



**TEKNIikka JA LIIKENNE**

**Tietotekniikka**

**Dataverkot**

**INSINÖÖRITYÖ**

**Distributed File System - käyttöönnotto**

**Työn tekijä: Antti Salovaara**  
**Työn ohjaajat: Jari Myllykangas**  
**Antti Rautalin**

**Työ hyväksytty: 25.05. 2010**

**Janne Salonen**  
**Yliopettaja**



## **ALKULAUSE**

Tämä insinööriyö tehtiin Napa Yhtiöiden toimeksiannosta Helsingissä. Kiitos kaikille projektissa avustaneille.

Helsingissä 20.5.2010

Antti Salovaara

## TIIVISTELMÄ

<b>Työn tekijä:</b> Antti Salovaara	
<b>Työn nimi:</b> Distributed File System - käyttöönotto	
<b>Päivämäärä:</b> 20.5.2010	<b>Sivumäärä:</b> 46 s.
<b>Koulutusohjelma:</b> Tietotekniikka	<b>Suuntautumisvaihtoehto:</b> Dataverkot
<b>Työn ohjaaja:</b> Yliopettaja Janne Salonen	
<b>Työn ohjaaja:</b> Jari Myllykangas (Napa Oy), Antti Rautalin (Napa Oy)	
<p>Työssä toteutetaan Distributed File System -ratkaisu käyttäen Microsoftin palvelimiin sisäänrakennettua ominaisuutta ja perehdytään tekniikoihin, joihin DFS perustuu kuten Active Directoryyn ja Domain Name Systemiin.</p> <p>Yrityksen tiedostokopioinnit on tähän asti hoidettu erilaisilla xcopy-pohjaisilla ratkaisuilla käyttämällä Windows-ajastetuilla tehtävillä ajastettuja .bat-tiedostoja. Bat-ratkaisu on tarkoitus korvata keskitetysti hallitulla Distributed File System -ratkaisulla, jolla voidaan helposti jakaa tiedostoja Helsingistä etäkonttoreihin ja päinvastoin. Lisäksi tavoitteena on muodostaa DFS-Mesh -yhteys Aasian konttoreiden välille helpottamaan konttorien välistä tiedonsiirtoa ja korvaamaan tähänastinen ratkaisu, jossa kaikki liikenne on kiertänyt Helsingin kautta.</p> <p>Yrityksen verkko on jakautunut usealla mantereelle, joten työssä tutustutaan lisäksi konttorien välisten VPN-yhteyksien toimintaan ja päivitetään yrityksen verkkoa vastaamaan nykypäivän haasteita.</p>	
<b>Avainsanat:</b> DFS, DNS, Active directory, tiedostojaot, hajautettu tiedostonjako	

## ABSTRACT

<b>Name:</b> Antti Salovaara	
<b>Title:</b> Distributed File System - implementation	
<b>Date:</b> 20.5.2010	<b>Number of pages:</b> 46
<b>Department:</b> Information Technology	<b>Study Programme:</b> Datane트워크s
<b>Instructor:</b> Principal Lecturer Janne Salonen	
<b>Supervisor:</b> Jari Myllykangas (Napa Oy), Antti Rautalin (Napa Oy)	
<p>In this work we implement a Distribute File System solution using the build in DFS solution included in Microsoft Server family. We will also be taking a look on the technologies DFS is built over, including Active Directory and Domain Name System.</p> <p>So far the Company's file distribution to foreign offices has been conducted using multiple xcopy based solutions utilising Scheduled Task service and bat files. The bat-solution is to be replaced by a centrally managed Distributed File System solution that can be easily utilised to distribute files from Helsinki office to company's foreign offices and vice versa. It is also intended to be used to create a DFS-Mesh between the Asian offices to facilitate data transfers between the offices and replace the current model of all data going trough Helsinki.</p> <p>The Company's network is divided over multiple continents so we are also looking at the VPN solutions at use between our offices and implementing updates where necessary, both in Helsinki's local network infrastructure and abroad.</p>	
<b>Keywords:</b> DFS, DNS, Active Directory, file shares, distributed file systems	

## SISÄLLYS

### ALKULAUSE

### TIIVISTELMÄ

### ABSTRACT

<b>1</b>	<b>JOHDANTO</b>	<b>1</b>
<b>2</b>	<b>TEORIA</b>	<b>1</b>
<b>2.1</b>	<b>DNS</b> .....	<b>1</b>
2.1.1	<i>Historia</i> .....	2
2.1.2	<i>Tekniikka</i> .....	2
<b>2.2</b>	<b>Active Directory</b> .....	<b>5</b>
2.2.1	<i>Historia</i> .....	6
2.2.2	<i>Tekniikka</i> .....	6
<b>2.3</b>	<b>Distributed File System</b> .....	<b>8</b>
2.3.1	<i>Teoria</i> .....	9
2.3.2	<i>DFS laitteisto- ja ohjelmistovaatimukset</i> .....	12
2.3.3	<i>Muutoksia 2008 versiossa</i> .....	13
2.3.4	<i>Skenaarioita</i> .....	15
<b>3</b>	<b>TESTIYMPÄRISTÖ</b>	<b>16</b>
<b>3.1</b>	<b>Testiympäristön pystytys</b> .....	<b>16</b>
<b>3.2</b>	<b>DFS ympäristön rakentaminen</b> .....	<b>16</b>
<b>4</b>	<b>TOTEUTUS</b>	<b>18</b>
<b>4.1</b>	<b>Verkkoinfrastruktuuri</b> .....	<b>19</b>
<b>4.2</b>	<b>VPN</b> .....	<b>22</b>
<b>4.3</b>	<b>Point-to-point DFS - Server 2008</b> .....	<b>24</b>
<b>4.4</b>	<b>Point-to-Point - Server 2003R2</b> .....	<b>31</b>
<b>4.5</b>	<b>Point-to-multipoint - Server 2003R2</b> .....	<b>36</b>
<b>4.6</b>	<b>Mesh topologia - Server 2003R2</b> .....	<b>38</b>
<b>4.7</b>	<b>Linkkitoteutus - Server 2003R2</b> .....	<b>39</b>
<b>4.8</b>	<b>Toteutuksessa ilmenneitä ongelmia ja virheilmoituksia</b> .....	<b>42</b>
<b>5</b>	<b>YHTEENVETO</b>	<b>43</b>
	<b>VIITELUETTELO</b>	<b>45</b>

## 1 JOHDANTO

Napa Yhtiöllä on tarve saada virtaviivaistettua ja nopeutettua yrityksen verkkolevyillä olevan datan käyttöä etäkonttoreissa toisissa maanosissa, pääosin Aasiassa. Tarkoituksena on toteuttaa Distributed File System -ratkaisu käyttäen Microsoft Server -käyttöjärjestelmissä sisäänrakennettua DFS-ominaisuutta. Tämän toteutuksen ohessa uusitaan myös yhtiöiden verkko-laitteistoa ja virtaviivaistetaan sekä toimistojen sisäistä verkkorakennetta että kehitetään toimistojen välistä VPN-tunneleiden rakennetta. Kyseessä olevat datamäärät eivät ole kovin suuria, mutta linkit etäkonttoreissa ovat hitaita ja kärsivät ajoittaisista katkoksista, jotka rajoittavat huomattavasti yöaikaan tapahtuvaa tiedostonsiirtoa. Etäkonttorit kuuluvat hallinnollisesti samaan Domain-ympäristöön ja VPN-yhteydet ovat laitteistopohjaisia, joten tarkoituksena on konfiguroida myös etäkonttorit Helsingistä käsin.

## 2 TEORIA

Distributed File System on toiminnallisuudeltaan vahvasti sidottu kahteen Windows-palvelinmaailman perusrakenteeseen. Domain Name System eli DNS ja Active directory eli AD mahdollistavat DFS:n toiminnallisuuden ja hallittavuuden laajoissakin palvelinkokonaisuuksissa.

### 2.1 DNS

Domain Name System eli lyhyemmin DNS on Internetissä ja tietoverkoissa (IP-verkot) käytössä oleva nimipalvelujärjestelmä. DNS muuntaa annetun Internet osoitteen osoitemuotoon, jonka perusteella selain tai jokin muu palvelu ymmärtää, minne osoite pitäisi ohjata. Ihmisten on helpompi muistaa osoitteet niminä, mutta liikenne toimii kuitenkin IP-osoitteiden varassa. DNS-palvelimissa tapahtuva häiriö saattaa hetkellisesti aiheuttaa vakavaa haittaa kokonaisen maanosan dataliikenteelle, koska se on tärkeä osa IP-verkkojen perusrakennetta. DNS-palvelimista onkin käytetty termiä "Internetin puhelinluettelo". [1]

### 2.1.1 Historia

DNS:n historia alkoi vuonna 1983, jolloin IP-verkoissa olevien tietokoneiden määrä ylitti niin sanotun kriittisen massan ja osoitteiden nimimuutoksia ei voitu enää tehdä pelkästään tietokonekohtaisten "hosts-tiedostojen" varassa. Hosts.txt on konekohtainen tiedosto, jonne oli käsin tai jonkin apuohjelman avulla lisätty käyttäjän tuntemia osoitteita ja niiden nimiä. Tiedoston sisältö rajoittui käyttäjän omaan tietoon muista koneista eikä se päivittynyt automaattisesti jos jonkin koneen osoite sattui muuttumaan. Alussa hosts.txt-sisältöä pidettiin ajan tasalla SRI NIC (Network Information Centre) tarjoamien hosts.txt-päivitysten avulla (Arpanet-aika). Verkottuneiden koneiden määrän ollessa suhteellisen pieni tämä ei ollut ongelma, mutta verkottuneiden koneiden lisääntyessä listan ylläpidosta tuli käytännössä mahdotonta. Host.txt-tiedosto on kuitenkin vieläkin mukana myös moderneissa käyttöjärjestelmissä sen tarjoamien osoitteiden paikallishallintaominaisuuksien takia.

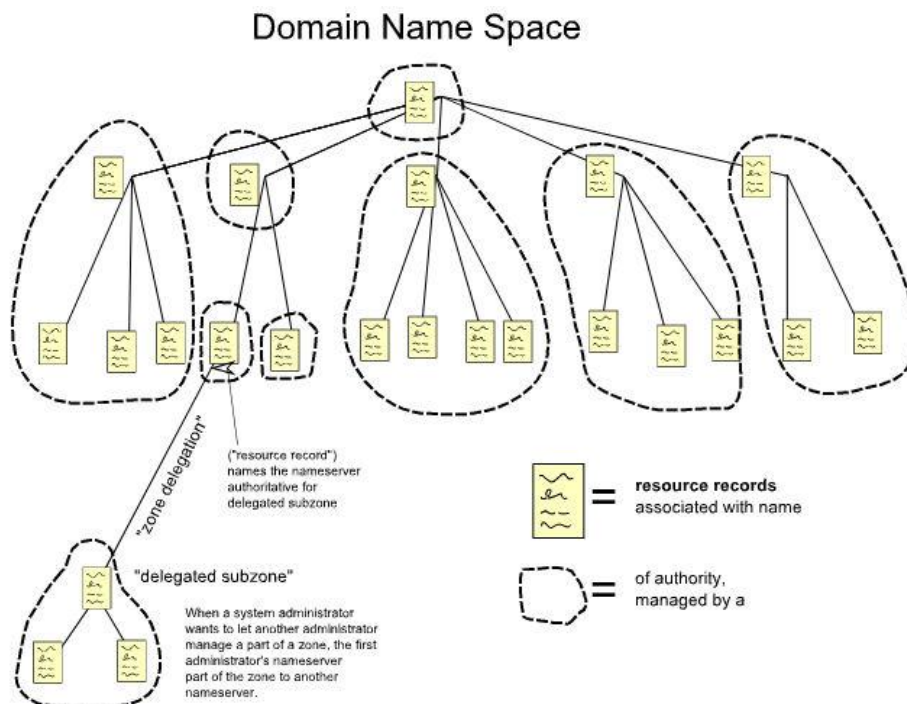
DNS-palvelun isä on Paul Mockapetris (1948-; Bachelors at MIT + PHD in University of California), joka kehitti automaattisen järjestelmän hoitamaan nimi-osoite käännöksiä vuonna 1983 (RFC 882 and RFC 883, 1987: RFC 1034 and RFC 1035). Ensimmäinen UNIX-toteutus, joka pohjautui DNS-malliin, oli vuonna 1984 neljän Berkleyn (University of California, Berkeley) opiskelijan toteutukseen. Berkleyn opiskelijoiden toteutuksen ja Kevin Dunlapin vuonna 1985 tekemän DNS:n laajan uudelleenkehityksen seurauksena syntyi BIND (Berkeley Internet Name Domain), joka on tällä hetkellä Internetissä valtaapitävä DNS-sovellus.

### 2.1.2 Tekniikka

DNS-nimiavaruus koostuu puumaisesta domainrakenteesta. Jokaisen solmukohdan (node) alla on joko nolla tai enemmän tietueita, jotka sisältävät informaatiota liittyen kyseiseen domain-nimeen. Jokainen näistä tietueista voi sisältää nolla tai enemmän tietueita, joka voi sisältää lisää tietueita ja niin edelleen.

DNS-palvelimia on yleensä Internet-palveluntarjoajilla. Valtio ja maanosatasoilla on niin sanottuja "pää" DNS -palvelimia, jotka yhdistelevät ja välittävät tietoa. Yrityksillä on myös DNS-palvelimia, mutta ne ovat yleisesti käy-

tössä yrityksen sisäistä käyttöä varten, eikä niihin ole pääsyä muilla kuin yrityksen sisäverkossa olevilla tietokoneilla.



Kuva 1. DNS:n rakenne

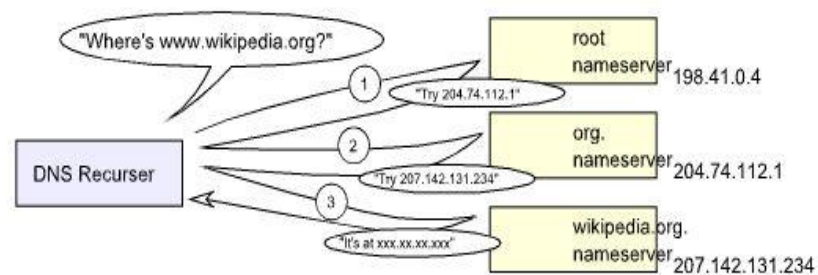
DNS-palvelua ylläpidetään hajautetun tietokantasysteemin avulla, joka käyttää asiakas-palvelin -mallia. Asiakas kyselee tietoa tutulta DNS-palvelimelta, ja jos kyseinen DNS palvelin (node) ei tunnista kysyttyä osoitetta, se lähettää kyselyn hierarkiassa ylemmällä tasolla olevalla palvelimelle, joka joko osaa vastata tiedusteluun, tai lähettää kyselyn eteenpäin ja niin edelleen, kunnes asiakkaalle on saatu vastaus tai tiedustelun TTL-arvo tulee vastaan (Time to live). TTL-arvo on aika, jolloin asiakas päättää ettei tietoa ole saatavilla tai paketti on hävinnyt. Jokaisen DNS-palvelimen tulee tietää ainakin yksi ylemmän tason palvelin (root nameserver), jolta tiedustella tuntematonta osoitetta.

