

Lauri Luoma

Microsoft SQL Server 2012 Business Intelligence ja sen tuomat uudistukset

Metropolia Ammattikorkeakoulu

Insinööri (AMK)

Tietoverkot

Insinööriytyö

16.12.2013

Tekijä(t) Otsikko Sivumäärä Aika	Lauri Luoma Microsoft SQL Server 2012 Business Intelligence ja sen tuomat uudistukset 34 sivua 16.12.2013
Tutkinto	Insinööri (AMK)
Koulutusohjelma	Tietotekniikka
Suuntautumisvaihtoehto	Tietoverkot
Ohjaaja(t)	Koulutusohjelmavastaava Janne Salonen
<p>Insinööriyö käsittelee Metropolia Ammattikorkeakoulun Microsoft Business Intelligence (BI) ratkaisut -kurssilla käytetyn laboratoriomanuaalin MCTS Self-Paced Training Kit (Exam 70-448) -harjoitusten siirtämistä SQL Serverille 2012.</p> <p>Työllä on tarkoitus osoittaa, että SQL Server 2012 BI -työkalut soveltuvat harjoituksiin. Työssä siirrytään käyttämään uudempaa työkalua, SQL Server Data Toolsia, joka korvaa SQL Server 2008 R2 Business Intelligence Development Studion.</p> <p>Työn alussa tutustutaan Business Intelligence -termiin, käydään läpi BI:n historiaa ja kerrotaan sen käytöstä nykyään. Työssä tutustutaan myös lyhyesti keskeisimpiin BI-menetelmiin, SQL Serveriin 2012 ja sen asentamiseen. Lopuksi työssä käydään läpi 43 harjoitusta laboratoriomanuaalista ja selvitetään poikkeavuudet harjoituksissa verrattuna SQL Server 2008:aan.</p> <p>Laboratoriomanuaalin harjoitukset todettiin yhteensopiviksi SQL Server 2012 BI -ympäristöön. Uusi asennettu SQL Server Data Tools (SSDT) osoittautui hyvin samantapaiseksi BI-ratkaisuiden kehitystyökaluksi kuin Microsoft SQL Business Intelligence (BI) -kurssilla käytetty Business Intelligence Development Studio (BIDS).</p>	
Avainsanat	Business Intelligence, SQL Server, SSDT

Author(s) Title Number of Pages Date	Lauri Luoma Microsoft SQL Server 2012 Business Intelligence and Its Implications 34 pages 16 December 2013
Degree	Bachelor of Engineering
Degree Programme	Information Technology
Specialisation option	Data Networks
Instructor(s)	Janne Salonen, Head of Degree Programme
<p>This thesis project transfers the exercises in the MCTS Self-Paced Training Kit (Exam 70-448) to SQL Server 2012. The training kit is used in the Microsoft Business Intelligence solutions class at the Helsinki Metropolia University of Applied Sciences.</p> <p>The purpose of this thesis is to demonstrate that the SQL Server 2012 BI Tools are suited for these exercises. This project uses the current tool, SQL Server Data Tools, which replaces the SQL Server 2008 R2 Business Intelligence Development Studio.</p> <p>First, we describe the term Business Intelligence, go through BI history and report on the current usage and benefits of Business Intelligence. We also explore the essential BI processes and the SQL Server 2012 and its installation. Lastly, the focus is on 43 exercises of the laboratory manual and the differences in them compared with SQL Server 2008.</p> <p>The laboratory manual exercises were found to be compatible with SQL Server 2012 Business Intelligence environment. The new SQL Server Data Tools (SSDT) turned out to be very similar as a development tool compared to the Business Intelligence Development Studio (BIDS) used in Microsoft SQL Business Intelligence course.</p>	
Keywords	Business Intelligence, SQL Server, SSDT

Sisällys

Lyhenteet

1	Johdanto	1
2	Business Intelligence	1
2.1	BI-termi	1
2.2	Business Intelligencen historia	2
2.3	BI ja kasvava tiedon määrä	3
2.4	BI ja päätöksenteko	3
2.5	BI ja tiedon integrointi	3
2.6	BI ja tavoitteet	4
3	BI:n keskeiset menetelmät	4
3.1	ETL-prosessi	4
3.2	OLAP, Data Mining ja rakenne	6
3.2.1	OLAP	6
3.2.2	Data Mining	6
3.2.3	Rakenne	6
3.3	Raportointi	7
4	SQL Server 2012	7
4.2	Saatavuus	8
4.3	Tietoturva	8
4.4	Versiot	8
4.5	Teknologia	9
5	Ympäristö ja asennus	10
5.1	SQL Serverin asennus	11
5.2	AdventureWorks2008-tietokannat	12
5.3	SQL Serverin hallinta- ja kehitystyökalut	12
5.4	SQL Server Data Tools	13
5.5	SSDT-asennus	13
6	Laboratoriomanuaalin harjoitukset	16
6.1	SQL Server Integration Services	16

6.2 SSIS-paketti	17
6.2.1 SSIS-paketin kehitys	21
6.2.2 Virheiden korjaus ja käsittely SSIS:ssä	23
6.2.3 SSIS-paketin asentaminen ja konfigurointi	24
6.2.4 Hallinnointi, suojaaminen ja suorittaminen SSIS-paketeissa	25
6.3 Yhteenveto SSIS-harjoituksista	26
6.4 SQL Server Analysis Services	26
6.4.1 SSAS-kuutioiden kehitys	27
6.4.2 SSAS-kuutioiden laajennus	29
6.5 Yhteenveto SSAS-harjoituksista	30
6.6 SQL Server Reporting Services	31
6.6.1 SSRS-raportin kehitys	31
6.7 Yhteenveto SSRS-harjoituksesta	32
7 Yhteenveto	32
Lähteet	33

Lyhenteet

BI	Business Intelligence. Liiketoimintatiedon hallinta.
BIDS	Business Intelligence Development Studio. Visual Studio 2008 kehitys - ympäristö BI-työkaluille.
BISM	BI Semantic Model. BI-semanttinen tietomalli.
Cube	Kuutio. Analysis Services -tietomallin moduuli.
DSS	Decision Support System. Tietojärjestelmä päätöksenteon tueksi.
DSV	Data Source View. Tietolähdenäkymä käyttäjälle.
ETL	Extract-Transform-Load. Termi tiedon poiminnalle, muokkaukselle ja lataukselle.
FTP	File Transfer Protocol. Tiedonsiirtoprotokolla.
HTML	Hypertext Markup Language. Avoin standardoitu kuvauskieli.
MDX	Multidimensional Expression. Moniulotteinen kyselykieli OLAP-tietokantoihin.
MOLAP	Multidimensional OLAP. OLAP-ratkaisu puhtaimmillaan, tiedot on tallennettu moniulotteiseen rakenteeseen.
MSDTC	Microsoft Distributed Transaction Coordinator. Windows-käyttöjärjestelmän palvelukomponentti.
OLAP	OnLine Analytical Processing. Moniulotteinen analysointi.
POWER-PIVOT	Microsoftin apuohjelma, jolla voidaan analysoida suuria tietomääriä.
SQL	Structured Query Language. IBM:n kehittämä standardoitu kyselykieli.

SSAS	SQL Server Analysis Services. BI-ohjelmiston komponentti analyyttiseen prosessointiin.
SSDT	SQL Server Data Tools. Microsoftin BI-kehitystyökalu.
SSIS	SQL Server Integration Services. BI-ohjelmiston komponentti ETL-prosessointia varten.
SSMS	SQL Server Management Studio. SQL Serverin kehitysympäristö.
SSRS	SQL Server Reporting Services. BI-ohjelmiston komponentti, jolla voi luoda, ottaa käyttöön ja hallita raportteja.
UDM	Unified Dimensional Model. Tietomalli.
XML	Extensible Markup Language. Standardoitu merkintäkieli.
XMLA	XML for Analysis. Analyyttinen merkintäkieli.
.NET	Microsoftin kehittämä ohjelmistokomponenttikirjasto.

1 Johdanto

Business Intelligencella eli yrityksen liiketoimintatiedon hallinnalla tarkoitetaan yrityksen suorittamaa sisäistä (kvantitatiivista) tietojen hankintaa, tallennusta ja analysointia. Sillä voidaan tarkoittaa myös yrityksen suorittamaa ulkoista (kvalitatiivista) tietojen hankkimista kilpailevien yritysten toiminnasta tai markkinoista ja lopuksi tietojen analysoimista.

SQL Server 2012:n BI-työkaluilla voidaan suorittaa mikä tahansa Business Intelligence -toiminto. Microsoftilla on erilaisia palvelinvaihtoehtoja pienten yritysten tarpeisiin ja suurien yritysten laajoja tietovarastoja varten.

Tämä insinööriö pohjautuu Metropolia Ammattikorkeakoulun Moodle-työtilassa olevaan Microsoft Business Intelligence (BI) ratkaisut -kurssiin. Kurssilla käytetään Microsoft SQL Server 2008 Business Intelligence Development and Maintenance -nimistä kirjaa, joka on tarkoitettu kurssi- ja itseopiskelukäyttöön.

Insinööriön tarkoitus oli sovittaa kurssin laboratoriomanuaalin harjoitustehtävät uudempaan ympäristöön käyttämällä Microsoftin SQL Server 2012 Business Intelligence -versiota aiemmin kurssilla käytetyn SQL Server 2008 R2:n tilalla. Samalla syntyi opas, jossa kerrotaan ohjeet poikkeuksiin, joita tulee eteen harjoituksissa. Työympäristö asennettiin suoraan työn tekijän kotikoneelle. Työssä käytettiin Microsoftin harjoituskäyttöön tarkoitettuja AdventureWorks2008-tietokantoja.

2 Business Intelligence

2.1 BI-termi

Business Intelligence -termille on yritetty pitkään etsiä hyvää suomennosta. Liiketoimintatiedon hallinta -termiä käytetään nykyisin yleisesti, vaikka tämäkään ei kuvaa täydellisesti englanninkielistä vastinetta. BI-termiä myös tulkitaan kahdesta eri näkökulmasta.

Kvantitatiivisella näkemyksellä ymmärretään organisaation sisäisesti hankkiman liiketoimintatiedon analyttistä käyttöä ja hallintaa. Lähteinä toimivat yrityksen sisäiset tietokannat ja tietojärjestelmät.

Kvalitatiivisella näkemyksellä tarkoitetaan yrityksen kilpailijoista ja markkinoista saatavan laadullisen tiedon hyödyntämistä ja hallintaa. Tieto on yleisesti asiakirja- ja dokumenttipohjaista aineistoa. Jälkimmäisestä on tosin alettu käyttämään englanninkielisiä termejä, kuten Market Intelligence ja Competitive Intelligence. [1.]

Business Intelligence -työkalujen avulla voidaan kerätä järjestelmistä systemaattisesti oleellista tietoa, tutkia ja analysoida sitä ja käyttää lopuksi päätöksenteon tukena.

2.2 Business Intelligencen historia

Decision Support Systemiä (DSS) eli yritysten päätöksentekoa tukevia tietojärjestelmiä on kehitetty siitä lähtien, kun ensimmäiset tietokoneet otettiin käyttöön yrityksissä. Jo 1960- ja 1970-luvuilla oli käytössä johdon informaatiojärjestelmä eli Management Information System (MIS) ja 1980-luvulla Executive Information System (EIS). Ensimmäinen verkkopohjainen BI-järjestelmä oli 1980-luvun lopulla rakennettu ryhmäpäätöksenteon järjestelmä (Group DSS).

Nykyään katsotaan, että Business Intelligencen historia alkaa vuodesta 1989, vaikka tietojärjestelmätieteilijä Hans Luhn kirjoitti älykkäistä järjestelmistä jo 1950-luvulla. Howard Dresner ryhtyi kuvaamaan vuonna 1989 nykyisin käsitteellä ymmärrettäviä menetelmiä ja teknologioita Business Intelligence -termillä. BI-termiä käytettiin kylläkin samaan aikaan Iso-Britanniassa kuvaamaan yritysmaailmassa tapahtuvaa tiedustelu- ja vakoilutoimintaa, jossa käytettiin tietolähteinä yritysten ulkopuolella olevia tietolähteitä.

