



Jemi Vuori

Microsoft SQL Server 2008:n ylläpito (70-433)

Metropolia Ammattikorkeakoulu
Insinööri (AMK)
Tietotekniikka
Insinöörityö
Päivämäärä 1.12.2011

Tekijä(t) Otsikko	Jemi Vuori Microsoft SQL Server 2008:n ylläpito (70-433)
Sivumäärä Aika	28 sivua 1.12.2011
Tutkinto	insinööri (AMK)
Koulutusohjelma	tietotekniikka
Suuntautumisvaihtoehto	tietoverkot
Ohjaaja(t)	yliopettaja Janne Salonen
<p>Insinööriytyössä oli tavoitteena tehdä kurssimateriaali ja virtuaalikoneen malli Microsoftin SQL Server 2008 -tietokannan kehittäjäkurssia varten Metropolia Ammattikorkeakoululle. Kurssi on valinnainen osa tietoverkkojen koulutusohjelmaa.</p> <p>Työn pohjana toimi Microsoft SQL Server 2008 - Database development -itseopiskelukirjan harjoitukset ja sen mukana tullut materiaali.</p> <p>Työssä käydään aluksi läpi, mitä uutta SQL Server 2008:ssa on verrattuna aikaisempiin versioihin ja mikä on SQL Azure. Sen jälkeen käydään läpi, mitä kurssi pitää sisällään ja miten levykuvan luominen onnistui.</p> <p>Insinööriytyössä tehtiin valmis levykuva, joka voidaan asentaa koulun virtuaaliympäristöön tai VirtualBoxiin. Kurssikirjan sisältämät tehtävät myös tarkistettiin, että ne toimivat syntyneellä alustalla.</p>	
Avainsanat	microsoft, sql, server, ylläpito, 70-433, azure

Author(s) Title	Jemi Vuori Microsoft SQL Server 2008 administration (70-433)
Number of Pages Date	28 pages 1 December 2011
Degree	Bachelor of Engineering
Degree Programme	Information Technology
Specialisation option	Data networks
Instructor(s)	Janne Salonen, Senior Lecturer
<p>The purpose of this final project was to make complete course material for Microsoft SQL Server 2008 database development course. The project was done for Metropolia University of Applied Sciences. The course is optional part of information technology programme.</p> <p>The project was based on Microsoft's self-paced training kit called Microsoft SQL Server 2008- Database Development (70-433).</p> <p>First this report discusses what is new in SQL Server 2008 version, compared to previous versions and what is SQL Azure. Next chapters take a look what this course contains and how successful was making the image of the training platform that Metropolia's virtual training environment needs.</p> <p>The platform consists of a operating system image that can be installed on the virtualization software. After that it was made sure that the training kits exercises work in that environment.</p>	
Keywords	microsoft, sql, server, administration, 70-433, azure

Sisällys

1	Johdanto	1
2	Tietokannat	2
2.1	Historia	2
2.2	Hyödyt	2
3	Microsoft SQL Server	3
3.1	Historia	3
3.2	SQL Server 2005	4
4	Microsoft SQL Server 2008	5
4.1	Turvallisuus	5
4.2	Tehokkuus	6
4.3	Microsoft SQL Server 2008 R2	7
5	ANSI-SQL ja T-SQL	8
5.1	ANSI-SQL	8
5.2	Transact-SQL	9
6	Microsoft SQL Azure	10
6.1	Pilvipalvelut	10
6.2	Microsoft Azure	11
6.3	Microsoft SQL Azure	11
7	Harjoitusympäristö	14
7.1	Eri asennukset ja ohjelmat	14
7.2	Koulun järjestelmä	14
7.3	VirtualBox	15
7.4	Kotilaboratorio	16
7.5	Harjoitustietokannat	20
7.6	Lopputulos	20
8	Kurssin sisältö	21
8.1	Luku 1. Data Retrieval	21

8.2	Luku 2. Modifying Data	21
8.3	Luku 3. Tables, Data Types and Declarative Data Integrity	22
8.4	Luku 4. Using Additional Query Techniques	23
8.5	Luku 5. Programming Microsoft SQL Server with T-SQL	23
8.6	Luku 6. Techniques to Improve Query Performance	24
8.7	Luku 7. Extending Microsoft SQL Server Functionality part 1	24
8.8	Luku 8. Extending Microsoft SQL Server Functionality part 2	25
8.9	Luku 9. An Introduction to Microsoft SQL Server Manageability Features	25
9	Yhteenveto	27
	Lähteet	28

Lyhenteitä ja käsitteitä

SQL	Structured Query Language. Standardoitu kyselykieli, jolla relaatiotietokantaan voi tehdä erilaisia hakuja, muutoksia ja lisäyksiä.
MCTS	Microsoft Certified Technology Specialist. Microsoftin sertifiointiohjelma.
IBM	International Business Machines. Kansainvälinen teknologiayritys.
RJ599	Research Journal 599. IBM:n tutkimusraportti.
ACM	Association for Computing Machinery. Tietotekniikan alan yhdistys.
T-SQL	Transact-SQL. Microsoftin laajennus SQL-kieleen.
ANSI	American National Standards Institute. Yhdysvaltalainen standardointiorganisaatio.
SQL CLR	SQL Common Language Runtime. Teknologia, mahdollistetaan Microsoft .NET:n ajaminen SQL Serverissä.
MARS	Multiple Active Result Sets. Mahdollistaa usean kyselyn tekemisen tietokantaan samaan aikaan.
CPU	Central Processing Unit. Tietokoneen suoritin.
.NET	Microsoftin kehittämä ohjelmistokomponenttikirjasto.
RDBMS	Relational Database Management System. Relaatiotietokannan hallintajärjestelmä.
TDS	Tabular Data Stream. Tiedonsiirtoprotokolla, jonka avulla asiakas- ja isäntäohjelma keskustelevat tietokantaympäristössä.
ISO	International Organization for Standardization. Kansainvälinen standardointijärjestö.

DDL	Data Definition Language.
DML	Data Manipulation Language.
DCL	Data Control Language.
iSCSI	Internet Small Computer System Interface. Internetprotokollaan perustuva standardi, jonka avulla liitetään tietoverkkojen välityksellä yhteen palvelinkeskuksia.
APIC	Advanced Programmable Interrupt Controller. Ohjelmoitava keskeytysohjain.
PAE/NX	Physical Address Extension. Ominaisuus, jolla voidaan laajentaa ohjelma käyttämään suurempaa osaa fyysisestä muistista.
CTE	Common Table Expression. Väliaikaisesti nimetty haun tulos.
API	Application programming interface. Ohjelmointirajapinta.
XML	Extensible Markup Language. Merkintäkieli tai standardi, jolla tiedon merkitys on kuvattavissa tiedon sekaan.
SMTP	Simple Mail Transfer Protocol. Sähköpostien välittämiseen tarkoitettu protokolla.

1 Johdanto

Jatkuva tiedon määrän lisääntyminen vaatii yhä monimutkaisempia, mutta helppokäyttöisiä tiedon tallentamiseen soveltuvia tietokannanhallintajärjestelmiä. Yksi näistä järjestelmistä, johon tämä opinnäytetyö perehtyy, on Microsoft SQL Server 2008.

Yritysten kasvun tarve aiheuttaa jatkuvasti uusia haasteita tietokannan ylläpitoon. Halutaan, että järjestelmät ovat saatavissa kellon ympäri unohtamatta kuitenkaan korkeaa tietoturvaa ja yhteensopivuutta. Näihin tarpeisiin on kehitetty erilaisia työkaluja ja Microsoft SQL Server 2008 yrittää vastata niihin järjestelmänsä luotettavuudella ja joustavuudella.

Työn tavoitteena on tutustua Microsoft SQL Server 2008:n ylläpidollisiin toimintoihin, kuten tiedon hakemiseen ja muokkaamiseen, datatyyppeihin ja taulukoihin, SQL Serverin ohjelmointiin ja tiedonhaun optimointiin.

Työ tehdään Metropolia Ammattikorkeakoulun virtuaaliympäristössä ja työn tuloksesta on tarkoitus syntyä kattava kurssimateriaali koulun käyttöön. Osana lopputulosta on myös tehdä valmis levykuva virtuaalikoneille.

