

TALLINNA TEHNIKAÜLIKOOL  
Infotehnoloogia teaduskond  
Informaatikainstituut

IDK40LT

Joonas Nurk 095824IABB

**ANDMEAIDA ARENDUSE RESSURSSIDE  
OPTIMEERIMINE MÄLUPÕHISE  
ANDMEBAASI KASUTUSELEVÕTUGA  
ERGO INSURANCE SE NÄITEL**

Bakalaureusetöö

Juhendaja: Tarvo Treier

Tallinn 2016

## **Autorideklaratsioon**

Kinnitan, et olen koostanud antud lõputöö iseseisvalt ning seda ei ole kellegi teise poolt varem kaitsmisele esitatud. Kõik töö koostamisel kasutatud teiste autorite tööd, olulised seisukohad, kirjandusallikatest ja mujalt pärinevad andmed on töös viidatud.

Autor: Joonas Nurk

22.05.2016

## **Annotatsioon**

Käesoleva bakalaureusetöö eesmärgiks on tuvastada mälu põhise andmebaasi kasutuselevõtuga kaasnevaid eeliseid keskmise suurusega ettevõtte aruandluskeskkonnale. Fookuses on aruandluskeskkonda täiendavate arenduste kulude piiramine ning täieliku andmekogumi pealt aruandluse teostamine. Autor juhib IT projekti andmebaasi migratsiooniks ning kirjeldab andmebaasi muudatustega seotud tegevusi ja ohte.

Töö käigus luuakse mälu põhisel andmebaasil põhineva aruandluskeskkonna prototüüp käsitluse all oleva ettevõtte, ERGO Insurance SE andmete baasil. Samuti viiakse läbi ERGO Insurance SE olemasoleva andmeida migratsioon mälu põhisesse andmebaasi ning luuakse võimalused reaalsajas aruandluseks.

Töö tulemusena leiab kinnitust mälu põhise andmebaasi kasutuselevõtuga kaasnev arenduskulude kokkuhoid, aruandluskeskkonna kiiruse kasv ning võimalus reaalsaja andmetel baseeruvaks aruandluseks. Autori hinnangul on saavutatud tulemusi võimalik realiseerida ka teistes sarnase suurusega ettevõtetes.

Lõputöö on kirjutatud eesti keeles ning sisaldab teksti 31 leheküljel, 6 peatükki, 3 joonist.

## **Abstract**

### **The Optimization of Data Warehouse Development Resources by the Use of In-Memory Database in ERGO Insurance SE**

The objective of this bachelor thesis is to describe the benefits of utilizing an in-memory database for a reporting system in an average sized company. The main focus is set on limiting the costs of future developments of the reporting system and the possibility of creating reports based on the full scope of the data available.

The author describes the steps needed to migrate from the existing database system to an in-memory database system in SAP landscape. The steps are covered from the perspective of the project manager as the author of the thesis fulfilled the role of IT project manager during the process of migration.

The outcome of the project is the migration of ERGO Insurance SE current reporting system from Oracle database to SAP HANA in-memory database and creation of a new data warehouse. The outcome is achieved by collecting the requirements, creating a working prototype in the cloud environment, migrating the existing data warehouse and creating a new data warehouse using the capabilities of an in-memory database.

The author concludes that the migration is able to reduce long-term costs, provides performance increases and enables the end-users access to the full scope of the data collected by the resource planning system. The results of this thesis can be replicated in other similar sized companies.

The thesis is written in Estonian and contains 31 pages of text, 6 chapters, 3 figures.

## Lühendite ja mõistete sõnastik

ETL/ELT	Extract Transform Load / Extract Load Transform - andmete laadimine ühest süsteemist teise, mille jooksul viiakse andmed nõutud kujule.
ERP süsteem	Enterprise Resource Planning – Ettevõtte ressursside planeerimise tarkvara, mis katab ettevõtte vajadused alates tootearendusest ja laovarude haldamisest kuni tarneahela juhtimise ja raamatupidamiseni.
OLTP	Online Transactional Processing – OLTP viitab transaktsioonilistele süsteemidele, mille põhitööks on andmete sisestus ning kuvamine.
OLAP	Online Analytical Processing – OLAP viitab andmete analüüsile, mille käigus andmemudelite baasil tehakse äriolulist vajalikke järeldusi.
Tähtskeem	Star schema – Tähtskeem on andmete hoidmise loogika, milles kesksed faktitabelid viitavad dimensioonidele. Kasutatakse peamiselt analüütiliste süsteemide puhul.
SAP SE	Rahvusvaheline tarkvarafirma, mis loob ettevõtetele suunatud majandustarkvara. Peakontor asub Walldorfis, Saksamaal.
HANA	SAP SE poolt loodud mälupõhine andmebaasisüsteem

## Sisukord

1 Sissejuhatus .....	8
2 Kasutusel oleva aruandluslahenduse kirjeldus .....	11
3 Prototüübi implementeerimine .....	14
3.1 Prototüübi eesmärgid .....	15
3.2 Prototüübi teostus ja tulem .....	16
4 HANA kasutusele võtmine andmeaida andmebaasina .....	20
4.1 Andmeaida andmemahutude hindamine ja riistvara hange .....	20
4.2 Laadimiste migratsioon HANA-le .....	21
4.3 Universumite uuendus ja aruannete migratsioon .....	23
4.4 Lisaväärtus HANA migratsioonist .....	24
5 Reaalajas aruandlus .....	25
5.1 ERP andmete reaalajas laadimine HANA-sse .....	25
5.2 Universumi loomine normaliseeritud andmebaasile .....	26
5.3 Edasised sammud .....	27
6 Kokkuvõte .....	29
Kasutatud kirjandus .....	31

## **Jooniste loetelu**

Joonis 1. Praegune baltikumi aruandlusloogika .....	11
Joonis 2. Andmelaadimise loogika algsüsteemist andmeaita .....	13
Joonis 3. Ridade võrdluse kood .....	23

# 1 Sissejuhatus

Mistahes organisatsiooni tähtsaimaks varaks lisaks inimestele on 21. sajandil tõusnud andmed, mida organisatsioon valdab. Äärmuslike näidetena võib välja tuua tehnoloogiafirmad Facebook, Twitter või WhatsApp, mille kogu väärtus seisnebki andmetes, mida need enda kasutajate kohta omavad. Siiski ei ole tehnoloogiafirmad ainsad mille jaoks on kõikvõimalike andmete kogumine ning töötlemine kriitilised.

Pikalt olid kõikvõimalikes tegevusharudes tegutsevate firmade peamiseks fookuseks ERP (Enterprise Resource Planning – ettevõtte ressursside planeerimine) süsteemid. ERP süsteem on tüüpiliselt rakendus või grupp rakendusi, mis võimaldab ettevõttel koguda, säilitada ning hallata andmeid erinevate äritegevuste kohta – näiteks tootearenduse, tootmisressursside, turunduse, laovarude, tarneahela jne [1]. Ettevõtete jaoks oli esmatähtis majandada laovarused, hinnastamist ning igapäevast müügitööd ja seda nii tootmis- kui ka teenindussektoris. Samas, andmemahtude kiire kasvu tõttu raskenes ERP süsteemide abil organisatsiooniüleste järelduste tegemine, sest need süsteemid on optimeeritud igapäevaseks ressursside planeerimiseks mitte mahukaks aruandluseks.

Aruandlusvajaduste katmiseks on loodud ERP süsteeme toetama spetsiaalselt selleks optimeeritud aruandlussüsteemid, mida kutsutakse andmeaitadeks või andmevakkadeks. Andmeaitade ja andmevakkade eesmärk on koguda kokku üleettevõttelist või mõne osakonna vajadusi kattev andmekogum ning võimaldada selle baasil järelduste tegemist. Sarnaselt ERP andmetega hoitakse ka andmeaitade ja -vakkade andmeid relatsioonilistes andmebaasides, mis kasutavad andmete hoidmiseks kõvakettaid, kuid struktuuri optimeeritakse selliselt, et oleks võimalik suurte andmemahtude kiire presenteerimine.

