%	16.15%	80.69%	11.24%	12.58%
Alpha Bank Srbija A.D. - Beograd	-170	2.34%	3.10%	6.28%	1.76%	13.80%	14.28%	1.31%	10.64%
Banca Intesa A.D. - Beograd	552	5.91%	3.28%	5.90%	1.18%	8.10%	50.57%	6.06%	9.72%
Banka Fortis banka Srbija A.D. - Beograd	143	12.02%	6.79%	18.94%	0.52%	2.86%	30.79%	3.25%	0.26%
Credit Agricole banka Srbija A.D. - Novi Sad	-89	5.22%	3.32%	5.31%	1.26%	7.24%	16.11%	1.65%	7.24%
Croby banka A.D. - Kragujevac	-19	2.93%	4.54%	12.14%	1.65%	9.89%	10.23%	0.64%	6.00%
Erste Bank A.D. - Novi Sad	-12	5.62%	3.32%	5.33%	1.49%	9.94%	24.67%	2.76%	13.26%
Eurobank EFG A.D. - Beograd	132	5.30%	3.37%	5.72%	1.41%	11.68%	32.14%	3.76%	17.75%
Fininvest banka A.D. - Beograd	3	9.01%	5.08%	8.44%	1.76%	12.17%	21.39%	2.89%	11.71%
Hype Alpe Adria Bank A.D. - Beograd	62	3.85%	2.82%	3.88%	1.59%	13.40%	39.62%	5.07%	17.82%
JUBMES banka A.D. - Beograd	21	6.48%	2.70%	3.51%	0.32%	0.57%	19.01%	2.53%	2.80%
Jugobanka Jugobanka A.D. - Kosovska Mitrovica	-2	0.33%	0.11%	16.88%	0.00%	11.68%	17.28%	1.40%	0.00%
KB C banka A.D. - Beograd	-50	4.29%	2.68%	2.82%	1.16%	13.53%	11.81%	1.28%	24.14%
Komercijalna banka A.D. - Beograd	253	3.16%	2.97%	4.02%	1.47%	10.52%	25.11%	2.52%	17.21%
Kosovske Motobank banka A.D. - Zvečan	-3	5.29%	4.21%	2.85%	0.00%	0.00%	11.57%	0.98%	0.00%
Martin Bank A.D. - Beograd	-36	3.97%	2.71%	3.95%	1.33%	10.05%	13.72%	1.88%	19.54%



Slika C.11. Izvedeni pokazatelji profitabilnosti na dan po bankama, sa grafičkim prikazom. Pokazatelji se delom čuvaju u agregiranoj tabeli činjenica, a delom izvide na samom pregledu primenom propisane metodologije

LITERATURA

- [1] W.H.Inmon, *Building the Data Warehouse*, Wiley Computer Publishing, Third Edition, 2002
- [2] R.Kimball, M.Ross, *The Data Warehouse Toolkit*, Wiley Computer Publishing, Second Edition, 2002
- [3] C.Adamson, *Mastering the Data Warehouse Agregates*, Wiley Computer Publishing, 2006
- [4] W.H.Inmon, D.Strauss and G.Neushloss, *DW 2.0 The Architecture for the Next Generation of Data Warehousing*, Morgan Kaufmann, 2007
- [5] Oracle University, *Data Warehousing Fundamentals and Data Warehouse Database Design*, Oracle Corporation, First Edition, 2002
- [6] M. Ross, R. Kimball, *Differences of Opinion - the Kimball Bus Architecture and the Corporate Information Factory: What are the Fundamental Differences?*, <http://www.intelligent-enterprise.informationweek.com>, 2004
- [7] J.Poole, D. Chang, D. Tolbert and D. Mellor, *Common Warehouse Metamodel*, Wiley Computer Publishing, 2003
- [8] B.Ćirić, *Poslovna inteligencija*, Data Status, 2006
- [9] Veb stranica Narodne banke Srbije <http://www.nbs.rs>
- [10] Veb stranica Centralnog registra za hartije od vrednosti <http://www.crhov.rs>