

Kotitehtävät vko 4

Puskurit. Transaktioidenhallinta. Suorituskyky. Monitorointi ja jäljitys. Hälytykset. IBM DB2:een tutustuminen:

- a) Asenna AdventureWorks tietokanta SQL Server:iin. AdventureWorks on Microsoftin harjoittelutietokanta SQL Server:iin. AdventureWorks tietokannasta on olemassa eri versiot melkein jokaiselle eri SQL Server versiolle (2005, 2008, 2012 versiot jne). Myös SQL Server Express:lle löytyy oma lite versionsa AdventureWorksLT. Tässä käytetään täys versiota SQL Server:stä, joten asenna täys versio myös AdventureWorks:sta. Tätä tietokantaa tarvitaan myöhemmissä harjoituksissa. Esitä mitä vaiheita tarvitaan kyseisen tietokannan asennuksessa. AdventureWorks tietokanta löytyy osoitteesta:
<http://msftdbprodsamples.codeplex.com/>.

Vastaus:

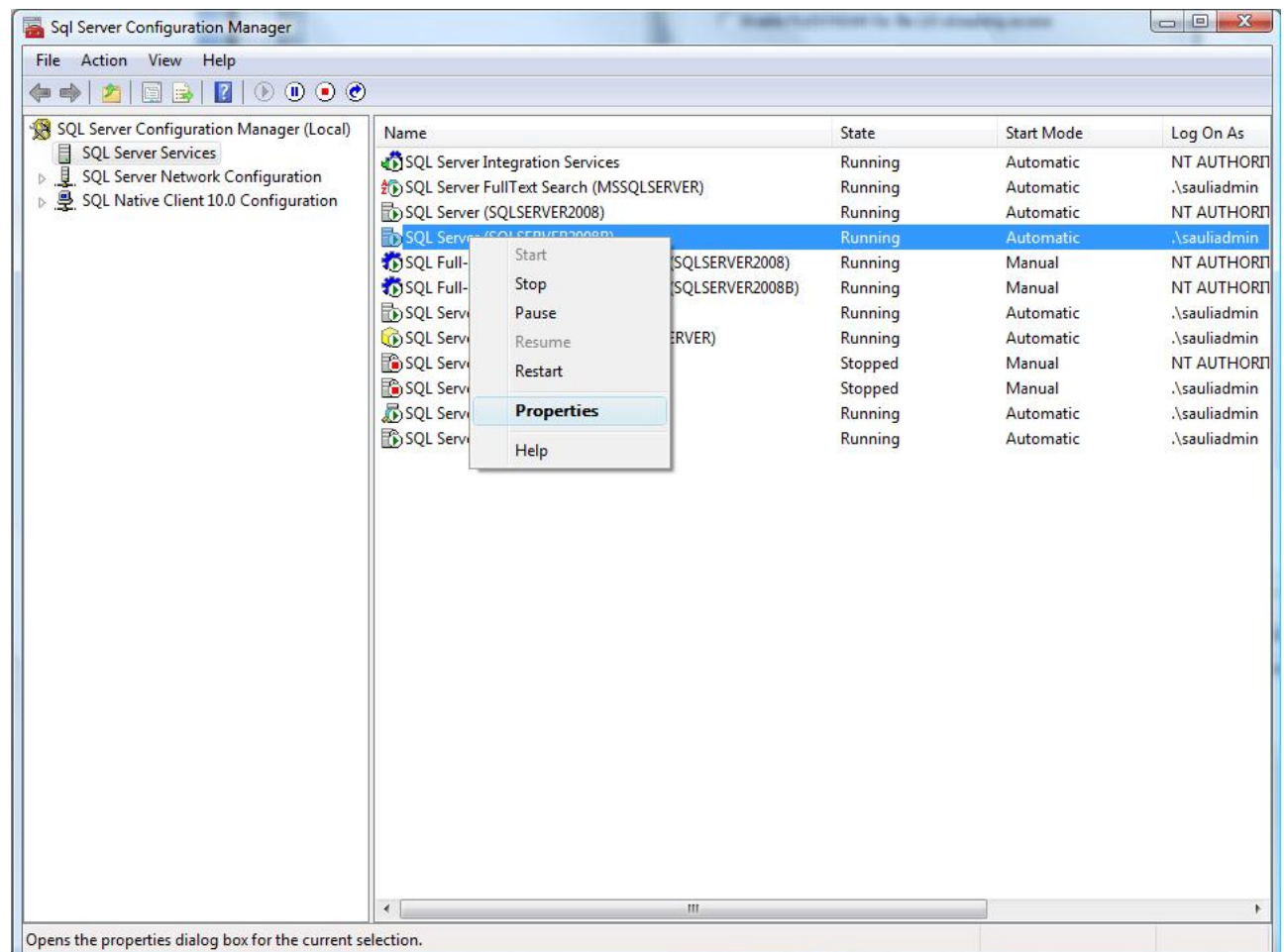
- AdventureWorks tietokanta on ladattava osoitteesta <http://msftdbprodsamples.codeplex.com/>. Sieltä valitaan AdventureWorks OLTP versio SQL Server 2005:lle. Vanhempi versio on mahdollista asentaa uudempaan versioon SQL Server:stä. Toisin päin se ei onnistu ilman SQL Server:in päivittämistä.
- Kopioi AdventureWorks_Data.mdf ja AdventureWorks_Log.ldf tiedostot SQL Server instanssiin käyttämään kansioon. Koulun koneilla esimerkiksi kansioon:

C:\Program Files\Microsoft SQL Server\MSSQL10.MSSQLSERVER\MSSQL\DATA

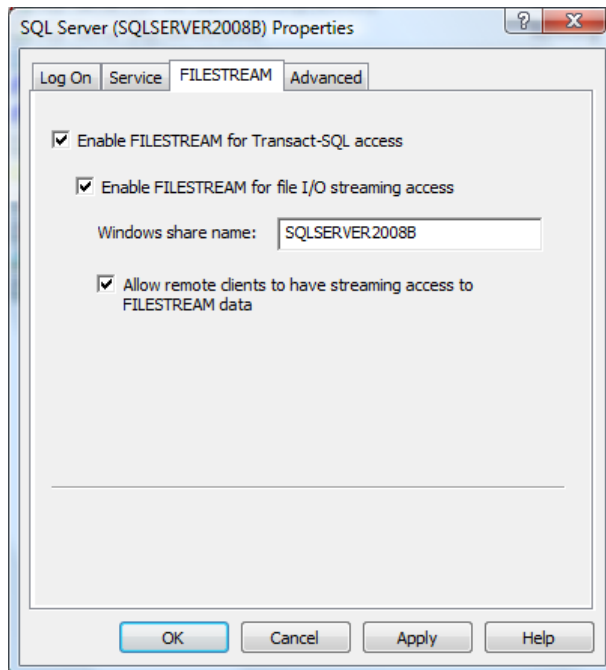
- AdventureWorks käyttää streamattuja tiedostoja tietokannassaan. Sen vuoksi SQL Server on konfiguroitava käyttämään niitä. SQL Server:in streamatuista tiedostoista on kerrottu lisää osoitteessa: <http://msdn.microsoft.com/en-us/library/bb895234.aspx> (huomaa valita se versio SQL Server:stä jota käytät sillä ohjeet eroavat toisistaan päivitysten vuoksi).

- Valitse SQL Server:in Configuration Manager:ista se SQL Server:in instanssi, jonka konfiguraatiota haluat muuttaa. Valitse siis:
Start → All Programs → Microsoft SQL Server 2008 → Configuration Tools → SQL Server Configuration Manager

Ja klikkaa alla olevan kuvan mukaisesti instanssia hiiren oikealla ja valitse Properties:

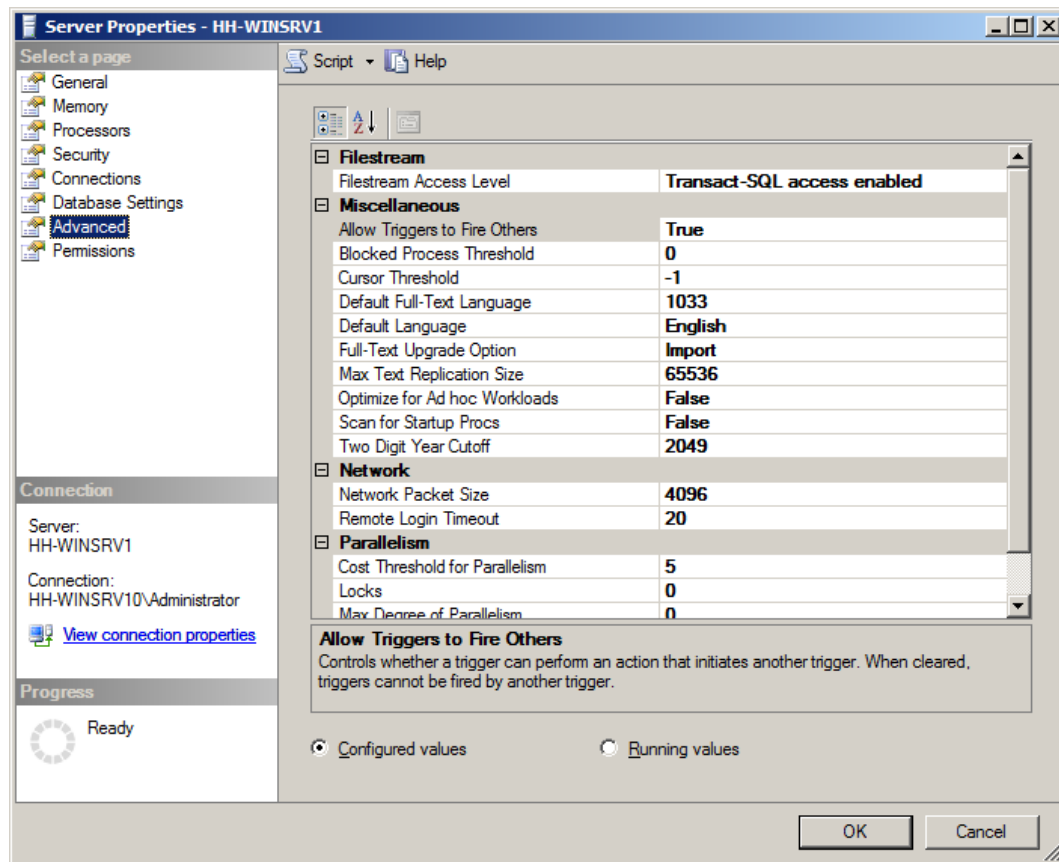


- Siirry Filestream välilehdelle ja laita Filestream päälle (ks. kuva alla):