Andmete kiireks kasutajatele kuvamiseks on seega vajalik saada andmed ERP süsteemist andmeaita või andmevakka, mis tähendab, et andmed on duplitseeritud ning ei pruugi olla igal ajahetkel mõlemas süsteemis identsed. Samuti võtab andmete liigutamine ühest süsteemist aega ning vajab ressursse, et andmelaadimise protsesse välja töötada ning hooldada.



Seoses tehnoloogia odavnemisega on aga tekkinud võimalus aeglasemate, kõvakettaid kasutavate andmebaaside kõrval võtta kasutusele mälu põhiseid ehk serveri mälu andmete hoidmiseks kasutavaid andmebaase, mis on kordades kiiremad. Kas selline tehnoloogia võiks tekitada võimalust ERP süsteeme ja andmeaitu kuidagi ühildada? Kas võiks äkki pääseda andmete laadimisest ühest süsteemist teise ja saada aruandeid koostada hoopis otse ERP süsteemist? Kas aruandlus muutuks sellest kiiremaks ja/või odavamaks?

Antud bakalaureusetöö eesmärk on uurida ERGO Insurance SE näitel eeliseid, mida annaks mälu põhise andmebaasi kasutuselevõtt aruandluskeskkonnale. Hüpoteesidena on tõstatatud järgnevad potentsiaalsed eelised:

- Aruandluskeskkonna arenduste kulude langemine
- Aruandluse kiiruse kasv
- Andmelaadimiste ning andmete liiasuse elimineerimine
- Võimalus ERP andmete baasil reaalaajas aruandluseks

Esmalt on plaanis luua mälu põhisel andmebaasil põhinev prototüüp, mille abil saaks tuvastada mälu põhise andmebaasi sobivus ERGO Insurance SE jaoks. Prototüübi loomise käigus peaks selguma andmemahu vajadused ning uue loogika alusel ehitatud andmeida kasutusmugavus.

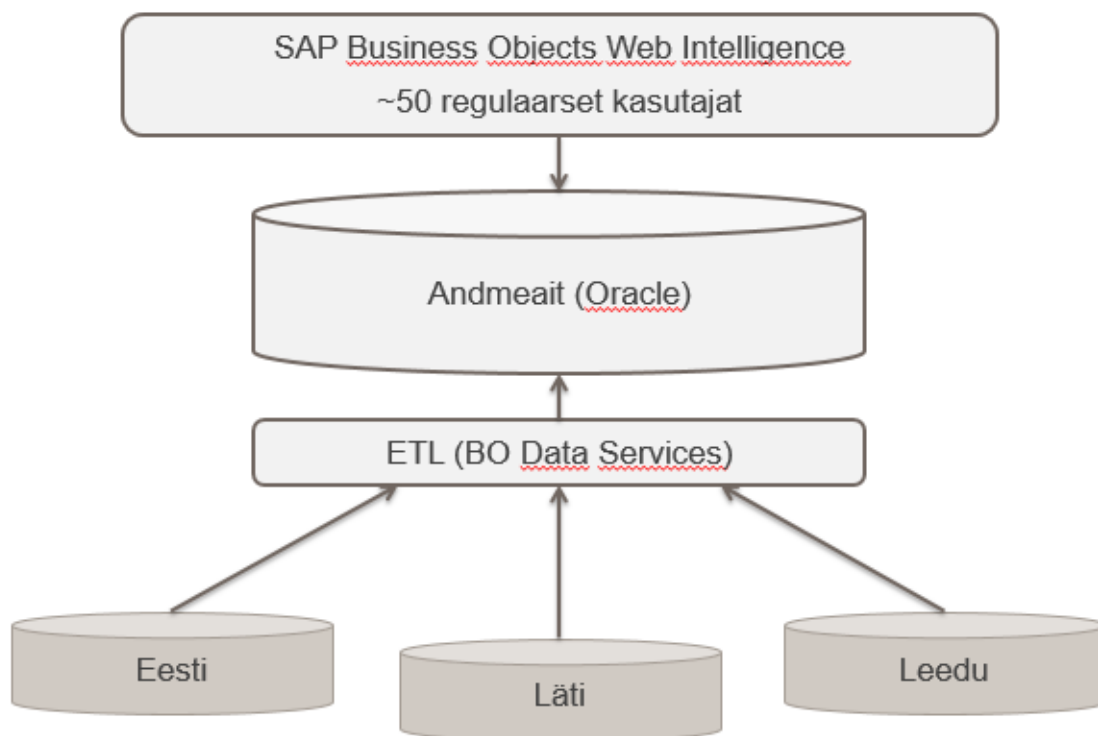
Prototüübi loomise käigus kogutud andmete baasil peaks saama otsustada ressursside üle, mis on vajalik andmeida migreerimiseks olemasolevast andmebaasist mälu põhisesse andmebaasi. Migratsiooni teostamise eelduseks on vajaduste kaardistamine ning hankeprotsessi läbi viimine. Mälu põhise andmebaasi kasutusele võtuga on võimalik anda vastused eelnevalt püstitatud küsimustele. Selle kasutuselevõttuga peaks vähenema arenduskulud, paranema aruandlussüsteemi kiirus ning lõppeesmärgina peaks olema võimalik aruandlus reaalaaja andmete baasil.

ERGO Insurance SE, olles ülebaltiline ettevõtte, suurus võimaldab antud töö autori hinnangul antud töö järeldusi üle kanda ka teistesse keskmise suurusega Eesti või Balti ettevõtetele. Projekti erinevate sammude teostamisel osaleb antud töö autor IT projektijuhi rollis. Töö koosneb kuuest peatükist - lisaks sissejuhatusele ja kokkuvõttele on 4 sisulist peatükki.

- Peatükk 2 kirjeldab ERGO Insurance SE aruandluskeskkonda ning tutvustab tehnilisi termineid. Tegemist on taustainfoga, et järgnevates peatükkides kirjeldatud tegevused oleksid sisuliselt arusaadavad.
- Peatükis 3 kirjeldatakse prototüübi teostamise protsessi, mille eesmärk on tuvastada mälu põhise andmebaasi võimekust normaliseeritud andmete baasil aruandluse teostamiseks.
- Peatükk 4 kirjeldab mälu põhise andmebaasi kasutusele võtmise protsessi. Antakse ülevaade hankeprotsessist ning olemasoleva lahenduse migratsioonist, mis on eelduseks projekti lõppeesmärkide saavutamiseks.
- Peatükis 5 kirjeldatakse reaalajas aruandluse realiseerimist, mis on üks antud töö põhieesmärke. Sellega seotult tuuakse välja arenduskulude ning aruandluskeskkonna kiiruse muutumine.

## 2 Kasutusel oleva aruandluslahenduse kirjeldus

ERGO baltikumi aruandluskeskkonnal on täna suurusjärgus 50 regulaarset kasutajat. Tarkvarana kasutatakse SAP BusinessObjects (edaspidi BO) tooteperekonda kuuluvat Web Intelligence lahendust, mis võimaldab andmetele ligipääsu nii töölauakliendi kui ka veebiliidese abil. Aruandluskeskkonna andmed asuvad Oracle andmebaasis kuhu need laetakse algsüsteemidest BO Data Services ETL tööriistaga Eesti, Läti ja Leedu algsüsteemidest (Joonis 1).



Joonis 1. Praegune baltikumi aruandlusloogika

Antud projekt ei hõlma kõigi kolme riigi erinevaid algallikaid vaid skoopt on võetud ainult peamine Eestis kasutusel olev ERP süsteem. Joonis 2 kirjeldab detailsemalt andmelaadimise protsessi ning aruandluse loogikat.