Relaatiotietokantojen ja SQL-kielen yleistyessä 1990-luvulla ryhdyttiin puhumaan ensimmäisistä liiketoimintaa tukevista tietovarastointiratkaisuista, joista ensimmäisiä olivat IBM:n Information Centerit. SQL-kieli mahdollisti liiketoimintalähtöisten kyselyjen teon, ja ensimmäisillä BI-ratkaisuilla pyrittiinkin tekemään SQL-kyselyjen tekeminen mahdollisimman helpoksi. Moniulotteista analysointia (OLAP) alettiin käyttää 1990-luvun alussa, ja ensimmäiset internetin kautta jaettavat ohjelmistot julkaistiin 1990-luvun lopussa. Client/Server-arkkitehtuuri ja tietovarastoratkaisujen toteuttaminen ETL-työkaluilla mahdollistivat tämän. Suurten yrityskauppojen myötä nykyisin isot ohjelmistotalot, ku-

ten Microsoft, Oracle, IBM ja SAP, omaavat isoimmat markkinaosuudet BI-markkinoista. [1.]

2.3 BI ja kasvava tiedon määrä

Tiedon syntyminen ei tarkoita vain tiedon syöttämistä tietokantoihin, vaan tietoa voi kertyä paljon automaattisesti. Esimerkiksi teleoperaattorit alkavat saada tietokantoihin tietoa jo pelkästään siitä, kun asiakas siirtyy matkapuhelin kädessä paikasta toiseen. Tätä dataa voidaan sitten hyödyntää BI-ratkaisuissa. Tieto tallennetaan ja siirretään analysointia varten tietovarastoihin (Data Warehouse, DW).

Kaupan alalla kuluttajien ostoskäyttäytymisiä tutkitaan keräämällä ostoksista kertyvää rivitason tietoa, jota voidaan hyödyntää eri tietotarpeisiin. Kauppojen kanta-asiakaskortit synnyttävät sellaista tietoa, jota voi analysoida ostoskäyttäytymistä tutkimalla.

Tiedon hankkijoiksi ovat viidentoista viime vuoden aikana ryhtyneet lähes kaikki toimialat. Niin kunnat kuin valtiohallintokin ovat julkisella puolella ryhtyneet keräämään tietoja. Julkisen puolen kustannustehokkuuden tavoittelu ja liikelaitosmainen toiminta ovat johtaneet johtamisjärjestelmien remonttiin ja sitä kautta BI-ratkaisujen tehokkaampaan hyödyntämiseen. [1.]

2.4 BI ja päätöksenteko

Yrityksissä on nykypäivänä entistä vähemmän aikaa tehdä päätöksiä, jotka liittyvät liiketoimintaan. Monissa yrityksissä vuositason suunnittelu on muuttunut kvartaalitalouden aikaan niin, että raportteja tuotetaan kuukausi-, viikko- tai jopa päivätasolla. Pankki- ja vakuutussektorilla päivän sisälläkin tehtävät ennustukset ovat yleisiä. BI-ratkaisujen näkökulmasta tällä tarkoitetaan raporttien nopeita latauksia. [1.]

2.5 BI ja tiedon integrointi

Tietovarastointiratkaisut ovat muuttuneet entistä haastavammiksi, koska informaatiota koostetaan yhä useammista tietolähteistä. Tiedon integrointi on keskeisessä osassa esimerkiksi yrityskauppojen yhteydessä. Tämä aiheuttaa integrointipaineita esimerkiksi

asiakasrekistereissä, myyntijärjestelmissä sekä muissa toiminnanohjausjärjestelmissä. ERP- eli Enterprise Resource Planning -ohjelmiston toimittajien strategisissa valinnoissa yhdeksi merkitsevimmäksi kehityksen kohteeksi on noussut BI. [1.]

2.6 BI ja tavoitteet

Yksi BI:n tavoitteista on parantaa ja nopeuttaa yritysten kykyä tehdä päätöksiä. Tämä on BI-ratkaisujen keskeisimpiä tavoitteita, ja se näkyy jo edellä mainituista varhaisempien ratkaisujen nimestä päätöksenteon tukijärjestelmä (Decision Support System). Edelleen tarkoitus on tuottaa laadukkaampaa informaatiota käyttäjille, jotta nämä pystyvät tekemään parempi päätöksiä myös silloin, kun informaatio toimintaympäristöstä on puutteellista.

BI-ratkaisujen tavoite on vastata yritysten tietotarpeisiin oikea-aikaisesti. Yksi BI-ratkaisujen keskeisistä tavoitteista on tuottaa informaatiota suunnitelmallisella tavalla ja mahdollistaa myös ennalta määrittelemättömien tietojen nouto.

BI-ratkaisut eivät ole poikkeus liike-elämässä tapahtuvien investointien kolmesta pää-tavoitteesta, jotka ovat myynnin lisäys, kustannuksien säästö ja uusien liiketoimintamahdollisuuksien löytyminen. [1.]

Business Intelligence -työkalut tulevat arvioiden mukaan ohittamaan kaikki muut ohjelmistoliiketoiminnan osa-alueet. Jo tällä hetkellä Business Intelligence on yritysten ykkösprioriteetti sekä uusilla että korvaavilla investoinneilla, uskovat analyytikot. [2.]

3 BI:n keskeiset menetelmät

3.1 ETL-prosessi

ETL tulee sanoista extract (poiminta), transform (muokkaus) ja load (lataus). Prosessissa tiedot luetaan operatiivisista järjestelmistä tai siirtotiedostoista, muokataan tietovarastotietokannan muotoon ja lopuksi kirjoitetaan tietovarastoon (Data Warehouse). Muokausvaiheeseen sisältyy virheiden tarkistusta, koodien muokkausta ja tietojen arkistointia.

Operatiivisten tietokantojen rakenteet ovat yleensä hyvin vaativia ja hankalia lukea [1.]. Poiminnan tarkoituksena on noutaa tieto, joka halutaan kohteesta, ja muuntaa se haluttuun muotoon seuraavaa muokkausvaihetta varten.

Muokkauksen ensimmäinen vaihe on aineiston tarkistus. Siinä tarkistetaan esimerkiksi, onko aineiston tauluissa tuplarivejä ja tarvittaessa niistä raportoidaan ja ne poistetaan. Virheellisten tietojen käsittelyyn käytetään useampaa vaihtoehtoa. Virheelliset tiedot voidaan hylätä tai kirjoittaa virhetauluihin. Jos virheellinen rivi otetaan tietovarastoon saakka, se pitää merkitä virheelliseksi. [1.]

Muokkauksen toinen vaihe on varsinainen tietojen muuntaminen, jossa yhdenmukaistetaan poimitut ja tarkistettut tiedot tietovarastotietokannan raporttia tukeviin rakenteisiin. Monasti tietosarakkeita ei tarvitse muokata ja ne kelpaavat sellaisenaan tietovarastoon saakka. Esimerkkinä muokkauksesta on tietojen täydentäminen, jossa postinumeron mukaan riville liitetään myös kaupunki ja maa. [1.]

Lataus voidaan suorittaa INSERT-lauseella SQL-kielessä, tietokannan omilla latausohjelmilla tai ETL-välineillä. Latausvaiheessa muokatut tiedot ladataan tietovarastoon (DW) [1.]. Latausvaiheessa tietoa voidaan kirjoittaa päälle tai lisätä. Vaihtoehtoja on käyttötarkoituksesta riippuen monia. Microsoft SQL Server 2012 BI tarjoaa ELT-toiminnot työkaluillaan. Työssäni käytettävä tietovarastot eli kohteet ovat Microsoftin tarjoamat tietokannat AdventureWorks2008DW ja AdventureWorksDW2008 R2.

ETL-ajaja olisi paras ajoittaa liiketoiminnan kannalta hiljaisimpiin hetkiin toteuttamalla esimerkiksi yöllisiä eräajolatauksia, jotta päiväaikaan työskentely oli helpompaa ja nopeampaa [1.]. Nykyisin globaalisti ympäri vuorokauden toimivissa suuryrityksissä ei hiljaisia hetkiä ole, joten palvelimilta vaaditaan korkeaa käytettävyyttä, skaalautuvuutta ja nopeaa toipumista, jotta vasteajat varmistettaisiin ETL-ajaja ja raportteja tehdessä. Tätä varten on kehitetty SQL Server 2012 AlwaysOn High Availability Group -teknologia, joka mahdollistaa joko synkronisesti tai asynkronisesti rakennettavien sekundääristen lukunoodien toteutuksen. [3.]

3.2 OLAP, Data Mining ja rakenne

3.2.1 OLAP

Moniulotteinen analysointi (Online Analytical Processing, OLAP) on yksi BI-ratkaisujen menetelmistä, joka tarjoaa moniulotteisen näkymän tietoon analysointia varten [1.]. OLAP-tietokannat on suunniteltu yritystietoja koskevien kyselyjen tekemiseen ja nopeuttamaan tietojen noutamista. OLAP-tiedot on järjestetty hierarkkisesti ja tallennettu kuutioihin. Tämä on tekniikka, jonka moniulotteisten rakenteiden ansiosta analysoitavat tiedot saadaan nopeasti. OLAP-tietokannat sisältävät mittatietoja, jotka ovat numeerisia tietoja, kuten määriä ja keskiarvoja, sekä ulottuvuustietoja, jotka ovat näiden mittojen järjestämisessä käytettäviä luokkia. [4.]

3.2.2 Data Mining

Tiedonlouhinnassa (Data Mining) on kyse BI:n yhdestä tiedon analysointiin tarkoitettusta menetelmästä. Se on kehitetty valtavan kokoisia tietokantoja varten, aina biotekniikan tietokannoista tavanomaisempiin kauppojen suuriin tietokantoihin.

Tiedonlouhintamenetelmä kehitettiin nimensä mukaan, jotta piilevä hyödyllinen informaatio saataisiin esiin tietovarastoista. Yleisesti on käytetty tilastollisia, matemaattisia ja tilastojenkäsittelyllisiä puoliautomaattisia analysointimenetelmiä louhimaan dataa. [1.]

3.2.3 Rakenne

Microsoft BI -kehitystyökalut tukevat sekä OLAP- että Data Mining -toimintoja. Analysis Services käyttää sekä palvelin- että työasemakomponentteja toimittamaan OLAP- ja Data Mining -toimintoja Business Intelligence -sovelluksiin.

- Palvelinkomponentti on toteutettu Microsoft Windows -palveluna. Käytössä voi olla useita instansseja samassa tietokoneessa, ja jokainen Analysis Services -instanssi toteutetaan erillisenä Windows-palvelun instanssina.
- Asiakasohjelmat kommunikoivat Analysis Servicesin kanssa käyttäen avoimen standardin mukaista XML for Analysis (XMLA) -standardia, SOAP-pohjaista

protokollaa, jolla annetaan komentoja ja vastaanotetaan vastauksia. Tämä toimii web-palveluna. Myös asiakasobjektin mallit annetaan XMLA:lla ja niitä voidaan käsitellä tavalliseen tapaan käyttämällä joko ulkopuolista palveluntarjoajaa, kuten ADOMD.NETiä tai natiivia OLE DB -rajapintaa.

- Kyselykomentoja voidaan antaa käyttämällä Data Mining Extensions (DMX) -kieltä, alan standardikieltä, joka on orientoitunut tiedonlouhintaan. Analysis Services -komentosarjakieltä (ASSL) voidaan myös käyttää hallitsemaan Analysis Services -tietokantaobjekteja. [5.]

3.3 Raportointi

BI-ratkaisuissa raportointia pidetään tärkeimpänä tiedon hyödyntämismuotona. BI-raportointimenetelmiä kehitetään jatkuvasti vastaamaan yritysmaailman kiihtyvää tiedonsaamistarvetta. BI-ratkaisuja on kehitetty niin, että kyselyjen tekeminen on mahdollisimman helppoa rakentamalla käyttäjille selkeitä käyttöliittymiä. [1.]

SQL Server Reporting Services -raportointityökaluilla voidaan luoda, muokata ja ottaa käyttöön raporteja.