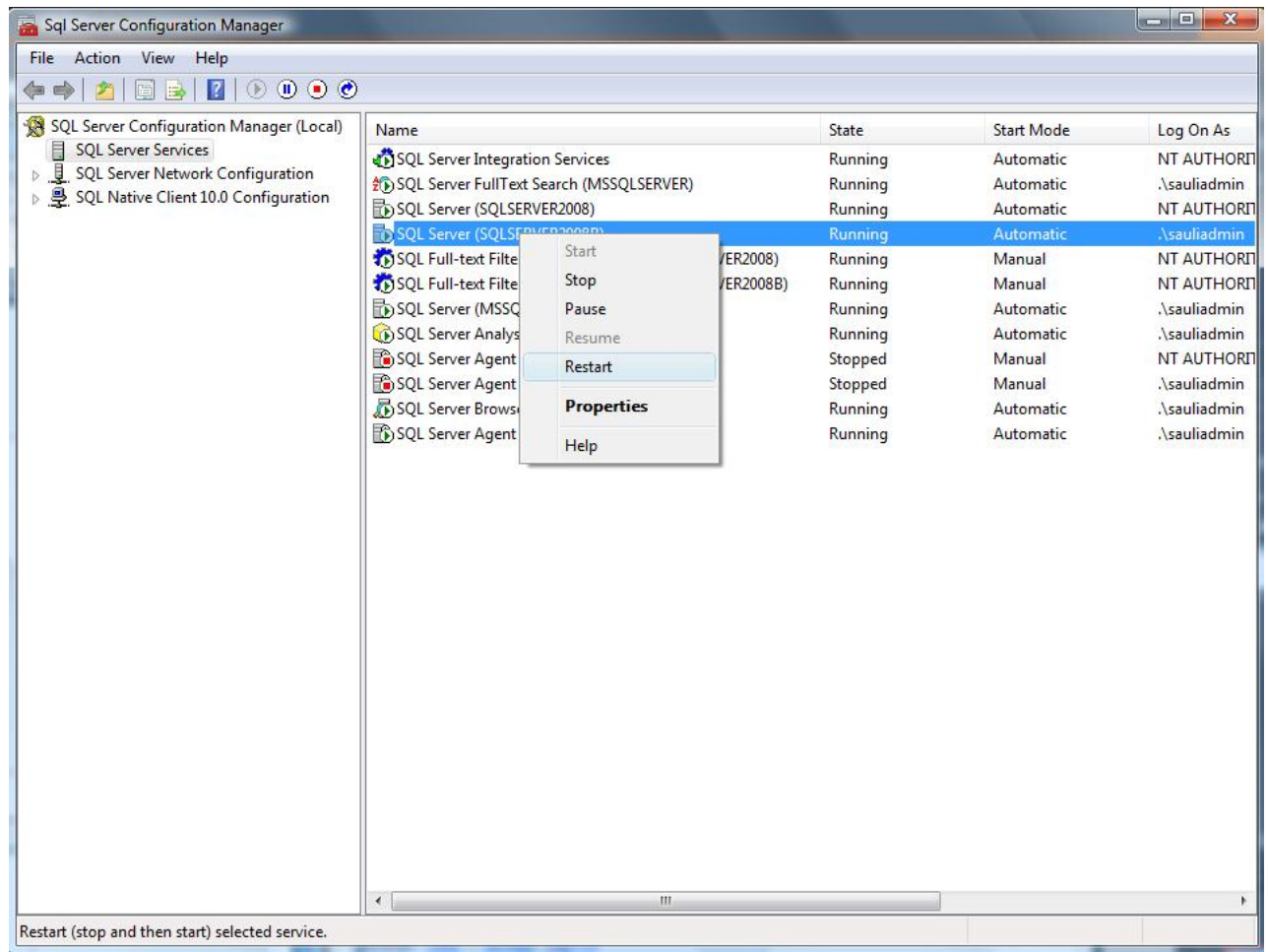


Hyväksy muutokset OK:lla.

- Aseta SQL Server Management Studioissa konfiguroitavalle instanssille Filestream Access Level alla olevan kuvan mukaiseksi, eli Transact-SQL access enabled:



- Käynnistä instanssi uudelleen SQL Server Configuration Managerin avulla.



- SQL Server Management Studiossa ajetaan seuraavaksi alla oleva T-SQL koodi (Query ikkunassa):

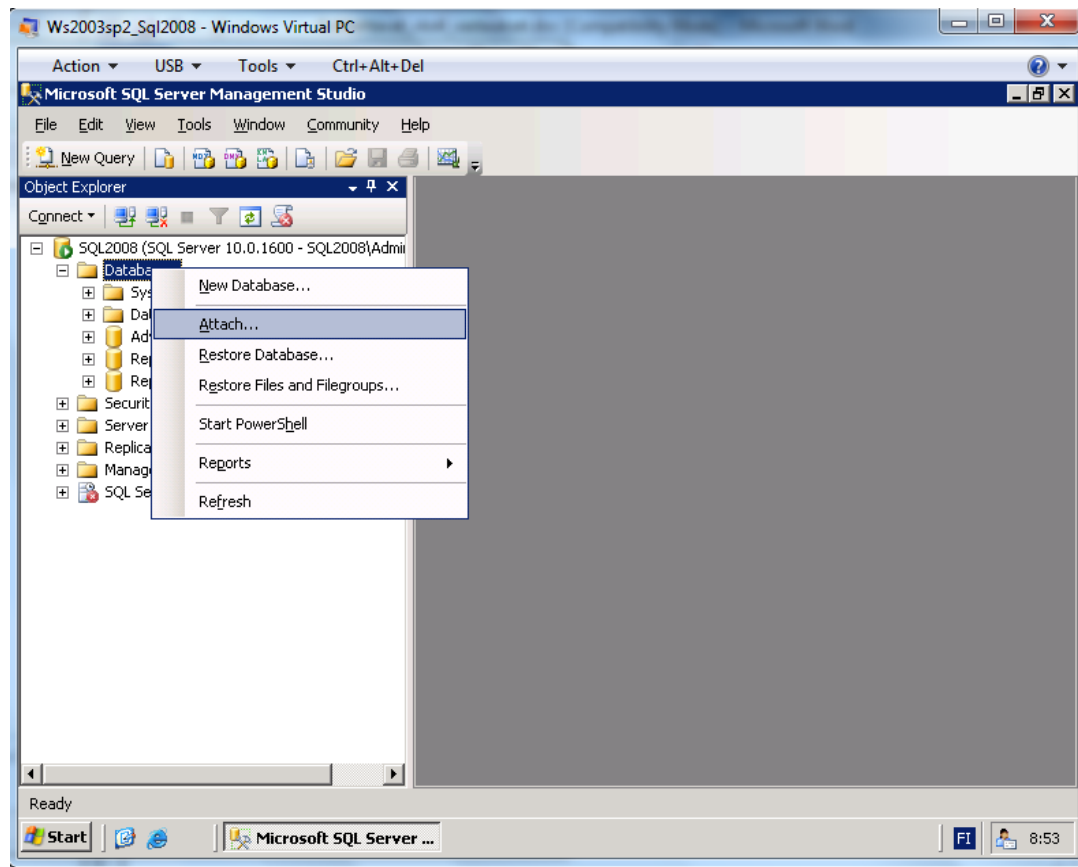
```
EXEC sp_configure filestream_access_level, 2  
RECONFIGURE
```

Onnistuneesta konfiguroinnin muutoksesta tulee alla oleva viesti:

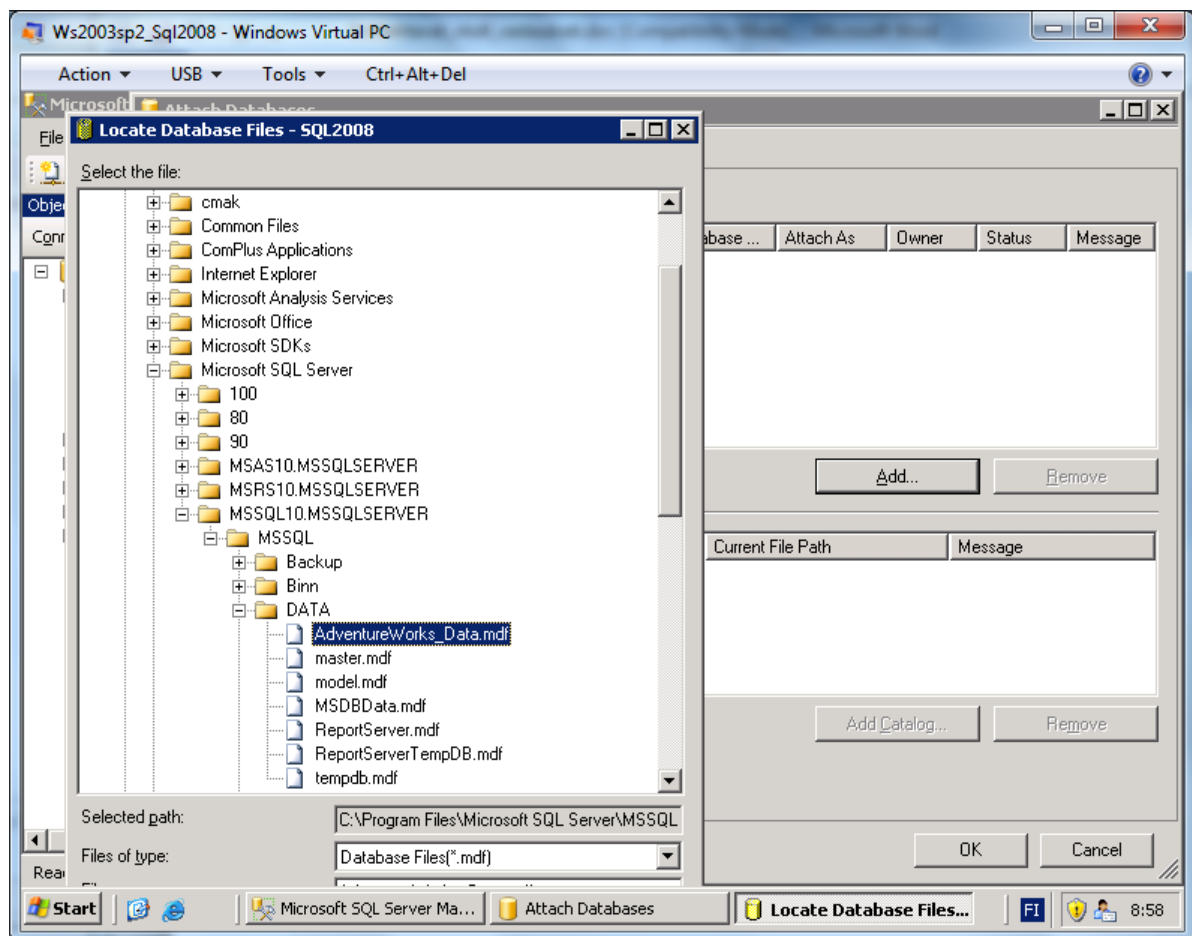
Configuration option 'filestream access level' changed from 0 to 2. Run the RECONFIGURE statement to install.

- Seuraavaksi SQL Server:in instanssi on käynnistettävä uudelleen. Se onnistuu suoraan SQL Server Configuration Manager:sta:

- Nyt päästään vihdoinkin liittämään eli attach:aamaan ladattu AdventureWorks_Data.mdf ja AdventureWorks_Log.ldf tiedostot. Näistä mdf on varsinainen tietokannan datatiedosto ja ldf on transaktiologin tiedosto. Siirry siis SQL Server Management Studioon ja klikkaa hiiren oikealla Databases:ia Object Explorer:ssa ja valitse Attach:



- Klikkaa Add painonappulaa ja valitse haluamasi tietokanta, jonka haluat liittää SQL Server:in instansiin.