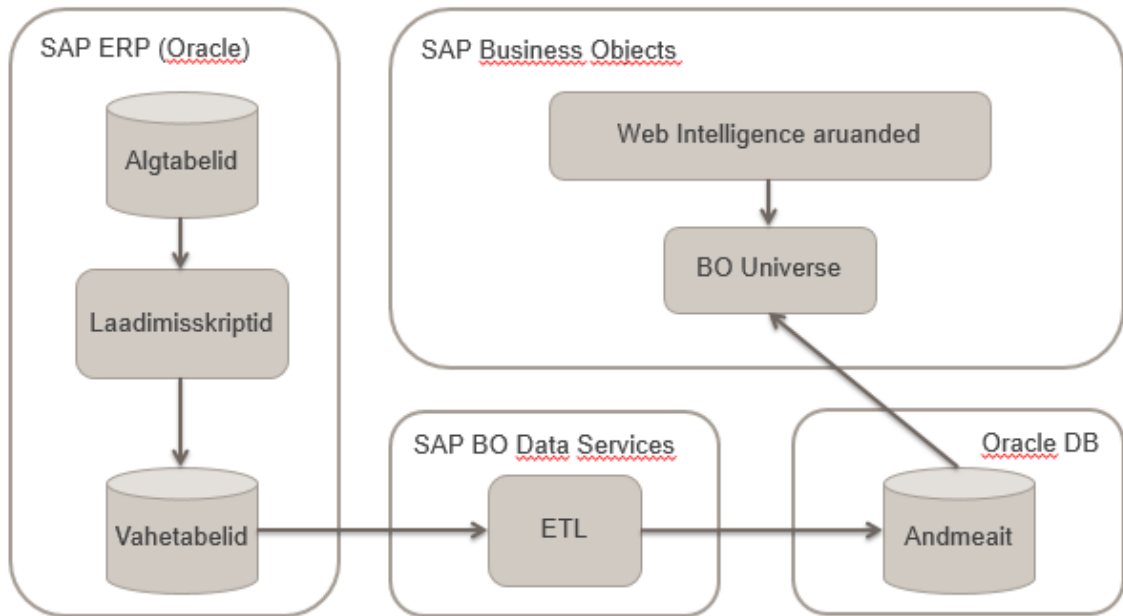
Algallikaks on SAP ERP süsteem, mis hoiab kahjukindlustusega seotud andmeid. Sinna hulka kuuluvad lepingute, klientide ja finantsandmed ning sellega nimekiri ei piirdu. ERP süsteem kasutab Oracle relatsioonilist andmebaasi ning andmed on normaliseeritud kujul. Selliselt on täidetud transaktsioonilise ehk OLTP (Online Transactional Processing)

süsteemi nõuded kõikvõimalike äriliste operatsioonide kiireks läbi viimiseks, mille käigus sisestatakse andmebaasi kirjeid [2]. Samas tekitab selline normaliseeritud andmete hoidmine probleeme suurte andmemahude kuvamisel aruandluseks.

Et viia andmeid kujule, mis on paremini sobilikud analüütiliseks ehk OLAP (Online Analytical Processing) [3] süsteemis viiakse praeguses süsteemis läbi mitmeid ETL (Extract Transform Load) [4] protsesse. Esimene selline protsess toimub ERP süsteemi siseselt kus analüüsiks vajalikud andmed laetakse vahetabelitesse. See protsess toimub üks kord öö jooksul, kui süsteem on kasutajale suletud, et erinevate tabelite vahel ei tekiks ebakõlasid. Laetakse vaid *delta* andmed ehk andmed, mis on süsteemi lisatud peale viimast laadimist. Vahetabelid kasutavad sama andmeruumi, milles asuvad ka ERP transaktsioonilised andmed.

Järgneva sammuna laetakse andmed ERP süsteemi vahetabelitest BO Data Services ETL vahendi abil andmeaida tähtskeemi. Tähtskeem on andmete andmebaasis hoidmise loogika, milles kesksel kohal asuvad niinimetatud faktitabelid, milles asuvad faktid ehk agregeeritavad andmed ning viited dimensioonidele, mis fakte detailandmetega täiendavad [1]. Näiteks kliendi lepingu number ja summa asuvad faktitabelis koos viidetega kliendi, müüja ja lepingu objekti viidetele. ETL protsessi abil viiakse niisiis andmed normaliseeritud kujult tähtskeemi denormaliseeritud kujule. Seda kõike selleks, et aruandlus oleks kiirem, kuigi sellega tekib andmete liiasus ning andmemahu kasv. Tehniliselt asuvad andmeaida andmed endiselt Oracle andmebaasis, aga erinevas skeemis.

Tähtskeemis hoitavatele andmetele pääseb lõppkasutaja ligi BO Web Intelligence töövahendi abil. Aruannete loomiseks on loodud semantiline kiht, mis kirjeldab tähtskeemi ärikasutajale arusaadavate terminitega ning võimaldab lõppkasutajal ilma tehnilisi teadmisi omamata aruandeid luua.



Joonis 2. Andmelaadimise loogika algüsteemist andmeaita

### 3 Prototüübi implementeerimine

Reaalsetele aruandlusvajadustele vastava süsteemi hindamiseks on otsustatud luua mälu põhise aruandluskeskkonna prototüüp. Prototüüp peab katma müügiaruandluse aspekte, mis on olemasolevas aruandluskeskkonnas puudu või puudulikud.

ERGO Insurance SE kasutab nii ERP kui ka aruandluskeskkonnana Saksa tarkvarafirma SAP SE tooteid. Vastavalt SAP R3 ning SAP BusinessObjects tooteperekonda kuuluvat tarkvara. Sellele baseerudes on võetud vastu otsus katsetada mälu põhise lahenduse realiseerimist SAP HANA riistvaral.

SAP HANA puhul on tegemist mälu põhise andmebaasisüsteemiga, mis on loodud tarkvarafirma SAP SE poolt. HANA andmebaasi eripäraks on selle optimeeritus andmete hoidmiseks vahemälu, mitte kõvaketastel olevates andmeplokkides. Samuti kasutab andmebaasisüsteem mitmeid andmete kompressioonimeetodeid selleks, et andme maht oleks võimalikult väike. Andmebaasisüsteemi loomisel on arvestatud nii OLTP kui ka OLAP vajadustega. Näiteks võimalus hoida andmeid veerupõhistes tabelites võimaldab kiirendada suuri analüütilisi päringuid [3].

Tehnilisest küljest on toetatud ühendused SAP ERP ja BusinessObjects tarkvaraga ning olles SAP klient on litsentsitasud mõistliku hinnaga. Eelkõige viimasele punktile baseerudes ei olnud konkureerivaid mälu põhiseid lahendusi mõtet kaaluda, sest kõikvõimaliku riistvara pideva odavnemise valguses omavad litsentsitasud aina suuremat kaalu ning potentsiaalne kokkuhoid on märgatav.

SAP HANA andmebaasisüsteemi eripärade tõttu võib andmebaasisüsteemi kasutada ainult serveritel, mis on SAP SE poolt sertifitseeritud. Ei piisa suvalisele serverile piisava mälu lisamisest, vaid HANA kasutuselevõtt eeldab spetsiifilise riistvara soetamist. Selline lahendus garanteerib HANA kasutajatele SAP poolt pakutud võimekuse.

Prototüübi teostamiseks on kaalutud erinevaid variante. Riistvara ning litsentsi soetamine on projektimeeskonna hinnangul esmaseid eesmärke arvestades liiga kulukas olgugi et neid saaks hilisema realisatsiooni puhul edasi kasutada. Riistvara ning litsentside soetamise puhul on kriitiline teada vajaminevat mälumahtu ning seda hinnangut soovib projektimeeskond tuvasta personaliseeritud prototüübiga, mitte otsustada kalli riistvara ja litsentside soetamist puhtalt reklaamtekstide põhjal.