4 SQL Server 2012

SQL Server 2012 on perinteinen asiakas/palvelin-relaatiotietokannan hallintajärjestelmä, joka suorittaa asiakkaan palvelimen välisen liikenteen SQL-kielellä. Sen päätehtäviin kuuluu varastoida tietoa ja hakea tallennetun tiedon joukosta se tieto, jota asiakasohjelmat haluavat. Asiakasohjelmat voivat sijaita eri paikoissa. Pyyntöjä voidaan joko tehdä verkon yli muilta koneilta, tai tieto voi sijaita palvelimella itsellään.

SQL Server 2012 on uusin Microsoftin pilvivalmis tietokantapalvelu, ja se toimii myös mobiililaitteilla, joko julkisessa (Public Cloud) tai yksityisessä (Private Cloud) pilvessä. [6.]

4.1 Asennus

SQL Server 2012 tarjoaa laajennetut mahdollisuudet virtuaalisissa ympäristöissä. SQL-palvelimen jokaista versiota voidaan käyttää joko 32- tai 64-bittisinä ja asentaminen voidaan suorittaa joko komentoriviltä, opastetusti tai Sysprepillä. Asennus on mahdollista suorittaa vanhan version kanssa rinnakkain (side-by-side migration) tai aiemman päälle (in-place upgrade). Tietoturvan ja käytettävyyssajan parantamiseksi SQL Server 2012 on mahdollista asentaa myös Windows Server Core -käyttöjärjestelmälle. [7.]

4.2 Saatavuus

SQL Serveriin 2012 on lisätty saatavuuden laajennuksia. Tämä parantaa tietokantojen käytettävyyssaikaa ja lisää niiden luotettavuutta. Uutena SQL Server -ominaisuutena on edellä mainittu AlwaysOn Availability Groups -teknologia, joka tarjoaa tietokantojen palautusta ja saatavuutta. Tämä ominaisuus suojaaa tietokantoja ottamalla niistä kopioita, ja se on integroitu SQL Server Management Studion (Object Explorer) kanssa. [7.]

4.3 Tietoturva

Palvelimen tietoturvahukien minimoimiseksi ja toiminnallisten virheiden vähentämiseksi pääsy palvelimelle rajataan valtuutettuihin käyttäjiin määrittelemällä roolit käyttäjämäärittelyissä palvelinrooleissa (user-defined server roles). Käyttäjämäärittelyt erottelvat oikeuksia nykyisin tietokantatasolla.

SQL-serveriin sisälletty tietokanta-autentikointi (Contained Database Authentication) tarjoaa tietokantatodennuksen, jossa ei ole tarvetta kirjautua tietokantakoneeseen (Database engine). Sisälletyn tietokantatodennuksen avulla käyttäjä voi halutessaan kirjautua suoraan käyttäjätietokantaan. [7.]

4.4 Versiot

SQL Server 2012:sta on julkaistu seuraavat pääversiot: Standard, Enterprise ja Business Intelligence.

Standard-versio on suunniteltu pienten ja keskikokoisten yritysten tietokantoihin ja rajoitettuun BI-sovelluksien tiedonhallintaan. Nämä ovat käytössä pk-ratkaisuissa, pienissä organisaatioissa tai osastokohtaisissa ratkaisuissa.

Enterprise-versio on suunniteltu suuria tietokeskuksia varten. Se tarjoaa suorituskykyä ja korkeamman tason saatavuutta sovelluksille. Enterprise-versio mahdollistaa käytön myös yksityisessä pilvessä. Siinä on tarjolla myös virtualisoidut ympäristöt ja keskitettyjä tai ulospäin suuntautuneita BI-ratkaisuja.

Microsoft julkaisi ensimmäistä kertaa SQL Server Business Intelligence -version. Versio on täysin uusi SQL-palvelin aiempiin versioihin nähden. Versio sisältää laajat BI-menetelmät, skaalatun raportoinnin, analysoinnin, Power View'n ja PowerPivotin. Tämä versio on räätälöity erityisesti organisaatioille, jotka haluavat saada yritykselle Business Intelligence- ja itsepalvelukapasiteettia, mutta jotka eivät tarvitse täyttä OLTP-suorituskykyä ja skaalautuvuutta, jonka Enterprise tarjoaa. SQL Server 2012 BI-asennus käydään läpi kappaleessa 5. [6.]

4.5 Teknologia

SQL Server 2012 sisältää useita tietojen hallintaan ja analysointiin tarkoitettuja teknologioita.

Database Engine on SQL Serverin tietokantakone ja ydinpalvelu, joka on tarkoitettu tiedon varastointiin, käsittelyyn ja suojaamiseen. Sen tarkoitus on mahdollistaa tietokannan hallittu käyttö niin, että tapahtumat voidaan käsitellä vaativimpienkin sovellusten tarvitsemalla nopeudella. Database Engine tarjoaa myös laajan tuen, jotta hyvää käytettävyyttä voidaan ylläpitää. [8.]

Data Quality Services (DQS) on tarkoitettu datan puhdistukseen. DQS:n avulla voidaan rakentaa tietokanta, jonka tietoa voi käyttää korjaamaan ja karsimaan dataa käyttäen apuna sekä tietokoneavusteisia että vuorovaikutteisia menetelmiä. Se mahdollistaa pilvipalveluiden käytön referenssidatapalveluissa. DQS:n avulla pystytään rakentamaan tiedonhallinnan ratkaisu, joka integroituu DQS:n SQL Server Integration Servicesin ja Master Data Servicesin kanssa. [8.]

Master Data Services (MDS) on SQL Serverin ydintietojen hallintaan kehitetty ratkaisu. Ominaisuudella varmistetaan, että raportointi ja analyysi perustuvat oikeaan tietoon. Master Data Services luo säilytyspaikan ydintiedolle ja ylläpitää tarkistettavissa olevia ja turvallisia rekistereitä seuraamalla niitä aika ajoin. [8.]

Replikoinnilla (Replication) tarkoitetaan joukkoa tekniikoita tietojen ja tietokannan kohteiden kopiointiin ja jakeluun yhdestä tietokannasta toiseen. Tekniikka keskittyy myös synkronointiin tietokantojen välillä, jotta näiden keskinäinen johdonmukaisuus säilyy. Replikoinnin avulla voidaan jakaa tietoa eri paikkoihin. Tämä on mahdollista toteuttaa sekä etänä että mobiilisti paikallisten ja alueverkkojen, puhelinyhteyksien, langattomien yhteyksien ja internetin kautta. [8.]

5 Ympäristö ja asennus

Insinööriyön työympäristö asennettiin suoraan kotikoneelle. Koneelle oli alun perin asennettu SQL Server 2008 R2, joten asennuksen olisi voinut halutessa tehdä joko rinnan yhdistämällä sovelluksia ja tietokantoja tai toinen vaihtoehto olisi ollut asentaa se vanhan päälle. Asennus päätettiin tehdä ”puhtaalta pöydältä” poistamalla aiempi Server.

Toinen vaihtoehto olisi ollut rakentaa työympäristö käyttämällä kurssin suosittamaa WMWARE:n virtualiohjelmistoa. Etuna virtualisoinnissa on resurssien säästö ja riippumattomuus.

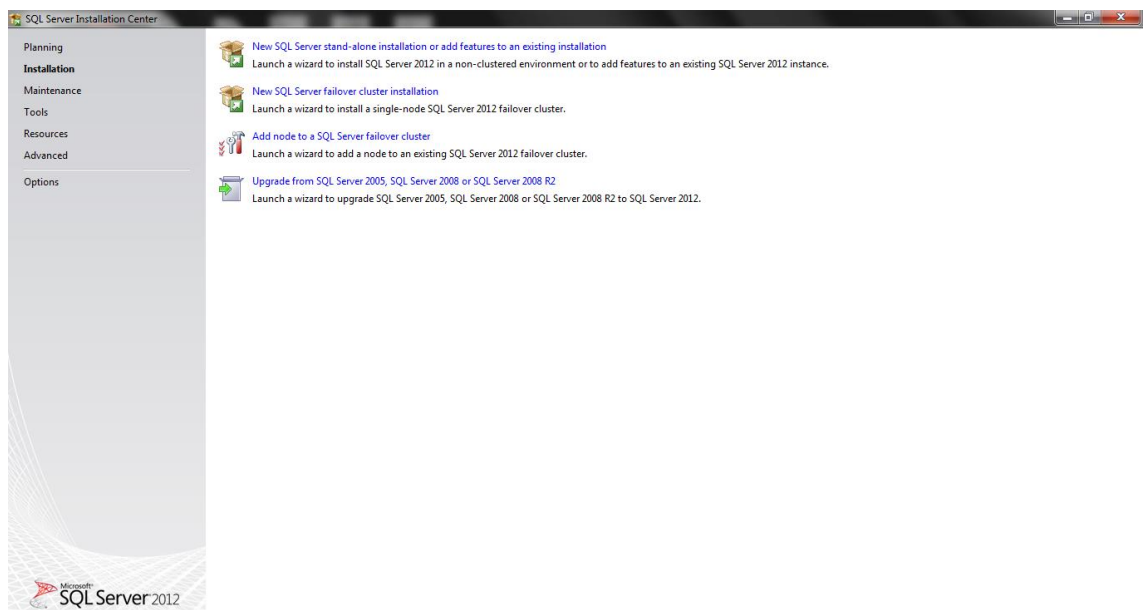
Tiedot koneesta, jolle asennus suoritettiin:

- Windows 7 Home Premium 64-bit.
- Intel® Core™ i3-2310M CPU @ 2.10GHz
- 4Gt RAM.

5.1 SQL Serverin asennus

Microsoft SQL Server 2012 Business Intelligence, Service Pack 1 (64-bit), asennettiin työn tekijän kotikoneelle Windows 7 -käyttöjärjestelmälle. Asennettava palvelin oli ilmainen Metropolia Ammattikorkeakoulun opiskelijoille ja ladattavissa Microsoft Store -sivustolta omilla opiskelijatunnuksilla.

Asennus käynnistettiin SQL Server Installation Centerissä valitsemalla New SQL Server stand-alone installation or add features to an existing installation.



Kuva 1. SQL Server Installation Center.

Asennusvaiheessa ei ilmennyt ongelmia. Apuna käytettiin Moodle-työtilassa olevaa itseopiskeluun tarkoitettua Microsoft SQL Business Intelligence (BI) ratkaisut -kurssin SQL Server 2008 R2 -asennusohjetta. Yhtä poikkeusta lukuun ottamatta SQL Server 2012 BI asennettiin oletusasetuksilla. Reporting Services Configuration -asennusvaiheessa asennettiin kohdat Reporting Services Native Mode ja Reporting Services SharePoint Integrated Mode asetuksella Install only.

5.2 AdventureWorks2008-tietokannat

SQL Server 2012:n asennuksen jälkeen asennettiin AdventureWorks2008-tietokannat. Nämä tietokannat ovat Microsoftin omia, kurssikäyttöön tarkoitettuja tietokantoja, jotka ladattiin Microsoft SQL Server Product Samples -sivustolta [9.]. Tietokantojen asennukseen oli sivustolla annettu tarkat ohjeet. Halutessaan voi käyttää AdventureWorks2008R2- tai AdventureWorks2012-tietokantoja. Nämä kannat sisältävät samat tiedot, mutta AdventureWorks2012-tietokantoja käyttämällä asennukset sujuvat paremmin harjoituksissa SQL Server 2012:n kanssa.

5.3 SQL Serverin hallinta- ja kehitystyökalut

SQL Server 2012 sisältää kaksi graafista käyttöliittymää, joiden avulla voi hallita, valvoa, ylläpitää ja kehittää SQL Server -ympäristöä [10.]. Nämä ovat myös työssä käytettävät kehitystyökalut.

SQL Server Management Studion (SSMS) avulla voidaan tehdä lähes mikä tahansa yllämainituista toiminnoista. SSMS:ssä yhdistyy Enterprise Managerin, Query Analysisin ja Analysis Managerin piirteet. Management Studio toimii kaikkien SQL Serverin komponenttien, kuten Integration Servicen, Reporting Servicen ja Analysis Servicen, kanssa. [11.]

SSMS on integroitu ympäristö, josta on pääsy SQL Serverin instansseihin. Siinä on SQL Serverin kehittäjille ja ylläpitäjille useita työkaluja, jotka yksinkertaistavat SQL Serverin instanssien luomista ja määrittämistä. [10.]