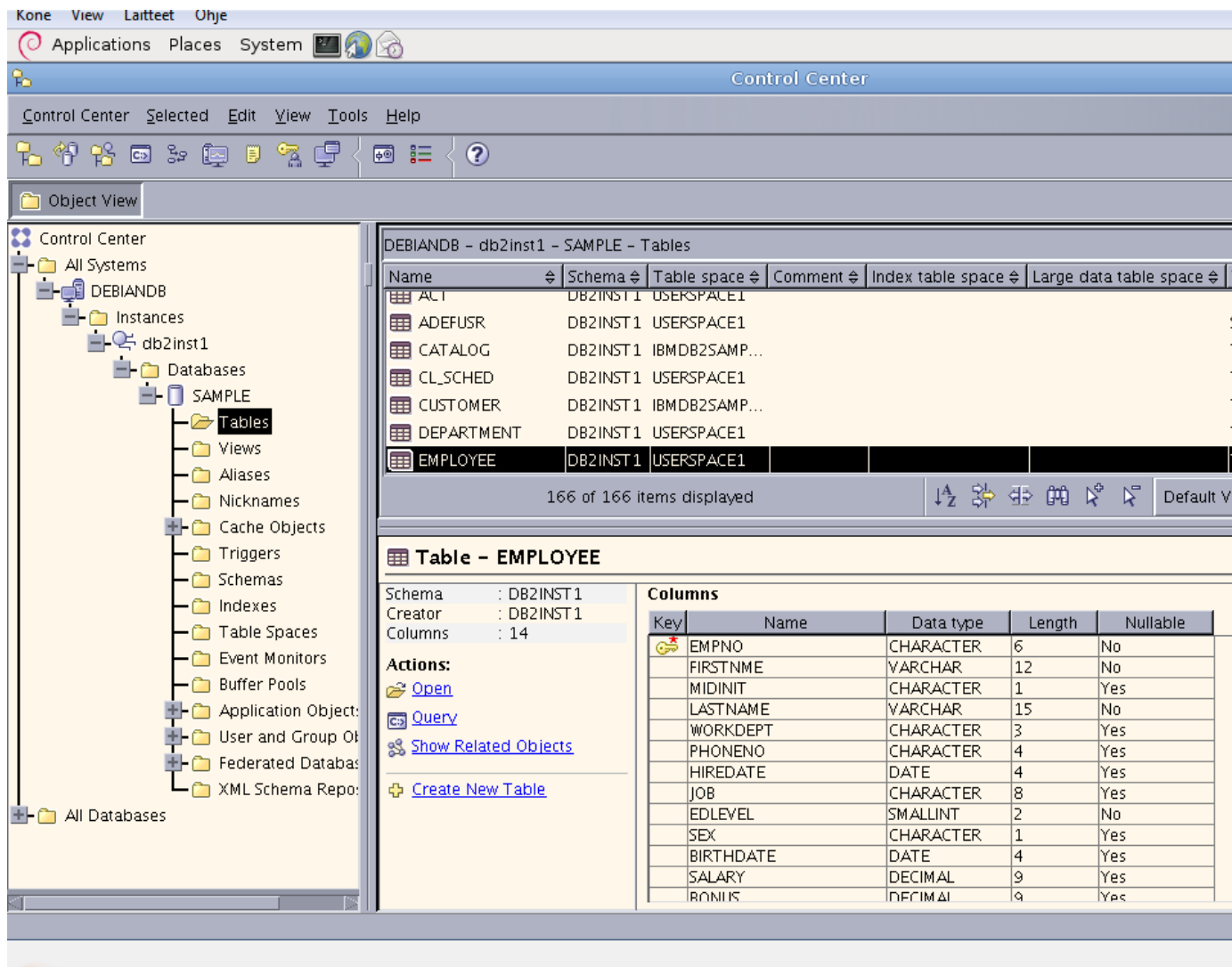
SQL Server kysyy Full Text Search Catalogia. Sitä ei välttämättä tarvita. AdventureWorks tietokannan voi liittää ilman sitä.

- b) Tutustu IBM:n DB2:een ja sen työkaluihin, tutkimalla esimerkiksi DB2:sta löytyvää SAMPLE tietokantaa. Tutki millainen tietokanta on taulujen, kenttien ja tietotyyppien, sekä indeksien osalta käyttäen IBM DB2 Control Center:iin. DB2 löytyy edellä jo tutuksi tulleesta VirtualBox:sta, johon Debian Linux on virtualisoitu. Kerro mitä näillä työkaluilla voi tehdä?