Prototüübi loomise hetkel, kevadel 2015, ei ole Eesti teenusepakkujatel mitte ühtegi SAP HANA installatsiooni, mida oleks võimalik prototüübi loomiseks kasutada ning seetõttu on lahenduseks valida erinevate rahvusvaheliste teenusepakkujate vahel. Koostööpartnerite soovitusel on kaalumisel järgnevaid teenusepakkujaid:

- Cloudshare
- Amazon Web Services

Hinda arvestades langeb esmane otsus esimese valiku ehk Cloudshare kasuks. Projektimeeskonnal puudub varasem kogemus HANA-ga töötamisel ning sisulist hinnangut erinevatele pilveteenustele on väga keeruline anda. Töö käigus ilmneb, et odavam teenus ei suuda pakkuda kogu vajaminevat funktsionaalsust ning töö kolitakse ümber Amazon Web Services poolt pakutud lahenduse peale.

Pilveteenuse kasutamise eeliseks on lihtsus ja odavus. Süsteem on juba eelnevalt seadistatud ja töökorras ning lõppkliendil on vaja teha vaid valik mälumahu ja vajadusel spetsiifiliste moodulite seast, mis ei kuulu põhipaketti. Peale valikute langetamist on süsteem tunnipõhise hinnaga kasutatav.

### **3.1 Prototüübi eesmärgid**

Prototüübi loomise eesmärgiks on tõestada järgnevalt nimetatud aspektid reaalsete andmete baasil. Eesmärgid on seatud küllaltki „pehmed,“ sest nii ERGO-l kui ka koostööpartneril puudub varasem põhjalik kogemus mälupõhiste andmebaasidega ning seetõttu on väga raske seada numbriliselt mõõdetavaid tulemeid. Positiivseks on plaanis lugeda kõiki tulemusi, mis oleks olemasolevast lahendusest mingilgi määral paremad ning annaks kinnituse, et mälupõhise andmebaasi reaalne kasutuselevõtt on õigustatud.

- Andmeaida ETL arenduse kulu vähenemine
- Andmeaida ETL protsessi lihtsustumine
- Andmete jalajälje vähenemine SAP poolt lubatud 7-10 korda
- Aruannete kiirem avanemine aruandluskeskkonnas

Plaan on loobuda algtabelitest andmete liigutamine vahetabelitesse ning sealt ETL abil andmeaita. Selle asemel liigutatakse andmed algtabelitest otse andmeaita ilma kõikvõimalike transformatsioonideta. Ainukene erinevus baaside vahel on selles, et

algtabelid on Oracle baasis reapõhiselt, aga andmeaida andmed on SAP HANA baasis veerupõhiselt. Selline protsess peaks kõigi eelduste kohaselt andmeaida arendust lihtsustama ning muutma seda ka soodsamaks.

Andmete hoidmine veerupõhiselt peaks SAP hinnangul vähendama andmete jalajälge 7-10 korda. See aspekt on kriitiline riistvara soetamisel. HANA riistvara hind on suurusjärgus 100 000 eurot<sup>1</sup> terabaidi eest, aga Oracle andmebaaside puhul räägime 15 000 dollarist [2] terabaidi eest. Eelnevast tuleneb, et hinnalt on SAP HANA mälu põhise andmebaasi riistvara iga bait pea 7 korda kallim kui Oracle relatsioonilise andmebaasi riistvara. Seega üks ühele andmebaasi vahetus ei tuleks enamustes organisatsioonides eelarveprobleemide tõttu kõne alla kui ei oleks SAP poolt reklaamitud andmemahu vähenemist. Prototüüp peab tõestama, et kas olukord on ka reaalse andmete puhul nii positiivne nagu reklaamitakse.

Aruannete kiirust on plaanis mõõta pisteliselt. Kasutajatelt palutakse subjektiivset hinnangut kasutuskogemuse kohta ning lisaks sellele mõõdetakse HANA andmebaasi aruannete kiirust võrreldes Oracle andmebaasil baseeruvate aruannetega.

### **3.2 Prototüübi teostus ja tulem**

Prototüübi realiseerimine koosneb järgnevatest sammudest:

- Skoobi defineerimine ja ärianalüüs
- Tehniline analüüs
- Riistvara valik ja seadistamine
- Prototüübi loomine
- Prototüübi tehniline testimine
- Prototüübi äriline testimine

Skoobi defineerimiseks kaasatakse äritelliija kes hiljem ka vastavalt skoobile arendust testib. Antud projekti raames on skoobiks alamosa müügiaruandlusest, mis võimaldaks hinnata uue lahenduse kiirust ilma liigse keerukuseta. Kõikvõimalikud agregeerimised

---

<sup>1</sup> HANA riistvara hinnad ei ole avalikult veebist leitavad, vaid põhinevad individuaalsetel pakkumistel. ERGOle tehtud pakkumiste konkreetseid summasid ei ole siinkohal võimalik avaldada.



ning arvutused on jäetud kõrvale ning kasutatakse vaid andmebaasis olevaid algandmeid. Seega mahukaks ärianalüüsiks vajadus puudub ning kirjeldatakse vaid nõuete kogumik, mille baasil hiljem äritesti teostada.

Tehnilise analüüsi käigus tuvastab projektimeeskond tabelid ning nendevahelised ühendused, mida prototüübi teostamiseks vaja läheb. Kõikvõimaliku seadistustega seotud info jäetakse kõrvale ning valimiks on vaid tabelid, mis sisaldavad lepingute, klientide ning kahjude detailset infot. Antud prototüübi puhul on vajalikud 22 tabelit, mille maht on Oracle relatsioonilises andmebaasis 15 gigabaiti.

Tehnilise analüüsi järgneva sammuna vaadeldakse võimalusi andmete liigutamiseks uude andmebaasi ehk Oracle andmebaasist SAP HANA peale. See sõltub otseselt uue riistvara valikust ja seadistusest. Antud prototüübi puhul on valitud Amazoni pilveteenus ning seega on vaja planeerida ühendused väljaspool kohtvõrku asuva serveriga. Nagu eespool mainitud, siis pilveteenus on eelseadistatud ning valida on vaja ainult mälu maht, sest Oracle andmebaasist SAP tarkvaraga andmeid laadides lisavõimalusi tellida ei ole vaja.

Prototüüpi arendava partneri soovitusel valitakse väikseima mahuga Amazon Web Services SAP HANA One pakett, mille mälumahuks on 60 gigabaiti [3]. Sellest 15 gigabaiti on kasutatud HANA rakenduse installatsiooniks ning ülejäänud mälu on vaba andmete jaoks. Arvestades, et tehnilise analüüsi käigus valitud tabelite maht on 15 gigabaiti, siis peaks antud mälumahust piisama nii andmete hoidmiseks kui nendega kalkulatsioonide tegemiseks ja aruannete loomiseks.

Andmete liigutamiseks Oracle andmebaasist HANA mälusse kasutatakse ERGO poolt kasutusel olevat, lokaalsele rakendusserverile installeeritud SAP BusinessObjects (edaspidi BO) Data Services ETL programmi. BO Data Services on mõeldud täislahendusena andmete algsüsteemist võtmiseks, keerukate transformatsioonide tegemiseks ning sihtandmebaasi paigutamiseks, aga antud juhul kasutatakse seda andmete transformatsioonita ühest süsteemist teise liigutamiseks. Erinevuseks jääb, et Oracle andmebaasis hoitakse andmeid reapoõhiselt kõvaketastel, aga HANA puhul veerupoõhiselt mälus. Andmete liigutamise kiirusele tähelepanu ei pöörata, sest see ei ole prototüübi puhul primaarne ning välisühendust mõjutavad välised tegurid, mida sisevõrgus ei ole.

Andmeid Oraclest HANA peale liigutades saadakse kohe vastuse ühele seatud eesmärkidest. Andmete jalajälg on HANA mälus 1,5 gigabaiti ehk SAP poolt välja reklaamitud 7-10 korda andmemahu vähenemist on tõestatud – andmemahut on ERGO andmete näitel 10 korda väiksem. Selle informatsiooni baasil on võimalik teha otsuseid SAP HANA riistvara soetamisel.