DNS-palvelimia on kahta päätyyppiä: autoratiivisia (Authorative) ja rekursiivisia/välimuistisia (Recursive/caching). Autoratiiviset palvelimet tarjoavat vain niihin asetettua hyväksytystä lähteestä saatua tietoa (original source), esimerkiksi domain pääkäyttäjän tai dynaamisten DNS-palveluiden kautta. Jokaisessa DNS-vyöhykkeessä tulee olla vähintään yksi autoratiivinen palvelin, joka täytyy olla ylemmän asteen DNS-hierarkian tiedossa. R/C DNS -palvelimet ovat yleisimmin käytössä Internet-ympäristössä, koska ne varastoivat kyselyiden mukana kertynyttä tietoa määrätyn ajan. Yleisesti tarvitta-



vaa tietoa ei aina tarvitse varmistaa ylemmältä hierarkia-asteelta, vaan nimikysely voidaan ratkaista palvelimeen varastoidun (cache) tiedon perusteella rekursiivisesti (recursive). Tämä nopeuttaa nimikyselyiden toimintaa huomattavasti. Cache-palvelimet mahdollistavat osoitteenmuutosten nopeuden. Jos kaikki kyselyt päättyisivät hierarkian ylimmille asteille asti, nämä palvelimet olisivat täysin käyttökeltottomia massiivisten kyselymäärien takia. DNS on kuitenkin rakennettu redundanttina systeeminä ja näin ollen yhden, tai useammankaan palvelimen kaatuminen ei ratkaisevasti heikennä verkon liikennettä.

Peruskäyttäjät eivät yleensä kommunikoi suoraan DNS-palvelinten kanssa vaan DNS toiminto tapahtuu ohjelmistojen välillä, esimerkiksi sähköpostiohjelmiston tai selaimen välityksellä. Käänteinen toiminto DNS-nimen selvitykselle on "DNS reverse lookup", jossa selvitetään nimi tai nimet, jotka johtavat tiettyyn IP-osoitteeseen. Haku toimii vastaavalla hierarkkisella rakenteella kuin normaali DNS-nimikysely.



Kuva 2. DNS-kysely

Domain-nimien muodostus on määritelty kolmessa eri dokumentissa (RFC 1035, RFC 1123, and RFC 2181). Domain-nimen tulee koostua yhdestä tai useammasta osasta esimerkiksi "esimerkki.osoite.fi". Domain nimen rakenne on hierarkkinen ja sitä tulee lukea oikealta vasemmalle. Oikeanpuoleisin pala kertoo mihin "top-level" domainiin osoite kuuluu, tässä esimerkissä *fi*. Seuraava osuus "osoite" on domainin *fi* alidomaini (subdomain) ja *esimerkki* taas *osoite* palan alidomaini, ja niin edelleen. Tämä puu voi koostua maksimissaan 127 osasta ja jokainen osa voi olla maksimissaan 63 merkkiä pitkä. Koko osoitteen pituus ei voi kuitenkaan ylittää 253 merkkiä. DNS sallii kaikkien niiden merkkien käytön, jotka voidaan esittää oktettimuodossa (01001001). Yleisesti on hyväksytty, että osoitteissa käytetään vain ASCII-

merkistön mukaisia merkkejä ja niin sanottua LDH-sääntöä (letters, digits, hyphen).

Verkkoteknisesti DNS toimii OSI-tasolla 7 (application layer). DNS käyttää liikenteessään UDP-porttia 53. DNS-kyselyt koostuvat yhdestä alle 512 tavun pyynnöstä ja palvelimen vastauksesta. Jos paketin koko ylittää 512 tavua käytetään TCP:tä. DNS kyselyt käyttävät standardoitua rakennetta.

*Taulukko 1. DNS-tietueen osat*

<i>Kenttä</i>	<i>Kuvaus</i>	<i>Pituus (oktettia)</i>
<i>NAME</i>	<i>Solmukohdan nimi johon tämä listaus viittaa (Fully qualified domain name).</i>	<i>(vaihteleva)</i>
<i>TYPE</i>	<i>Resurssimerkinnän tyyppi.</i>	2
<i>CLASS</i>	<i>Luokkakoodi.</i>	2
<i>TTL</i>	<i>Elossaoloaika, maksimi 2147483647 sekuntia.</i>	4
<i>RDLENGTH</i>	<i>RDATA kentän pituus.</i>	2
<i>RDATA</i>	<i>Resurssirekisterin lisätietoja.</i>	<i>(vaihteleva)</i>

[1;2;3]

## 2.2 Active Directory

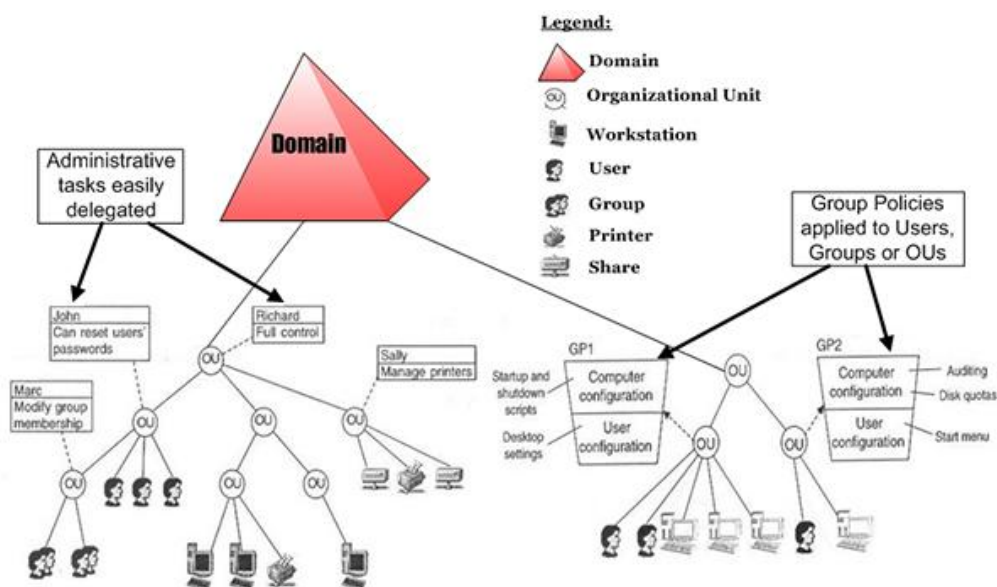
Active Directory eli AD on nykyaikaisen palvelinrakenteen kantava voima. AD sisältää informaatiota verkon käyttäjistä, tietokoneista ja verkon muista resursseista. AD mahdollistaa näiden resurssien keskitetyn hallinnan, valvonnan ja suojauksen. AD on Microsoftin palvelinkäyttöjärjestelmissä oleva keskitetty hallintaratkaisu, joka on ollut yrityksen palvelintuotteissa Windows 2000 Server -käyttöjärjestelmästä lähtien.

### 2.2.1 Historia

Active Directory otettiin käyttöön vuonna 1999, kun Windows 2000 Server julkistettiin korvaamaan vanhaa NT4-pohjaista ratkaisua. Sitä on myös kehitetty uusien Windows-palvelinkäyttöjärjestelmien julkaisujen myötä. Server 2008 julkaisun yhteydessä AD:ta muokattiin enemmän ja se uudelleennimettiin Active Directory Domain Services:ksi. Joissain yhteyksissä AD:sta on myös käytetty nimitystä NTDS (NT Directory Service). Erilaisia luetteloratkaisuja, jollaisena Active Directory:kin voidaan pitää, on ollut olemassa 1980-luvun lopulta lähtien. Ensimmäisenä voidaan pitää 1988 ITU:n (International Telecommunication Union) julkaisemaa X.500-standardikokoelmaa sähköisistä hakemistopalveluista, jota hyödynnettiin sähköisissä postipalveluissa ja nimipalveluissa. Suuri merkkipaalu hakemistopalveluissa oli LDAP:n (Lightweight Directory Access Protocol) synty Michiganin yliopistossa 1992. Sitä kehitetään vielä tänäkin päivänä. Uusin virallinen LDAP-versio on 3, mutta sen pohjalta on kehitetty monia avoimia ratkaisuja. LDAP:in kohtaa vielä tänäkin päivänä erilaisissa integraatoratkaisussa Active Directoryn ja avoimen lähdekoodin ratkaisujen välissä mahdollistamassa tiedonkulkua Active Directoryn ja palvelun välillä.

### 2.2.2 Tekniikka

Active Directory koostuu objekteista. Objekteja ovat esimerkiksi käyttäjä, tietokone, palvelu tai resurssi. AD:n tehtävänä on valvoa näitä resursseja. AD noudattaa hierarkkista kehysrakennetta, kuten yleisimmät tiedostorakenteet ja DNS. Jokainen objekti edustaa jotain yksittäistä oliota, oli se sitten käyttäjä, palvelu tai muu objekti. Tietyt objektit voivat olla myös säilöjä (container) muille objekteille. Active Directory rakentuu Domain controllereista (DC), jotka ohjaavat domainin toimintaa ja rakennetta sekä muista domainiin liitettyistä laitteista kuten palvelimista, työasemista ja tulostimista.



Kuva 3. Domainin rakenne

Domain Controllerit ovat Domainin hallintakoneita, joiden Active Directoryyn varastoidaan kaikki domainin toiminnan kannalta tärkeä tieto. Domain controllerit välittävät eli replikoivat domainin sisällä tietoa keskenään jotta kaikki laitteet pysyvät ajan tasalla mahdollisen laitevian tai muun vastaavan ongelman takia. Onkin suositeltavaa, että jokaisessa domainissa olisi käytössä vähintään kaksi domain controlleria, jotta ratkaisu olisi redundantti eli yhden DC:n kaatuminen ei lamauttaisi koko järjestelmää. Replikaatiot eri domain controllereiden välillä on hoidettu automatisoidulla ratkaisulla, joka välittää tietoa controllereiden välillä aina 5 minuuttia muutosten tapahtumisen jälkeen (oletusaika). Tämä tehdään siksi, että järjestelmänvalvoja ehtii tehdä kaikki muutoksensa ennen kuin ne replikoituvat muille palvelimille. Tämä vähentää tarvittavien siirtotapahtumien määrää. Jos muutoksia ei tapahdu, controllerit replikoivat oletuksena 3 tunnin välein varmistaakseen datan oikeellisuuden. Paikallisissa controllereiden välisissä replikaatioissa käytetään RPC-protokollaa (Remote procedure call), mutta maantieteellisesti kauempana oleville controllereille on myös mahdollistettu SMTP-protokollan (Simple Mail Transfer Protocol) käyttö. AD käyttää niin sanottua multimaster replikaatiota, joka mahdollistaa muutosten teon millä tahansa Domain Controllerilla.

Domain on aina osa jotain DNS-nimiavaruutta. Siksi domainin sisällä voidaan tehdä nimikyselyjä ja ohjata palvelukyselyitä oikeille palvelimille. Domain on myös yksi hallittava yksikkö, johon eri käyttäjillä on eritasoisia oikeuksia erilaisiin resursseihin. Active Directoryssä asetetut oikeudet vaikuttavat oikeuksiin koko domainissa ja näin ollen tarjoavat keskitetyn hallinnan ratkaisun verkon oikeuksien hallintaan. Domain-oikeuksia voidaan myös luottamussuhteiden (Trust Relationship) avulla laajentaa muihin domaineihin järjestelmänvalvojien niin halutessa.

AD:n rakenne on määritelty objektien avulla. Näitä objekteja voidaan ryhmitellä erilaisiin hallinnollisiin ja teknisiin ryhmiin. Yleensä domain on puumainen rakennelma, joka jaotellaan erilaisiin organisaatioyksiköihin (OU - Organizational Unit), joita voidaan hallita erilaisilta hierarkkisilta asteilta. Jotkut asetukset ja oikeudet vaikuttavat vain tiettyyn OU:hun ja jotkut kaikkiin domainin objekteihin. Administraativisia tehtäviä voidaan myös jakaa (delegate) OU kohtaisesti. Tämä mahdollistaa maantieteellisesti laajan domainin tiettyjen osien hallinnan paikallisesti.

OU:den avulla ei voida kuitenkaan jakaa käyttöoikeuksia erilaisiin objekteihin. Tähän käytetään erilaisia ryhmiä. Ryhmillä on Active Directory:ssä kaksi keskeistä tehtävää; turvallisuusryhmät (Security Group), joilla toteutetaan käyttöoikeuksien jako monelle käyttäjälle ja jakeluryhmät (Distribution Group), joilla mahdollistetaan yleisesti ryhmäsähköpostien lähettäminen AD-pohjaisissa sähköpostijärjestelmissä kuten Exchange. Turvallisuusryhmiä voidaan käyttää myös jakeluryhminä. Jokaiselle ryhmälle määritellään tietty laajuus (scope), jonka sisällä ryhmä toimii. Näitä laajuuksia on kolme erilaista, globaali (Global), universaali (Universal) ja domain lokaali (Domain local). Näillä kolmella joukolla määritellään ryhmien toimivuus luottamussuhteiden avulla liitettyjen domainien kanssa. Lokaali ryhmä toimii vain siinä domainissa missä se on luotu. Globaaleja ja universaaleja ryhmiä voidaan hyödyntää myös ristiin domainien välillä.

[4;5;6]

## 2.3 Distributed File System

Distributed File System eli DFS on Microsoftin uudemmissa palvelinkäyttöjärjestelmissä mukana oleva ominaisuus, jonka avulla voidaan muodostaa erilaisia tiedostojakoja ja replikointeja useiden sijaintien välillä. Rakenteel-

taan DFS on hierarkkinen ratkaisu, jolla voidaan muodostaa keskitetty ratkaisu useista tiedostopalvelimista ja jaoista verkossa. DFS:n avulla kyetään luomaan toteutus, jossa käyttäjä ohjataan lähimmälle palvelimelle, missä käyttäjän hakema tieto sijaitsee. Sama tieto voi sijaita useissa paikoissa identtisenä, ja käyttäjä ohjautuu aina lähimmälle palvelimelle, josta tarvittava tieto on saatavilla. Jos käyttäjän sijainti muuttuu lähemmäs toista palvelinta DFS-palvelin alkaa ohjata samaa käyttäjää uudelle palvelimelle, koska siellä oleva tieto on nopeammin saatavilla käyttäjälle. Palvelimen valinta pohjautuu Active Directoryssä sijaitseviin "site metric" -tietoihin. DFS-palveluista on suurta hyötyä yrityksissä, joilla on konttoreita maantieteellisesti keskenään kaukaisissa pisteissä, mutta kaikilla konttoreilla on tarve päästä käsiksi samaan dataan.

### 2.3.1 Teoria

DFS-käytössä on olemassa kolme keskeistä termiä, joiden ymmärtäminen on suotavaa, jotta DFS-rakenne on ymmärrettävää.

- DFS Root (Juuri): Verkkoon näkyvä jako. Se sisältää alikansioita ja tiedostoja. Juuri näkyy käyttäjällä kuten mikä tahansa normaali tiedostojako sillä pienellä poikkeuksella, että jakonimi ei sisällä tietokoneen nimeä vaan on mallia *domainnimi.fi/jaonnimi*. Juuri on säiliö (container), jota käytetään tiedon varastointiin Active Directory -ympäristössä. Server 2008 -toteutuksissa juuret ovat muuttuneet nimiavaruuksiksi (DFS Namespace).
- DFS Link (Linkki): Juuren sisällä oleva linkki ohjaa toiselle jaolle jossain verkon sisällä. Server 2008 toteutuksissa käytetään termiä kansio (folder).
- DFS Target/Replica (Kohde/Replica(kopio)): DFS-replikoinnin kohde, eli kansio joka sisältää samat tiedot kuin DFS-juuri. Nämä kaksi jakoa löytyvät saman juuren alta ja käyttäjän sijainti vaikuttaa siihen kummasta juuresta tiedot haetaan. Server 2008 -toteutuksissa kutsutaan kansio-kohteiksi (folder targets).