Vastaus:

IBM DB2 Control Center:

- Tutkitaan ensin DB2:n Control Center:iä. IBM DB2 Control Center löytyy Debianista valitsemalla: Applications -> IBM DB2 -> Control Center



- Control Center:stä löytyy työkaluja tietokannan system catalogin tarkasteluun, taulujen luomiseen ja muuttamiseen. Myös DB2 SQL kielisiä lauseita voi ajaa Command Editorissa, valitsemalla Control Center:ssä:
Tools -> Command Editor

Tähän työkaluun tutustutaan seuraavaksi, koska suurin osa ylläpidosta tehdään DB2 SQL kieltä käyttäen.

- Command Editorissa täytyy ensin kirjautua haluamaansa tietokantaan DB2:ssä DB2 SQL kieltä käyttäen. Kielen syntaksista löytyy tietoa osoitteesta:

<http://publib.boulder.ibm.com/infocenter/dzichelp/v2r2/index.jsp?topic=%2Fcom.ibm.db2z9.doc.sqlref>

[%2Fsrc%2Ftpc%2Fdb2z_sql_connect.htm.](#)

Ensin on siis kirjaututtava DB2:sta löytyvään tietokantaan. Esimerkiksi alla kirjaututaan SAMPLE nimiseen tietokantaan, koska sellainen löytyy sieltä. Käyttäjätunnus on db2inst1 ja perässä seuraa sen salasana.

```
connect to sample user db2inst1 using linstdb2;
```

- Kun tietokantaan on päästy onnistuneesti kirjautumaan, voidaan kokeilla mitä tahansa DB2 SQL:n lausetta. Alla on listattu sample tietokannasta löytyvän employee taulun kaikki data rivit:

```
select * from employee;
```

Alla olevassa kuvassa on esitetty millaiselta Command Editor näyttää:

The screenshot shows the IBM Command Editor 2 window. The title bar reads 'Command Editor 2'. The menu bar includes 'Command Editor', 'Selected', 'Edit', 'View', 'Tools', and 'Help'. The toolbar contains various icons for file operations and execution. The 'Commands' tab is active, showing the following text:

```
connect to sample user db2inst1 using linstdb2;
select * from employee;
```

Below the command input area, the results of the query are displayed in a table with 11 columns. The first few rows are visible:

EMPNO	NAME	SEX	EMPID	DEPTNO	EMPNO	NAME	SEX	EMPID	DEPTNO	EMPNO
000300	PHILIP	X	SMITH	E11	2095	08/19/2002	OPERATOR	14	M	10/27/1978
000310	MAUDE	F	SETRIGHT	E11	3332	09/12/1994	OPERATOR	12	F	04/21/1961
000320	RAMLAL	V	MEHTA	E21	9990	07/07/1995	FIELDREP	16	M	08/11/1962

At the bottom of the results area, it says '42 record(s) selected.' The status bar at the very bottom shows 'Control Center' and 'Command Editor 2'.

c) Mitä Buffer Pool:illa tarkoitetaan tietokantapalvelimella?

Vastaus:

Buffer Pool on tietokantapalvelimen käynnistyksen yhteydessä palvelintietokoneen muistista tietokantapalvelimen käyttöön varattu muistialue. Sinne luetaan levyllä olevat tietokannan datat, joita käyttäjät tarvitsevat. Esimerkiksi, jos käyttäjä haluaa muuttaa tietokannassa olevan henkilön tietoja, tarkistaa tietokantapalvelin, että mistä kyseisen henkilön tiedot löytyvät. Ovatko ne jo valmiiksi luettu Buffer Pool:in aikaisemmin vai onko ne haettava levyltä. Tämä tietysti vie enemmän aikaa kuin suoraan muistista lukeminen. Kun Buffer Pool:ssa olevaa tietoa muutetaan merkitään se muuttuneeksi, jotta tietokantapalvelin tietää, että levyllä oleva vastaava tieto on vanhentunut ja se pitää päivittää buffer pool:ssa olevalla päivitetyllä tiedolla. Asiasta on kerrottu tarkemmin SQL Server:in osalta osoitteessa: [http://technet.microsoft.com/en-us/library/aa337525\(v=sql.105\).aspx](http://technet.microsoft.com/en-us/library/aa337525(v=sql.105).aspx)

d) Mikä on keskeisin asia miksi transaktioidenhallintaa tarvitaan?