Kurssin tarkoitus on valmentaa opiskelija Microsoft Certified Technology Specialist 70-433 (MCTS) -sertifikaatin suorittamiseen. Kurssi edellyttää, että osallistujalla on aikaisempaa kokemusta tietokantapalvelimien käytöstä tai suoritettuna MCTS 70-432 -kurssi.

2 Tietokannat

Tietokanta on tietotekniikassa käytetty termi tietovarastolle. Se on kokoelma tietoja, joilla on yhteys toisiinsa. Tietokannan ei välttämättä tarvitse olla sähköisessä muodossa, vaan esimerkiksi puhelinluettelokin on tietokanta.

Nykyään suurin osa tietokannoista on relaatiotietokantoja. Relaatiotietokanta yhdistää tietokokonaisuuksia niiden yhteisten attribuuttien mukaan. Tämä tapa helpottaa tiedon ymmärtämistä ja hakua.

2.1 Historia

Relaatiotietokantamallin kehitti tohtori Edgar Frank Codd. Malli mainittiin ensimmäisen kerran vuonna 1969 IBM:n tutkimusraportissa RJ599. Tästä huolimatta artikkeli, jonka katsotaan olevan relaatiotietokantojen perustana, on nimeltään "A Relational Model of Data for Large Shared Data Banks". Artikkelin kirjoitti jo edellä mainittu Edgar Codd, mutta se julkaistiin vasta vuonna 1970 julkaisussa "Communications of the ACM". [1.]

Edgar Codd jatkoi artikkelien julkaisemista seuraavien kahden vuosikymmenen aikana, ja niissä määriteltiin tarkemmin relaatiotietokantojen toteutusta. Näiden vuosien aikana Edgar Codd määrittä yhteensä 333 sääntöä relaatiotietokannoille ja nämä on kuvattu vuonna 1990 julkaistussa kirjassa "The Relational Model for Database Management". [1.]

2.2 Hyödyt

Relaatiotietokanta tallentaa tiedot erillisiin taulukoihin, jotka koostuvat riveistä ja sarakkeista. Jokainen taulukko perustuu erityyppisiin attribuutteihin. Taulukkojen väliset suhteet voidaan määritellä rivin tunnisteiden (*key*) avulla.

Tietojen tallentaminen erillisiin taulukoihin ja niiden riippuvuus toisista taulukoista pienentävät tietokannan kokoa, helpottaa tiedonhakua, vähentävät virheiden todennäköisyyttä ja parantavat tiedon varmistamista. [2.]

3 Microsoft SQL Server

Microsoft SQL Server on relaatiotietokantapalvelin. Sen päätehtävä on varastoida tietoa ja hakea varastoidun tiedon joukosta ne tiedot, joita asiakasohjelmat pyytävät. Pyyntöjä esittävät asiakasohjelmat voivat sijaita joko palvelimella itsellään tai muilla tietokoneilla, jolloin pyynnöt tehdään tietoverkon yli. Microsoft SQL Server käyttää tietojen hakemiseen pääasiallisesti T-SQL- ja ANSI SQL -kieliä.

3.1 Historia

Vuonna 1988 Microsoft julkaisi ensimmäisen version SQL Serveristä. Se suunniteltiin OS/2-alustalle yhteistyössä Sybasen kanssa. 1990-luvun alussa Microsoft alkoi kehittää uutta versiota SQL Serveristä NT-alustalle. Kehitystyön aikana Microsoft päätti, että ohjelmisto tulisi sitoa tiiviisti yhteen NT-käyttäjärjestelmän kanssa. Vuonna 1993 julkaistiin Windows NT 3.1 ja SQL Server 4.2 NT-alustalle. [3.]

Microsoft loi helppokäyttöisen, mutta korkean suorituskyvyn omaavan tietokantaohjelmiston. Tämä ajatusmalli osoittautui hyvin menestyksekkääksi. Microsoftista tuli nopeasti toiseksi suosituin relaatiotietokantaohjelmistojen toimittaja. Vuonna 1994 Microsoft ja Sybase lopettivat virallisen yhteistyönsä. Seuraavana vuonna Microsoft julkaisi version 6.0 SQL Serveristä, jossa käytännössä rakennettiin uudelleen koko ohjelmiston ydin. Versio 6.0 paransi huomattavasti ohjelmiston suorituskykyä, mahdollisti sisäänrakennettuna tietokantojen replikoinnit ja helpotti keskitettyä ylläpitoa. Vuonna 1996 Microsoft julkaisi ohjelmistosta jo version 6.5. Tätä seuraavina vuosina Microsoft julkaisi 6.5-versiosta Enterprise-painoksen ja itse ohjelmasta version 7.0, jossa ohjelmiston tietokantamoottori oli rakennettu jälleen kerran kokonaan uudelleen. [3.]

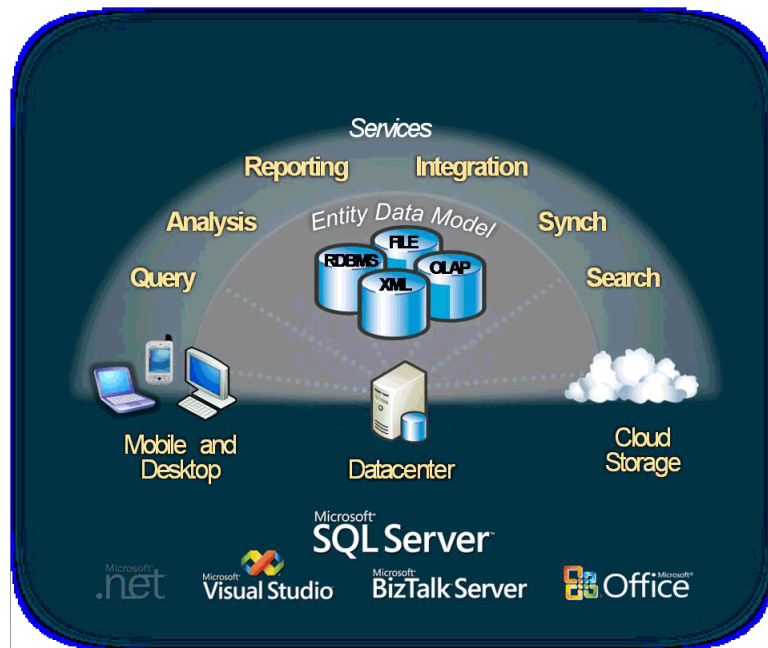
Uutta versiota tietokantaohjelmistosta saatiin odottaa vuoteen 2000 asti ja se nimettiin julkaisuvuotensa mukaan Microsoft SQL Server 2000:ksi. Vaikka ohjelmisto perustui pitkälti edelliseen 7.0-versioon, ohjelma sisälsi suuria muutoksia, joista suurimpana voidaan mainita raportointipalvelut. [3.]

3.2 SQL Server 2005

SQL Server 2005 (koodinimi Yukon) julkaistiin lokakuussa 2005. Se sisälsi tuen niin relaatiotiedolle kuin XML-tiedolle. SQL CLR esiteltiin osana SQL Server 2005:tä, ja se mahdollisti ohjelmiston integroinnin .NET Frameworkin kanssa. Uutena ominaisuutena oli myös MARS, jonka avulla tietokantaan muodostettavia yhteyksiä voitiin käyttää moneen eri tarkoitukseen. Ohjelmiston ensimmäisen päivityksen myötä (SP1) voitiin ottaa käyttöön myös tietokantojen peilaus. Julkaisuversiossa tämä ominaisuus oli jo mukana, mutta se ei ollut tuettuna. [3.]

4 Microsoft SQL Server 2008

Microsoft SQL Server 2008 julkaistiin elokuussa 2008, ja se perustuu pitkälti version SQL Server 2005 vahvuuksille. Tavoitteena ohjelman tuotekehityksessä on ollut vastata Microsoft Data Platform Visionin vaatimuksiin. [4.]



Kuva 1. Microsoft Data Platform Vision

4.1 Turvallisuus

Nykyaikana yritykset tarvitsevat pääsyä järjestelmiensä tietoihin kellon ympäri. Tämän takia tarvitaan vahvoja turvallisuusominaisuuksia, jotka parantavat tietokantaohjelmiston luotettavuutta ja skaalautuvuutta. [4.]