Kuva 4. DFS rakenne

Nykymuotoinen DFS-ratkaisu kehitettiin Server 2003 -käyttöjärjestelmää varten, mutta se on vain kehittyneempi versio Server 2000 olleesta "File Replication service" -ratkaisusta. 2003 versio sisältää vikasietoisuusratkaisuja, nopeamman toiminnan, erilaisia kuormanjakoratkaisuja ja käytössä olevan kaistan rajoittamisen. 2003 mahdollistaa myös useamman kuin yhden juuren muodostamisen domainia kohden.

DFS-replikaatio on tilapohjainen "multimaster" -replikaatoratkaisu. Replikaatiossa käytetään Remote Differential Comparison (RDC) -toteutusta, joka mahdollistaa tehokkaan replikoinnin hitaiden yhteyksien yli. RDC havaitsee tiedon lisäyksen, poiston ja uudelleenjärjestelyn tiedostojen sisällä. Näin mahdollistuu vain muuttuneen datan siirto ja siten säästetään kaistaa, koska kaikkia dataa ei tarvitse siirtää uudestaan.

RDC on yleinen termi, joka voi viitata useaan pakkausalgoritmiin joita voidaan käyttää vertaamaan kahta tiedostoa, jotka sijaitsevat eri paikoissa ilman että toinen tiedostoista täytyisi siirtää toisen tiedoston sijaintiin. [lainaus,7]

Käytännössä RDC jakaa käsittelemänsä tiedostot pieniksi osioiksi ja määrittelee mitkä osioista ovat jo valmiiksi kohdesijainnissa, jotta oikea kopio tiedostosta voidaan kasata kohdesijainnissa. Tämän saavuttamiseksi on tarpeen siirtää vain ne osiot, jotka eivät ole vielä kohdesijainnissa. Jos suuri osa osioista on jo valmiiksi kohdesijainnissa, nopeutuu tiedonsiirto huomattavasti. Osioiden tunnistuksen avulla tiedoston muuttuessa, vain muuttuneet osiot joudutaan siirtämään. Microsoftin tarjoamassa esimerkkitapauksessa alkuperäisen datan replikoinnin jälkeen käytettiin datan replikointiin noin 63 % vähemmän kaistaa kuin normaalin tiedostonkopiointiin olisi käytetty.

Koska DFS on tilapohjainen ratkaisu, se kykenee replikoimaan useamman kohteen kanssa kuin edeltäjänsä FRS. Microsoft lupaa DFS:n kykenevän alla esitettyihin arvoihin.

- Jokainen palvelin voi kuulua maksimissaan 256 replikaatioryhmään.
- Jokainen replikaatioryhmä voi sisältää 256 replikoituvaa kansiota.
- Jokainen palvelin voi käsitellä 256 yhtäaikaista yhteydenottoa (esimerkiksi 128 sisääntulevaa ja 128 ulosmenevää).
- Jokainen replikaatioryhmä voi sisältää 256 jäsentä.
- Jaettu juuri voi sisältää 8 miljoonaa replikoituvaa tiedostoa ja maksimissaan teratavun replikoituvaa dataa.

Replikointia varten Distributed File System sisältää kaksi vaihtoehtoa. Automaattinen replikointi on mahdollista vain domain DFS käytössä. Se replikoi dataa aina muutosten sattuessa ja seuraa jatkuvasti jakojen tilaa mahdollisten muutosten varalta. Manuaalinen replikointi on ainoa vaihtoehto itsenäisessä DFS-ratkaisussa. Se vaatii käyttäjän toimia aloittaakseen replikaation. Manuaalinen ratkaisu ei seuraa jaoissa tapahtuvaa tiedostoliikennettä vaan aloittaa replikoinnin vasta, kun sille niin ilmoitetaan. Kahden tai useamman jaon välisissä replikaatioissa synkronoidaan tiedostot käyttäjän määrittämän topologian mukaan. Topologiat ovat perinteisiä verkkotopologiamalleja; rengas (ring), säteittäinen malli (hub & spoke) ja täyskytketty malli (mesh ratkaisu). Lisäksi on mahdollista käyttää kehittyneempää "Custom"-metodia, jossa pisteiden väliset linkit voidaan määrittellä linkki linkiltä vastaamaan käyttäjän tarpeita, aikakohtaisesti 16 kilobitin ja 256 megabitin välillä. Liikenteen voidaan antaa kulkea myös rajoittamattomana. Aikaero-ongelmien välttämiseksi ajastukset voidaan suorittaa joko käyttämällä palvelimen paikallista aikaa tai UTC-aikaa (Coordinated Universal Time).

Ongelmatilanteiden varalta DFS-replikaatioissa on niin sanottu "Self-healing"-ominaisuus mahdollisen synkronointitietokannan korruptoinnin takia. Tiedostojen synkronointi lakkaa toiminnon ajaksi, mutta mitään käyttäjän toimia ei vaadita ongelman korjaamiseksi. Optimaalisessa tilanteessa käyttäjä tai järjestelmävalvoja ei edes huomaa ongelmaa. Korjauksen aikana tietokanta synkronoituu toisen replikaatioryhmän jäsenen tietokannan kanssa. Tietokannan koko määräytyy jaossa olevien tiedostojen lukumäärän mukaan. DFS varastoi globaalit konfiguraatietonsa Active Directoryssä ja paikallisesti jokaisen DFS-jäsenen kovalevyllä .xml-tiedostona. Esimerkiksi laiterikon aiheuttamasta korruptiosta voidaan toipua Active Directoryssä olevan informaation avulla ja uudelleenrakentaa jäsenkohtainen replikaatietokanta.



Kun ollaan lisäämässä uutta palvelinta replikaatioryhmään tai kun ollaan luomassa täysin uutta replikaatioryhmää, jäsenpalvelimia voidaan valmistella (staging) tähän rooliin kopioimalla osa replikoitavasta datasta etukäteen. Tämä on tehokasta tilanteissa, joissa tiedon on hyvä olla molemmissa paikoissa, mutta linkki on hidas ja tapahtuvat muutokset eivät ole suuria. Kun replikaatioryhmä luodaan, isäntäpalvelin tunnistaa toisessa päässä jo valmiiksi olevat tiedostot eikä turhaan kopio niitä vaan päivittää vain muutokset. Yhteensopimattomat tiedostot siirretään PreExisting-piilokansioon, jossa niitä voi tarkastella myöhemmin.

DFS-replikaatiot ja nimiavaruudet luodaan käyttämällä joko Windowsin hallintakonsolia ja DFS-management lisäosaa tai käyttäen komentorivipohjaisia työkaluja Dfsradmin.exe ja Dfsrdiag.exe. DFS-asetuksia pystytään tekemään oletuksena vain Domain Admin -tasoisella tunnuksella, mutta oikeudet voidaan myös delegoida muille käyttäjille tarpeen vaatiessa. Delegaatio voidaan toteuttaa käyttämällä edellä mainittuja työkaluja. Server 2003 R2 lisättiin perus DFS-asennukseen uutena työkaluna Dfsutil.exe, joka oli jo aikaisemmin ollut osa erillistä Support tools -pakettia. Lisäksi pakettiin lisättiin graafinen DFS management -lisäosa, joka mahdollistaa erityisomaisuuksien konfiguroinnin. Se oli aikaisemmin mahdollista vain komentorivipohjaisissa ratkaisuissa.

DFS -toimintaa voidaan seurata WMI-ratkaisun avulla (Windows Management Instrumentation), jossa on sisäänrakennettuja seurantatyökaluja. DFS-replikaatiota voidaan monitoroida myös "Windows DFS Replication Management Pack for Microsoft Operations Manager" (MOM) -avulla. Monitorointia voidaan seurata tosiaikaisena jäsenkohtaisesti. Monitoroinnissa seurataan replikaatiopalveluiden tilaa, ryhmiä, replikoituneita kansioita ja levyjä, joille replikoidut tiedostot on varastoitu.

### 2.3.2 DFS laitteisto- ja ohjelmistovaatimukset

DFS replikaatio 2003:

- Active Directory schema täytyy päivittää sisältämään DFS -määritykset, jotka ovat saatavissa 2003R2 toiselta asennuslevyltä.
- Kaikkien palvelinten, jotka ajavat DFS replikaatio palvelua, tulee olla vähintään Server 2003R2 tasoa. Lisäksi DFS Management -applikaation tulisi olla asennettuna.

- Replikaatioryhmän kaikkien palvelinten tulee sijaita samassa metsässä (forest).
- Kaikkien replikoitavien kansioden tulee sijaita NTFS-levyjaoilla ja palvelin klustereilla. Kansioden tulisi sijaita paikallisella levyllä koska DFS-replikaatio ei ole klusteri yhteensopiva.

DFS nimiavaruus 2003:

- Active Directory schema täytyy päivittää sisältämään DFS-määritykset, jotka ovat saatavissa 2003R2 toiselta asennuslevyltä.
- Kaikkein palvelinten, jotka ajavat DFS replikaatiopalvelua ja domain kontrollereiden, tulee olla vähintään tasoa Server 2003 SP1 tai Server 2003R2. Lisäksi DFS Management -applikaation tulisi olla asennettuna.
- Asiakaskoneilla tulisi olla Client Failback -päivitys asennettuna.

DFS-toimintoja on myös mahdollista ajaa vaikka kaikki edellä mainitut ehdot eivät täytyisikään, mutta silloin toiminta on rajoittunutta. Täyden Server 2003 tasoisen DFS-palvelun ajaminen vaatii kaikkien ehtojen täyttämistä.

### 2.3.3 Muutoksia 2008 versiossa

Distributed File System ratkaisun Server 2008 versiossa käytetty palvelu on jakautunut kahdeksi erilliseksi palveluksi, DFS Namespace's ja DFS Replication, joiden nimet selittävätkin miten toiminta on jakautunut. Toinen palvelu hoitaa vain replikoinnit ja toinen hoitaa nimiavaruuden hallinnan ja viittaukset. Server 2008 esiteltyjä uusia ominaisuuksia ovat:

- "Access-based enumeration" parantaa käyttäjäkohtaista oikeuksienhallintaa DFS jaoilla.
- "Cluster-support" tuo DFS ratkaisuihin "failover cluster" tuen. Ominaisuus toimii kuitenkin vain "stand alone" (yksittäisessä) nimiavaruudessa.
- Komentorivityökaluja Dfsutil.exe ja Dfsdiag.exe on kehitetty vastaamaan uusia ominaisuuksia.
- Graafiseen hallintaliittymään on lisätty etsintätoiminto helpottamaan suurien kokonaisuuksien hallintaa.
- "Content Freshness" -ominaisuus selkeyttää toimintaa tilanteissa joissa tietty palvelin on ollut pitkään irti verkosta.
- Nopeutettu toipuminen odottamattomista katkoksista.

Uusien ominaisuuksien käyttöönotto edellyttää seuraavien ehtojen täyttymistä.

- Metsän täytyy käyttää vähintään Server 2003 toimintatasoa.
- Domainin täytyy olla vähintään Server 2008 toimintatasolla.
- Kaikkien nimiavaruuspalvelimien tulee olla Server 2008 tasoa.

Lisäksi Server 2008:ssa on parannettu synkronointiprotokolla, jonka avulla saadaan aikaiseksi nopeampi replikointi sekä isoilla että pienillä tiedostoilla ja alkusynkronointi valmistuu nopeammin. Päivitys mahdollistaa myös tehokkaamman kaistanleveyden hyödyntämisen korkean latenssin verkoissa ja LAN-verkoissa. Replikointiin on lisätty myös "Replicate now" -ominaisuus, jolla voidaan pakottaa replikointi alkamaan heti riippumatta ajastuksista ja mahdollisista muutoksista. Myös diagnostiikkaominaisuuksia on laajennettu.

Taulukko 2. Server 2003R2 ja 2008 synkronointieroja

<b>Windows Server 2003 R2</b>	<b>Windows Server 2008</b>
<i>Multiple RPC calls</i>	<i>RPC Async Pipes (when replicating with other servers running Windows Server 2008)</i>
<i>Synchronous inputs/outputs (I/Os)</i>	<i>Asynchronous I/Os</i>
<i>Buffered I/Os</i>	<i>Unbuffered I/Os</i>
<i>Normal Priority I/Os</i>	<i>Low Priority I/Os (this reduces the load on the system as a result of replication)</i>
<i>4 concurrent file downloads</i>	<i>16 concurrent file downloads</i>

### 2.3.4 Skenaarioita

Distributed File System on suunniteltu käytettäväksi erityisesti kolmessa alla esitettyssä skenaariossa, mutta myös muunlaiset ratkaisut ovat mahdollisia. Tässä vaiheessa on hyvä huomioida, että DFS ei sovellu käytettäväksi ja-oissa, joissa tiedostot ovat jatkuvasti käytössä, kuten tietokannat tai vastaavat tiedostot. Toisena ongelmallisena ominaisuutena voidaan pitää tilannetta, jossa replikoitua tiedostoa muokataan useassa sijainnissa samanaikaisesti. Replikaatio pohjautuu heuristiseen "viimeinen kirjoittaja voittaa" -malliin eli jos tiedosto avataan usealta juurelta samanaikaisesti on mahdollisesta että jotain muokattua tietoa voidaan menettää.

DFS-ratkaisuja voidaan käyttää dokumenttien, ohjelmistojen ja business datan julkaisuun organisaation sisällä. Julkaistu DFS-replikaatioryhmä mahdollistaa suurien tietomäärien keskitetyn mutta samanaikaisesti hajautetun jaon. Nimiavaruus voi sisältää useita palvelimia, joilla kaikilla sijaitsee identtinen data. Kun käyttäjä hakee tietoa, hänet ohjataan käyttäjän kannalta nopeimmalle palvelimelle. Se nopeuttaa ja helpottaa yrityksen sisäistä tiedonhakua. Erilaisia käyttäjien ohjausprioriteetteja voidaan määrittellä Active Directoryn avulla.

Suuremmissa organisaatioissa on usein tarvetta jakaa tietoa useiden etäkonttoreiden kanssa ja mahdollistaa pääkonttorissa sijaitsevan datan nopea käyttö. DFS-replikaatiolla mahdollistetaan etäkonttorin pääsy käsiksi tiedostoihin vaikka toimistojen välinen linkki olisikin hetkellisesti poikki. Linkin palautuessa DFS tarkistaa tiedostojen tilan ja replikoi muutokset konttoreiden välillä. Ratkaisu vaatii kuitenkin paljon levytilaa etäkonttoreiden palvelimilta ja riippuen käytössä olevan datan määrästä, tiedon replikoitiin kuluva aika voi olla huomattava. Ratkaisu helpottaa myös paljon matkustavien työntekijöiden elämää, koska tiedostot ovat aina saatavilla lähimmästä konttorista.