SSMS sisältää Object Explorer-, Solution Explorer- ja Templates Explorer -työkaluikkunat.

- Object Explorer on hierarkkinen käyttöliittymä, jolla voi tarkastella ja hallinnoida kaikkia objekteja yhdessä tai useamassa instanssissa SQL Serverillä.
- Solution Explorer -työkalu aukeaa Ctrl Alt L -komennolla, ja sillä voidaan rakentaa projekteja, joilla voidaan hallinnoida skriptejä ja kyselyitä.

- Templates Explorer -työkalu avautuu Ctrl Alt T -komennolla, ja sen avulla voi rakentaa ja hallinnoida tiedostoja, mikä nopeuttaa skriptien ja kyselyiden kehitystä. [12.]

SSMS:n lisäksi SQL Server 2012 sisältää SQL Server Data Tools (SSDT) -kehitystyökalut. SSDT on toinen integroitu ympäristö, mutta se on suunniteltu erityisesti tietokannan kehittäjille. Tietokantaa ja tietokannan kohteita voi tutkia käyttämällä Data Toolsin SQL Server Object Exploreria. SSDT:n ominaisuuksia ovat mahdollisuus luoda tai muokata tietokannan kohteita ja tietoja ja tehdä kyselyjä suoraan käyttöliittymästä. [10.]

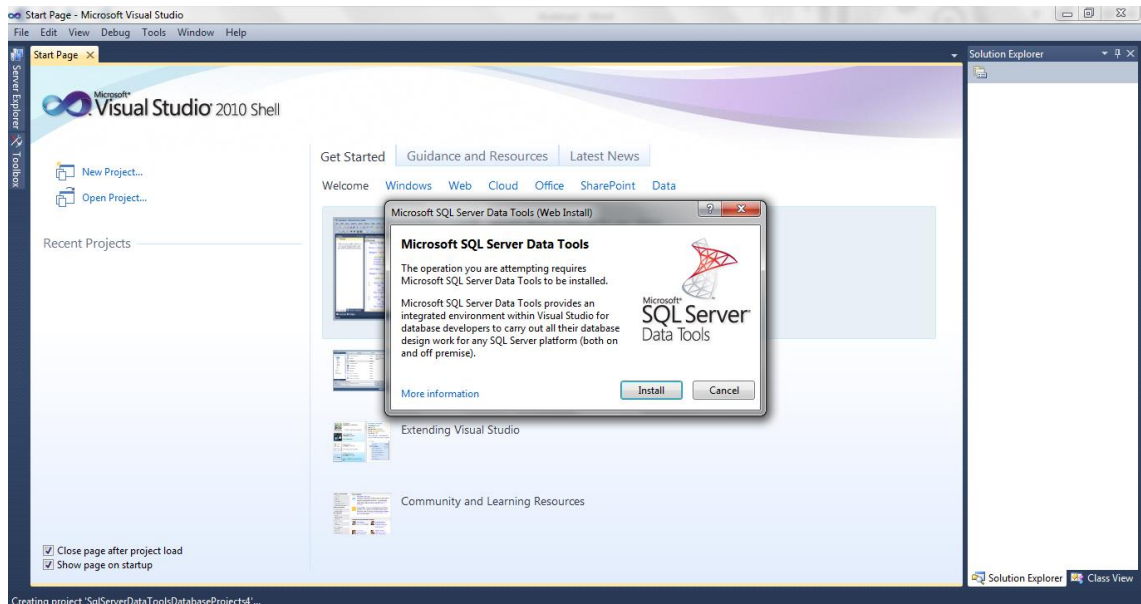
5.4 SQL Server Data Tools

SQL Server 2008 R2:n Business Intelligence Development Studio (BIDS) on korvattu SQL Server 2012 BI-versiossa SQL Server Data Tools (SSDT) -kehitystyökalulla. Tämä on pohjaltaan Microsoft Visual Studio 2010:n ympäristö. Ympäristö on integroitu SQL Serverin komponenttien kanssa ja siihen on yhdistetty BI-projektimalleja.

SQL Server 2012 BI:n asennuksen yhteydessä asentuu Visual Studio 2010. Ensimmäisen Visual Studio 2010:n käynnistyksen yhteydessä täytyy asentaa SQL Server Data Tools.

5.5 SSDT-asennus

SSDT-ympäristön asennusvaiheeseen päästiin, kun avattiin Visual Studio 2010 ensimmäistä kertaa asennuksen jälkeen ja valittiin kohta New Project. Seuraavassa kohdassa valittiin Business Intelligence SQL Server, ja näin päästiin aloittamaan Microsoft SQL Server Data Toolsin asennus Visual Studio 2010 -ympäristöön.



Kuva 2. Microsoft SQL Server Data Tools (Web Install).

Asennus suoritettiin loppuun valitsemalla Business Intelligence Settings ja käynnistämällä Visual Studio 2010 SQL Server Data Tools.

Ympäristön asetuksia voi halutessaan vaihtaa SSDT-valikon kautta valitsemalla Tools menu, Import and Export Settings ja valitsemalla Reset all settings.

Samalla asennettiin .NET Framework 4 -ominaisuus. Tämä on ohjelmistokomponenttikirjasto, jota Microsoftin Visual Studio.NET -ympäristössä kehitetyt ohjelmiston käyttäjät.

5.6 SSDT-ympäristö

SQL Server Data Tools Microsoft Visual Studio 2010 -ympäristö sisältää muutamia rakenteellisia uudistuksia käytettävyyden yksinkertaistamiseksi.

- SSIS-työkalulaatikko (SSIS Toolbox) on uudistus, jolla voidaan lisätä tehtävä- ja data flow -komponentteja pakettiin sen sijaan, että käytettäisiin Visual Studion työkalulaatikkoa, joka oli käytössä vielä SQL Server Integration Servicesin aiemmassa versiossa. Työkalupakissa on myös päivitetty kuvakkeet toiminnoille.

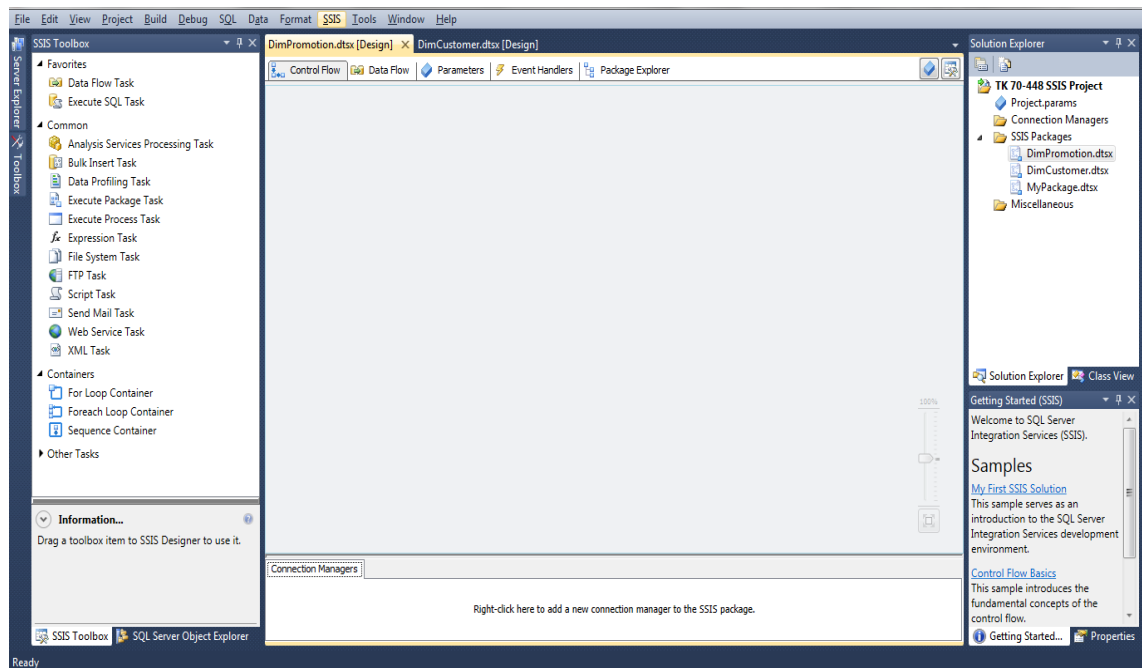
- Paketin suunnittelutyökalu sisältää uuden parametri-napin (Parameters tab), jolla voi avata paketin parametri-ikkuna. Parametrilla voi määrittellä ajonaikaiset arvot paketille, containereille, tehtäväominaisuuksille tai taulukoille.
- Paketin suunnittelutyökalun työkalupalkissa oleva muuttujat-nappi (variable button) avaa tälle ikkunan.
- SSIS Toolbox -nappi on uusi nappi paketin suunnittelutyökalun käyttöliittymässä ja se avaa SSIS:n työkalulaatikon silloin, kun se ei ole näkyvissä.
- Getting Started on aloitusopas, joka opastaa työskentelemään SSIS-työkalulla. Tämän ikkunan kautta pääsee linkkeihin, jotka vievät videoihin ja näytteisiin.
- Control flow- ja data flow -suunnittelunäkymät sisältävät nyt zoomauksen ohjauksen (Zoom control). [6.]

SSDT sisältää erilaisia wizardeja, työkaluja ja kehittämiä. SSDT sisältää myös Control Flow- ja Data Flow -designerin sekä Solution Explorer-, Connection Manager- ja Properties-ikkunat.

Solution Explorerissa ei ole enää Data Sources- ja Data Sources View -kansioita, vaan ne on korvattu uudella tiedostolla Project.params ja uudella kansiolla Connection Managers. [6.]

SQL Server Data Toolsissa on uutena ominaisuutena SQL Server Object Explorer, jota ei aiemmassa SQL Server 2008 R2 -versioissa ollut. Tämän avulla voidaan rajoitetusti hallinnoida tietokantoja ja tehdä suunnittelutyötä. Object Explorerilla pystytään luomaan, muokkaamaan, nimeämään uudelleen ja poistamaan tauluja. Sillä voidaan myös suorittaa kyselyitä suoraan käyttöliittymästä.

Kuvakaappauksessa uusi SSIS -ympäristö:



Kuva 3. SQL Server Data Tools (SSIS).

6 Laboratoriomanuaalin harjoitukset

Tässä luvussa käydään läpi 43 harjoitusta laboratoriomanuaalista, joka pohjautuu kirjaan MCTS Self-paced Training Kit (Exam 70-448) Microsoft SQL Server 2008 Business Intelligence Development Maintenance. Harjoitukset ja poikkeukset kuvataan työssä geneerisesti ilman syvempää tarkastelua. Harjoitusten kuvaamisella on tarkoitus antaa käsitys, että harjoitukset voidaan suorittaa SQL Serverillä 2012. Laboratoriomanuaalin komponentit ovat SQL Server Integration Services (SSIS), SQL Server Analysis Services (SSAS) ja SQL Server Reporting Services (SSRS).

6.1 SQL Server Integration Services

SQL Server Integration Services -projekteissa voidaan rakentaa tiedon integrointiratkaisuja, mukaan luettuna SSIS-paketit (package), jotka tarjoavat ETL-prosessointia tietojen varastointia varten. [8.]

Integration Services -uudistukset SQL Server 2012:ssa:

- muutoksia suunnitteluosan (designer) käytettävyyteen tuottavuuden parantamiseksi.
- parannuksia ETL-projektien käyttöönottoon, määrittämiseen, raportointiin, vianmäärittämiseen ja valvontaan.
- ETL-pakettien käynnistys etänä.
- SQL Server Integration Servicesin ajaminen erillisellä SQL Server -instanssilla.
- Change Data Capture (CDC) SQL Serverille usein toistuvan prosessoinnin tehokkaaseen hallintaan [13.]. Tämän avulla voidaan muun muassa jäljittää ja toimittaa tietokannan muuttuneet tietueet. [1.]