SQL Server 2008 mahdollistaa kokonaisten tietokantojen, tiedostojen ja lokitiedostojen salaamisen ilman muutoksien tekemistä ohjelmistoon. Salaus mahdollistaa tiedon yksityisyyden säilymisen ja mahdolliset organisaatioiden tietoturvasäännösten noudattamisen. [4.]

External Key Management

SQL Server 2008 tarjoaa kattavan ratkaisun salausavaimien hallintaan. Organisaatiot sijoittavat rahaa erillisiin palveluihin, jotka vastaavat eri avaimien turvallisuudesta, mutta SQL Server 2008 vastaa tähän tarpeeseen tukemalla kolmansien osapuolten avainten hallintaa. [4.]

Enhanced Auditing

SQL Server 2008 parantaa turvallisuutta myös mahdollistamalla tehokkaan tiedon auditoinnin. Auditointi perustuu tietoihin siitä, milloin tietoa on luettu tai milloin tietoa on muokattu millään tavalla. Auditoinnin määreet voivat olla myös erilaisia eri tietokantojen välillä ja auditointituloksien suodattaminen onnistuu tarkasti. [4.]

Enhanced Database Mirroring

SQL Server 2008 pakkaa uloslähtevän lokivirran vähentääkseen tietoliikenteen määrää, jonka tietokannan kahdentaminen vaatii. Kahdentamisen edistymisestä on saatavilla myös aikaisempaa tarkempaa tietoa Dynamic Management Viewin ansiosta. [4.]

Hot Add CPU

Aikaisemmin SQL Server mahdollisti muistin lisäämisen laitteistoon sen käynnissä ollessa, mutta SQL Server 2008:n myötä on mahdollista myös lisätä prosessoriresursseja jos palvelinlaitteisto tätä vain tukee. [4.]

Resource Governor

Resource Governor mahdollistaa palvelimen resurssien ohjaamisen ruuhka-aikoina niihin palveluihin, jotka on asetettu etusijalle. Tällöin voidaan tuottaa esimerkiksi loppukäyttäjälle miellyttävän mahdollinen käyttökokemus. [4.]

Performance Data Collection

Suorituskyvyn hienosäätö ja ongelmanratkaisu voivat olla aikaa kuluttavia tehtäviä ylläpitäjälle. Tästä syystä SQL Server 2008:ssa on laaja kokoelma erilaisia suorituskykyä mittaavia määreitä: uusi keskitetty säilytyspaikka tälle tiedolle ja uudet raportointi- ja valvontatyökalut. [4.]

4.2 Tehokkuus

Microsoft .NET Framework ja Visual Studio Team System mahdollistavat ohjelmoijien kehittävän suorituskykyisiä seuraavan sukupolven tietokantasovelluksia. [4.]

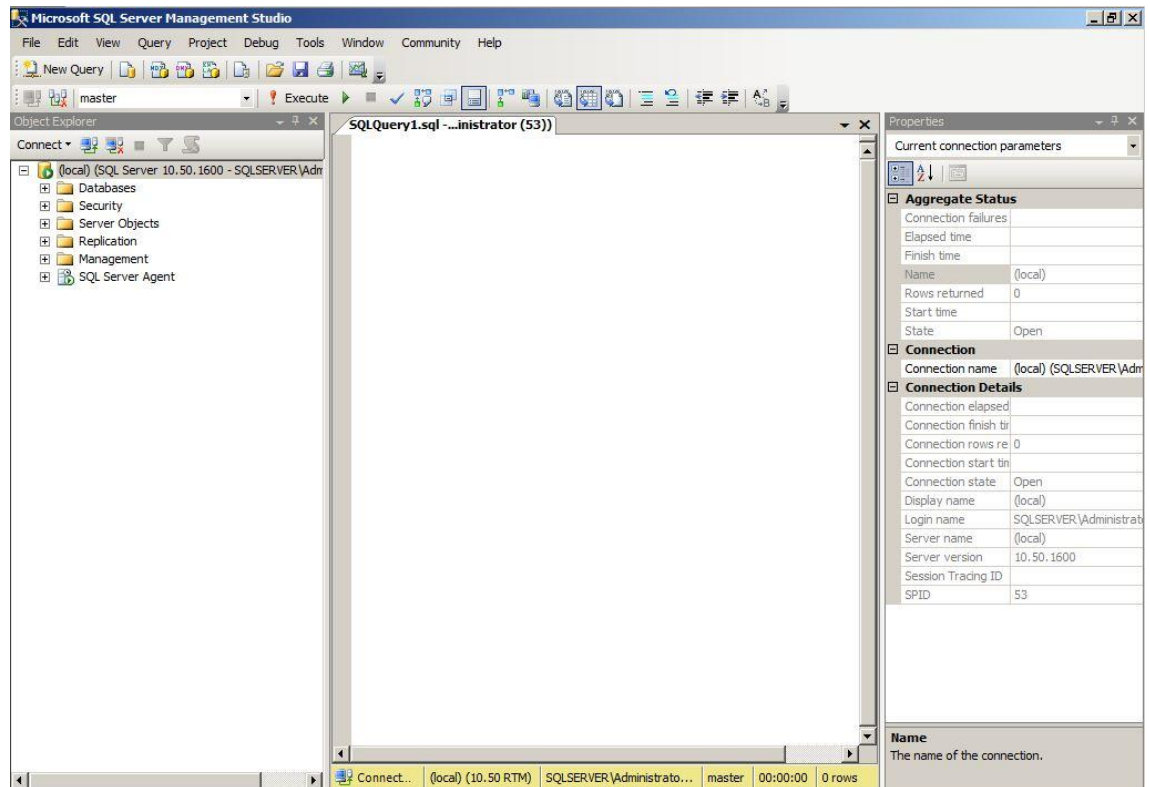
Tavoitteenaan vähentää tuotteen kokonaiskustannuksia asiakkaalle, SQL Server 2008 uutena ominaisuuksena on Policy-Based Framework. Sen etuna on mahdollisuus hallinnoida yhtä tai useampaa instanssia SQL-palvelimesta. [4.]

4.3 Microsoft SQL Server 2008 R2

SQL Server 2008 R2 julkaistiin TechEd 2009 tapahtumassa ja sitä alettiin valmistaa huhtikuussa 2010. Uusina ominaisuuksina R2-versio tarjoaa mm. keskitetyn konsolin, jolla voidaan hallita SQL Serverin instansseja ja palveluja kuten relaatiotietokantoja, raportointipalveluita, analyysipalveluita ja integraatiopalveluita. [4.]

SQL Server 2008:n uusia palveluita ovat:

- PowerPivot for Excel
- PowerPivot for SharePoint
- Master Data Services
- StreamInsight
- Report Builder 3.0
- Reporting Services Add-in for SharePoint
- SQL Server Utility.



Kuva 2. Microsoft SQL Server Management Studion käyttöliittymä

5 ANSI-SQL ja T-SQL

5.1 ANSI-SQL

SQL kehitettiin samaan aikaan relaatiotietokantamallin kanssa IBM:n laboratoriossa San Josessa. Projektin menestyksekkään lopputuloksen johdosta vuonna 1979 SQL:stä tuli suosittu Relational Database Management System (RDBMS) -toimittajien keskuudessa. Tämän seurauksena alkoi markkinoille ilmaantua monia eri yrityksiä tietokantasovelluksiansa kanssa, esimerkiksi Oracle, INGRES ja Sybase. [5.]

Vuonna 1986 SQL sai sertifiointin ANSI ja ISO -standardointi-instituuteilta. Tätä SQL-86 -standardia laajennettiin edelleen uuteen SQL-92 standardiin vuonna 1992. Standardisoidun kielen käyttämisestä on monia eri hyötyjä, kuten pienemmän koulutuskustannukset, riippumattomuus toimittajasta ja yhteensopivuus eri alustojen kanssa. [5.]

SQL-komentoja on olemassa kolmea eri tyyppiä:

- Data definition language (DDL) -komennoilla joko luodaan, muutetaan tai poistetaan taulukoita. Yleensä näitä komentoja voi suorittaa vain tietokannan ylläpitäjä, jolla estetään vahingossa tapahtuvat tai tahallisesti tehdyt vahingolliset komennot
- Data manipulation language (DML) -komentoja käytetään lisäämään, päivittämään tai poistamaan tietoja tauluista
- Data control language (DCL) -komennoilla määritetään tarvittavat oikeudet tietokannan rakenteen muuttamiseen.