Kolmantena tyypillisenä skenaariona voidaan pitää eriasteista tiedostojen keräämistä, yleisesti tiedostojen varmuuskopiointia. Etäkonttoreissa sijaitseva data replikoidaan tiettyyn sijaintiin, jossa suoritetaan keskitetty varmuuskopiointi. Tällä ratkaisulla eliminoidaan esimerkiksi tarve paikallisille, usein kalliille, varmuuskopiointiratkaisuille. RDC:N avulla replikoidaan aina vain muutokset, joten alkuperäisen datasiirron jälkeen tietomäärien tulisi olla helposti käsiteltäviä. Ratkaisua voidaan myös kehittää käyttämällä ajastettua replikointia ja käytettävän kaistan rajoittamista jotta datasiirto ei aiheuta hita-

utta business-kriittisissä sovelluksissa. Nimiavaruuksien avulla voidaan myös määritellä DFS siten että etäkonttorien ongelmatapauksissa, esimerkiksi tiedostopalvelimen rikkoutuessa, käyttäjät hakevat automaattisesti tarvitsemansa tiedot suoraan pääkonttorin palvelimelta. Käyttäjät eivät huomaa muutosta muuten kuin linkin mahdollisena hitautena.

[8;9;10;11;12]

### **3 TESTIYMPÄRISTÖ**

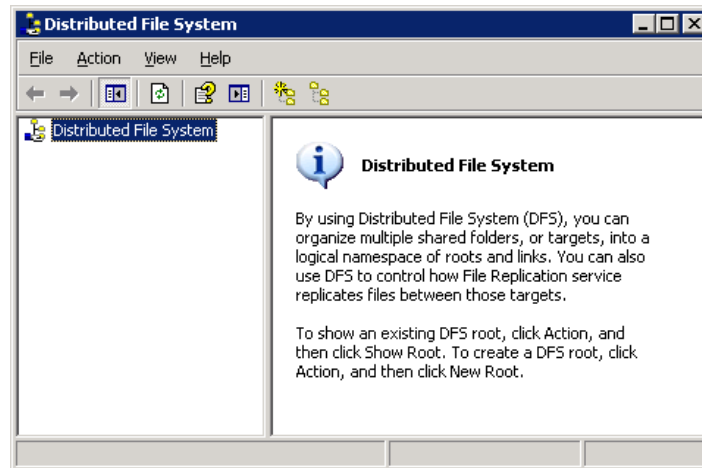
#### **3.1 Testiympäristön pystytys**

Ennen Distributed File System -ratkaisun rakentamista yritysympäristössä todettiin, että olisi hyvä kokeilla ratkaisua ympäristössä, jossa se ei aiheuttaisi vahinkoa olemassa olevalle infrastruktuurille. Päätettiin rakentaa kahden testikoneen välille yksinkertainen domain-ratkaisu, jossa DFS-ratkaisua voitaisiin kokeilla ja tarvittaessa myös harjoittaa ongelmanratkointia hallituksessa ympäristössä.

Molemmille koneille asennettiin Windows Server 2003R2 -käyttöjärjestelmä ja kaikki sillä hetkellä jakelussa olleet päivitykset ja yrityksen lisensoima tietoturvaratkaisu eli virustorjunta ja palomuri. Tarkoituksena oli luoda mahdollisimman samankaltainen asennusympäristö kuin yritysverkossakin. Ensimmäinen palvelin muutettiin Domain Controlleriksi käyttäen "Configure Server" -velhoa. Käyttöön otettiin täysin uusi domainpuu. Lisäksi palvelin määriteltiin toimimaan DNS-palvelimena. DNS-palvelimelle luotiin sekä forward ja reverse lookup vyöhykkeet, koska samat ratkaisut ovat käytössä myös yritysverkossa.

#### **3.2 DFS ympäristön rakentaminen**

Aluksi asennetaan File Server rooli "Configure Server" -velhon avulla. Asetuksista määriteltiin asennettavaksi myös kaikki DFS-työkalut. Tämän jälkeen käytettävien koneiden palomureista tulee avata portti 445, joka sisältyy File and Print Sharing -sääntöön Windows-palomuurissa. Lisäksi "Distributed File System" -palvelun (Service) tila tulee muuttaa siten, että se käynnistyy automaattisesti.



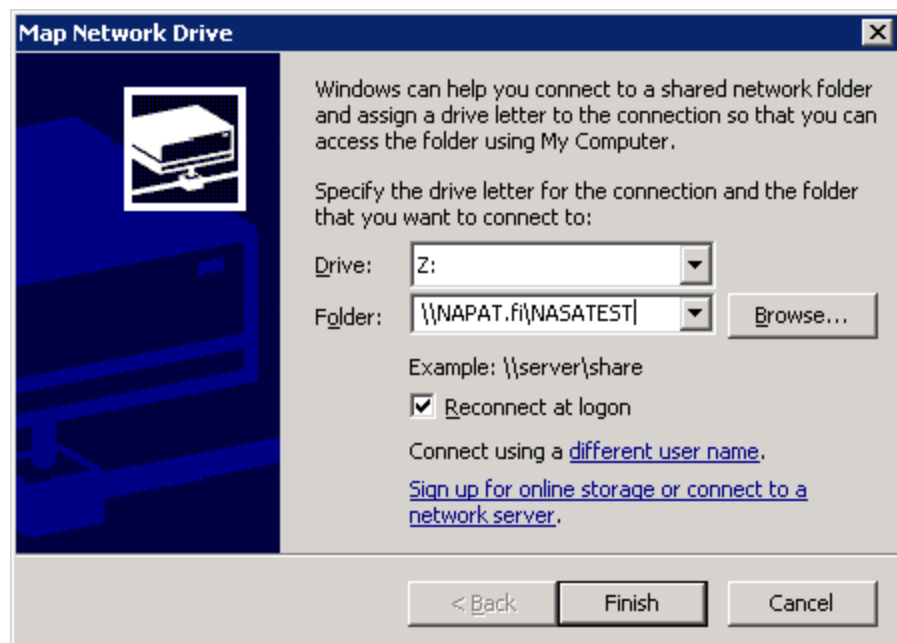
Kuva 5. Server 2003R2 DFS hallintakonsoli

Seuraavaksi käynnistetään DFS management -konsoli, josta voidaan luoda ja hallita uusia DFS-ryhmiä. Vasemmassa palkissa näkyvät olemassa olevat DFS-juuret ja -linkit. Oikealla puolella näkyvät valitun juuren tai linkin jäsenet. Uusia juuria voidaan lisätä ja muokata joko Action-valikosta tai Distributed File System -linkin alta aukeavasta valikosta. Luodaan siis uusi juuri (Root) Action-valikon alta valitsemalla Create New Root, joka käynnistää velhon joka ohjaa kohta kohdalta prosessin läpi.

1. Aluksi tulee määrittellä juuresta joko Domain vai Stand-alone root. Stand-alone -vaihtoehtoa ei kannata käyttää kuin silloin, kun muuta vaihtoehtoa ei ole, eli kun domainia ei jostain syystä ole.
2. Kun valitaan domain vaihtoehto, seuraavaksi määrittellään mihin domainiin halutaan liittyä. Todennäköisesti vaihtoehtoja on vain yksi, mutta suuremmissa yritysympäristöissä vaihtoehtoja voi olla useitakin erilaisten domainien välisten luottamussuhteiden takia (trust-relationship).
3. Host Server toimii kyseessä olevan DFS-juuren isäntäpalvelimena eli hallinnoivana jäsenenä. Hallintaa voidaan kuitenkin suorittaa miltä tahansa domainissa olevalta palvelimelta tai työasemalta käsin (Admin Pack).
4. Juuren nimeksi tulee määrittellä jokin lyhyt ja helposti muistettava selkeyden takia.

5. Jaoksi tulee määritellä jokin Host Server koneella oleva kansio. Kansion ei tarvitse olla valmiiksi jaettu ja se voi sisältää dataa.
6. Lopuksi hyväksytään asetukset.

Juuri on nyt valmis mutta sisältää vain yhden kohteen Seuraavaksi voidaan lisätä uusia "Root Targetteja" yllä olevan kaavan mukaisesti tai lisätä linkkejä eli "al kansiota" tähän juureen. Kun kaikki halutut komponentit on lisätty, mennään juuren ominaisuusvalikkoon, josta juurelle määritellään replikaatioasetukset ja se julkaistaan. Jako on tämän jälkeen näkyvissä domainissa kuten mikä tahansa verkkoresurssi.



Kuva 6. Esimerkki jakoon yhdistämisestä

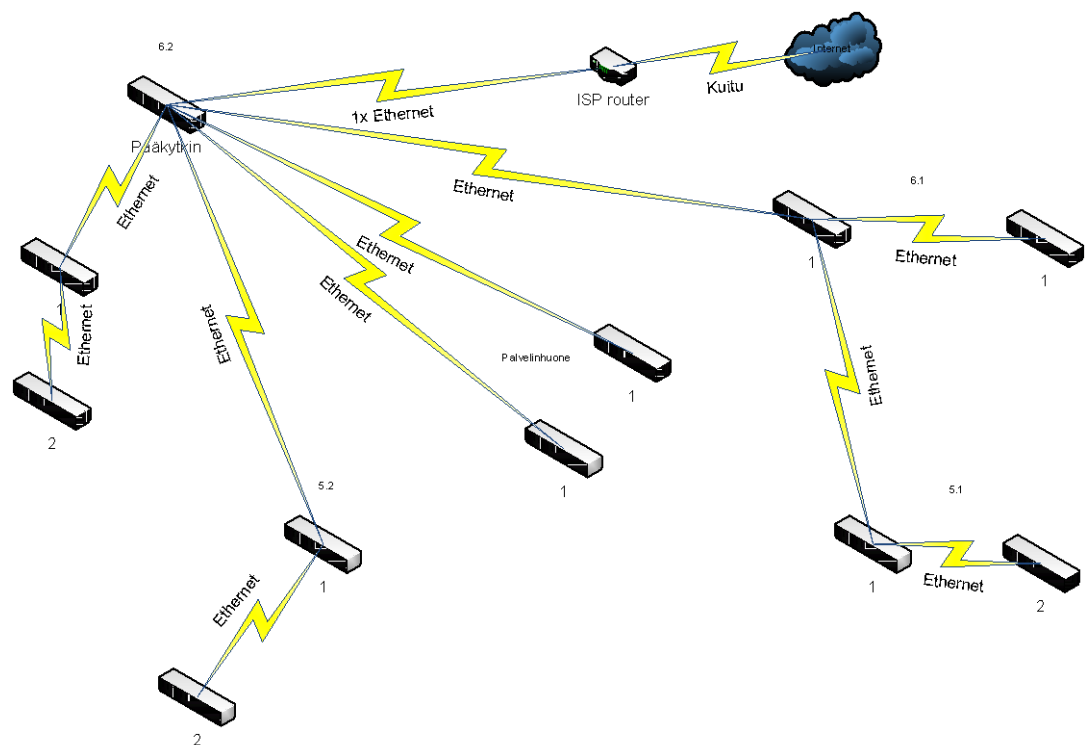
#### 4 TOTEUTUS

Projektin tarkoituksena on muodostaa vakaa ja helposti hallinnoitava ratkaisu, jolla kyetään siirtämään dataa yrityksen Helsingin ja Aasian etäkonttorien välillä. Ratkaisu tulee koostumaan useasta rinnakkaisesta DFS-replikaattoriratkaisusta, joita voidaan kuitenkin hallinnoida keskitetysti DFS-Management konsolin kautta.

## 4.1 Verkkoinfrastruktuuri

Yrityksen sisällä verkolla on tärkeä osa tiedonkulun riipeydessä. Erilaiset konfiguroinnit vaikuttavat suuresti liikenteen sujuvuuteen yrityksen sisällä. Moniporttisesta gigabitin kytkimistä ei ole suurta hyötyä, jos kytkinten väliset linkit hoidetaan myös gigabittisillä porteilla. Pääkytkimenä tulisi olla myös yrityksen monipuolisin kytkin, sillä kytkinten kyvyissä käsitellä dataliikennettä on suuria eroja. Näin ollen yrityksessä päädyttiin Distributed File Systemin käyttöönoton yhteydessä myös samalla päivittämään Helsingin toimiston verkkoinfrastruktuuria. Kaikki yrityksen verkossa olevat hallittavat kytkimet ovat HP:n Procurve -kytkimiä.

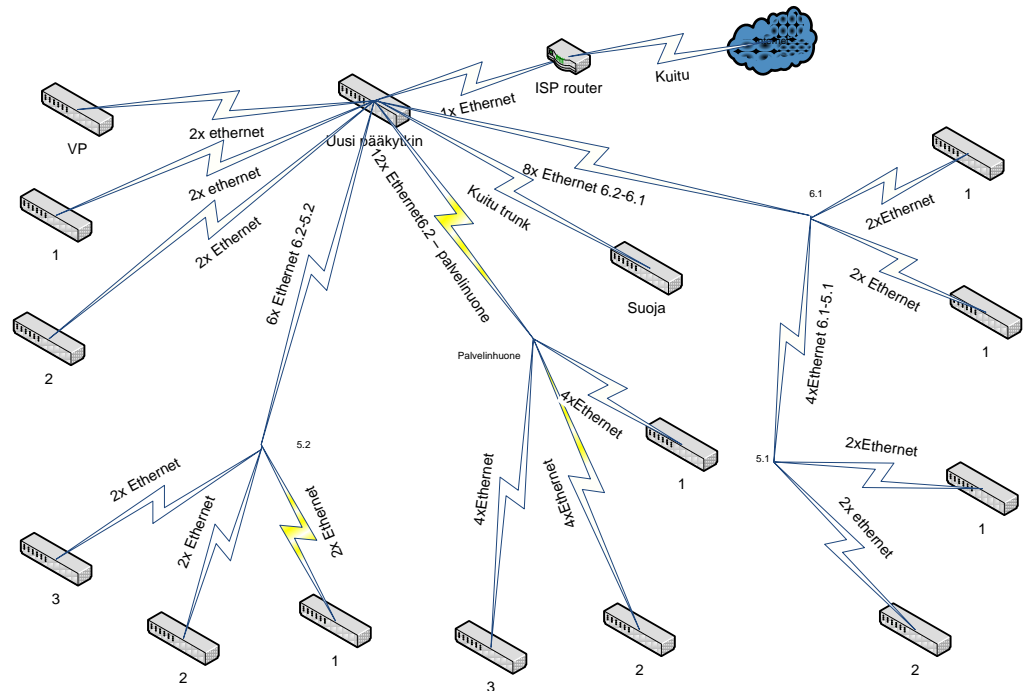
Vanha Helsingin verkkorakenne oli sekoitus eri-ikäisiä kytkimiä, joista vanhimmassa, joka oli vielä projektin alussa käytössä, oli vain kymmenen megabitin portit. Rakenteeltaan verkko oli lähinnä hajanainen puu, jonka keskipisteenä oli vanha pääkytkin.