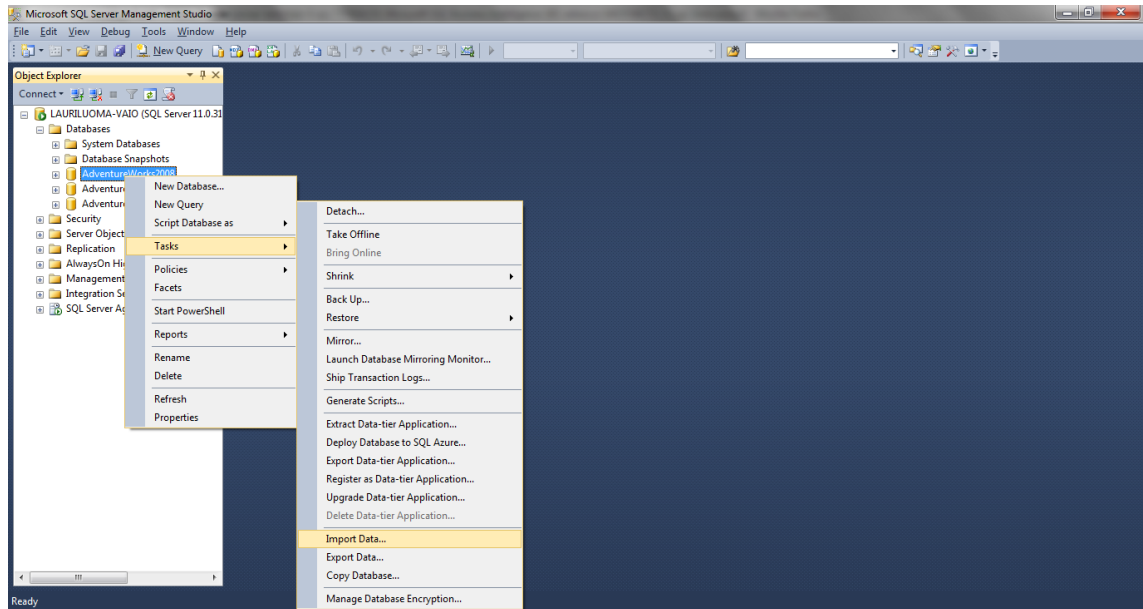
6.2 SSIS-paketti

SQL Server Integration Servicesin pääobjekti on paketti (package), joka sisältää tiedon käsittelyyn, poimintaan ja muokkaamiseen tarvittavan logiikan, jonka avulla data siirretään kohteeseensa. Paketit sisältävät myös tärkeitä work flow -elementtejä datan prosessointiin. Kun paketti suoritetaan (execute), sen sisäinen logiikka suorittaa suunnitellut vaiheet. [14.]

Paketit sisältävät yhteystiedot datalähteistä (data source) ja datakohteista (data destinations). Nämä yhteydet rakennetaan erilaisiin ulkoisiin järjestelmiin, kuten tietokantoihin, tiedostoihin, File Transfer Protocol (FTP) -palvelimiin, Simple Mail Transfer Protocol (SMTP) -palvelimiin ja niin edelleen. Yhteyksiä käytetään SSIS:n data flow -koneessa ja SSIS:n control flow -koneessa. [14.]

Opetuskäytössä oleva kirja neuvoo paketin luomisen kahdella eri tavalla. Paketin voi luoda nopeasti ja ohjeistetusti käyttämällä Microsoft SQL Server Management Studioon Import And Export Wizardia. SSMS on edellä kuvattu kehitystyökalu, jolla päästään hallinnoimaan SQL Serverin komponentteja.

SSMS:n käynnistyksen jälkeen Object Explorerin kautta voi varsin pian aloittaa paketin luonnin. Jos tietokanta on sen datan lähde, joka halutaan lähettää eri järjestelmään, valitaan Export Data. Jos halutaan lähettää parhaillaan järjestelmän ulkopuolella olevia tiedostoja tietokantaan, valitaan Import Data.



Kuva 4. SQL Server Management Studio.

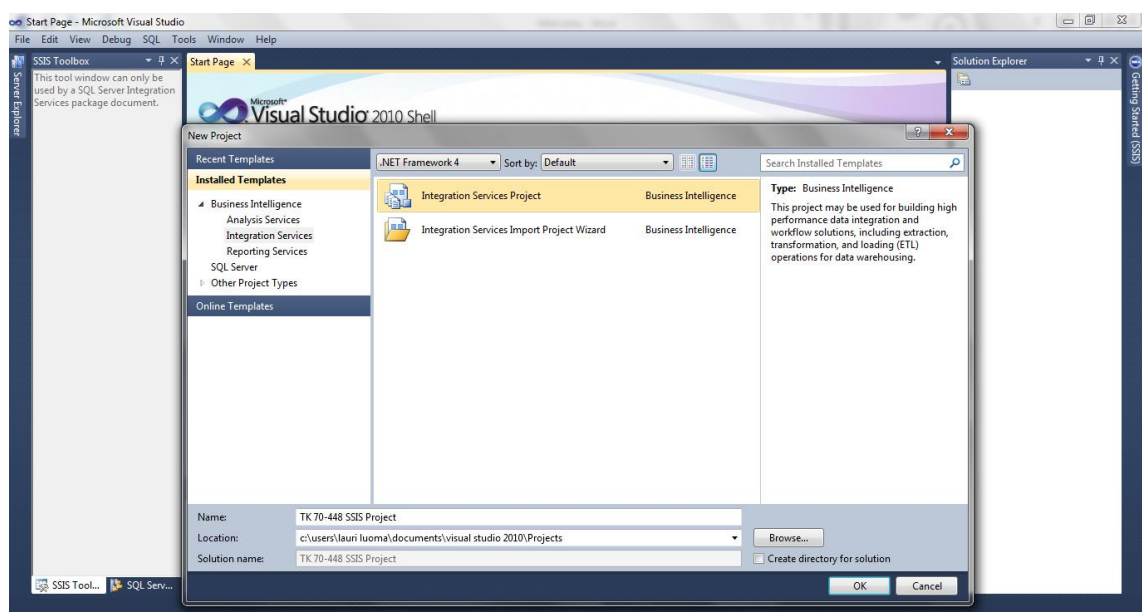
Wizardilla luotu paketti voidaan suorittaa heti, ajastaa se, tai tallentaa ja yhdistää se SSIS-projektiin. Wizardilla luotuja paketteja voidaan tosin käyttää rajoitetusti, koska niissä voidaan määrittää vain yksi lähde ja kohde. Wizardi ei myöskään jaa datalähteitä muiden pakettien kanssa.

Wizardi pystyy kopioimaan dataa mille tahansa ja mistä tahansa tietolähteestä, jolle on olemassa hallinnoitu .NET Framework-tietopalvelu tai natiivi OLE DB -palvelu. Käytävissä olevat palveluntarjoajat sisältävät seuraavat tietolähteet:

- SQL-palvelin
- yksitasoinen tiedosto
- MS Office Access
- Excel-taulukko.

Import and Export Wizard on hyödyllinen, kun pitää luoda tietoja yhdestä lähteestä yhteen kohteeseen siirtävä nopea paketti. Nämä paketit ovat kuitenkin usein vain lähtökohta. Useimmiten pitää joko kehittää paketti, jossa on monimutkaisempia vaatimuksia, tai luoda koordinoitujen pakettien joukko. Näissä tapauksissa pitää ensin luoda uusi SSIS-projekti SQL Server Data Toolsissa.

BI-komponenttien projektit (SSIS, SSAS ja SSRS) luodaan SSDT:ssä alussa samalla tapaa. SSIS:llä työskennellessä käytetään Integration Services -mallia (Templates) ja nimetään projekti laboratoriomanaalin ohjeen mukaan TK 70-448 SSIS -projektiksi. Tämän jälkeen päästiin luomaan ensimmäinen paketti.



Kuva 5. Integration Services Project.

Kun luodaan paketteja SSDT:ssä, paketti tallennetaan tiedostojärjestelmään dtsx-tiedostotunnisteella. Tähän XML-tiedostoon tallentuvat paketin logiikka ja asetukset.

Koska SSIS:n pääasiallinen tarkoitus on siirtää tietoa lähteestä kohteeseen, seuraava vaihe on lisätä lähde- ja kohdepointterit. Nämä pointterit ovat datalähteet (Data Sources) ja yhteydet (Connections). Datalähteet tallennetaan projektitasolla, ja yhteydet määritellään pakettien sisällä.

Datalähde on SSIS-projektin kohde. Datalähteissä on yhteysmerkkijonot, jotka viittaavat tiedostoihin tai tietokantoihin, ja niihin voi viitata yhdessä tai useammassa paketis-

sa. Datalähteet ovat valinnaisia SSIS:n sisällä, mutta ne ovat hyödyllisiä kehittämisen aikana, jos on iso määrä paketteja, joiden pitää käyttää samaa tietokantaa tai tiedostoa. Datalähteen käyttö auttaa myös, jos pitää vaihtaa monen paketin käyttämä yhteys. Silloin voi vaihtaa datalähteen kerran ja avata sen jälkeen projektin jokainen paketti, mikä automaattisesti synkronoi pakettiin tietolähteen kanssa tallennetun yhteysmerkkijonon.

Paketin yhteysmanageri (Package Connection Manager), joka tunnetaan myös nimellä paketin yhteys (Package Connection), on riippumaton tietolähteistä. Paketin yhteydet voivat kuitenkin viitata projektin tietolähteeseen. Paketin yhteydet mahdollistavat sen, että SSIS:n eri komponentit ovat yhteydessä kohteisiin pakettien ulkopuolella (tietokanta, tiedosto tai palvelin). Paketin yhteyksiä voi käyttää lähtheadapterina, FTP- tai sähköpostipalvelimena tai yksitasoisena tiedostona.

Control Flow Task on SSIS-komponentti, joka suorittaa tehtäviä kuten sähköpostin lähettämisen, SQL-ohjeen suorittamisen tai tiedoston kopioimisen FTP-palvelimelta. Kun Control Flow Task on suoritettu, se joko onnistuu tai epäonnistuu. Control Flow'ta käytetään koordinoimaan tehtävien suorittaminen rinnakkain tai sen avulla voi määrittää etuoikeusjärjestyksen rajoitteet sen mukaan, kuinka paljon tehtävistä on saatu päätökseen.

Control Flow Containerin avulla voidaan ryhmittää tehtäviä yhteen kontrolloimaan, miten tehtäviä, tilauksia, kirjauksia ja tapahtumia suoritetaan rinnakkain. Säiliöt (Containerit) voivat myös suorittaa tehtäviä niiden puitteissa useaan kertaan perustuen toistuviin vaatimuksiin.

Yksi tärkeimmistä Control Flow Taskeista on Data Flow Task. Paketissa voi olla nolla, yksi tai useampia tietovirtoja. Data Flow Taskilla voidaan työskennellä kuljettamalla Data Flow Task Control Flow -työkalupakista työpöydälle ja kaksoisklikkaamalla sitä tai klikkaamalla Data Flow -nappia SSIS-suunnittelunäkymässä. Kun Data Flow -nappia on painettu, avautuu Data Flow -työpöytä, missä Data Flow'ta voidaan käyttää tietojoukkojen muuntamiseen ja käsittelyyn.

6.2.1 SSIS-paketin kehitys

Ensimmäisessä varsinaisessa harjoitusluvussa käytiin läpi SSIS-paketin kehitys (Developing SSIS Packages).

SSIS-paketin ja tietolähteiden luonti:

Ensimmäisessä harjoituksessa käytettiin SQL Server Integration Services Project -mallia, luotiin kolme SSIS-pakettia ja nimettiin ne. Seuraava vaihe oli luoda datalähteet ja pakettien yhteydet. Harjoituksessa käytettiin asennettuja AdventureWork2008- ja Adventureworks2008DW-tietokantoja. Datalähteiden lisäämisen jälkeen tallennettiin tiedot.

Poikkeuksena aiempaan SSDT:ssä projektin datalähteet lisättiin Solution Explorerissa Connection Managers -kansion kautta. SQL Server 2012:ssa on poistettu Data Sources -kansio, jonka kautta projektin datalähteet aiemmin lisättiin.