5.2 Transact-SQL

T-SQL on Microsoftin ja Sybasen kehittämä oma laajennus SQL-kieleen.

Laajennus sisältää:

- proseduraalisen ohjelmoinnin
- paikalliset muuttujat
- eri funktioita merkkijonojen prosessointiin
- päivämääräprosessoinnin
- laskutoimitukset
- pieniä muutoksia *DELETE*-ja *UPDATE*-lauseisiin.

Nämä lisäominaisuudet tekevät T-SQL:stä turing-vahvan. Kaikki sovellukset, jotka kommunikoivat Microsoft SQL Serverin kanssa, tekevät niin lähettämällä sille T-SQL-komentoja. [6.]

6 Microsoft SQL Azure

6.1 Pilvipalvelut

Pilvipalvelut ovat seuraava iso kehitysaskel tietotekniikan alalla. Vastaavanlaisia palveluita on ollut olemassa aikaisemminkin, mutta vasta tietoliikenneyhteyksien nopeuduttua ja kiintolevykapasiteetin hinnan laskettua niistä on tullut mahdollista tehdä arkipäivää. Käsitettä "pilvi" käytetään kielikuvana, jolla viitataan palvelun monimutkaiseen toteutukseen.

Pilvellä tarkoitetaan, että vaadittavat resurssit eivät sijaitse asiakkaan omissa tiloissa vaan jossain palveluntarjoajan ylläpitämässä tilassa (palvelinkeskus), jonka fyysinen sijainti ei edes ole käyttäjän tiedossa. Tämä mahdollistaa sen, että palvelut ovat saatavissa silloin, kun käyttäjä niitä vaatii ja esimerkiksi lisätehoja tarvittaessa ne saadaan käyttöön välittömästi. Palvelinkeskuksen resurssit jaetaan muiden käyttäjien kesken tarpeen mukaan. Pilvipalveluiden pääluokat ovat IaaS (Infrastructure as a Service), PaaS (Platform as a Service) ja SaaS (Software as a Service).

IaaS (Infrastructure as a Service)

IaaS-luokan pilvipalvelua käytettäessä valitaan ensin käyttäjän tarpeisiin sopiva virtuaalikonene levykuva, johon asennetaan tarvittavat sovellukset - esimerkiksi tietokantaohjelmisto tai web-palvelin. Etuna tämän luokan palvelussa on alustan korkea muokattavuus, mutta tämän ominaisuuden myötä järjestelmän ylläpidosta tulee monimutkaisempaa.

PaaS (Platform as a Service)

PaaS-luokan pilvipalveluissa palveluntarjoajat tarjoavat valmiin alustan, jolle käyttäjän tarvitsee vaan asentaa haluamansa sovellukset ja vastata niiden ylläpidosta. Itse käyttöjärjestelmää ja palvelimen kuormitusta hallinnoi palveluntarjoaja. Tämä vähentää huomattavasti käyttäjän työmäärää ylläpidon suhteen.

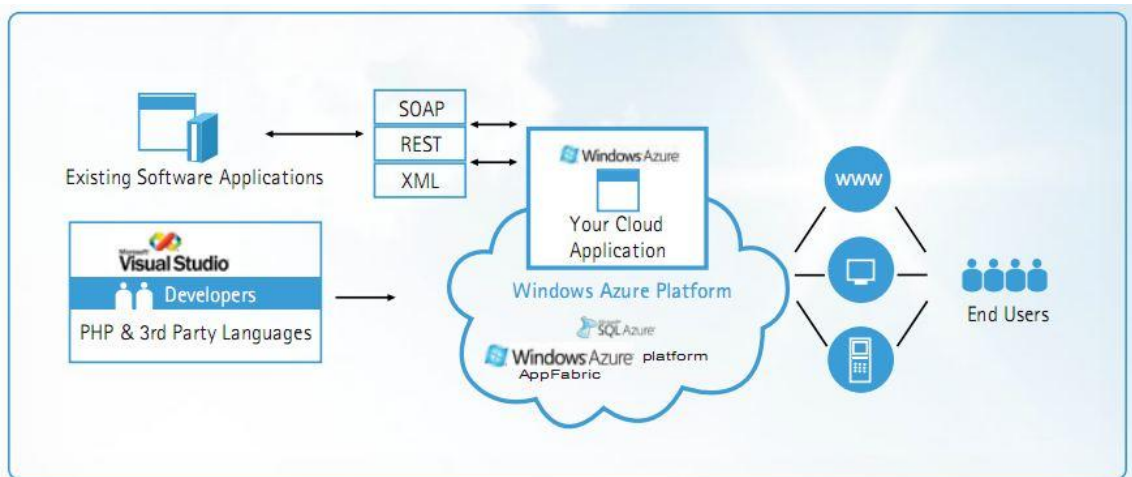
SaaS (Software as a Service)

SaaS-luokan pilvipalvelussa käyttäjälle annetaan oikeus käyttää valmiiksi ylläpidettyä ohjelmistoa esimerkiksi sähköpostipalvelinta ja kaikesta ylläpidollisista toimista vastaa

palveluntarjoaja. Tämän kaltaisia palveluita on ollut tarjolla jo yli vuosikymmenen verran, ja tästä hyvänä esimerkkinä on Hotmail.com-sähköpostisovellus.

6.2 Microsoft Azure

Windows Azure tarjoaa Windows-pohjaisen ympäristön sovelluksien suorittamiseen ja isäntänä toimimiseen. Ideana on, että fyysiset palvelimet sijaitsevat Microsoftin ylläpitämissä palvelinkeskuksissa, ja Windows Azure toimii näiden palvelimien päällä. Voitaisiin melkein sanoa, että Windows Azure on käyttöjärjestelmä palvelinkeskuksille samaan tapaan kuin Windows Server yksittäiselle tietokoneelle. Suuri palvelimien määrä ja niiden resurssien yhdistäminen takaa korkean tason suorituskyvyn isäntäohjelmille. Windows Azure on PaaS-luokan pilvipalvelu. [7.]



Kuva 3. Microsoft Azure alustan kolme pääkomponenttia

6.3 Microsoft SQL Azure

SQL Azure on DaaS (Data as a Service) -komponentti, joka on osa Azure-alustaa, ja se mahdollistaa RDBMS-palvelut pilvessä. Käytännössä se on SQL-palvelin pilven sisällä, joten käyttäjän ei tarvitse sitä erikseen asentaa eikä ylläpitää, vaan voidaan keskittyä olennaiseen, eli tietokannan rakentamiseen ja käyttöön. [7.]

Yhteys tietokantaan

SQL Azure mahdollistaa tietokannan käytön kahdella eri tavalla. Ensimmäisessä tapauksessa vain tietokanta sijaitsee Azure-ympäristössä ja itse tietokantasovellus asiakkaan omissa tiloissa. Tällöin tietokantaan päästään käsiksi asiakkaan valitsemalla ohjelmalla, joka osaa hyödyntää TDS (Tabular Data Stream) tiedonsiirtoprotokollaa. Tässä tapauksessa tiedonsiirto salataan SSL-yhteydellä. [7.]

Toisessa tapauksessa sekä tietokanta, että tietokantasovellus sijaitsevat kummatkin Azure-ympäristössä ja tietokantaan yhdistämiseen käytetään web-asiakasohjelmaa. [7.]

Ominaisuudet

Pilvessä työskennellessä tietoturva ja suorituskyky ovat päähuolenaiheet. Tästä syystä SQL Azureen on integroitu palomuri ja se pitää konfiguroida niin että yhteys sallitaan vain halutuilta IP-osoitteilta. Yhteys tulee olla myös aina salattu. Yhteyden automaattiseen katkaisuun on myös kiinnitetty paljon huomiota väärinkäytösten estämiseksi.

SQL Azureen siirtyminen vanhasta tietokantaympäristöstä on helppoa, koska olemassaoleva tieto voidaan helposti siirtää SQL Azureen SQL Server Management Studion kautta skriptien avulla tai käyttämällä *Bulk Copy* työkalua, joka siirtää tiedon rivi kerrallaan SQL palvelimelta SQL Azureen. [7.]