Vastaus:

Transaktioidenhallinnassa halutaan varmistaa, että operaatio onnistuu virheettömästi, tietokannan eheys pysyy kunnossa ja virhetilanteissa voidaan peruuttaa operaation sql kieliset lauseet virheettömästi siten että tietokannan data palautuu aikaisempaan tilaansa. Kyse on siis siitä, että halutaan ns. ACID ominaisuudet. Hyvä käytännön esimerkki tästä voisi olla esimerkiksi maksun maksaminen nettiyhteyttä tms. käyttäen. Kun asiakas haluaa maksaa vaikkapa sähkölaskunsa on hänen tililtään vähennettävä maksun summa ja heti perään lisättävä sähkölaitoksen pankkitilille. Tämä vaatii siis kaksi update lausetta. Niiden ajon ollessa kesken, ei tietenkään ole hyvä asia, jos vaikkapa sähkölaitoksen controlleri pääsee katsomaan tilannetta, jolloin asiakkaan tililtä on kyllä vähennetty maksun summa, mutta sitä ei ole vielä ehditty lisätä sähkölaitoksen tilille. Tilisiirron operaatio pitää siis eristää muista operaatioista siten, että ne eivät pääse näkemään kesken olevia operaatioita. Myös virhetilanteissa on pystyttävä peruuttamaan kesken jäänyt tilisiirto virheettömästi alkuperäiseen tilaansa. Tämän vuoksi esimerkiksi asiakkaan operaatio eristetään omaksi transaktiokseen ja sähkölaitoksen controllerin operaatio omaksi transaktiokseen. Niistä ajetaan ensimmäisenä se joka ehtii ensimmäisenä ajoon. Toisen on odotettava, jos se tarvitsee samoja resursseja, kuten esimerkiksi sähkölaitoksen controlleri halusi nähdä asiakkaamme tilitiedot samaan aikaan kuin asiakas oli tekemässä maksua. Kun tilisiirto on ehditty tehdä loppuun asti virheettömästi, pääsevät seuraavat transaktiot käyttämään vapautuneita resursseja. Ks. <http://msdn.microsoft.com/en-us/library/ms190612%28v=sql.105%29.aspx> ja <http://msdn.microsoft.com/en-us/library/ms188929.aspx>.

e) Selitä lyhyesti mihin yleensä tarvitaan tietokantaympäristössä **monitorointeja** ja mitä tarkoituksia on erilaisia jäljityksillä (**tracing**). Voit tarkastella tätä yleisesti tai tietyn DBMS-tuotteen kannalta.

Vastaus:

Monitoroinneista on kerrottu kerrottu ehkä selkeimmin SQL Server:in sivuilla: <http://msdn.microsoft.com/en-us/library/ms191199%28v=sql.105%29.aspx> ja <http://msdn.microsoft.com/en-us/library/bb510705%28v=sql.105%29.aspx>. Tietokantapalvelimella halutaan siis monitoroida tietokantainstanssiin yhteydessä olevien käyttäjien yhtyeiksi, prosessien numeroita ja niiden tilaan. Myös käyttäjien suorittamat sql kieliset lauseet ovat kiinnostuksen kohteena tässä ja objektien (taulut, proseurit, indeksit, triggerit jne) lukitukset ja millaisia nämä lukot ovat (read vai write lukkoja esimerkiksi). Transaktiot voivat päätyä myös "pattitilanteeseen" eli deadlock tilanteeseen. Tietokantapalvelin ratkaisee tämän yleensä siten, että transaktiosta rollback:taan se jos menetetään vähintään tehtyä työtä, tai se jolla on eniten työtä vielä tekemättä. Tässä on tietysti seurattava sitä, ettei käy niin, että jotain transaktiota ei saada koskaan ajettua loppuun. Myös tätä halutaan monitoroida.

IBM DB2:n osalta monitoroinnista on kerrottu osoitteessa: <http://my.haaga-helia.fi/~dbms/ict4tn006/DB2PerfTuneTroubleshoot-db2d3e972.pdf> kappaleessa 2 ja <http://pic.dhe.ibm.com/infocenter/db2luw/v10r1/index.jsp?topic=%2Fcom.ibm.db2.luw.admin.dm.doc%2Fdoc%2Ft0057364.html>. Tavoitteet DB2:n monitoroinnissa ovat samat kuin minkä tahansa tietokantapalvelimen, esimerkiksi SQL Server:in.

IBM DB2:n dokumentaatiossa on ehkä selkeimmin kerottu trace:istä:

<http://publib.boulder.ibm.com/infocenter/db2luw/v9r5/index.jsp?topic=%2Fcom.ibm.db2.luw.admin.trb.doc%2Fdoc%2Fc0020800.html>. Trace:en avulla voi siis tallentaa tietoa tietokantapalvelimessa suoritettavista operaatioista ja tapahtumista. Edellä esiteltyjä monitoroitavia asioita voi siis automatisoida seurantaan ja tallentaa syntyvät "raportit" trace:n avulla.

Microsoft:in SQL Server trace:stä on kerrottu osoitteessa: <http://msdn.microsoft.com/en-us/library/hh245121.aspx> ja <http://msdn.microsoft.com/en-us/library/ms188662.aspx>