Kuva 7. Yrityksen vanha verkkorakenne

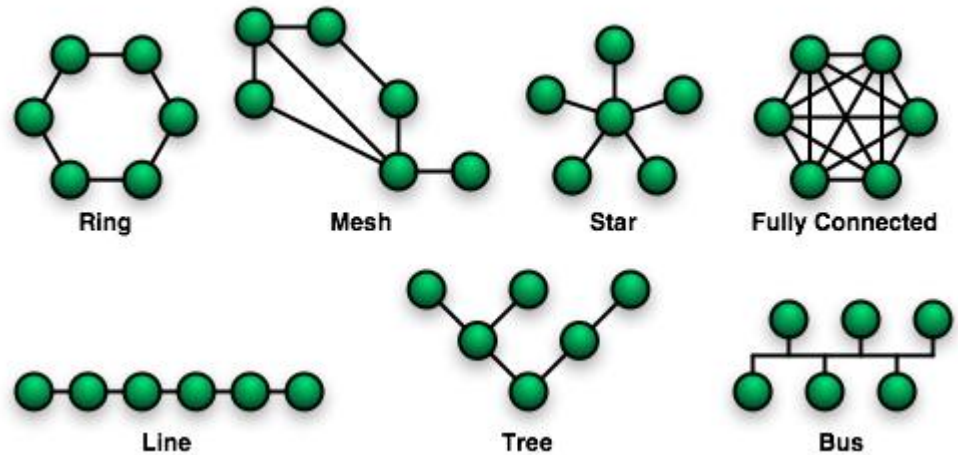


Laitteistoa uusittiin lisäämällä jokaiseen kytkinkaappiin uusi 24 porttinen gigainen kytkin, palvelinhuoneeseen kolme uutta 24 porttista gigaista kytkintä ja lisäksi pääkytkinkaappiin uusi 48 porttinen kytkin uudeksi pääkytkimeksi. Tästä pääkytkimestä muodostettiin myös uusi kuitulinkki toisessa tilassa olevaan varmuuskopiointiratkaisuun.



Kuva 8. Yrityksen uusi verkkorakenne

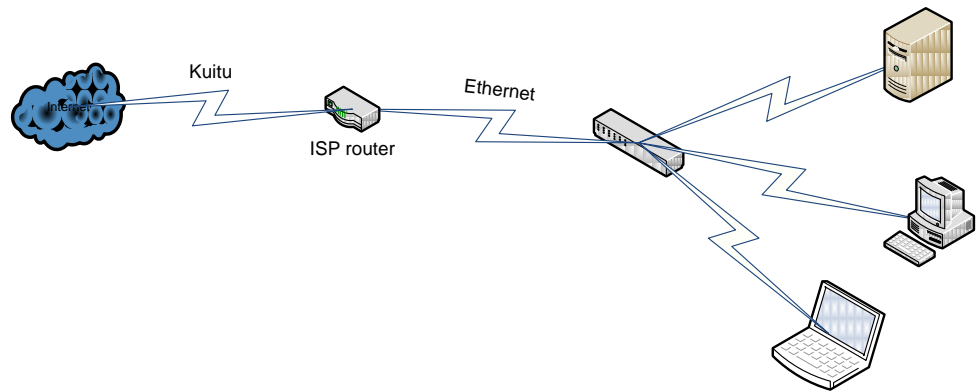
Kytöntien väliset linkit muutettiin vanhoista yksittäisistä ethernet linkeistä 2-4 ethernet kaapelin trunkeiksi. Kaikki muut kytkimet kytkettiin näillä trunkeilla suoraan pääkytkimeen. Rakenteessa käytettiin siis niin sanottua tähtimallia. Tämän rakenteen avulla kyetään hyödyntämään paremmin kytkinten yksittäisten porttien kapasiteettia ja yrityksen sisäisessä liikenteessä säästetään pitkiltä välimatkoilta (hop-count).



Kuva 9. Verkkotopologioita

Trunk on Ciscon käyttämä termi ratkaisusta, jossa kahta tai useampaa fyysistä linkkiä käytetään tiedonsiirtämiseen kahden laitteen välillä. Ratkaisussa kyetään hyödyntämään linkkien yhteenlaskettua kapasiteettia tiedonsiirtoon. Kytöinten väliset trunkit muodostettiin käyttäen LACP:tä (Link Aggregation Control Protocol) joka on IEEE (Institute of Electrical and Electronics Engineers) määritelmä eli standardi, jonka tulisi toimia myös eri valmistajien kytöinten välillä. Ratkaisuun päädyttiin sen yleisyyden ja helppouden takia. Uudessa verkkoratkaisussa muodostettiin lyhyessä ajassa 15 trunkkia, joten asennuksen helppous ja nopeus olivat merkittävä tekijä. Trunkkeihin määritettiin myös "Flow control" eli liikenteen hallinta ja multicast-pakettien suodatus parantamaan verkon toimivuutta suurien liikennemäärien aikana. Haluttiin myös välttää "Broadcast storm" -tapauksia, joita voi ilmetä virhekytkentöjen seurauksena, joten kytkimissä oli kytketty päälle "spanning tree" ja "loop guard" -ratkaisu.

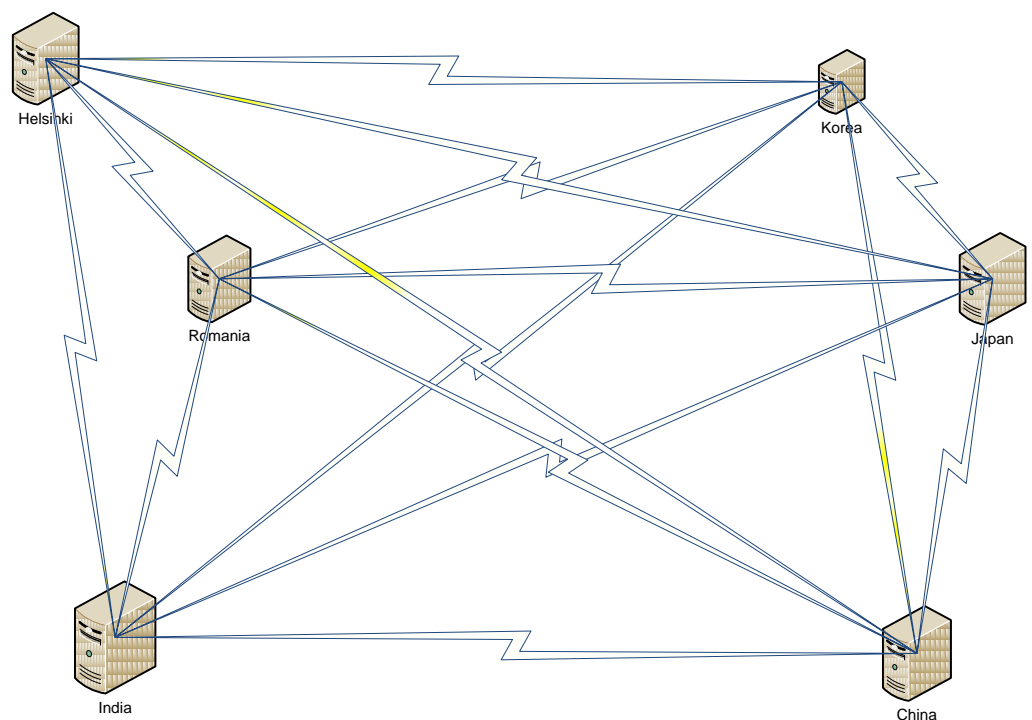
Etäkonttoreiden verkkorakenteisiin ei tehty muutoksia niiden yksinkertaisuuksien takia. Etäkonttorit koostuvat yleisesti yhdessä HP:N 24 porttisesta kytkimestä, VPN-laitteesta, muutamasta palvelimesta ja työasemasta.



Kuva 10. Etäkonttorin perusrakenne

## 4.2 VPN

Jokaisen yrityksen konttorin välille on muodostettu VPN-tunneli jotta kaikilla konttoreilla on pääsy Helsingissä oleville palvelimille ja jotta etäkonttoreiden kommunikoidessa keskenään kaiken liikenteen ei tarvitse kulkea Helsingin pääkonttorin kautta.



Kuva 11. Konttorien väliset VPN tunnelit

VPN-tunneli eli Virtual Private Network on linkki jonka sisällä eri toimipisteiden välistä liikennettä voidaan kuljettaa WAN-verkon (Wide area network) läpi suojattuna ja salattuna siten, että se pystytään purkamaan vain linkin

toisessa päässä oikeilla salausavaimilla. Jokainen VPN-linkin läpi kulkeva paketti paketoitetaan (encapsulate) VPN-kehyksiin, jotka poistetaan linkin toisessa päässä ja paketti muuttuu takaisen normaaliverkossa luettavaan muotoon. Käyttäjälle VPN-käyttö näkyy pienentyneenä linkkinopeutena kasvaneen pakettikoon takia. Linkki voidaan muodostaa joko "gateway to gateway" tai "gateway to host". Gateway to host -mallia käytetään yleisesti, kun halutaan muodostaa linkki ulkoverkossa olevalta koneelta yrityksen verkkoon ja gateway to gateway, kun halutaan muodostaa pysyvämpi linkki eri toimipisteiden välille. Liikenteen suojaamiseen on olemassa useita erilaisia salausvaihtoehtoja, joista yleisimpiä ovat:

- IPsec (Internet Protocol Security) -protokolla: Standardipohjainen kokonaisuus TCP/IP pohjaisia tietoliikenneprotokollia Internet-yhteyksien turvaamiseen. Pakollinen osa IPv6 mutta yleisesti käytössä myös IPv4-verkoissa. Ei vaadi IPsec-tukea välissä olevilta reitittimiltä.
- TLS (Transport Layer Security) -protokolla: Yleisesti käytössä oleva protokolla jolla suojataan koneen ja WWW-sivun välistä liikennettä, HTTPS.
- SSH (Secure Shell) -protokolla: Käytetään salaamaan päätepohjaisia (console) etäyhteyksiä. Mahdollistaa myös laajemman VPN-käytön mutta sisältää rajoittuneet tunnistautumisominaisuudet.

Yrityksen verkoissa käytetään IPsec-protokollaa, joka koostuu useista avoimista standardeista. VPN:n avulla kaikki yrityksen etäkonttorit kuuluvat samaan domainiin ja eri konttorien domain controllerit kykenevät replikoitumaan näiden linkkien yli. Hallinnollisesti palvelinten sijainnilla ei ole väliä vaan kaikkia laitteita pystytään hallitsemaan Helsingistä. Kaikki palvelimet sijaitsevat myös samassa sisäisessä /16 aliverkossa. Linkit muodostetaan konttoreissa olevien Ciscon gateway laitteiden välille. Helsingissä on käytössä raskaampi Cisco ASA 5510 ja etäkonttoreihin on otettu käyttöön kevyempi Cisco ASA 5505, jonka kautta voidaan luoda maksimissaan 30 yhtäaikaista linkkiä eri IP-osoitteita käyttäen.

Näitä linkkejä hyödynnetään tiedostojakojen lisäksi muun muassa versionhallinnassa ja videoneuvotteluyhteyksissä. Linkkien nopeus vaihtelee suuresti, mutta latenssit pysyvät keskimäärin samoina. Näitä arvoja tarkkailemalla saadaan selville optimaaliset polut konttorien välisille tiedostosynkronaatioille.

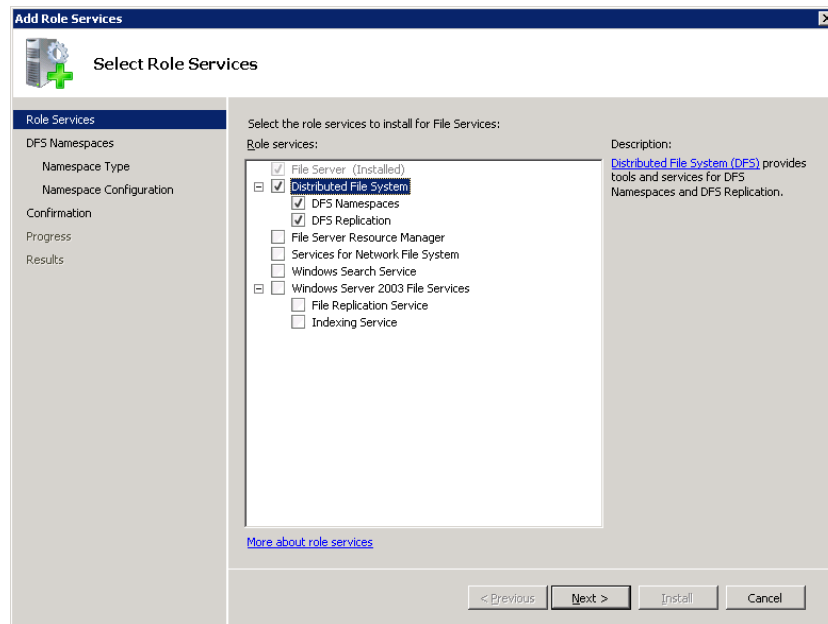
Taulukko 3. Konttorien välisiä latensseja

	Helsinki	Romania	Intia	Korea	Japani	Kiina
Helsinki	X	67ms	319ms	380ms	314ms	488ms
Romania	67ms	X	-	-	320ms	480ms
Intia	319ms	-	X	-	300ms	470ms
Korea	380ms	-	-	X	200ms	95ms
Japani	314ms	320ms	300ms	200ms	X	100ms
Kiina	488ms	480ms	470ms	95ms	100ms	X

Latenssi on aika, joka paketilta kuluu matkata lähettäjältä vastaanottajalle ja takaisin. Latenssin mittaamiseen käytetään yleensä kaikissa käyttöjärjestelmissä olevaa ping-ohjelmaa, jolla lähetetään paketteja määritetylle vastaanottajalle ja pakettien toimitusajoista lasketan keskimääräinen latenssiarvo. Yrityksen sisäisissä verkoissa latenssin tulisi olla maksimissaan parin millisekunnin luokkaa. Normaaleissa kotiyhteyksissä, kohteesta riippuen, latenssi on parikymmentä millisekuntia. Mannertenvälisissä yhteyksissä päästään satoihin millisekunteihin. Toteutuksen kannalta kaistanleveydellä ei ole paljonkaan merkitystä niin kauan kun yhteys on toiminnassa. Suuri latenssi aiheuttaa kuitenkin ongelmia edestakaisten pakettikyselyiden takia.

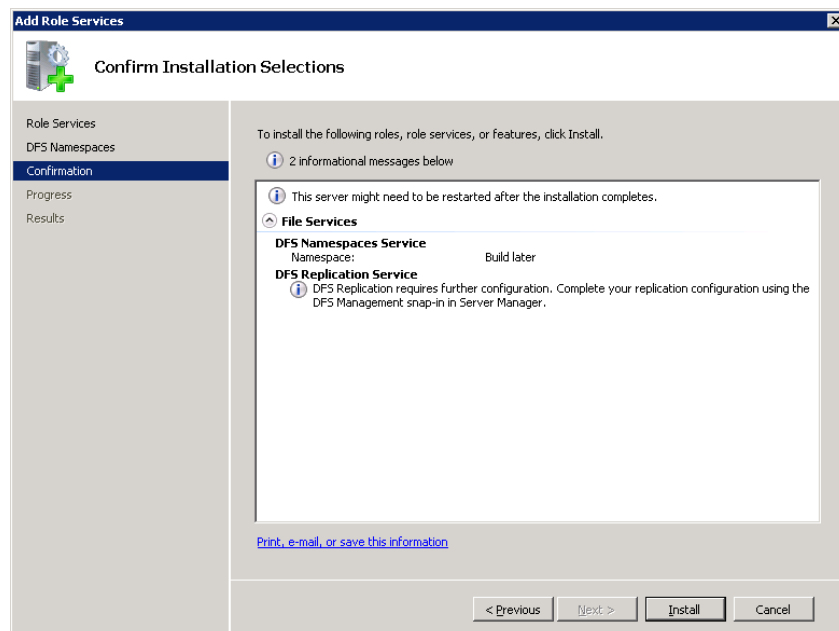
#### 4.3 Point-to-point DFS - Server 2008

Ensimmäiseksi toteutettavaksi yhteydeksi valittiin Point to point -kansion sisällön kopioiminen etäkonttoriin (Helsinki-Kiina). Yhteys on ongelmallinen ja pätkii paljon. Tavoitteena on saada noin 700 megatavun tiedosto siirtymään Helsingistä Kiinaan vähintään kerran vuorokaudessa. Optimaalinen tulos olisi uuden version siirtyminen parin tunnin välein. Yhteys voi toimia hyvin päiviä, mutta ajoittain yhteys pätkii jatkuvasti ja on myös joskus pitkiäkin aikoja täysin poikki. Helsingistä katsottuna Kiinassa on kaikkein heikoimmat yhteydet, joten se valittiin ensimmäiseksi point to point -testipisteeksi.