	SQL Azure Web Edition	SQL Azure Business Edition	SQL Server 2008 Edition	SQL Server 2008 Standard Edition
Pricing	\$ 9.99/month/DB*	\$ 99.99/month/DB*	Retail \$13,969 with 25 CALs	Retail : \$1,849with5CALs
Size Limitation	1GB	10 GB	Limited by Hardware (upto524TB)	Limited by Hardware (upto524TB)
T-SQL Support	Partial	Partial	Complete	Complete
SSIS	Supported (Data Transfer and loading)	Supported (Data Transfer and loading)	All features Supported	All features Supported
SSAS	Partial(Data Access)	Partial(Data Access)	All features Supported	All features Supported
SSRS	Partial(Data Access)	Partial(Data Access)	All features Supported	All features Supported

Kuva 4. SQL Server vs SQL Azure

Käyttötarkoitus

SQL Azurea ei ole tehty korvaamaan SQL palvelinta. Suorituskyvystään huolimatta SQL Azuren tämän hetkiset rajoitukset tarkoittavat, että se sopii parhaiten tietynlaisiin skenaarioihin. Koska SQL Azure on vasta ensimmäinen painos, on sen pääkäyttötarkoitus lähinnä sellaisissa sovelluksissa, jotka eivät vaadi kovin suuria resursseja ja joiden käyttöaste ei ole korkea. Tällainen voisi esimerkiksi olla yrityksen tietokanta, jota normaalisti käytettäisiin pöytätietokoneen tasoisella palvelimella. [7.]

Vaikka SQL Azure ei tämän hetkisessä muodossaan ole kovinkaan käyttökelpoinen vastaamaan suurten yritysten tietokantapalvelintarpeisiin on se osoitus siitä, mihin suuntaan Microsoft haluaa pilvipalveluitaan viedä. SQL Azurella on myös potentiaalia kasvaa erittäin käyttökelpoiseksi tietokantapalveluksi.

7 Harjoitusympäristö

7.1 Eri asennukset ja ohjelmat

Tarkoituksena oli, että työ tehdään koulun virtuaaliympäristössä, jotta saadaan suoraan testattua, toimiiko ympäristö Microsoftin materiaalin kanssa. Tämä ei onnistunut ajanpuutteen vuoksi ja päädyttiin ratkaisuun tehdä työ valmiiksi suppeammassa virtuaaliympäristössä ja testata toimivuus myöhemmin koulun laboratoriossa.

7.2 Koulun järjestelmä

Kurssiin liittyvät käytännön harjoitukset tehdään virtuaalisessa VMware-ympäristössä koulun tiloissa. Metropolian tietoverkkojen palvelinryhmä koostuu kuudesta korttipalvelimesta. Palvelinryhmään on asennettu VMware ESX -järjestelmä, jota käytetään työympäristönä. Kolme kuudesta korttipalvelimesta ovat Dell PowerEdge 1855 -koneita ja loput kolme ovat Dell PowerEdge 1955 -koneita. 1855-sarjan palvelimissa on kahdeksan gigatavua muistia ja jokaisessa on kaksi Xeon-prosessoria, jotka toimivat 3 GHz:n kellotaajuudella. Levytilaa näille palvelimille on varattu 1,3 teratavua. 1955 -sarjan palvelimissa on 16 gigatavua muistia ja jokaisessa kaksi Xeon E5345 -prosessoria, jotka toimivat 2,33 GHz:n kellotaajuudella. Näille palvelimille on varattu 1,4 teratavua levytilaa.

Ylimääräistä muistia on varattu 900 Gt palvelinryhmien välille ja palvelimien yhteysprotokollana levyjärjestelmiin toimii Internet Small Computer System Interface (iSCSI). Jokaisessa palvelimessa on kaksi verkkokorttia. Ne ovat yhteydessä kahteen kytkimeen, joista toinen toimii verkkoliikenteen ylläpitäjänä ja toinen on liitetty Dell AX 150i -levyjärjestelmään. Palvelinryhmälle on varattuna neljä virtalähdettä, joita yksi on varavirtalähde ongelmatilanteita varten.

7.3 VirtualBox

VirtualBox on virtualisointiohjelma, joka tukee Windows-, Mac-, Linux- ja Solaris-käyttöjärjestelmiä. Sen avulla on mahdollista ajaa montaa eri käyttöjärjestelmää jo olemassa olevan käyttöjärjestelmän sisällä. Esimerkiksi voidaan käyttää Linux-käyttöjärjestelmää Windows-koneella samaan tapaan kuin muitakin tietokoneen ohjelmia. Virtuaalikoneiden määrää ei ole rajoitettu ohjelmallisesti vaan rajana on lähinnä tietokoneen fyysiset tehot. [8.]

Ohjelman kehittäjänä toimi alunperin Innotek GmbH -yhtiö, jonka Sun Microsystems osti vuonna 2008. Kaksi vuotta myöhemmin Oracle Corporation taas osti Sun Microsystemsin ja nykyään ohjelmaa kehittää Oracle Corporation. Tästä syystä ohjelman nimi-kin vaihdettiin Oracle VM VirtualBoxiksi. VirtualBox on General Public Licensen (GPL) alainen ja ohjelman käyttäminen on ilmaista. [8.]

Ohjelman käyttöönotto tapahtuu ensin asentamalla VirtualBox isäntäkoneelle. Ohjelman käynnistämisen jälkeen uusi virtuaalikone luodaan *New*-nappia painamalla, jolloin avautuu *New Virtual Machine Wizard*, jonka avulla määritetään virtuaalikoneen fyysiset resurssit ja mikä käyttöjärjestelmä halutaan asentaa. Ohjelman käyttö on hyvin suoraviivaista ja helppoa, mutta asetuksia pääsee halutessaan muokkaamaan melko yksityiskohtaisestikin.

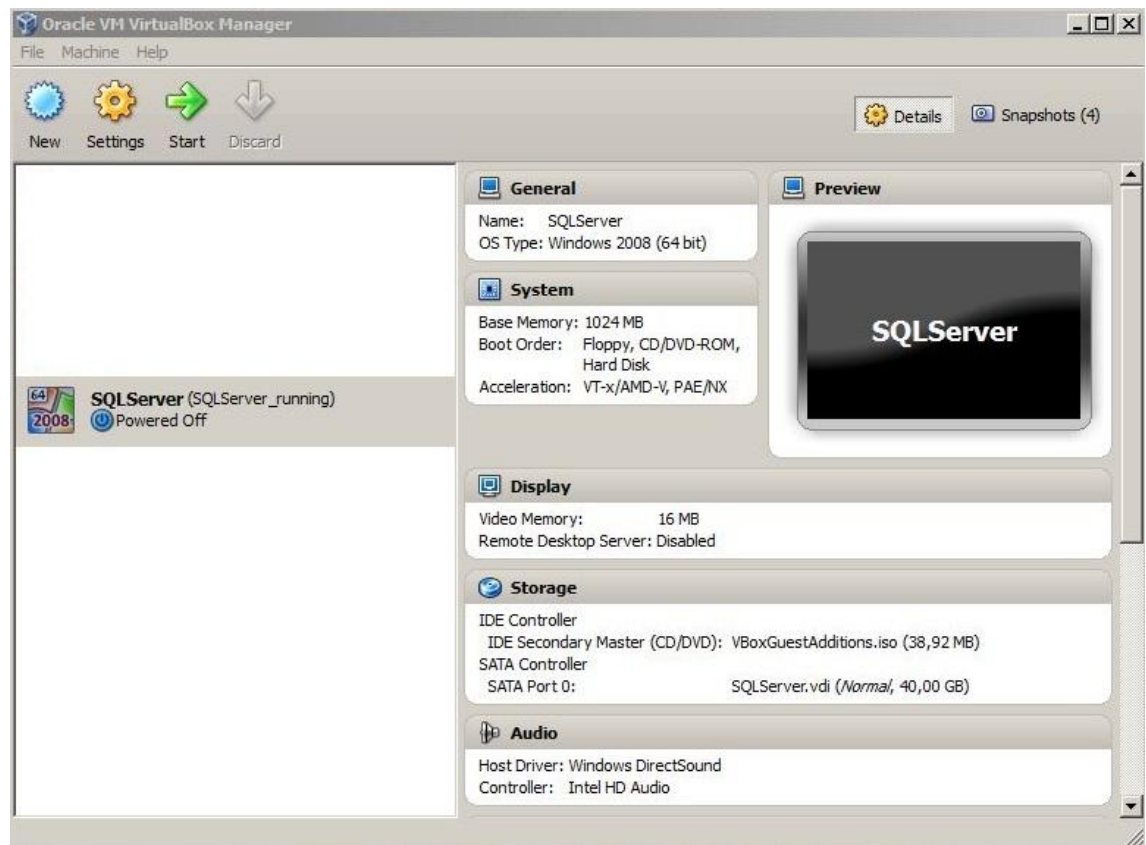
Oracle VM VirtualBox on erittäin käyttökelpoinen ohjelma kotikäyttäjille helppokäyttöisyytensä ja ilmaisuutensa takia. Ohjelmaan tulee myös säännöllisesti päivityksiä, joten uudet käyttöjärjestelmäversiot ja tietokoneiden uudet fyysiset ominaisuudet saa käyttöön mahdollisimman nopeasti niiden ilmestyttyä.