Control Flow -objektien luonti ja muokkaus:

Harjoituksessa käydään läpi objektien luonti ja muokkaus Control Flow'ssa. SSIS-työkalupakista siirrettiin Control Flow -työpöydälle SQL-tehtävä. Tehtävään määriteltiin päivityskomento Update ja annettiin tehtävälle nimi Update ProductLine. Seuraavaksi siirrettiin Control Flow -työpöydälle Sequence Container -toiminto ja tämän sisään siirrettiin päivitetty ja uudelleen nimetty SQL-tehtävä. Lopuksi testattiin paketin suoritus uudessa ympäristössä. Harjoitus onnistui.

Paketin voi suorittaa kolmella tavalla: valitsemalla Start Debugging suunnitteluvalikosta, painamalla F5 tai Solution Explorerissa valitsemalla suoritettava paketti ja Execute Package.

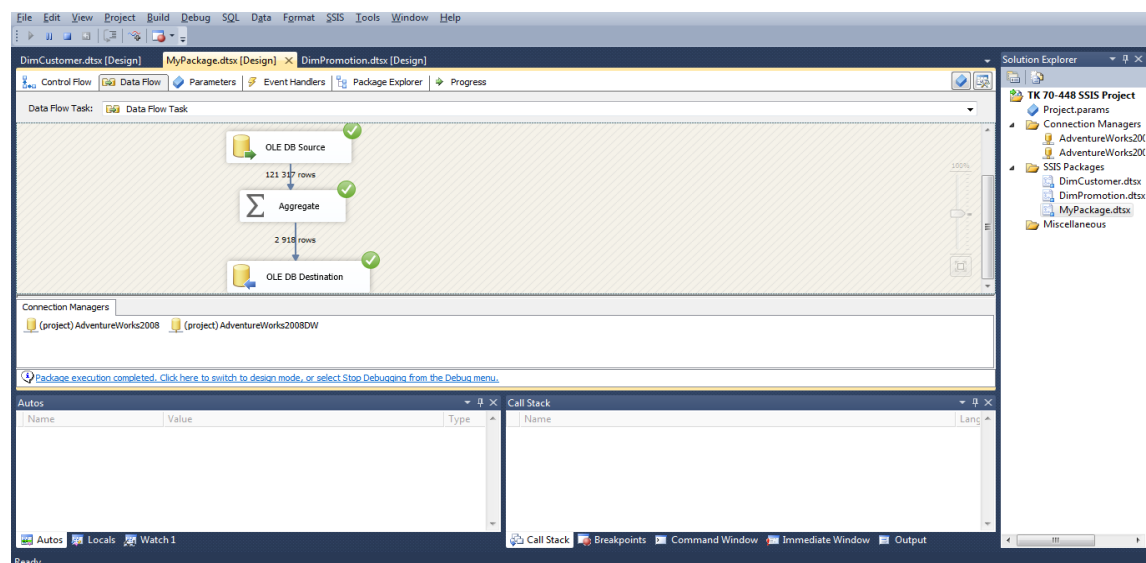
Seuraavassa vaiheessa muokattiin DimCustomer ETL-pakettia Control Flow'ssa. Harjoituksessa luotiin SQL Server Management Studioissa päivityksille uusi tietokantataulu komennolla Create Table ja annettiin päivityskomento paketille.

Poikkeuksena aiempaan Control Flow -työkalupakki sisältää harjoituksessa kaksi uutta kategoriaa toiminnolle ja päivitettyt kuvakkeet useimmille toiminnolleen. Apuna käyttä-

jälle työkalupakki sisältää kuvaukset toiminnolle, joka on sillä hetkellä valittuna. Lisäksi samat päivitykset on tehty Data Flow'n työkalupakkiin, jota seuraavassa harjoituksessa siirrytään käyttämään. Nämä kaksi edellä mainittua työkalupakkia sisältävät eri toiminnot, mutta niihin ei tarkemmin pureuduta tässä työssä.

Data Flow -adaptereiden ja -transformaatioiden käyttö:

Harjoitus jatkaa aiemmin luodun paketin (MyPackage) kehitystä. Control Flow -työpöydälle tuotiin Data Flow Task ja tähän yhdistettiin Update ProductLine -tehtävä. Muokkausta jatkettiin Data Flow -työpöydällä. Data Flow -työpöydälle siirrettiin OLE DB Source (lähde) ja lisäksi tietokantaan tehtiin Select-kysely. Kyselyn muokkaamiseen ja ryhmittelyyn käytettiin Aggregate-toimintoa. Lopuksi Data Flow -työpöydälle siirrettiin OLE DB Destination (kohde) ja luotiin tähän uusi taulu. Tehtävä suoritettiin ja haku tallentui.



Kuva 6. MyPackage suoritettiin.

Seuraavassa vaiheessa jatkettiin DimCustomer-paketin kehitystä ja luotiin Data Flow -kohde. Data Flow -työpöydälle siirrettiin kaksi OLE DB -lähettä, jotka nimettiin Customer Sourceksi ja Customer Dim Sourceksi. Nämä sisältävät kaksi erillistä tulosta, jotka yhdistettiin Merge Join -tehtävällä. Välissä Conditional Split Transformationia editoitiin (Output). Harjoituksen lopuksi tehtiin sarakkeiden kartoitus (mapping). Sarakkeet liitettiin manuaalisesti toisiinsa vetämällä "nuoli" näiden välille. Kartoitus tehtiin Column-

nimen mukaan: Education Column liitettiin EnglishEducation Columniin ja niin edelleen. SSDT ei osannut tehdä tätä automaattisesti.

6.2.2 Virheiden korjaus ja käsittely SSIS:ssä

Toisessa luvussa käytiin läpi virheiden korjaus ja käsittely SSIS:ssä (Debugging and Error Handling in SSIS).

Pakettitransaktioiden ja tarkistuspisteiden konfigurointi:

Harjoituksessa täytyi käynnistää ensin MSDTC-palvelu, eli Microsoft Distributed Transaction Coordinator. Tämä palvelu koordinoi tapahtumia, jotka hajautetaan useisiin tietokantoihin, sanomajonoihin, tiedostojärjestelmiin ja muihin tapahtumasuojattuihin resurssinhallintaohjelmiin.

Harjoituksessa TransactionOption-ominaisuudeksi asetettiin required. Control Flow -työpöydälle luotiin Execute SQL Task. SQL-tehtävän suoritus epäonnistui. Tämä johtuu TransactionOptionin required-asetuksesta, koska tietokantaan ei ole mennyt päivityksiä. Transaction-komento palauttaa datan ennalleen, jos käsky epäonnistuu.

Harjoituksessa Properties-ikkunan kautta asetettiin paketille tarkistuspistetiedosto. Tarkistuspiste määritty XML-tiedostoon, joka sisältää paketin informaation, kun paketti epäonnistuu. Harjoituksessa määriteltiin FailPackageOnFailure-ominaisuudeksi true ja vielä Data Flow Task Objectin ForceExecutionResult-ominaisuuden arvoksi failure. Tällä simuloitiin epäonnistuneen paketin tarkistuspiste. Lopuksi testataan, miten checkpoint.xml-tiedosto asettuu ja poistuu automaattisesti ForceExecutionResult-ominaisuuden arvon ollessa none.

Paketin tilan tarkistus, logien määrittely ja virhetehtävien käsittely:

Harjoituksessa luotiin etuoikeusjärjestys (precedence constraint), joka määrittää, milloin Control Flow -tehtävät suoritetaan. Harjoituksessa käytiin läpi login kirjaus tietokantaan. Lopuksi katsottiin SSMS:stä login kirjaukset SQL-kyselyllä `SELECT * FROM dbo.sysssislog`.

Tietopolkujen käsittely ja virheenetsintä:

Luvun viimeisessä harjoituksessa tutustutaan datapolkujen käsittelyyn ja virheenetsintään. Polut (paths) käsittelevät tietorivejä ja ne havainnollistetaan kahdella eri tavalla: punaisella virhepolut (error paths) ja vihreällä tietopolut (data paths). Harjoituksessa määriteltiin muuttuja ja virhepolku, ja tämän jälkeen Data Viewerin avulla tulkittiin virheet ulostuloriveistä (Output rows).

Poikkeuksena aiemmasta versiosta Data Viewerin käyttöä on yksinkertaistettu SSDT:ssa, kun käsitellään virhepolkuja.

6.2.3 SSIS-paketin asentaminen ja konfigurointi

Kolmannessa luvussa käytiin läpi SSIS-paketin asentaminen ja konfigurointi (Deploying and Configuring SSIS Packages).

Pakettikonfiguraation ja lausekkeen käyttö:

Pakettien konfiguraatiot ovat oletusarvona pois päältä. Harjoituksessa valittiin XML-konfiguraatiotiedosto ja tämä lisättiin kahteen pakettiin. Lisäksi valittiin SQL-konfiguraatiotyyppi, joka lisättiin seuraavaksi luotuun pakettiin. SQL-skriptillä luotiin tietokantaan uusi SSIS Configurations -taulu, ja tämän sisältö katsottiin SQL Server Management Studiossa.

Harjoituksessa käytiin lisäksi läpi lausekkeiden (Expressions) käyttö yhteyksien päivittämiseksi.

SSIS-paketin käyttöönotto:

Harjoituksessa tutustuttiin paketin käyttöönottoon konfiguraatioiden ja lausekkeiden asentamisen jälkeen.

Paketti voidaan siirtää tiedostojärjestelmään tai SQL Serverille. Harjoituksessa käsiteltiin DTUtil-apuohjelman käyttöä.

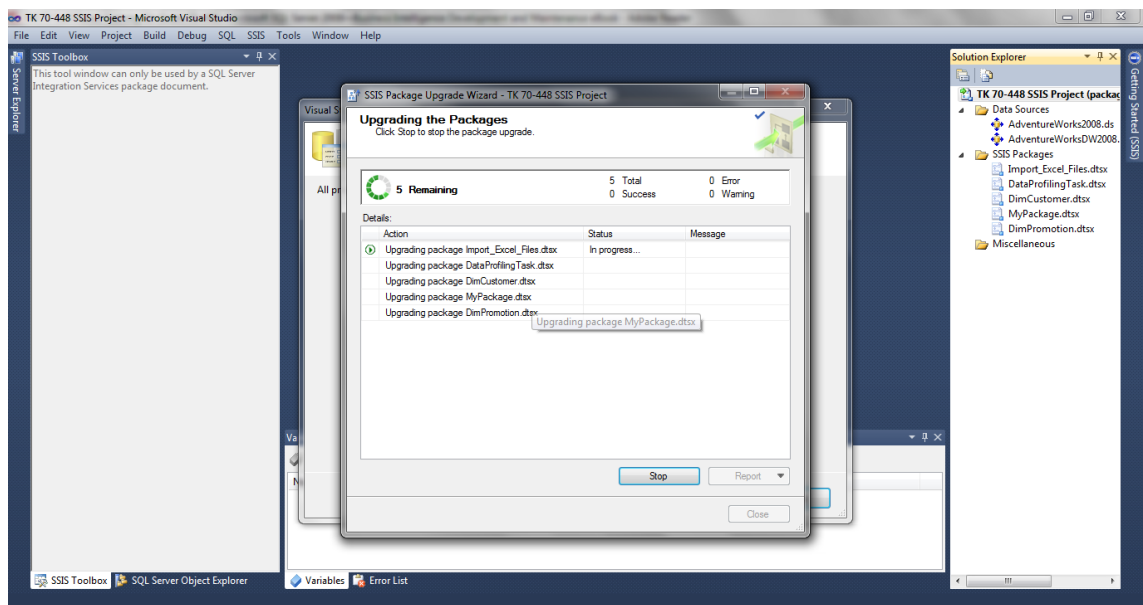
Harjoituksen läpiviemiseksi piti Windows-komponenttipalveluista käydä määrittämässä paikalliselle käyttäjälle (käyttäjätunnus Lauri) pääsyoikeudet MS SQL Server Integration Services -palvelulle, jotta tämän sai aukaistua SQL Server Management Studioissa.

6.2.4 Hallinnointi, suojaaminen ja suorittaminen SSIS-paketeissa

SSIS-osiota käsittelevässä viimeisessä luvussa käydään läpi hallinnointi, suojaus ja suorittaminen SSIS-paketeissa (Administering, Securing and Executing SSIS Packages).