Aruandluse kiiruse võrdlemiseks luuakse HANA-s paiknevate andmete kuvamiseks semantiline kiht ehk universum, mida saab kasutada SAP BO tarkvaras aruannete loomiseks. Universum on sisuliselt äri vaade andmetele. Ärikasutaja ei pea tundma SQL-i ning andmebaasi loogikat, vaid saab kasutada aruannete loomiseks objekti, millel on äriliselt arusaadavad nimed. SQL on ärikasutaja eest peidetud ning samuti on tegemist ainult lugemise õigusega ehk ärikasutajal ei ole õigusi ega ka mingit võimalust andmeidas olevaid andmeid manipuleerida [4]. Prototüübi jaoks loodud universum on sarnane olemasolevale ning kasutajate jaoks puudub vajadus uue keskkonna või loogika õppimiseks. Kahjuks aga ei anna aruannete sellisel kujul võrdlemine täit pilti kiirusest.

Seoses universumite erinevustega, eelkõige vanas universumis puuduvate andmetega, ei ole võimalik üks ühele võrdlust luua. Samuti asuvad Oracle andmed kohtvõrgus, aga pilves asuvate andmete puhul on piiranguks ka välisvõrgu kiirus. Testijate subjektiivne tagasiside hindab tulemust pigem kiiremaks kui „vana“ süsteem, aga konkreetseid mõõdetavaid tulemusi sellises olukorras tekitada ei ole võimalik.

Siiski tuleb tähele panna, et HANA-s on andmed algsel ehk normaliseeritud kujul ning kiirust hinnatakse paremaks kui Oracle andmeidal kuhu andmeid laadides toimub mitmeid transformatsioone. Sellest tagasisidest lähtuvalt võib HANA puhul ETL protsessist kõikvõimalikud Oracle andmeaida puhul vajalikud kiirust parandavad transformatsioonid ära jätta. See lihtsustab ja muudab arendust oluliselt odavamaks, sest piisab vaid andmete üks-ühele andmaita laadimisest. Tähele tuleb panna, et ETL protsessi käigus toimuvad ärioloogilised kalkulatsioonid tuleb realiseerida HANA-s või universumis.

Prototüübi teostamisega on neljast seatud eesmärgist kolm saavutatud ning üks eesmärk on subjektiivselt hinnatud saavutatuks. Andmete jalajälg vähenes reklaamitud kujul ning sellega seoses võetakse projektimeeskonna pool vastu otsuse, et ERGO jaoks vajaliku andmemahuga HANA riistvara on tänases olukorras sobiva hinnaga. Andmeaida

arenduse keerukuse vähenemisega tekib rahaline sääst tulevastes arendustes ning aruandluskeskkond on kasutajate jaoks kasutatav vähemalt sama mugavalt ja kiirelt kui olemasolev süsteem.

## 4 HANA kasutusele võtmine andmeaida andmebaasina

Prototüübi teostamise tulemusena on tekkinud arusaam, et andmeaidas mälu põhise andmebaasi kasutamine on õigustatud. Eelduseks on, et täna kulukam riistvara on tulevikukindlam ning teenib ennast ka tagasi vähendades andmeaida täienduste keerukust ja hinda. Sellega seoses on ERGO-s võetud vastu otsus soetada HANA riistvara ning viia olemasolev aruandluskeskkond üle Oracle andmebaasilt HANA-le. Migratsiooni puhul ei ole küll tegemist antud bakalaureusetöös püstitatud põhieesmärkide täitmisega, kuid antud see on eelduseks järgnevatele arendustele ning vajalik taustainfo.

HANA kasutusele võtmiseks ERGO-s on vaja teha järgnevad sammud:

- Andmeaida andmemahtude hindamine
- HANA riistvara hange
- Olemasolevate andmelaadimiste suunamine HANA andmebaasi
- Olemasolevate Oracle põhiste ETL skriptide ümber kirjutamine HANA-le sobivaks
- Olemasolevate universumite uuendus viimasele versioonile
- Olemasolevate aruannete migratsioon uutele universumitele
- Uute aruandlusvajaduste kaardistamine
- Andmete algsüsteemist andmaita reaajas liigutamise seadistamine
- Uue universumi loomine

### 4.1 Andmeaida andmemahtude hindamine ja riistvara hange

Selleks, et vähendada administreerimisele kuluvat ressursi on plaan viia ka olemasolev andmeait Oracle baasist HANA-le. Alternatiivsel juhul tekiks oodatud lihtsuse asemel süsteemiadministraatoritele ainult tööd juurde, sest igapäevaselt peaks hoolt kandma kahe andmebaasi eest. Viies kõik andmeaida andmed üle HANA-le vabaneb Oracle baasi ressurss teiste süsteemide jaoks ning andmeaida süsteemiadministraatorid peavad tegelema ainult HANA-ga.

Seega tuleb enne HANA riistvara hanget hinnata lisaks alliksüsteemi andmemahtudele ka olemasoleva andmeaida andmemahte, et HANA mälu maht kataks kõik vajadused.

Andmemahutude suurusjärk Oracle baasis on järgnev:

- SAP ERP ~700GB
- Andmeaida andmed ~250GB

SAP soovib HANA mälu mahu hindamiseks valem: 50% kettapõhise andmebaasi andmete mahust + 20% puhvrit + 50GB HANA tarkvaraliste vajaduste jaoks [5]. Selle valemi järgi on ERGO vajadus HANA mälu mahuks:

$$(700\text{GB} + 250\text{GB}) * 0,5 * 1,2 + 50\text{GB} = 620\text{GB}.$$

SAP soovitude kohaselt oleks seega vaja ERGO vajaduste katmiseks HANA riistvara, mille mälu maht oleks vähemalt 620GB. Samas SAP ise reklaamib 7 kordset andmete jalajälje vähenemist ning ERGO prototüübi puhul tuvastati 10 kordne vähenemine. Sellega seoses võtab projektmeeskond vastu otsuse, et esmalt piisab ka 512GB mälu serverist, mis oleks vajadusel laiendatav 1 terabaidini. Lisaks produktiivkeskkonna serverile on vaja ka arendus-/testserverit.

Hankel saavad osaleda ainult firmad, mis on SAP poolt sertifitseeritud HANA riistvara pakkuma. Kuna tegemist on SAP poolt sertifitseeritud riistvara paketi, siis riistvaraliselt erinevatel pakkujatel suuri vahesid ei ole. Eristuda püütakse lisaväärtuse müümise. Kui mõni pakkuja tarnib vaid riistvara ning teostab ka ainult riistvara hooldust, siis on ka konkurente, kes tarnivad riistvara, paigaldavad selle ning on ka üheks kontaktiks nii riistvaraliste kui tarkvaraliste hoolduste puhul (tarkvara uuendused jne.). ERGO jaoks ongi valikukriteeriumite kriitiline osa lisaks hinnale lisaväärtus, mida üks või teine pakkuja pakub.

## **4.2 Laadimiste migratsioon HANA-le**

Seni on andmelaadimiste sihtbaasiks olnud Oracle andmebaas, mida jagatakse ka teiste Eestis kasutatavate infosüsteemidega. Liigutades olemasoleva andmeaida Oracle pealt HANA-le lihtsustatakse andmeaida süsteemiadministraatorite tööd, sest nad ei pea hakkama haldama kõrvuti kahte erinevat baasi ning vabaneb ka Oracle ressursid teiste infosüsteemide jaoks.

Kokku on 4 erinevat andmelaadimise protsessi, millest 2 käivad igapäevaselt ning 2 iga nädal. Igapäevased on andmelaadimised Eesti andmeallikatest ning iganädalaselt käivad

Baltikumiülesed andmelaadimised, mille puhul võetakse andmeid Eesti, Läti ja Leedu müügi- ja raamatupidamissüsteemidest. Kõikide andmelaadimiste realisatsioon on sarnane ning samu samme tuleb kasutada kõikide muutmiseks.