Kuva 5. Windows 7 -käyttöjärjestelmä Mac-tietokoneen sisällä ajettuna

7.4 Kotilaboratorio

Itse työ tehtiin Oracle VM Virtualbox virtuaaliympäristössä tavallisella kotikoneella, koska pääsyä koulun laboratorioon ei kesällä ollut. VirtualBox asennettiin koneeseen, jossa oli AMD Phenom II X4 -prosessori 3 GHz:n kellotaajuudella, 4 Gt muistia ja 64 bitin Windows 7 -käyttöjärjestelmä.

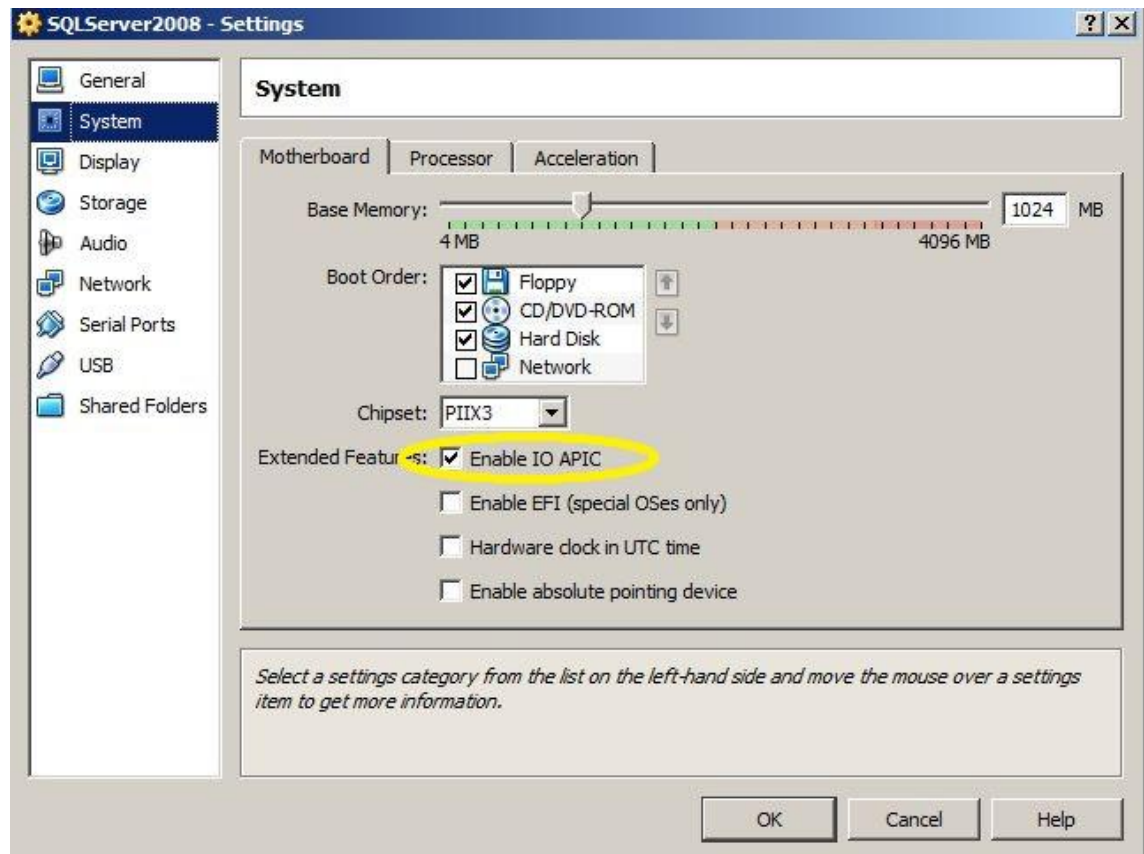


Kuva 6. Virtuaalikoneen ominaisuudet

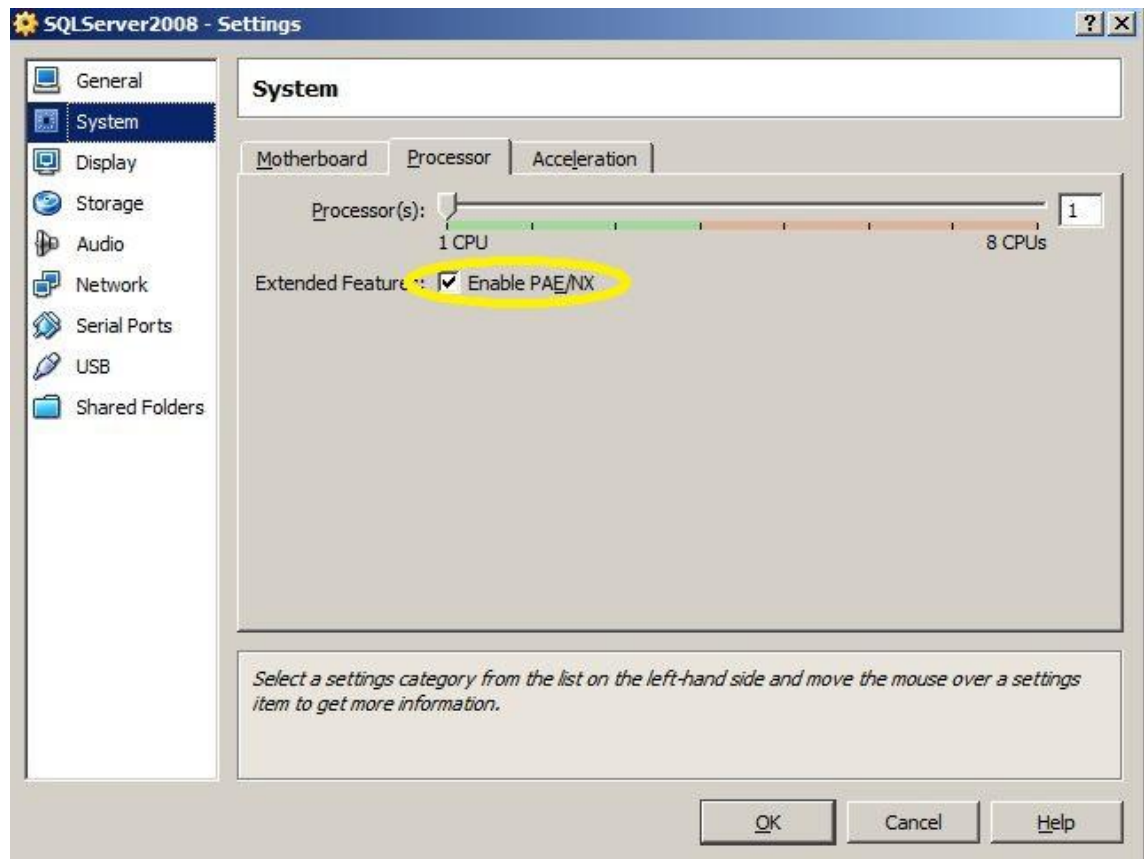
Tavoitteena oli tehdä harjoitusympäristöstä identtinen koulun järjestelmän kanssa, joten ensin tyhjälle virtuaalikoneelle oli asennettava Microsoft Server 2008 R2. Asentaminen ei sujunut ongelmitta, ja se piti tehdä kahteen kertaan. Syy löytyi virtuaalikoneen asetuksista, ja ongelmat korjaantuivat muuttamalla virtuaalikoneen asetuksia.