Kuva 12. Server 2008 Role Services -valikko

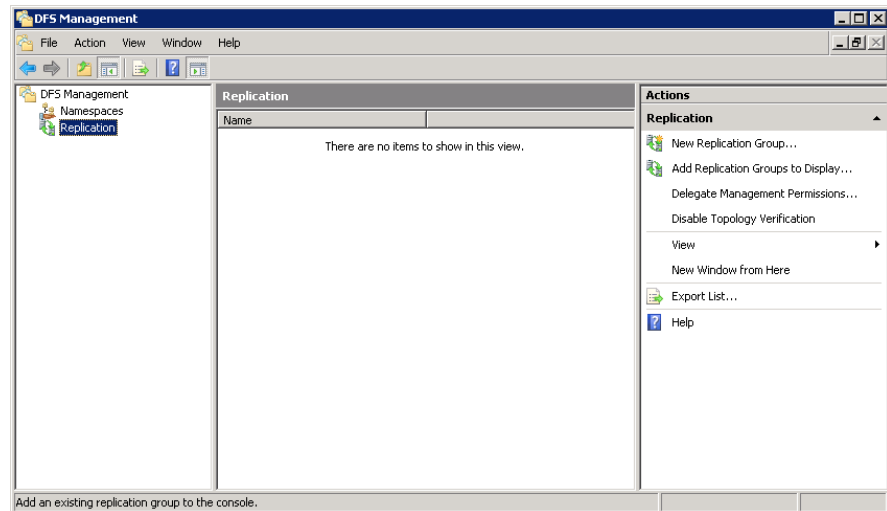
Palvelimena toimii Windows 2008R2 -kone, joten aluksi joudutaan asentamaan Distributed File System -roolin koneelle (osa File Server -ryhmää). Tämä palvelu on oletuksena asennettu Server 2003 ja Server 2003 R2 järjestelmissä, mutta Server 2008 ja Server 2008 R2 se tulee asentaa erikseen.



Kuva 13. DFS-roolin lisääminen

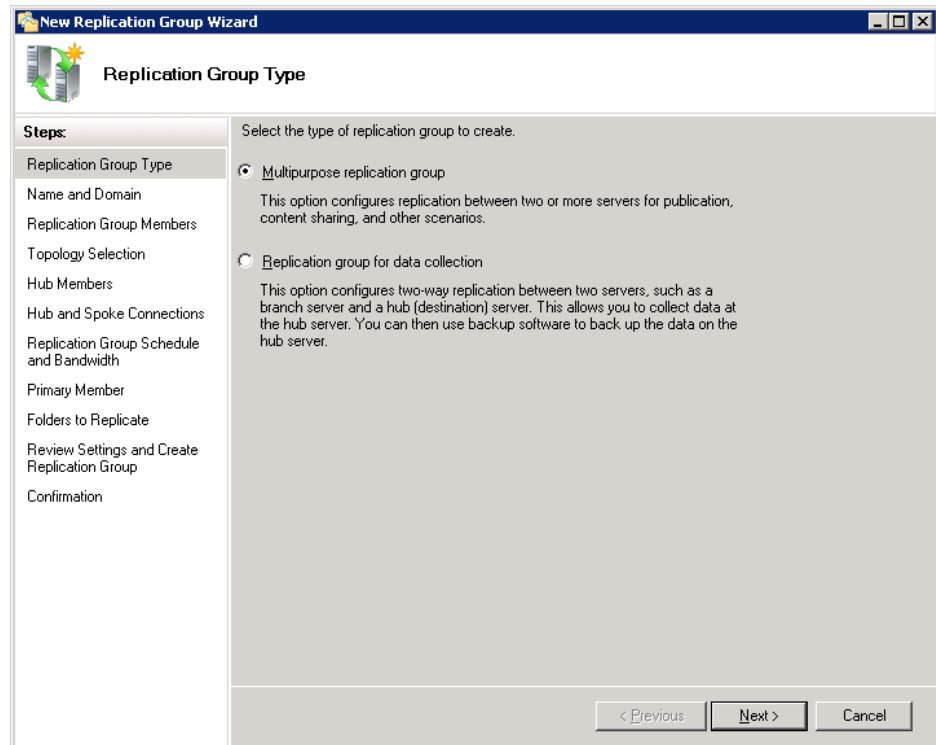
Seuraavaksi tarkistetaan palvelulistauksesta Distributed File System (2003 palvelimissa. DFS Replication ja DFS Namespace 2008 palvelinperheessä) -

palvelun automaattisen käynnistymisen ja tarkistamme että palomuurista on portti 445 auki (Windows-palomuurissa tämä portti sisältyy File and Print Sharing -pakettiin).



Kuva 14. Server 2008 DFS -hallintaliittymä

DFS management -ikkunassa näkyvät kaikki olemassa olevat DFS-nimiavaruudet ja konfiguroidut replikaatiot. Kaikkien domainissa olevien DFS jakojen toimintoja voidaan hallita keskitetysti tästä liittymästä. Olemassa olevia DFS-ryhmiä lisätään "Add replication groups to display" -velhon kautta. Velholla löytyvät tietysti vain domainiin tai luottamussuhteiden kautta muissa domaineissa olevat DFS-nimiavaruudet ja jaot. Hallinta voidaan delegoida ja konfiguraatiot viedä (export) varmuuskopiointia varten.



Kuva 15. Replikaatioryhmän luonti

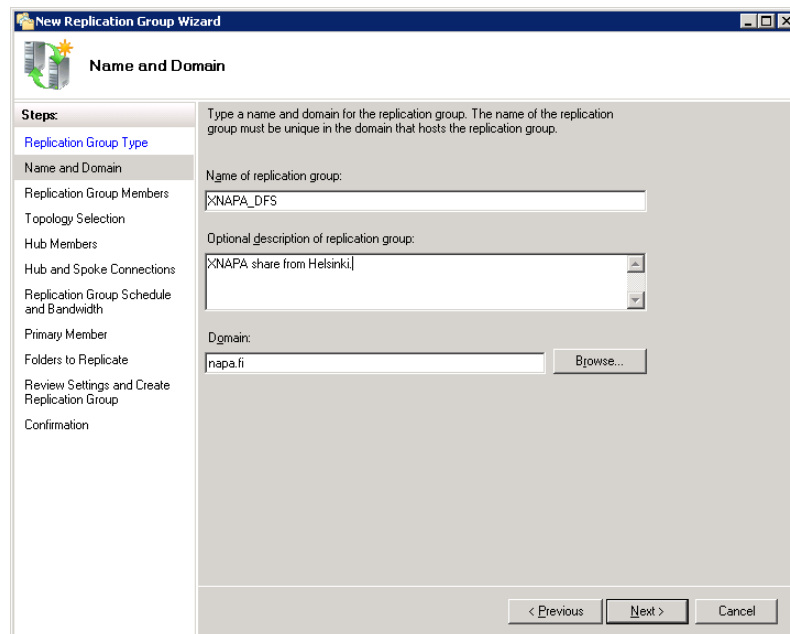
Uusi replikaatioryhmä luodaan käyttäen hallintakonsolin New Replication Group -velhoa, jossa määritellään kohta kohdalta replikaatioryhmän ominaisuudet. Tämä eroaa huomattavasti 2003R2 hallinnasta, joka on paljon rajoituneempi. 2008R2 hallintaliittymä onkin suositeltavin tapa määrittellä ja hallita DFS ryhmiä domainissa, jossa on käytössä myös vanhempia palvelinkäyttöjärjestelmiä. Hallinta on takaisinpäin yhteensopiva 2003-palvelinperheen tuotteiden kanssa.

1. Aluksi määritellään replikaatioryhmän tyyppi. Vaihtoehtoja on kaksi. ”Multipurpose replication group”, jolla voidaan määrittellä kahden tai useammanvälistä tiedostojenjulkaisua, sisällönjakoa ja muita skenaarioita varten. Toisena vaihtoehtona ”Replication group for Data Collecting”, joka on rajoittunut kahdenväliseen liikenteeseen. Tätä vaihtoehtoa suositellaan erityisesti etäkonttoreiden varmuuskopiointiin. Kaikki toisen vaihtoehdon ominaisuudet voidaan luoda myös ensimmäisen ryhmän avulla, mutta vaihtoehto kaksi tarjoaa yksinkertaisemman käyttöönoton erikoistuneemmalle ratkaisulle (etävarmuuskopioinnille). Helsinki-Kiina -ratkaisussa vaihtoehto kaksi olisi täysin riittävä, mutta mahdollisten tulevaisuuden laajennuksien takia päädyttiin käyttämään monipuolisempaa vaihtoehtoa eli ”Multipurpose replication Group”. Aluksi toteutusta käy-



tään vain tiedoston jakamiseen Helsingistä Kiinaan, mutta jos ratkaisu todetaan hyvin toimivaksi, laajennetaan sitä jakamaan tiedostoja myös muihin konttoreihin.

2. Replikaatioryhmän nimen tulee olla lyhyt ja ytimekäs, jotta laajemman käyttöönoton yhteydessä se on helppo erottaa muista jaoista. Replikaatioryhmä julkistetaan näkyväksi yrityksen domainissa helpomman yhteydenoton ja hallinnan takia (oikeuksien hallinta ja hallittavuus).

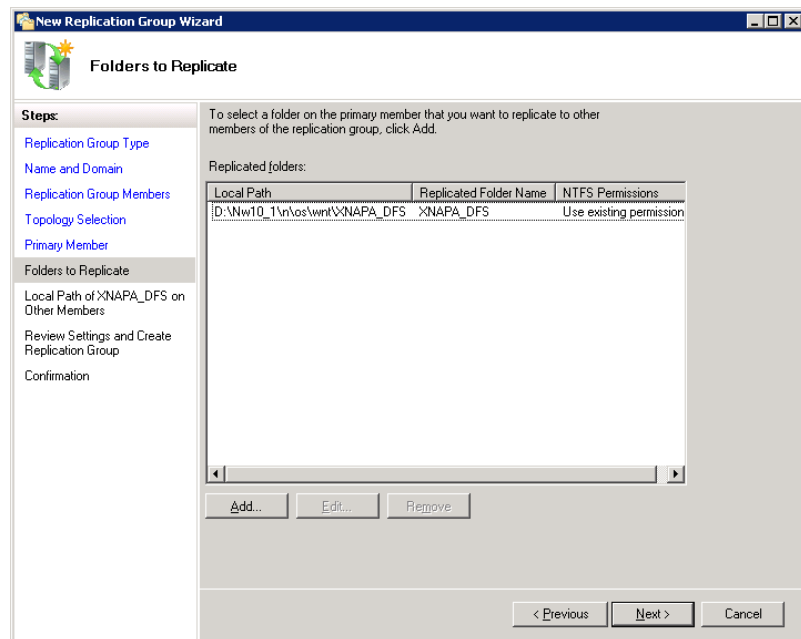


Kuva 16. Uuden replikaatioryhmän tiedot

3. Seuraavaksi lisätään replikaatioryhmään kuuluvat palvelimet. Tässä vaiheessa on hyvä huomata, että kaikissa palvelimissa, jotka aiotaan liittää ryhmään, tulee olla DFS-palvelu käynnissä ja palomuurissa oikeat portit avattuna. Lisäksi tulee huomata, että kaikki alustavat konfiguraatiot tulee tehdä Domain Admin tasoisella käyttäjätunnuksella. Myöhempi hallinta voidaan delegoida alempiarvoisillekin tunnuksille. Aluksi liitämme replikaatioryhmään kahteen palvelimeen Helsinkiin ja Kiinaan.
4. Tässä vaiheessa valitaan topologiamalliksi ”No topology” – vaihtoehdon ja konfiguroidaan se käsin velhon valmistumisen jälkeen.
5. Primary member määritellään palvelimeksi, jolla alkuperäinen data, eli replikoinnin kohde sijaitsee. Velhon jälkeen voidaan määritellä niin sanot-

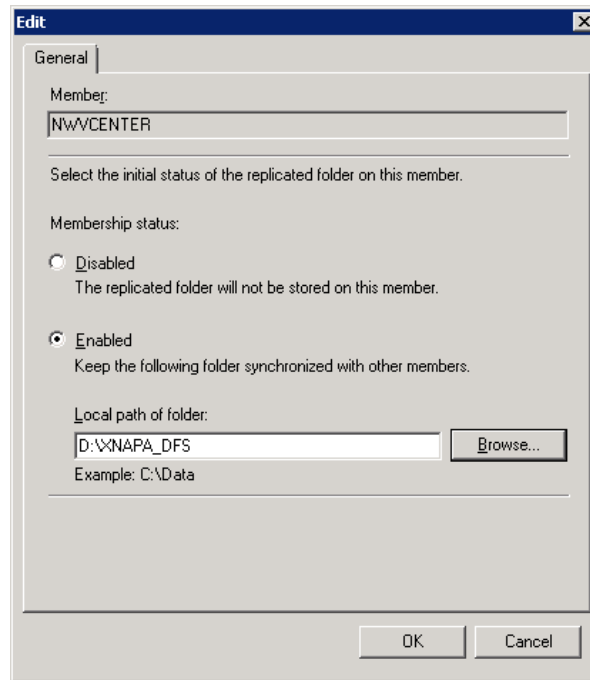
tu staging vaihtoehto, jossa palvelin vertaa eli sijainneissa olevia tiedostoja ja synkronoi vain muutokset.

- "Folders to replicate" kohdassa määritellään replikoitavat kansiot Primary member palvelimella. Aluksi määritellään vain yksi kansio replikoitavaksi. Kansioita pystyy määrittelemään myöhemmin lisää.



Kuva 17. Replikoitavat kansiot

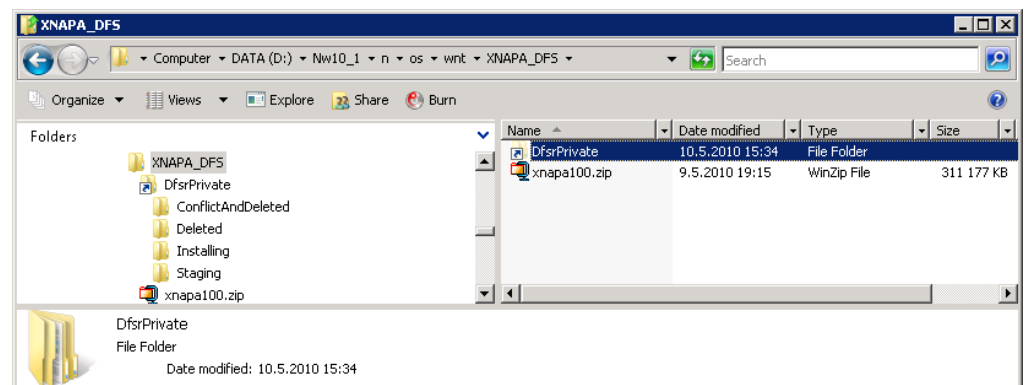
- Local Path valikossa määritellään replikaatioiden sijoituspaikat muissa ryhmän koneissa. Tässä tapauksessa luodaan uusi kansio olemassa olevan kansiorakenteeseen.