SSIS:n suojaamisella pyritään salamaan sen skeema- ja yhteystiedot. SSIS-paketin voi suojata joko SQL Serverin turvallisuusrooleissa, SQL Server Data Toolsissa tai käytämällä tiedostotason suojausta.

Tehtävän alussa Visual Studio 2008 -projekti täytyi konvertoida Visual Studio 2010:n kanssa yhteensopivaksi, koska harjoituksessa käytettiin valmista SSIS-projektia.



Kuva 7. SSIS Package Upgrade Wisard.

Hallinta SSIS-palvelussa ja pakettisuojausten konfigurointi:

Harjoituksessa käytiin läpi paketin salaus valitsemalla EncryptAllWithPassword ja testattiin paketin suojaus salasanalla SSDT:ssä. Lisäksi harjoituksessa tutustuttiin SSIS Serviceen ja määriteltiin turvallisuusroolit ja salasanasuojaus pakettiin SSMS:ssä.

SSIS-pakettien suorittaminen ja aikatauluttaminen:

Harjoituksessa käytettiin komentorivityökalua ja lisäksi käytettiin SQL Server Agentia, kun SSIS-paketteja suoritettiin ja aikataulutettiin.

Harjoituksessa paketti ladattiin SQL Serverille käyttämällä DTEXECUI-työkalua ja komentoriviä. Lopuksi paketin suoritus aikataulutettiin Management Studioon SQL Server Agentilla.

6.3 Yhteenveto SSIS-harjoituksista

Harjoitukset pystyttiin käymään läpi ilman suurempia poikkeavuuksia SQL Server 2012 BI-kehitystyökaluilla. Laboratoriomanuaali oli tehty alun perin SQL Serverille 2008, AdventureWorks2008-tietokannoille ja BIDS-kehitystyökalulle, joten poikkeavuuksia oli odotettavissa. Erot näkyivät Microsoftin mainostamissa SSDT-ympäristön pienissä rakenteellisissa uudistuksissa ja käytettävyyden parantumisessa aikaisempaan BIDS-ympäristöön nähden. Näkyvin ero kehitystyökalujen versioiden välillä oli projektin datalähteiden lisääminen harjoituksen alussa, koska Solution Explorer -ikkunasta oli poistettu Data Sources- ja Data Source View -kansiot.

6.4 SQL Server Analysis Services

SQL Server Analysis Services on Microsoftin SQL Server -komponentti, jonka työkalut on tarkoitettu henkilökohtaiseen, ryhmän ja yritystason Business Intelligence -käyttöön. Palvelimet ja client designers tukevat perinteisiä reaaliaikaisia analyttisiä prosesseja (kuten OLAP-ratkaisuja), uusia taulukkomuotoisia mallintamisratkaisuja sekä itsepalveluanalysointia. Analysis Services sisältää myös tiedonlouhintaominaisuuden, jotta isoista tietomääristä voi löytää ja paljastaa kaavoja ja suhteita. [8.]

Kahdessa harjoitusluvussa käydään läpi SQL Server Analysis Services -kuution kehitys ja laajennus. Harjoitukset käydään läpi yleiskuvallisesti ja selvitetään poikkeavuudet harjoituksissa, kuten edellisessä SSIS-lukua käsittelevässä osassa.

SSAS-projekti luodaan käyttämällä SQL Server Data Toolsin Analysis Services Multi-dimensional and Data Mining -mallia. Tätä mallia käytetään kehittämään perinteisenkaltaista Analysis Services -projektiä, joka tunnetaan nyt moniulotteisena mallina. Se on ainoa malli, joka tukee Analysis Services -komponentin tiedonlouhintaominaisuuksia.

SQL Server 2012:n BI semanttinen datamalli (BI Semantic Model, BISM) on UDM:n eli Unified Dimensional Modelin seuraaja. BISM tukee sekä entiteettilähestymistapaa käyttäen taulukoita ja entiteetin suhteita että moniulotteista lähestymistapaa käyttämällä hierarkioita ja aggregointia. [6.]

SQL Server 2012:n myötä on siis mahdollista tehdä mallinnusta kahdella eri tavalla: joko tabulaarisessa eli sarakepohjaisessa moodissa tai perinteisessä ”kuutiomoodissa”, joka nykyään totelee nimeä multidimensional. Näin ollen BISM ei tarkoita pelkkää tabulaarista mallia, vaan molempia instansseja. [15.]

6.4.1 SSAS-kuutioiden kehitys

Luvussa käytiin läpi SSAS-kuution kehitys (Developing SSAS Cubes).

Datalähteiden ja datalähdenäkymien luonti:

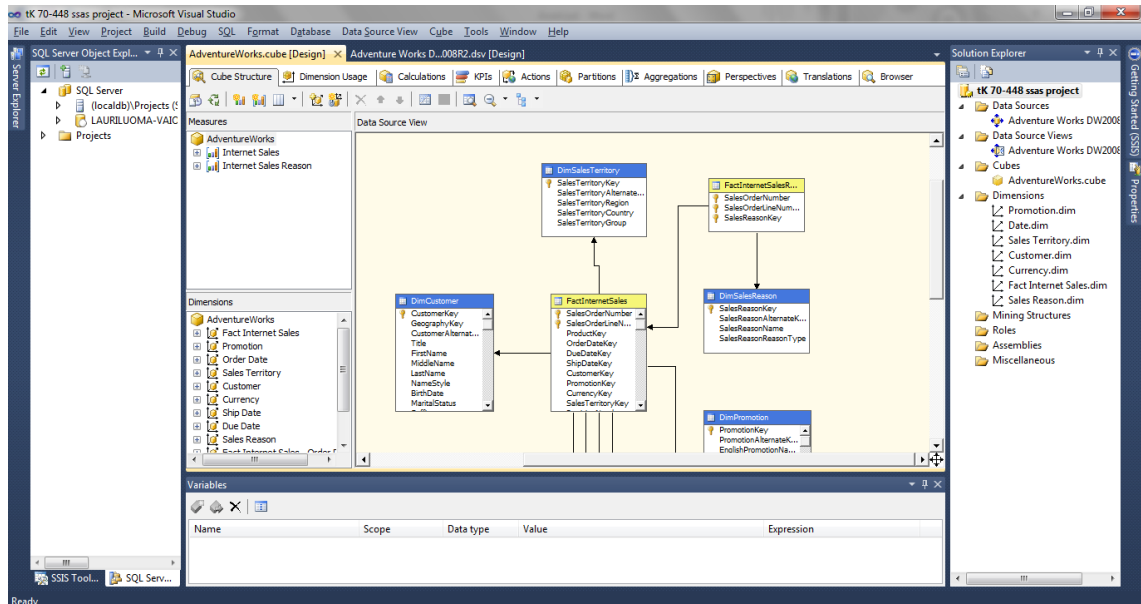
SSAS:llä työskennellessä SSDT:n Solution Explorer sisältää luonnollisesti Data Sources- ja Data Source View -kansiot.

Harjoituksessa luotiin datalähde ja lisäksi käytiin läpi datalähdenäkymän (Data Source View) luonti. Datalähdenäkymään lisättiin uudet nimetyt laskelmat (named calculation). Laskelmaa käytetään mm. kun dimensioiden tiedot sijaitsevat useammassa kuin yhdessä taulussa.

Harjoista tehdessä huomattiin, että AdventureWorks2008DW-tietokannasta puuttui DimDate-taulu, joten tietokanta piti vaihtaa AdventureWorksDW2008R2-tietokantaan.

SSAS-kuution luonti ja muokkaus:

Harjoituksissa SSAS-kuutio luodaan käyttämällä SSDT:n Cube Wizardia. Kuution luonnin jälkeen sitä voidaan muokata ja kehittää käyttämällä SSDT:n Cube Designeria, josta löytyy erilaisia työkaluja tätä varten.



Kuva 8. SSAS Cube Designer sisältää Cube Structure-, Dimension Usage-, Calculations-, KPIs-, Actions-, Partitions-, Aggregations-, Perspectives-, Translations- ja Browser-työkalut.

Harjoituksessa laboratoriomanaalista poiketen piti luoda uusi login nimeltään "NT Service\MSSQLServerOLAPService" SQL Server Database Enginein kautta. Uudelle loginille annettiin lukuoikeus kantaan AdventureWorksDW2008R2 määrittämällä loginille rooli db_datareader.

Dimensioiden luonti ja muokkaus:

Harjoituksessa luodaan dimensiot käyttämällä Dimension Wizardia. Tämän jälkeen muokkausta jatkettiin käyttämällä SSDT:n Dimension Designeria ja Cube Designeria.

Mittausryhmien luonti:

Harjoituksessa käytiin läpi uusien mittausryhmien (Measure Groups) ja niiden mittauspaikkojen luominen, määrittäminen ja konfigurointi. Harjoituksessa käytettiin SSDT:n Data Source View Designeria.

6.4.2 SSAS-kuutioiden laajennus

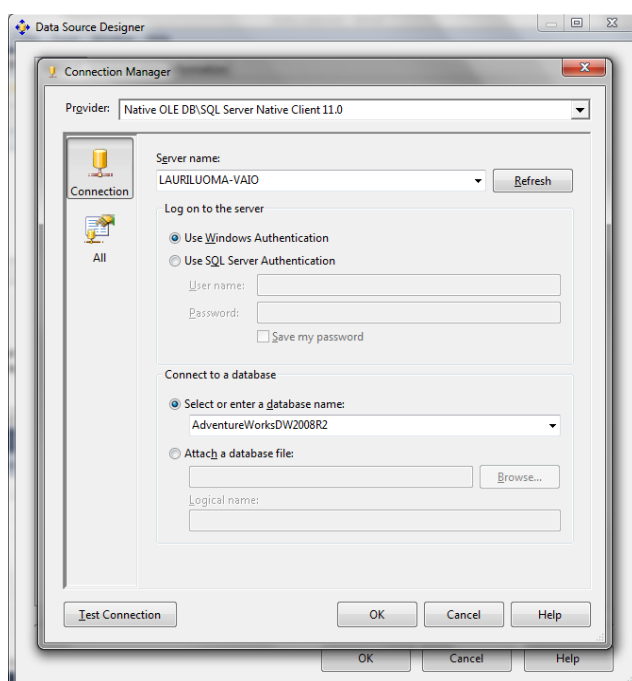
Toinen luku käsittelee SSAS-kuutioiden laajentamista (Extending SSAS Cubes). Tämä on myös viimeinen SSAS-lukuja käsittelevä harjoitus.

Käyttäjähierarkioiden ja dimensioiden määrittely:

Harjoituksessa luotiin dimensioiden hierarkioita. Lisäksi määriteltiin suhteita dimensioiden ja mittaryhmien välillä. Nämä suhteet oli määriteltävä tarkasti ohjeita seuraten, jotta harjoitus onnistui.

Avainsuoritusosoittimien, toimenpiteiden, käännösten ja perspektiivien luonti:

Harjoituksen alussa suoritettiin konvertointi, koska harjoituksessa käytettiin valmista SSAS-projektia. Connection Managerissa vaihdettiin palveluntarjoaja (Provider) ja lisättiin oikea datalähde, koska asetukset eivät olleet oikeat. Asetuksia päästiin vaihtamaan Data Sources -kansion kautta.