Põhiline loogika on ehitatud SAP BO Data Services ETL tööriista kasutades, mis asub eraldi Windowsi serveril. Alates viimasest andmelaadimisest muutunud andmed laetakse algsüsteemide alliktabelitest ja/või algsüsteemi vaadetest vahetabelitesse ning seejärel toimub ETL protsessi transformatsiooni osa. Selle käigus andmeid agregeeritakse ning denormaliseeritakse, et andmeid kasutamine oleks suurte andmehulkade puhul kiirem. Transformatsioonid, mille puhul ei ole mõistlik andmehulka algsüsteemist Data Services serveri mällu lugeda, teostatakse Oracle andmebaasi skriptide abil. Andmebaasi skriptid annavad võimaluse rakendada andmebaasi jõudlust transformatsiooni tegemiseks misjärel saab laadida andmetest ainult vajamineva alamosa.

Oracle pealt HANA peale migreerimise kriitilisim osa ongi seotud Oracle andmebaasi skriptidega. SAP HANA Data Services loogikat muuta ei ole vaja. See töötab erinevate alg- ja sihtandmebaasidega üheselt. Seevastu spetsiifiliselt Oracle jaoks kirjutatud skriptid ei ole üheselt üle kantavad, sest Oracle kasutab PL/SQL süntaksit, aga HANA kasutab SQL Scripti. Süntaksisse sügavamalt laskumata võib öelda, et mõlemal skriptimiskeelel on omapärasid, mis võivad, aga ei pruugi tekitada konverteerimisel tõrkeid.

Kõikide skriptide täielik läbi töötamine on antud töö autori hinnangul liigselt ajamahukas. Projektimeeskond otsustab kaardistada kriitilisemad ning teada olevad murekohad ning ülejäänud skriptid võtta kasutusele üks-ühele. Selliselt tulevad vead välja andmelaadimiste testimise käigus, sest võimaliku süntaksi vea korral andmelaadimine katkeb. Testimiseks eraldatakse tavapärasest pikem ajaperiood, 2 kuud, mille käigus andmelaadimisi testida ning vajadusel süntaksiprobleemid lahendada.

Kõikvõimalikud arendused teostatakse arenduskeskkonnas, misjärel tarnitakse need testkeskkonda. Testkeskkonna andmekooslus on piisav, et teostada andmelaadimise teste ning tuvastada süntaksiprobleeme. Samuti saab testkeskkonnas teostada esmase tehnilise testi. Selle käigus võrreldakse kirjete arvu kõikides Oracle ja HANA andmeaitade tabelites Joonis 3 esitatud päringuga:

```
SELECT COUNT(*) FROM TABLE;
```

Joonis 3. Ridade võrdluse kood

Selline kirjete arvu võrdlus annab kinnituse sellest, et peale migratsiooni ei ole skriptides fundamentaalseid probleeme. Peale tehniliste testide edukat läbimist võimaldatakse ligipääs ärikasutajatele, kes testivad andmete sisulist korrasolekut, aga seda saab teostada alles peale ärikasutajate semantilise kihi, universumi loomist ning aruannete migratsiooni.

### 4.3 Universumite uuendus ja aruannete migratsioon

Laadimiste sihtbaasi muutmisega on vaja korrigeerida ka universumite ühendusi andmebaasiga ning vajadusel korrigeerida SQL päringuid. Sarnaselt SQL skriptidega võib ka universumite SQL-s esineda andmebaasi põhiseid päringuid, mis on vaja PL/SQL asemel SQL Scripti spetsiifikat arvestades ümber kirjutada.

Piisaks ka lihtsalt olemasolevate universumite andmeallika vahetamisest, aga projektimeeskond otsustab projekti skoopi võtta ka universumite uuendamise. Senini kasutusel olnud *unv* laiendiga BO universumid on iganenud ning need konverteeritakse uuemat tüüpi *unx* laiendiga failideks. Sisuliselt on tegemist sarnase muutusega nagu Microsoft Wordi kolimine *doc* laiendi pealt *docx* peale. Uuem versioon pakub laiemat ühilduvust SAP toodetega.

Seega on vaja konverteerida olemasolevad 4 universumit SAP BusinessObjects Information Design Tool'i (edaspidi IDT) abiga *unv* formaadist *unx* formaati, vahetada ära andmeallikas ning kontrollida SQL päringuid. IDT võimaldab kontrollida standardset SQL süntaksit, aga ei suuda tuvastada erinevusi PL/SQL ja SQL Scripti vahel. Sellest tulenevalt peavad äritestijad potentsiaalsed probleemikohad testi käigus ning arendajale on eraldatud ressurss potentsiaalsete probleemide lahendamiseks.

Selleks, et äritestijad saaksid uusi universumeid testida on neile vaja tekitada ligipääs ning migreerida olemasolevad aruanded. Aruannete migratsioon teostatakse selliselt, et kõik olemasolevad aruanded kopeeritakse ning aruande koopiatel vahetatakse seotud universumi(d) ära. See annab võimaluse nii-öelda vana ja uut süsteemi võrrelda identsete aruannete baasil. Kui mõne aruande puhul esineb erinevusi, siis on testijatel võimalus universumite abil spetsiifilisi kirjeid võrrelda ning seejärel arendaja abi kaasata.

#### **4.4 Lisaväärtus HANA migratsioonist**

Olemasolevate andmeaitade Oracle andmebaasist HANA-le migreerimise peamine eesmärk on vähendada andmebaasiadministraatorite tuleviku ressursikulu sellega, et nad ei pea haldama korraga mitut erinevat baasi. Lisaks seatud eesmärgi täitmisele kaasneb migratsiooniga ka kasv aruandluskeskkonna jõudluses ning ressursi vabanemine teiste Oracle andmebaasi kasutavate süsteemide jaoks.

Senini kasutasid nii SAP ERP kui ka andmeait sama Oracle andmebaasi. Kasutuse tippaegadel võis esineda olukordi kus mõlemad süsteemid olid maksimaalsest võimalikust aeglasemad, sest mõlemad süsteemid kasutasid sama ressursi. Kolides andmeaidad Oracle baasist HANA-le ei mõjuta aruandluskeskkonna kasutajad enam igapäevast ERP süsteemi kasutamist ja vastupidi. Samuti on võimalik kõikvõimalike tõrgete korral lihtsamalt eristada probleemi sisu ning ühe süsteemi hooldused ei mõjuta teist.

Endiselt jäävad käiku igapäevased ja -nädalased andmelaadimised algallikatest andmeaita. Sihtandmebaasi vahetusega kaasneb ka kõikide laadimiste kiiruse kasv andmebaasi skriptide arvelt. 90 päevase võrdlusperioodi käigus oli SAP ERP-st andmete laadimine HANA-sse keskmiselt 26% kiirem kui Oracle andmeaita. Andmelaadimised käivad küll tavaliselt töövälisel ajal ning absoluutne kiirus ei ole esmatähtis, aga ressursikulu vähenemine võimaldab IT osakonnal seda aega kasutada muudeks vajaminevateks hooldustöödeks.

Lisaks andmelaadimise kiiruse kasvule on märgatavalt kasvanud ka aruannete laadimise kiirus. Eelkõige võidavad aruanded, mis vajavad pikema perioodi summade arvutamist. Näiteks kindlustuse puhul kvartali ja aasta kindlustuspreemiate või väljamaksete arvutamine. Kuna andmed on andmebaasi mälus, mitte ei loeta kõvakettalt ning neid hoitakse veerupõhiselt, siis on selliste aruannete puhul näha enam kui 50 kordset kiiruse kasvu. Kvartali kahjusuhte aruanne, mis eelnevalt laadis 180 sekundit laeb HANA andmeaidast ainult 3 sekundiga. Samas aruanded, mille puhul laetakse palju kirjeid nagu aktiivsete klientide kontaktandmed, ei saavuta märgatavat kiiruse eelist, sest sellises olukorras ei tule veerupõhise andmebaasi eelised esile.