Kuva 18. DFS replikaatioryhmän jäsenen ominaisuuksia

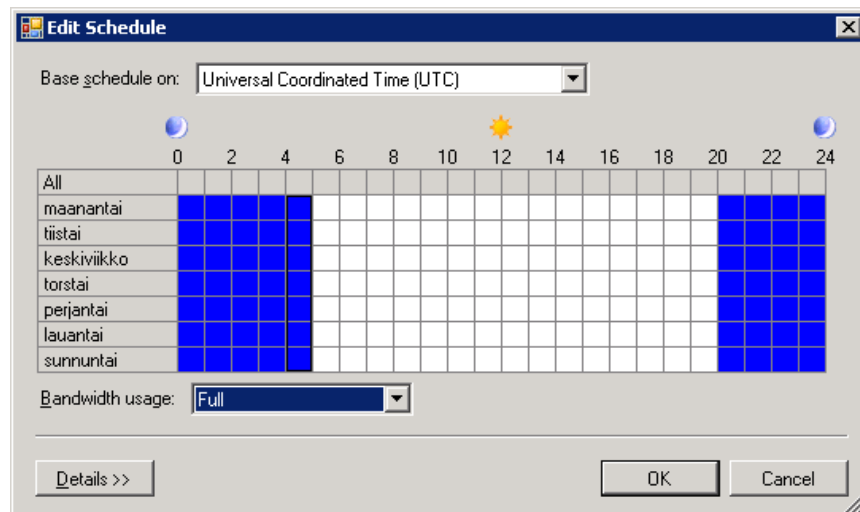
## 8. Lopuksi varmistetaan asetukset ja tarkistetaan toteutus.

Koska valituissa kansioissa voi teoriassa olla, ja todennäköisesti onkin olemassa olevaa dataa, on DFS-palveluun sisäänrakennettu varoitus estämään olemassa olevan datan häviötä. Kaikkiin replikaatioryhmän kohdekansioihin on luotu piilotettu DfsrPrivate kansio, joka sisältää kansion alkuperäisen datan, jos kansio on määritetty vain vastaanottavaksi osapuoleksi ja/tai staging määrittystä ei ole tehty olemassa olevalle datalle.



Kuva 19. Jaettu kansio joka sisältää DfsrPrivate -piilokansion

Koska topology kohdassa määriteltiin "No Topology", joudutaan replikaatioryhmän topologia määrittelemään vielä käsin ennen kuin replikaatio käynnistyy. Tämä tapahtuu DFS management -konsolista käyttämällä "Create New topology" -velhoa. Ryhmäksi valitaan "Full Mesh" vaikka sillä ei kahden koneen välisessä liikenteessä ole suurta merkitystä. Full Mesh on hyvä valinta monisuuntaiseen liikenteeseen kymmenen tai vähemmän koneen replikaatioryhmässä. Replikaatioille voidaan määrittää tietty rajattu kaista (vaihtoehdot ovat välillä 16 Kbps ja 64 Mbps ja lisäksi Full), jota protokolla voi käyttää jotta replikaatio ei hidastaisi muuta liikennettä kohtuuttomasti suurien tietomäärien ollessa kyseessä. Lisäksi voidaan määrittellä tietyt ajat joina palvelimet replikoivat tietoja keskenään. Eri kellonajoille voidaan myös määrittää eri kaistanleveyksiä. Helsinki-Kiina -ratkaisussa määritellään replikaatiot tapahtumaan aluksi vain öisin linkin hitauden takia.



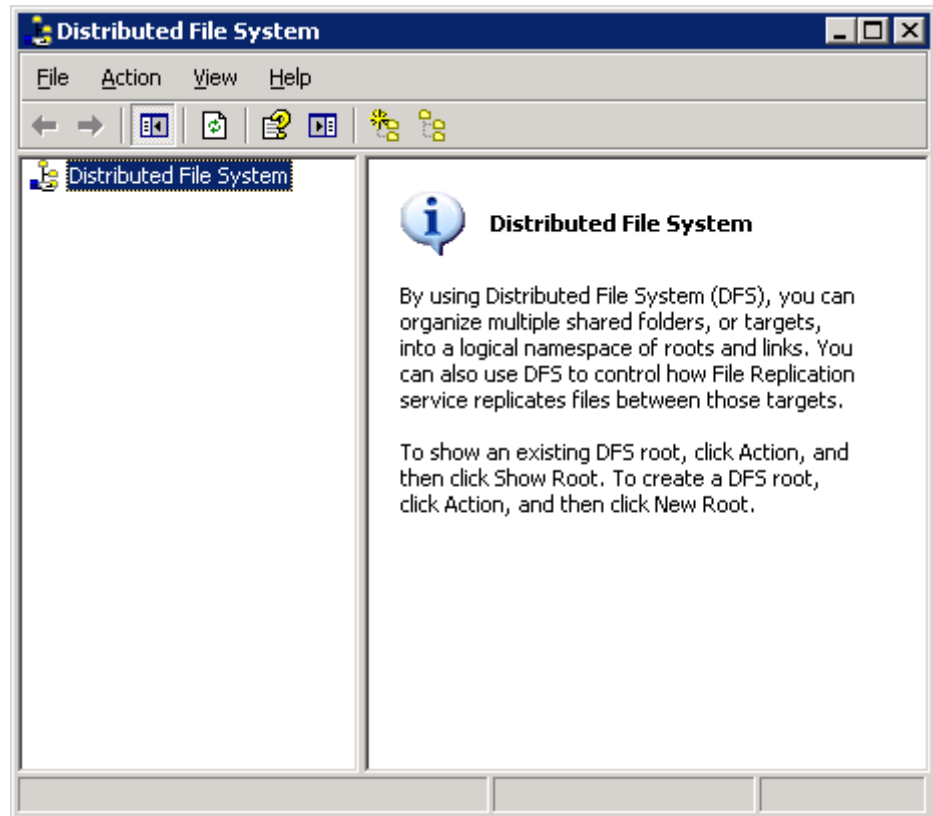
Kuva 20. DFS-replikoinnin ajastus

Näiden konfiguraatioiden lisäksi tulisi määrittellä DFS Namespace, mutta se ei onnistunut- koska yrityksen Active Directory schema on Server 2003 -pohjainen ja sen päivitys ei ole mahdollista tällä hetkellä. Uudempaan versioon aiotaan kuitenkin päivittää heti sopivan tilaisuuden ilmetessä.

#### 4.4 Point-to-Point - Server 2003R2

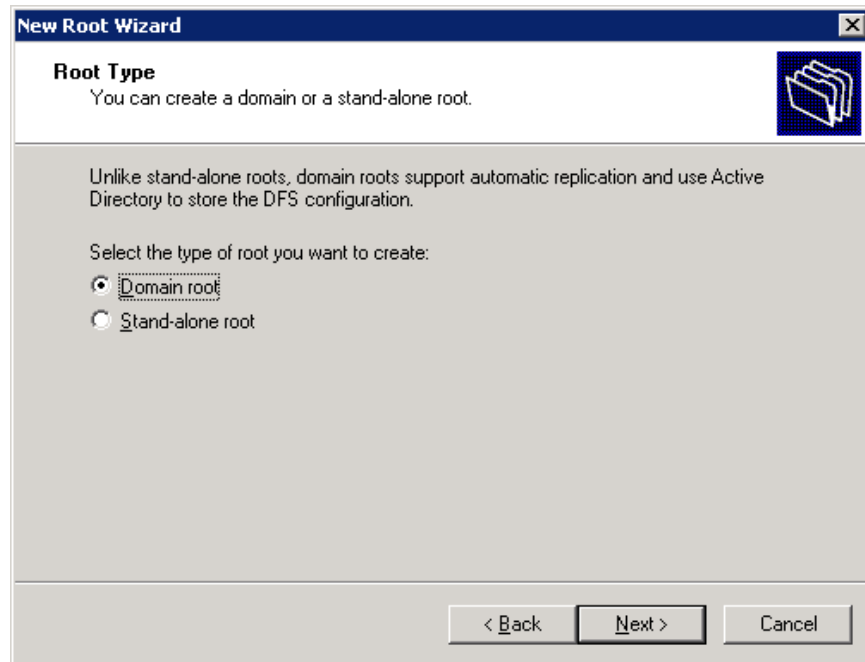
Kiinan ja muiden etäkonttoreiden (Korea, Japani) sekä Helsingin välillä muodostetaan useita point to point -tiedostonsiirtoja. Kaikissa etäkonttoreissa on näillä palvelimilla käytössä Server 2003R2 joten ryhmien muodostamiseen joudutaan käyttämään vanhempaa DFS hallintaliittymää. Tämä johtuu

yrittäjän Active Directory Scheman tasosta, joka on vain Server 2003:n tasolla.



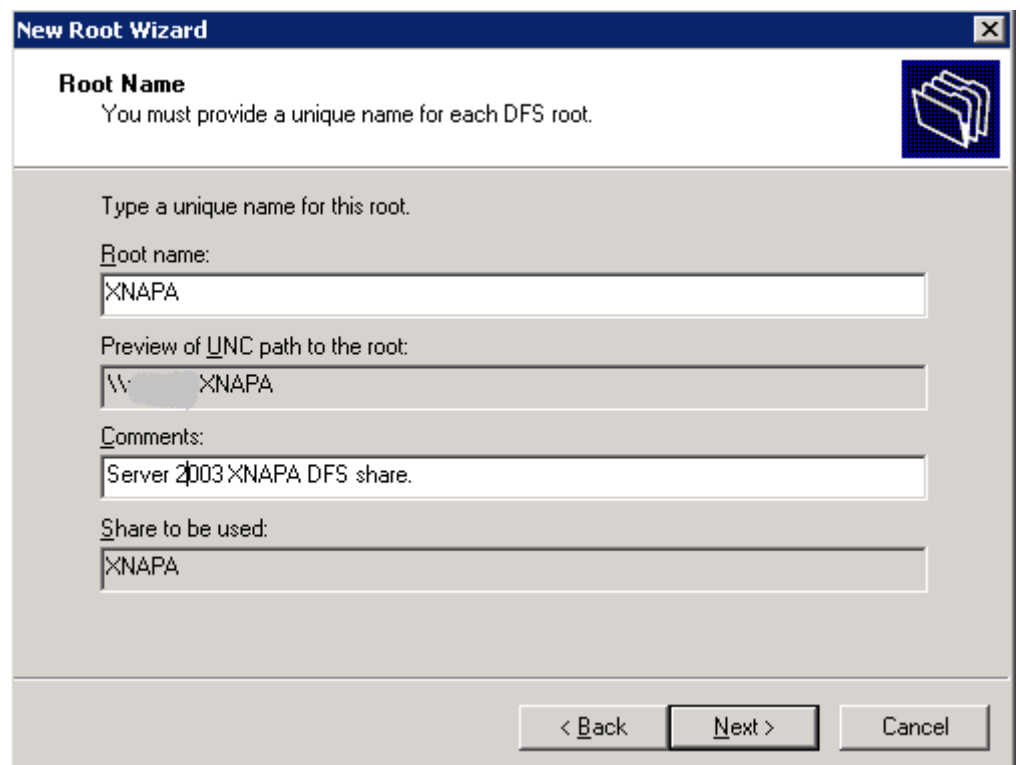
Kuva 21. Server 2003 DFS -hallintaliittymä

Aluksi luodaan uusi juuri-objekti (Root). Siihen lisätään halutut jaot. Juuresta luodaan Domain root, domainissa julkaistava "tietue", jota voidaan etsiä ja hallita domainin kautta. Se varastoi asetukset Active Directoryyn ja hallinnointi onnistuu miltä tahansa yhteensopivalta domain-koneelta. Ajastettu ja/tai automaattinen replikointi toimii vain Domainiin liitetyissä DFS-replikaatioissa.



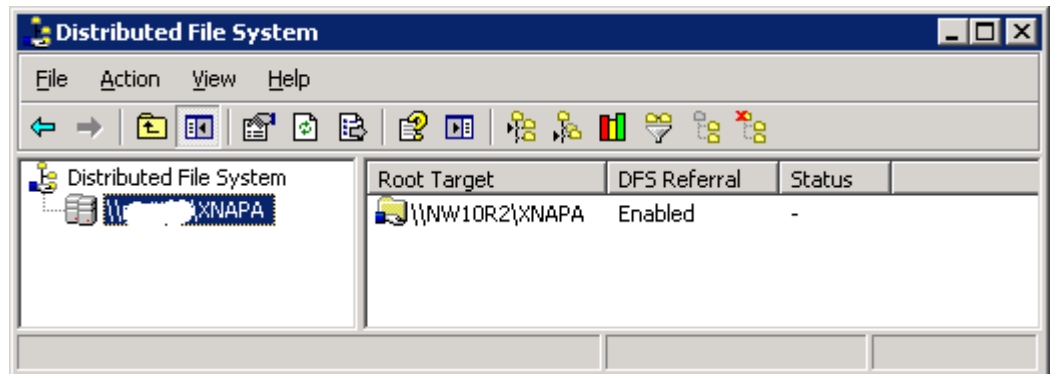
Kuva 22. DFS-juuren tyypin määrittely

Domainin valinnan yhteydessä voidaan liittää jako luottamussuhteiden yli useaan domainiin. Juuren isäntäpalvelimeksi valitaan Helsingissä sijaitseva tiedostonjakopalvelin. Nimeksi valitaan lyhyt ja kuvaava nimi hallinnan selkeyden ja oikeuden resurssien löytämisen helpottamiseksi.



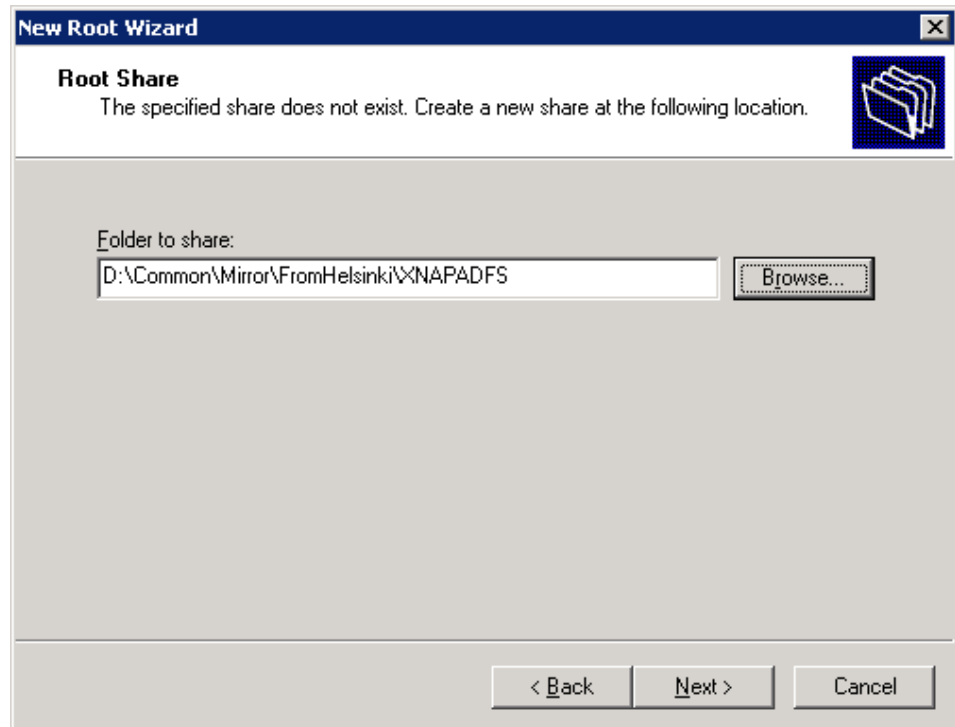
Kuva 23. Juuren tiedot

Jakokansioksi määritellään juuripalvelimella olemassa oleva kansio, joka sisältää Aasian toimipisteille tietyin väliajoin siirrettävät tiedostot.