Enable I/O APIC mahdollistaa virtuaalikoneen käyttämisen 64 bitin tilassa. Ilman tätä optiota kone ei suostu edes käynnistymään.

Enable PAE/NX -option avulla voidaan virtuaalikoneessa käyttää suurempaa muistimäärää kuin 4 GB, jos on tarvetta.



Kuva 7. Enable IO APIC



Kuva 8. Enable PAE/NX

Microsoft Server 2008:n asentamisen jälkeen yritettiin puhtaaseen palvelimeen asentaa Microsoft SQL Server 2008 R2. Ensimmäisessä asennuksessa käytettiin kurssin 70-432-materiaaleja, jotka hiukan poikkesivat tämän opinnäytetyön kurssin tarpeista. Asennuksessa tehtiin monta eri käyttäjätiliä, koska tarkoituksena oli harjoitella asiakaskoneen kanssakäymistä palvelinkoneen kanssa. Asennuksen valmistuttua palvelimen tietokantoihin ei saatu yhteyttä, ja useiden eri kokeilujen jälkeen päätettiin poistaa asennus ja yrittää uudelleen.

Seuraavaksi Microsoft SQL Server 2008 R2 asennettiin vain ylläpitäjä käyttäjätilin kanssa, ja tämän jälkeen ohjelmistot toimivat moitteetta.

7.5 Harjoitustietokannat

Harjoitustietokantoina käytettiin AdventureWorks 2008 R2 -tietokantoja, jotka ovat Microsoftin kehittämiä juuri kurssin vaatimaan tarkoitukseen. Tietokannat ovat vapaasti ladattavissa Microsoft Database Product Samples -sivustolta, ja ne löytyvät myös Microsoft SQL Server 2008 -Database Development -kirjan mukana olevalla CD:llä.

Virtuaaliympäristöä rakennettaessa käytettiin tietokantojen asentamiseen tiedostoa AdventureWorks2008R2_SR1.exe, jonka avulla harjoitusten vaatimat tiedostot asentuivat automaattisesti oikeisiin kansioihin SQL Server -asennuksessa.

Koska saatavilla ei ollut alkuperäistä kirjan mukana toimitettavaa CD-levyä, kirjan vaatimat muut tietokannat täytyi etsiä internethakukoneen avulla, mutta ne löytyivät helposti. Yhteensopivuuden takaamiseksi tiedostot lisättiin valmiiksi virtuaalikoneille asennettavaan levykuvaan.

7.6 Lopputulos

Tuloksena syntyi toimiva harjoitusympäristö kotilaboratorioon, jossa kirjan kaikki harjoitukset testattiin ja tämän perusteella laadittiin opiskelijoille laboratoriomanuaali, josta selviää, onko jossain kappaleessa odotettavissa ongelmia ja miten ne voidaan ratkaista.

Tämän jälkeen tehtiin koulun virtuaaliympäristössä sinne soveltuva harjoitusympäristö, josta otettiin levykuva, joka voidaan asentaa virtuaalikoneille tarpeen mukaan.

8 Kurssin sisältö

Tämän luvun tavoitteena on käydä läpi, mitä kirjan kappaleet pitävät sisällään ja mihin asioihin oppilaan tulisi keskittyä. Tarkoituksena ei ole käydä jokaista kappaletta yksityiskohtaisesti läpi, koska nämä asiat löytyvät oppikirjasta.

8.1 Luku 1. Data Retrieval

Tiedon hakeminen on yksi ensisijaisista tehtävistä, jota tietokannoille tehdään, joten on luonnollista aloittaa siitä. Tiedon hakemiseen käytetään *SELECT*-komentoa useiden eri operaattoreiden kanssa, jotka laajentavat huomattavasti komennon käyttöä.

Yksinkertaisin *SELECT*-komento sisältää *SELECT*-lauseen ja *FROM*-lauseen. *SELECT* määrittää, mitä haetaan, ja *FROM*, mistä haetaan. *WHERE*-lauseella voidaan rajoittaa kyselyn tuloksien määrää.

JOIN-operaattoria käytetään kolumnien ja eri taulukoiden yhdistelyyn, ja ne lisäävät käytettävyyttä ja selkeyttä *SELECT*-komentoon.

GROUP BY-lauseella järjestetään haun tulos haluttuun järjestykseen.

Taulukko 1. Luvun 1 sisältö

Lesson 1:	Querying Data
Lesson 2:	Joining Related Tables
Lesson 3:	Implementing Aggregate Queries
Lesson 4:	Combining Datasets
Lesson 5:	Applying Built-in Scalar Functions

8.2 Luku 2. Modifying Data

Tietokantojen perusominaisuuksiin kuuluu kyky lisätä, poistaa ja muokata sen sisältämää tietoa. *INSERT*, *DELETE* ja *UPDATE* -komennot mahdollistavat kyvyn manipuloida tietoa riippuen käyttäjän antamasta syötteestä tai tietokannassa sijaitsevasta tiedosta.

Microsoft SQL Server 2005 esitteli *OUTPUT*-lauseen, jonka avulla tietokannan taulukon rivejä muokattaessa *INSERT*-, *DELETE*- ja *UPDATE*-komennoina voitiin samalla myös palauttaa tietoa käyttäjälle. SQL Server 2008 myötä voidaan käyttää *MERGE*-määrettä yhdessä *INSERT*-, *DELETE*- ja *UPDATE*-komentojen kanssa, joka mahdollistaa taulukoiden muokkaamisen perustuen muista siihen liitetystä taulukoista haetun tiedon perusteella.

SQL Server mahdollistaa myös lauseiden suorittamisen osana transaktiota, koska se lisää tiedon yhtenäisyyttä ja luotettavuutta. Transaktio on ryhmä komentoja, joiden suorittamisen onnistuminen on riippuvainen siitä pystytäänkö niistä jokainen suorittamaan. Jos yksikin niistä epäonnistuu, kaikki epäonnistuvat. Esimerkkinä tästä ovat pankkien rahansiirrot: jos järjestelmässä tulee jokin häiriö transaktion aikana ja vain osa komennoina on suoritettu, kaikki komennot perutaan.

Taulukko 2. Luvun 2 sisältö

Lesson 1:	Modifying Data by Using INSERT, UPDATE and DELETE statements
Lesson 2:	Enhancing DML Functionality with the OUTPUT clause and MERGE statement
Lesson 3:	Managing Transactions

8.3 Luku 3. Tables, Data Types and Declarative Data Integrity

Yksinkertaisin relaatiotietokannanhallintajärjestelmän (RDBMS) konsepti on taulukko, mutta nykyajan taulukot ovat kehittyneet niin paljon ensimmäisistä RDBMS-versioista, että taulukoita suunniteltaessa otetaan huomioon monia yksityiskohtia.

Taulukon sisällön tyyppiä voi olla vaikea muuttaa tulevaisuudessa, joten oikeaa tyyppiä on pohdittava tarkkaan. On mahdollista määrittellä taulukon tyyppi myös sellaiseksi, että käyttäjä itse voi sen määrittellä haluamukseen, jos tarve sitä vaatii.

Taulukoiden, sarakkeiden ja rivien nimeämisessä ollaan yhdenmukaisia. Todella suuria taulukoita voidaan pakata pienemmäksi.

Taulukko 3. Luvun 3 sisältö

Lesson 1:	Working with Tables and Data Types
Lesson 2:	Declarative Data Integrity

8.4 Luku 4. Using Additional Query Techniques

Monissa ohjelmissa tarvitaan monimutkaisempia tiedonhakumenetelmiä kuin pelkkä *SELECT*-lause. Esimerkiksi haluttaessa tietää palautettavien rivien määrä kyselyn yhteydessä tai tiedon esittämisjärjestystä tietoa näytölle tulostettaessa.