## 5 Reaalajas aruandlus

Eelnevalt sai mainitud, et andmeaitade kontseptsioon tuleneb sellest, et ERP süsteemis hoitakse andmeid kujul, mis raskendavad suuremamahulist aruandlust. Selleks, et need fundamentaalselt erinevad nõudmised täidetud saaks on algsüsteemist laetud andmed andmeaita ning aruandeid luuakse andmeaita baasil. Andmelaadimise protsessid seevastu on ajamahukad ning seetõttu tehakse neid ERGO-s vastavalt kas kord päevas või kord nädalas, mis tähendab, et igal ajahetkel on aruanne vähemalt mingil määral iganenud andmetega. Samas näiteks aktiivse müügitöö puhul on kriitiline omada andmeid nii kiiresti kui võimalik, et omada ülevaadet müügi tulemustest või pakkuda kliendi ajaloo põhjal kliendile soodustusi jne. Tänapäevane lahendus ei ole selleks sobiv, aga mälu põhiste andmebaaside võidukäik muudab olukorda.

### 5.1 ERP andmete reaalajas laadimine HANA-sse

ERGO ERP andmed asuvad relatsioonilises Oracle andmebaasis ning aruandluseks laetakse andmeid kord päevas ja kord nädalas andmeaitadesse. Igapäevaste laadimiste asemel, mis puhul andmed on alati eelmise päeva seisuga, soovivad äri kasutajad näha aruandlust reaalajas. Eelkõige müügi seisukohalt on kriitiline, et tulemused oleks võimalikult aktuaalsed ning ajakohaste andmetega.

SAP pakub välja võimaluse andmete reaalajas HANA-sse laadimiseks Oracle andmebaasilogide abil. SAP Replication Agent for Oracle võimaldab andmebaasi logide põhjal kopeerida kõik Oracle baasis valitud tabelites tehtud muudatused reaalajas HANA-sse [6]. Peale replikatsiooniserveri püsti panekut ja seadistamist jääb üle vaid valida tabelid, mille andmeid HANA-sse soovitakse ning SAP tarkvara teeb kõik ülejäänud.

Replikatsiooni eeliseks on see, et andmed on identsel kujul nii Oracle andmebaasis kui ka HANA-s ning andmete laadimise viide sisuliselt puudub. Koheselt on võimalik samade andmete aruannete koostamine ilma Oracle andmebaasi koormamata. Samas tuleb arvestada, et andmed on ühest baasi teise kopeeritud identselt ehk tegemist on normaliseeritud kujuga. Andmeaita andmete laadimise puhul kasutatakse ETL protsessi lisaks andmete laadimiseks ka transformatsioonideks, mida antud juhul ei tehta.

Seega on HANA peale laetud Oracle andmebaasiga identsed normaliseeritud andmed, milles puuduvad kõikvõimalikud agregeerimised ning aruandluseks vajalikud täiendused. Tekib küsimus, et mis kasu on sellistest andmetest aruandluses jaoks isegi kui need on reaalselt kättesaadavad. Siinkohal tuleb esile mälu põhise andmebaasi võimsus, mida aruandluse jaoks rakendada saab.

Selle asemel, et aega võtva ETL protsessi käigus andmete transformatsiooniga tegeleda on mälu olevate andmete võimalik sarnane protsess läbida jooksvalt aruande loomise hetkel. Andmete liiasust saab vältida reaalselt teostatavate kalkulatsioonide ning vaadete kasutamise abil. Terve andmeida ülesehitus tähtskeemide abil, milles kesksel kohal on faktid ning viited dimensioonidele muutub ebavajalikuks, sest mälu põhise andmebaasi jõudlus võimaldab päringuid teha piisava kiirusega ka normaliseeritud tabelitest.

## **5.2 Universumi loomine normaliseeritud andmebaasile**

Kasutusel oleva andmeida lahenduse puhul matkib universumi loogika ETL protsessi käigus loodud tähtskeemi. Universumi SQL päringutesse pole liigset loogikat vaja sisse ehitada, sest enamuse loogikat on realiseeritud ETL protsessi käigus. Kõikvõimalik keerulisem loogika on jäetud ETL protsessi hoolde, sest selle jooksvalt realiseerimine oleks liiga ajamahukas.

Siiski ei saa piisava jõudluse olemasolul ehk mälu põhise andmebaasi puhul samuti optimeerimisest mööda vaadata. ETL mõte on lisaks aruandluse kiirendamisele ka aruandluse lihtsustamine. Kui tähtskeem on ehitatud üles ärinõuetele vastavalt, siis universumis seda peegeldades on andmed äri kasutajatele arusaadavad. Seevastu otse normaliseeritud andmetele universumi loomine oleks justkui lohista-sikuta (drag and drop) kasutajaliidesega SQL päringute loomine, mis ei oleks äri kasutajatele arusaadav ning oleks kindlasti olemasolevast lahendusest ebamugavam.

Eelnev kokkuvõtte ETL arenduselt avaldub seega siin faasis. Kõikvõimalik äri loogika tuleb realiseerida andmebaasi tasandil või BO universumis. ERGO puhul otsustab projektmeeskond kasutada osaliselt andmebaasi vaateid ja universumisse ehitatud loogikat. Kokkuvõtte ETL protsesside arenduselt on töö autori hinnangul suurem kui andmebaasi vaadete ja BO universumi arendamisele kuluks.

ETL protsessi loomiseks on vaja arendajat kes on SAP BO Data Services spetsialist. Selliseid arendajaid on turult väga raske leida ning teenuse sisse ostmine on kallis. Seevastu SQL keele spetsialiste on turul märgatavalt rohkem ning andmebaasi võimsuse rakendamine on ka parim praktika, mida tänases lahenduses laialdaselt realiseeritud ei ole. SAP ERP puhul on kasutusel ka ERP sisesed vahetabelid, mida uue lahenduse puhul vaja ei ole ning mille arenduse kulu täielikult säästetakse. BO universumi arendajat on vaja nii olemasoleva lahenduse kui ka uue arenduse jaoks ehk siinkohal arendusmahud küll kasvavad, aga mitte märgatavalt.

Ärikasutajate vaade universumile jääb sarnaseks olemasoleva lahendusega. Terminid on äriliselt arusaadavad ning kõikvõimalikud päringud tehakse ärikasutaja eest peidetult. Samuti on universumil ainult andmete lugemisõigus ning andmeid andmebaasis muuta ei saa. Märgatavalt aga kiireneb täiendavate aruandlusvajaduste rahuldamine. Selle asemel, et luua vahetabelid SAP ERP-s, ETL protsess andmete andmeaita laadimiseks ning BO universum piirduakse ainult järgnevate sammudega:

1. Andmete reaajas Oracle baasist HANA-sse kopeerimine – piisab ühekordsest seadistamisest, misjärel andmete laadimist enam korrigeerida ei ole vaja. Uute väljade laadimine toimub automaatselt, uue tabeli laadimine nõuab minimaalset muudatust.
2. HANA-s andmebaasi vaadete loomine – ETL asemel piisab andmebaasi vaadete loomisest, sest andmebaasi jõudlus võimaldab teha kalkulatsioone jooksvalt. Arenduse kulu on soodsam kui ETL protsessil, sest arendajad on kättesaadavamad ning arendust on vaja ainult agregeeritud andmete jaoks. Otse baasist võetavate andmete jaoks täiendavat arendust vaja ei ole.
3. BO universumi arendus – maht jääb sarnaseks olemasoleva arendusega. Mõningase loogika lisandumine universumi kihti võib arenduse hinda tõsta.

### **5.3 Edasised sammud**

SAP HANA kasutuselevõtt ei piirdu ainult andmeaidaga vaid see avab tulevikus võimaluse ka ERP süsteemi andmebaasi vahetuseks. Kuna ERGO kasutab ERP süsteemina SAP toodangut, siis on sellel SAP HANA toetus ning võimalus migratsiooniks olemas.