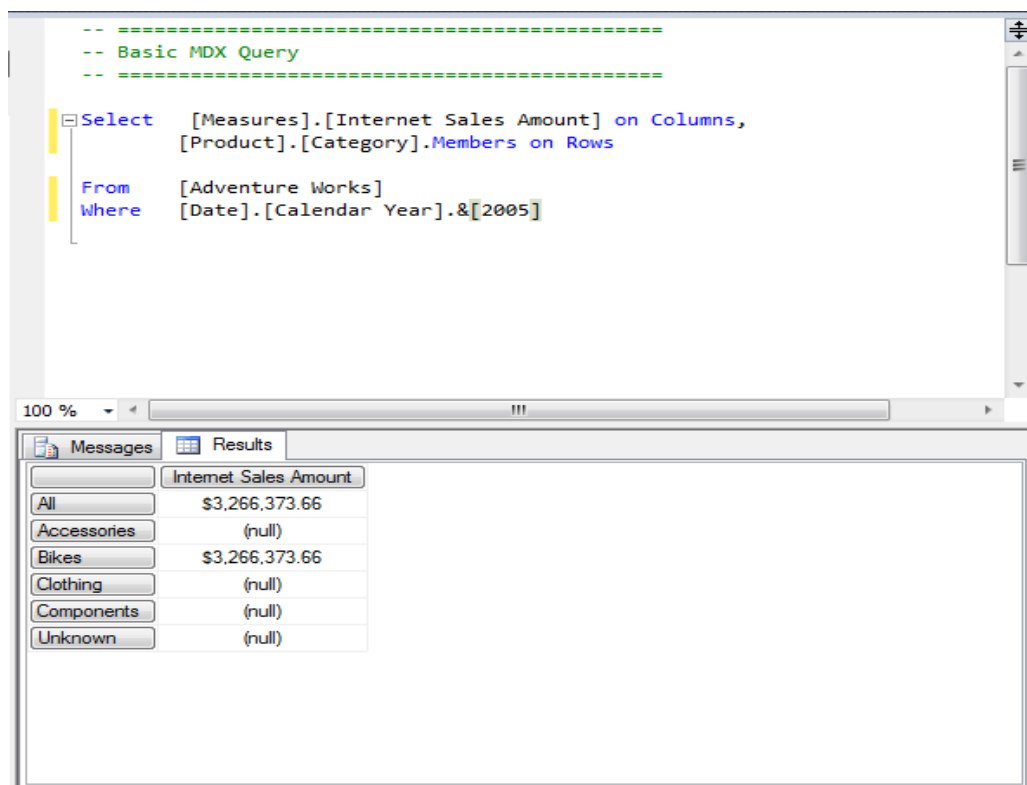


Kuva 9. Connection Managerissa testattiin yhteys asetusten jälkeen.

Harjoituksessa luotiin KPI (Key Performance Indicators), jonka tarkoitus on mitata suorituskykyä. Luotiin uusi Drillthrough-toimenpide ja käsiteltiin dimensioiden tasoja. Lisäksi tehtiin kuution sisällä käännös (Translations) espanjankieliselle käyttäjälle. Lopuksi luotiin perspektiivit, jotka näyttävät vain valitut kuution objektit.

Laskelmien ja kyselyiden luominen MDX-kyselykielellä:

Harjoituksessa luotiin SSMS:ssä Multidimensional Expression (MDX) -kysely.



Kuva 10. MDX-kysely SSMS:ssä.

Lisäksi harjoituksessa toteutettiin laskettuja jäseniä (Calculated Member) ja nimettyjä joukkoja (Named Set) käyttämällä SSDT:ssä Cube Designer -työkaluja.

6.5 Yhteenveto SSAS-harjoituksista

SSAS-lukujen harjoitukset pystyttiin viemään läpi SQL Server 2012 BI-työkaluilla. Laboratoriomanuaalin harjoituksia seuraamalla ei huomattu merkittäviä eroavaisuuksia uudessa ympäristössä.

6.6 SQL Server Reporting Services

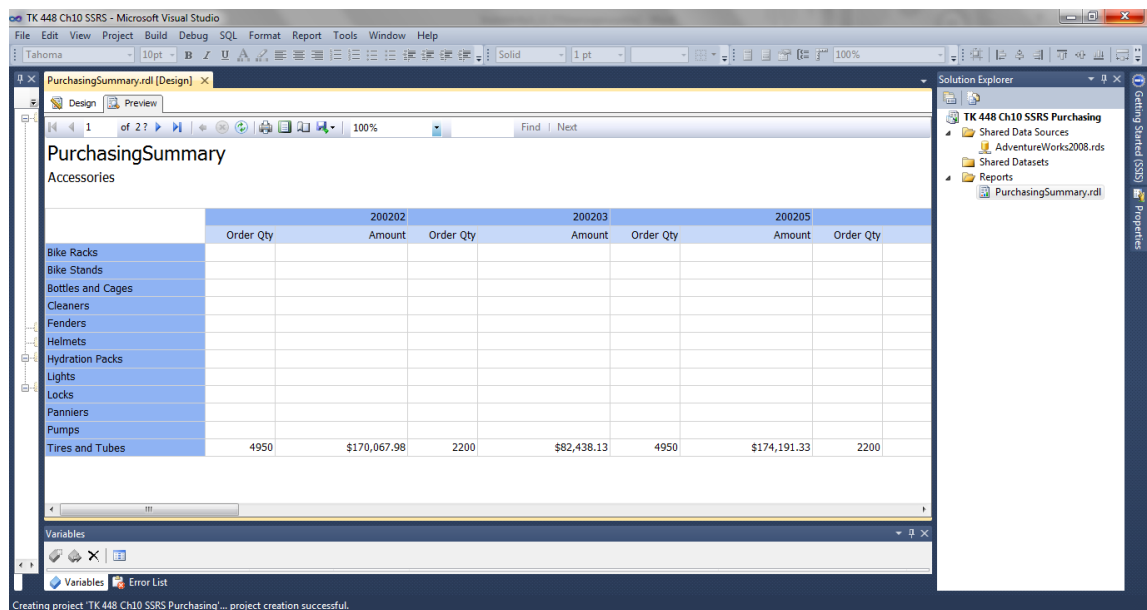
SQL Server Reporting Service tuottaa www-pohjaisia raportointitoimintoja, joilla voidaan luoda raportteja useista tietolähteistä, julkaista raportteja eri formaateissa ja hallinnoida keskitetysti turvallisuutta ja tilauksia. SSRS on myös palvelinpohjainen raportoinnin alusta, jossa on raportointitoiminnot useille tietolähteille. Reporting Services -työkaluilla voidaan luoda, hallita ja toimittaa raportteja. Se sisältää myös API-rajapintoja, joiden avulla voidaan integroida tai laajentaa tietojen ja raporttien käsittelyä mukautetuissa sovelluksissa. [8.]

6.6.1 SSRS-raportin kehitys

Lopuksi testattiin yhdellä harjoituksella SSRS-komponentin soveltuvuutta SSDT-ympäristöön.

SSRS-projektin ja raportin luonti SSDT:ssä:

SSRS-projekti luotiin alussa käyttämällä Report Wizardia. Tämän jälkeen luotua raporttia muokattiin Report Designerilla luettavampaan muotoon. Harjoituksessa PurchasingSummary-raportti tarkistettiin PDF-muodossa.



The screenshot shows the SSRS Report Designer interface. The report is titled 'PurchasingSummary' and is categorized under 'Accessories'. The data is presented in a table with the following structure:

	200202		200203		200205	
	Order Qty	Amount	Order Qty	Amount	Order Qty	Amount
Bike Racks						
Bike Stands						
Bottles and Cages						
Cleaners						
Fenders						
Helmets						
Hydration Packs						
Lights						
Locks						
Panniers						
Pumps						
Tires and Tubes	4950	\$170,067.98	2200	\$82,438.13	4950	\$174,191.33

Kuva 11. SQL Server Reporting Services

6.7 Yhteenveto SSRS-harjoituksesta

SSRS-harjoitus voitiin suorittaa uudessa SSDT-ympäristössä. Uusi SSDT-kehitystyökalu vaikutti harjoituksen perusteella hyvin samanlaiselta kuin edeltäjänsä BIDS.

7 Yhteenveto

Insinööriyön tarkoitus saavutettiin suunnitelman mukaisesti. Kurssin laboratoriomanuaalin läpikäytyt harjoitukset todettiin yhteensopiviksi työssä käytössä olleeseen SQL Server 2012 BI -ympäristöön. Uusi asennettu SQL Server Data Tools (SSDT) osoittautui hyvin samantapaiseksi BI-ratkaisuiden kehitystyökaluksi harjoituksia läpikäydessä kuin Microsoft SQL Business Intelligence (BI) -kurssilla käytetty Business Intelligence Development Studio (BIDS).

Insinööriyössä kohdattiin myös joitakin pieniä ongelmia. Harjoituksia läpikäydessä huomattiin, että AdventureWorks2008DW-tietokannasta puuttui kokonaan tehtävissä tarvittava taulu. Tämän vuoksi jouduttiin asentamaan AdventureWorksDW2008R2-tietokanta tilalle. Harjoituksissa tuli eteen myös tilanteita, joissa huomattiin SQL-kielen tuntemuksen olevan avuksi ja aiempi laajempi käyttökokemus SQL Servereistä olisi ollut merkittävä apu.

BI-kehitystyökalujen erojen selvittäminen aiempaan versioon vei paljon aikaa, vaikka käytännössä suuria muutoksia ei ilmennytäkään. Näistä selvittiin käyttämällä mm. Microsoftin TechNet-sivustoa, joka todettiin nopeimmaksi keinoksi hakea tietoa.

Insinööriyö antoi laajennetun käsityksen BI-ratkaisuissa käytetyistä työkaluista ja niiden käyttötarkoituksista.

Lähteet

- 1 Tietovarastot ja Business Intelligence. Jyväskylä: Docendo, 2009. Hovi A, Her-
vonen H, Koistinen H. Luettu 01.12.2013.
- 2 Tietoviikko, Markku Pervilä, 9.8.2013. Verkkodokumentti.
<<http://www.tietoviikko.fi/cio/forrester+bi+kiilaa+ohjelmistojen+ykkoseksi/a919726>>. Luettu 07.10.2013.
- 3 Technet-sivusto. Verkkodokumentti. <
<http://blogs.technet.com/b/markohot/archive/2012/11/06/reaaliaikainen-raportointi-ja-etl-operatiivisesta-sovelluksesta-ja-tietokannasta.aspx>> Luettu 29.11.2013.
- 4 Microsoft Office tuki -sivusto. Verkkodokumentti. <http://office.microsoft.com/fi-fi/excel-help/olap-online-analytical-processing-yleiskatsaus-HP010177437.aspx#BMwhat_is_on-line_analytical_processing> Luettu 10.10.2013.
- 5 Logical Architecture. Verkkodokumentti. < <http://technet.microsoft.com/en-us/library/bb522595.aspx>> Luettu 03.12.2013.
- 6 Introducing Microsoft SQL Server 2012. Microsoft Press, 2012. Ross Mistry, Stacia Misnes. Luettu 03.12.2013.
- 7 Sivukokoelman luonti SharePoint 2013 -ympäristöön. Insinööriyö. Tekijä: Elina Bergman. <
http://publications.theseus.fi/bitstream/handle/10024/61583/Elina_Bergman.pdf?sequence=1> Luettu 15.10.2013.
- 8 Books Online for SQL Server 2012. Verkkodokumentti.
<<http://technet.microsoft.com/en-us/library/ms130214.aspx>>. Luettu 7.10.2013.
- 9 Microsoft SQL Server Product Samples-sivusto. Microsoft, Inc.
<<http://msftdbprodsamples.codeplex.com/releases/view/93587>>. Luettu 08.10.2013.
- 10 Microsoft SQL Server 2012, Step by Step. O'Reilly Media, Inc. 2012. Patrick LeBlanc. Luettu 30.11.2013.
- 11 Use SQL Server Management Studio. Verkkodokumentti. <[tech-net.microsoft.com/en-us/library/ms174173.aspx](http://technet.microsoft.com/en-us/library/ms174173.aspx)>. Luettu 15.10.2013.
- 12 Tool Windows in SQL Server Management Studio. Verkkodokumentti. <
<http://technet.microsoft.com/en-us/library/ms174168.aspx>>. Luettu 10.10.2013.

- 13 Microsoft: SQL Server Integration Services. Verkkodokumentti.
<<http://www.microsoft.com/en-us/sqlserver/solutions-technologies/enterprise-information-management/integration-services.aspx>>. Luettu 15.10.2013.
- 14 Microsoft SQL Server 2008- Business Intelligence Development and Maintenance. Self-paced Training Kit. Microsoft Press, 2009. E. Veerman, T. Lachev, D. Sarka. Luettu 01.12.2013.
- 15 Business Intelligence for Better Information. Verkkodokumentti.
<<http://aureolisblogi.blogspot.fi/2012/12/microsoft-businessintelligence-with.html>>. Luettu 15.12.2013.