Tämän ongelman ratkaisemiseksi luvussa käydään läpi rekursiivisia Common Table Expression (CTE) -kyselyitä, alakyselyitä sekä kyselyjen tuloksien arvojärjestyksen muokkaamista.

CTE-kyselyt tuottavat väliaikaisen kyselytuloksen, johon voidaan viitata ja täten antaa lopullinen tulos halutulla tavalla. Alakyselyt ovat vaihtoehtoinen tapa tälle samalle asialle.

Taulukko 4. Luvun 4 sisältö

Lesson 1:	Building Recursive Queries with CTEs
Lesson 2:	Implementing Subqueries
Lesson 3:	Applying Ranking Functions

8.5 Luku 5. Programming Microsoft SQL Server with T-SQL

Microsoft SQL Server 2008 sallii kolmen erilaisen ohjelmoidun objektin luomisen: varastoidun proseduurin, funktion ja herättimen. Sen sijaan, että käytettäisiin yhtä komentoa tai lausetta kerrallaan voidaan rakentaa objekteja, jotka sisältävät korkean hierarkian ohjelmointikieltä. Tämä mahdollistaa monimutkaisten, mutta helppokäyttöisten funktioiden käytön.

SQL Serverin mukana tulee satoja valmiita funktioita ja sallii ylläpitäjän tehdä myös omia funktioitaan. Varastoidut proseduurit tarjoavat abstraktin API:n, joka mahdollistaa

funktioiden tekemisen sellaisille käyttäjille, jotka eivät ymmärrä monimutkaista tietokantojen suunnittelutyötä.

Herättimet ovat varastoitujen proseduurien muoto, jotka suoritetaan silloin, kun DML- tai DDL-komennot sitä vaativat.

Ohjelmointikielenä tässä luvussa käytetään T-SQL-kieltä.

Taulukko 5. Luvun 5 sisältö

Lesson 1:	Stored Procedures
Lesson 2:	User-Defined Functions
Lesson 3:	Triggers
Lesson 4:	Views

8.6 Luku 6. Techniques to Improve Query Performance

Tietokannan ja käyttöliittymän suunnittelun jälkeen on hyvä varmistaa, että tietokannan suorituskyky on käyttäjien haluamalla tasolla. Tässä luvussa käydään läpi miten löytää mahdollisia ongelmia suorituskyvyssä ja miten korjata ne.

Microsoft SQL Server 2008 mahdollistaa erittäin hyvän suorituskyvyn, mutta ongelmia voi esiintyä jos tietokanta on suuri, kyselyt ovat monimutkaisia tai kyselyjen määrä on suuri.

Taulukko 6. Luvun 6 sisältö

Lesson 1:	Tuning Queries
Lesson 2:	Creating Indexes

8.7 Luku 7. Extending Microsoft SQL Server Functionality part 1

Microsoft SQL Server 2008 on tehokas työkalu, jossa on enemmän käytettävyyttä kuin tyyppillisessä RDBMS:ssä. Ohjelman tähän versioon on lisätty tuki käsitellä isoja binääri-dataosia niin, ettei niitä tarvitse tallentaa tietokannan tiedostojen sisään. Myös XML-kielen tukea on paranneltu.

Tässä luvussa käydään läpi miten näitä uusia ja paranneltuja kykyjä päästään hyödyntämään.

Taulukko 7. Luvun 7 sisältö

Lesson 1:	Working with XML
Lesson 2:	Using SQLCLR and Filestream

8.8 Luku 8. Extending Microsoft SQL Server Functionality part 2

Microsoft SQL Server 2008:n uudet geometriset datatyypit mahdollistavat sijaintidatan tallentamisen, jonka myötä voidaan työskennellä sijainnin ja geometrisen tiedon kanssa. Tällöin voidaan käyttää *geography*- ja *geometry*-datatyyppejä, joilla voidaan seurata z-, y- ja z -koordinaatteja kuin myös ellipsoidisia kuvauksia, kuten leveys- ja pituusasteita. Yritykset voivat käyttää tätä keinoa myyntialueiden seuraamiseen, määrittämään logistiikan reittejä ja niin edelleen.

SQL Server 2008 sisältää työkalut, joilla tätä tietoa voidaan käyttää tehokkaasti. *Full-text Search* ominaisuus mahdollistaa monimutkaiset kieleen perustuvat kyselyt, jotka tarjoavat paljon enemmän mahdollisuuksia kuin *LIKE*-määre.

Lisäksi SQL Server Service Broker sallii sellaisten sovellusten rakentamisen, jotka hyödyntävät asynkronista tiedonsiirtoa.

Taulukko 8. Luvun 8 sisältö

Lesson 1:	Implementing Spatial Data Types
Lesson 2:	Implementing Full-Text Search
Lesson 3:	Implementing Service Broker Solutions

8.9 Luku 9. An Introduction to Microsoft SQL Server Manageability Features

Tietokannan ylläpitäjän tehtäviin kuuluu yleensä muutakin kuin vain tehdä ratkaisut tarvittavan tiedon noutamisesta käyttäjälle. Tässä luvussa tutustutaan eri ominaisuuks-

siin, johon Microsoft SQL Server 2008 pystyy, kun halutaan lisätä ja hallita tietokantaa tehokkaasti.

Luvun ensimmäinen kappale jää itse kurssilaisilta kokeilematta, koska harjoitusympäristössä ei tätä tehdessä ollut siihen mahdollisuutta Simple Mail Transfer Protocol (SMTP) -palvelimen puuttuessa.

Taulukko 9. Luvun 9 sisältö

Lesson 1:	Integrating Database Mail
Lesson 2:	Implementing Scripts by Using Windows PowerShell
Lesson 3:	Tracking Data Changes

9 Yhteenveto

Työn tekeminen alkoi kurssimateriaaliin tutustumalla, johon kuului vain Microsoft SQL Server 2008 -Database Development -kirja. Kirja sisälsi teoria- ja käytännönsuuden ja oli sisällöltään laaja. Suurin osa kappaleista sisälsi käytännönsuudessaan paljon SQL:llä ohjelmointia. Tästä syystä tehtävien toiminnan tarkastaminen oli hidasta.

Ensin työ oli tarkoitus tehdä suoraan koulun tarjoamassa virtuaaliympäristössä, mutta virtuaalikoneella ei ollut tarpeeksi kiintolevytilaa ohjelmistojen asentamiseen eikä oikeuksia tämän ongelman korjaamiseen. Tämän takia päädyttiin ohjelmistot ja harjoitus-tehtävät testaamaan ensin omassa kotivirtuaaliympäristössä, johon on nykytekniikalla kustannustehokas mahdollisuus. Tämä tuotti aluksi ongelmia, mutta yrityksen ja erehdyksen kautta ohjelmistot saatiin lopulta toimimaan.

Työn tekemiseen ja raportointiin meni odotettua enemmän aikaa juuri näistä alun ongelmista johtuen. Aihe on myös hyvin laaja, joten oleellisten asioiden poiminta vei aikaa.

Lopputuloksena syntyi hyvä valinnainen kurssi tietokantapalvelimien ylläpidosta kiinnostuneille tietotekniikan opiskelijoille.

Lähteet

- 1 Spenik Mark. 2001. Verkkodokumentti.
<<http://www.informit.com/articles/article.aspx?p=130898>>. Luettu 1.12.2011.
- 2 Coronel Carlos, Morris Steven, Rob Peter. 2009. Database systems: design, implementation, and management. Cengage Learning.
- 3 Microsoft SQL Server. Verkkodokumentti.
<http://en.wikipedia.org/wiki/Microsoft_SQL_Server>. Luettu 1.12.2011.
- 4 Microsoft SQL Server 2008 Product Overview. 2007. Microsoft Corporation.
- 5 Ruwanpura Supun. 2000. Verkkodokumentti.
<<http://www.csse.monash.edu.au/hons/projects/2000/Supun.Ruwanpura/sql.htm>>. Luettu 1.12.2011.
- 6 Transact-SQL. Verkkodokumentti. <<http://en.wikipedia.org/wiki/Transact-SQL>>. Luettu 1.12.2011.
- 7 Wipro SQL Azure, Moving Business Intelligence to Cloud. 2010.
- 8 VirtualBox manual. Verkkodokumentti.
<<https://www.virtualbox.org/manual/ch01.html>>. Luettu 1.12.2011.