Tänases olukorras on ERGO otsustanud katsetada uut tehnoloogiat aruandluskeskkonna peal, sest aruandluskeskkonnad saavad töötada ka paralleelselt, sest kasutavad andmeid algallikatest. ERP puhul on tegemist organisatsioonikriitilise süsteemiga, mille andmebaasi migratsiooni ei saa võtta kergekäeliselt. Potentsiaalsete probleemide korral ei ole nii-öelda vana süsteemi peale tagasi keeramine lihtne või üldse võimalik. Seetõttu on aruandluskeskkonna migratsioon hea võimalus ERGO süsteemiadministraatoritel SAP HANA tundma õppida ning mälu põhiste andmebaaside iseärasustega kohaneda.

ERP süsteemi migratsiooni peale tuleb siiski aktiivselt mõelda, sest sellel on rohkelt eeliseid. Paratamatult on mälu põhine andmebaas igapäevaseks tööks kiirem kui kõvakettal andmeid hoidev andmebaas. ERP süsteem, kasutades mälu põhise andmebaasi ei vajataks enam eraldi andmeaita ehk kõikvõimalikud andmelaadimised ja andmete reaajas kopeerimised jääksid ära. Sama andmekogumit saaks kasutada nii ERP süsteemi tööks kui ka aruandluseks. Replikatsiooniserveri ning eraldiseisva andmeaita andmebaasi hooldamise ressursid vabaneks täielikult ning ka riistvara saaks alternatiivseteks otstarveteks kasutada.

## 6 Kokkuvõte

Antud bakalaureusetöö eesmärk oli uurida ERGO Insurance SE näitel eeliseid, mida annaks mälu põhise andmebaasi kasutuselevõtt aruandluskeskkonnale. Hüpoteesidena oli tõstatatud järgnevad potentsiaalsed eelised:

- Aruandluskeskkonna arenduste kulude langemine
- Aruandluse kiiruse kasv
- Andmelaadimiste ning andmete liiasuse elimineerimine
- Võimalus ERP andmete baasil reaalajas aruandluseks

ERGO-s antud töö autori poolt juhitud projekti järelalusena võib töö autor väita, et algselt seatud hüpoteesid leidsid kinnitust.

Aruandluskeskkonna jätkuvaks toimimiseks vajaminevad arendused olid olemasoleva andmeaia puhul vägagi kompleksed, ajamahukad ning kulukad. Arendada oli vaja ERP süsteemi, andmeaia ETL protsessi ning ka lõppkasutajale käepärast universumit. Projekti käigus tuvastati, et algsüsteemist reaalajas andmeid mälu põhisesse SAP HANA andmebaasi liigutades on võimalik kulude kokkuhoid. Arenduse elementidest kaob täielikult ERP sisene vahetabelite ning ETL protsesside arenduse vajadus. Andmete replikatsioon vajab esmast üles seadmist ning HANA vaadete loomiseks on võimalik kaasata odavamad ressursid ning pakkujate ring on laiem kui SAP BO Data Services arendajate puhul. BO universumite arenduse kulud seevastu jäid sarnaseks.

Kulude ning arenduseks kuluva aja vähendamisele lisaks on aruandlusüsteem lõppkasutaja jaoks ka märgatavalt kiirem. Veerupõhistes tabelites mälu põhises andmebaasis päringuid tehes ei ole vaja andmeid eelnevalt agregeerida vaid kasutajad saavad päringuid teha otse algandmetelt ning kiiremini kui varem. Selline lähenemine annab kasutajate kätte rohkem võimalusi olemasolevate andmete baasil järelalusete ja otsuste tegemiseks.

Pikemas perspektiivis on võimalik ka ERP süsteemi kasutada mälu põhise andmebaasiga ning see kustutab viimse lõhe igapäevatoeks ning aruandluseks kasutatavate andmete vahel. Sellises situatsioonis on võimalik igal ajahetkel tehtavaid tegevusi kohe aruandluses kajastada ning kõikvõimalikud andmete laadimised ning agregeerimised saaks ära jätta. Sellisel juhul jõutakse olukorda kus nii ERP kui aruandlusüsteem

kasutavad sama andmekogumit. ERGO täna veel nii kaugel ei ole, aga kogemus aruandlussüsteemiga sisestab usku, et see on enam kui võimalik.

Sissejuhatuses püstitatud finantseesmärgid said prototüübi teostamisega ning andmeaida andmebaasi vahetusega täidetud. Lisaks sellele sai ka kinnitust hüpotees, et mälu põhised andmebaasid võimaldavad kaotada või vähemalt ähmastada piiri ERP ning aruandlussüsteemide vahel. ERGO näitel saab väita, et sama andmekogumi pealt on võimalik nii igapäevane töö kui ka laiem aruandlus. Seatud eesmärkide kõrval tõi andmebaasi vahetus kaasa ka olemasoleva andmeaida kiiruse kasvu nii andmelaadimiste kui ka lõppkasutajate aruannete avamisel, mis annab lahendusele lisaväärtuse.

## Kasutatud kirjandus

- [1] „Wikipedia - Enterprise Resource Planning,“ [Võrgumaterjal]. Available: [https://en.wikipedia.org/wiki/Enterprise\\_resource\\_planning](https://en.wikipedia.org/wiki/Enterprise_resource_planning). [Kasutatud 19 mai 2016].
- [2] „Wikipedia - Online Transaction Processing,“ [Võrgumaterjal]. Available: [https://en.wikipedia.org/wiki/Online\\_transaction\\_processing](https://en.wikipedia.org/wiki/Online_transaction_processing). [Kasutatud 19 mai 2016].
- [3] „Wikipedia - Online Analytical Processing,“ [Võrgumaterjal]. Available: [https://en.wikipedia.org/wiki/Online\\_analytical\\_processing](https://en.wikipedia.org/wiki/Online_analytical_processing). [Kasutatud 19 mai 2016].
- [4] „Wikipedia - ETL,“ [Võrgumaterjal]. Available: [https://en.wikipedia.org/wiki/Extract,\\_transform,\\_load](https://en.wikipedia.org/wiki/Extract,_transform,_load). [Kasutatud 19 mai 2016].
- [5] „Wikipedia - Star Schema,“ [Võrgumaterjal]. Available: [https://en.wikipedia.org/wiki/Star\\_schema](https://en.wikipedia.org/wiki/Star_schema). [Kasutatud 13 mai 2016].
- [6] F. Färber, N. May, W. Lehner, P. Große, I. Müller, H. Rauhe ja J. Dees, „The SAP HANA Database – An Architecture Overview,“ *Bulletin of the Technical Committee on Data Engineering*, kd. 35, nr 1, pp. 28-33, March 2012.
- [7] „Oracle,“ [Võrgumaterjal]. Available: <http://www.oracle.com/us/media/calculator/tieredstorage/index.html>. [Kasutatud 18 aprill 2016].
- [8] „Amazon Web Services,“ [Võrgumaterjal]. Available: <https://aws.amazon.com/sap/solutions/saphana/>. [Kasutatud 19 aprill 2016].
- [9] „Altek Solutions - What Is a BusinessObjects Universe,“ [Võrgumaterjal]. Available: <http://alteksolutions.com/wp/index.php/2009/03/what-is-a-business-objects-universe/>. [Kasutatud 13 mai 2016].
- [10] „HANA Sizing - SAP Insider,“ [Võrgumaterjal]. Available: <http://sapinsider.wispubs.com/Assets/Articles/2014/July/SPI-how-to-properly-size-an-SAP-in-memory-database>. [Kasutatud 12 mai 2016].
- [11] „Oracle-to-SAP HANA Database Replication Setup,“ [Võrgumaterjal]. Available: <http://infocenter.sybase.com/help/index.jsp?topic=/com.sybase.infocenter.dc01974.1571110/doc/html/bde1355001071038.html>. [Kasutatud 13 mai 2016].