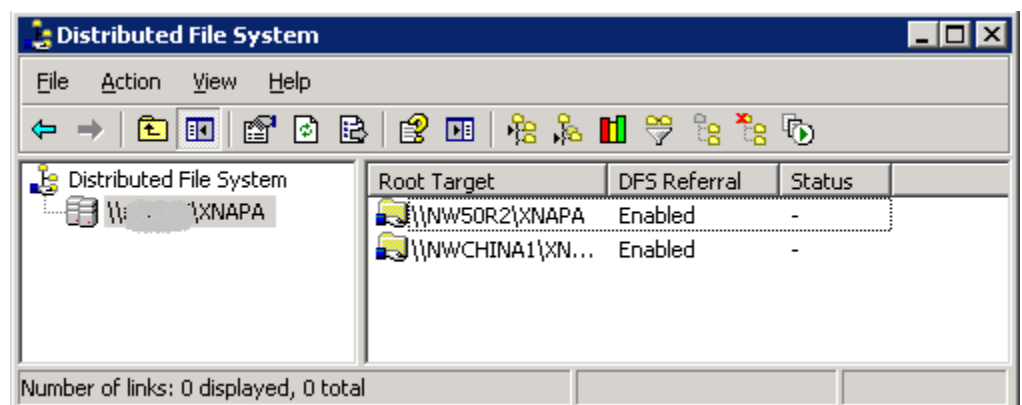
Kuva 24. Uusi juuri

Nyt on luotu palvelimen ensimmäinen DFS-juuri, jossa on jäsenenä aluksi vain luontijuuri. Toisin kuin Server 2008 pohjaisissa ratkaisuisa, joissa jo luontivaiheessa velholla määritellään vähintään kaksi osapuolta, 2003 ympäristössä luodaan ensin juuri jossa on vain yksi sijainti, isäntäkoneella. Juuren kohteita (Root Target) voidaan lisätä vasta kun juuri on menestyksekkäästi luotu. Uusia juuria lisätään käyttäen "New root" -velhoa ja ne liitetään olemassa oleviin jakoihin, tässä tapauksessa XNAPA jakoon. Tässä vaiheessa liitetään Kiinan palvelin ryhmään ja määritellään DFS-jaon kohdesijainti.



Kuva 25. Juuren lisääminen olemassa olevaan ryhmään

Kohteeksi luodaan olemassa olevassa verkkojoassa oleva kansio, jonne luodaan tätä jakoa varten uusi kansio, jotta ei sotketa olemassa olevia tiedostonsiirtoja. Kun uusi jako on lisätty juureen päästään tilanteeseen, jossa DFS-ryhmään kuuluu kaksi eri koneilla sijaitsevaa kansiota. Seuraavaksi tulee määrittellä millaisen politiikan ja topologian mukaan kansioiden välinen replikointi tapahtuu.

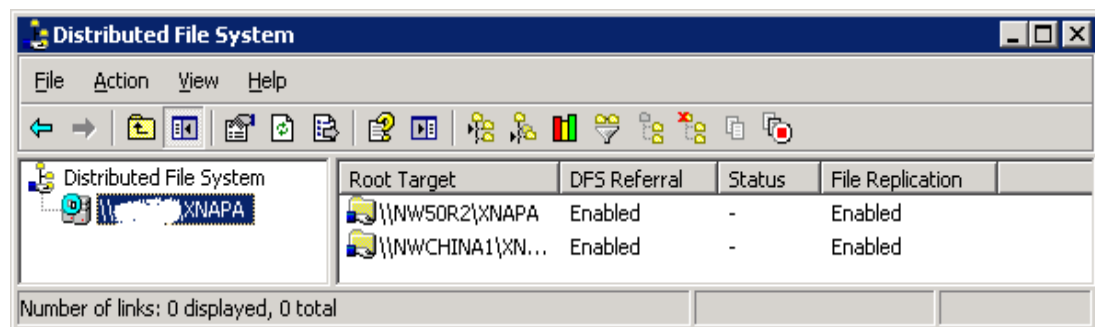


Kuva 26. Ryhmä jossa on kaksi juurta

"Configure replication" -velholla määritellään miten replikoidaan ja minkälaisista topologiaa käytetään. Aluksi määritellään "Initial master" eli kansio, jossa on jaettavaa dataa muille kohteille. Tässä vaiheessa on myös mahdollista



määrittellä niin sanotut staging optiot, joiden avulla voidaan säilyttää muissa sijainneissa olevat datat ja jakaa niitä ristiin. Initial master on tässä tapauksessa Helsingin palvelin. Topologiassa määritellään tässä vaiheessa "Custom", jotta voidaan määrittellä kytkennät käsin. Seuraavaksi määritellään topologia käsin "ryhmän ominaisuus" -valikosta. Valikossa voidaan käsin määrittellä eri pisteiden väliset yhteydet ja niiden välillä kulkeva liikenne ja liikenteen suunta, yksi vai kaksisuuntaisena. Tässä ratkaisussa määriteltiin siten, että liikenne kulkee vain Helsingistä Kiinaan mutta ei päinvastoin. Tämä on riittävää, koska tarkoituksena on siirtää uusia versioita isoista tiedostoista Helsingistä Kiinaan testattavaksi. Kiinasta ei kuitenkaan tarvitse lähettää niitä takaisin. Synkronointiprioriteetiksi ryhmälle laitetaan medium eri normaali. Lisäksi samasta valikosta julkaistaan jako näkymään domainissa, jotta sitä voidaan käyttää kuten normaalia fyysisellä tietokoneella olevaa verkkolevyä. Lisäksi samassa valikossa voidaan määrittellä erilaisia avainsanoja, jotka helpottavat jaon löytämistä isommasta domain-avaruudesta. Ajustuksessa määriteltiin, että replikointi on aktiivisena 17–07 joka päivä Kiinan paikallista aikaa.



Kuva 27. Valmis replikaatioryhmä

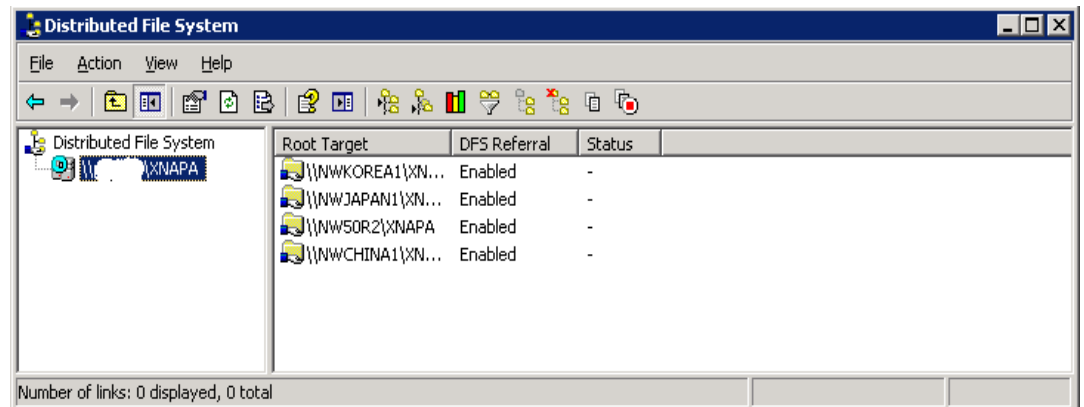
Replication-palvelu on nyt käynnissä ja julkistettu näkymään domainissa, jonka avulla siihen saadaan yhteys kuten mihin tahansa levyjakoon.

#### 4.5 Point-to-multipoint - Server 2003R2

Point-to-Multipoint ratkaisussa laajennetaan aikaisempaa Helsinki-Kiina -linkkiä lisäämällä Kiinan rinnalle Japani ja Korea. Konfiguraatiot hoidetaan Japanin palvelimelta, josta Active Directoryn kautta hallitaan Distributed File System jakoja. Tarkoituksena on muodostaa verkko, jossa jakajana toimii Helsingin konttori ja vastaanottajina kolme Aasian toimistoa. Aasian toimisto-

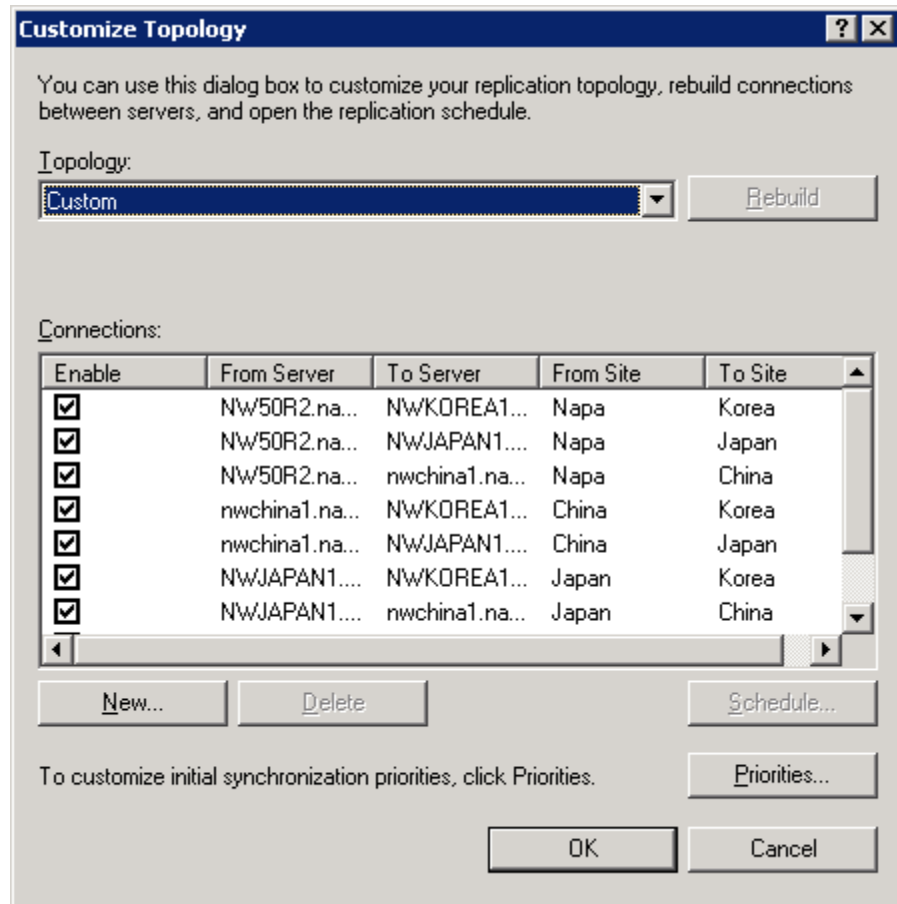
jen välille lisätään lisäksi keskinäiset molempisuuntaiset linkit, jotta hitaammista yhteyksistä kärsivät palvelimet voivat ladata nopeammalta Japanin palvelimelta.

Aluksi muutetaan ajastusta siten että replikointi on päällä ympäri vuorokauden jotta kasvanut datamäärä saadaan varmasti siirrettyä. Korean ja Japanin palvelimet lisätään juuriksi replikaatioryhmään käyttäen ”New Root Target” -velhoa.



Kuva 28. Valmis ryhmä jossa neljä jäsentä

Seuraavaksi määritellään juurien välinen synkronointitopologia. Liikenteen tulee siirtyä Helsingistä kaikkiin kolmeen Aasian konttoriin, mutta Aasian konttoreista ei tarvitse siirtää dataa Helsinkiin päin. Lisäksi Aasian konttoreiden tulee kommunikoida keskenään, jotta data siirtyy kaikille konttoreille vaikka Helsingin ja yhden konttorin välinen linkki olisikin poikki. Jotta topologiaa voidaan muuttaa, tulee kaikki ryhmässä kesken olevat synkronoinnit keskeyttää. Configure replication -velholla muodostetaan uudestaan topologia uuden mallin mukaisesti.

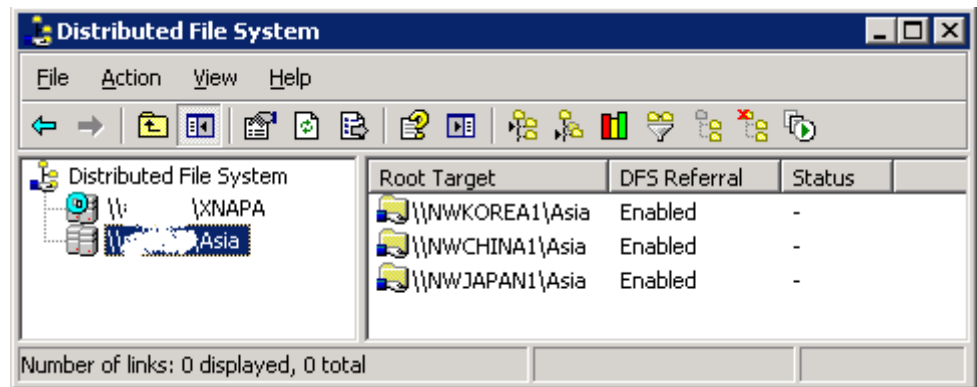


Kuva 29. Topologia-asetuksia

Ratkaisu on nyt uudestaan julkistettu domainiin ja nyt siihen kuuluu 4 jäsentä. Helsinki toimii Master-juurena ja Aasian konttorit saavat datan Helsingistä. Aasian konttorit synkronoivat myös dataa keskenään, mutta mitään dataa ei siirretä tässä jaossa Helsinkiin päin.

#### 4.6 Mesh topologia - Server 2003R2

Mesh ratkaisussa kaikki osapuolet jakavat tietoa tasapuolisesti, liikenne liikkuu kaikkien pisteiden välillä. Tässä ratkaisussa luodaan Aasian konttoreille oma yhteinen replikaatioryhmä, johon kaikista kolmesta konttorista on pääsy ja missä tahansa muokatut tiedostot synkronoituvat muille ryhmän jäsenille.



Kuva 30. Kaksi erillistä ryhmää

Heti juurien lisäämisen jälkeen määritellään replikoinnin ominaisuudet ”Configure Replication” -velhoa käyttäen. Tässä toteutuksessa ei ole väliä, mikä kohde valitaan ”Initial Master” -kohteeksi, koska kaikki kansiot ovat tyhjiä. Topologiaksi määritellään mesh, jossa kaikki osapuolet kommunikoivat keskenään.

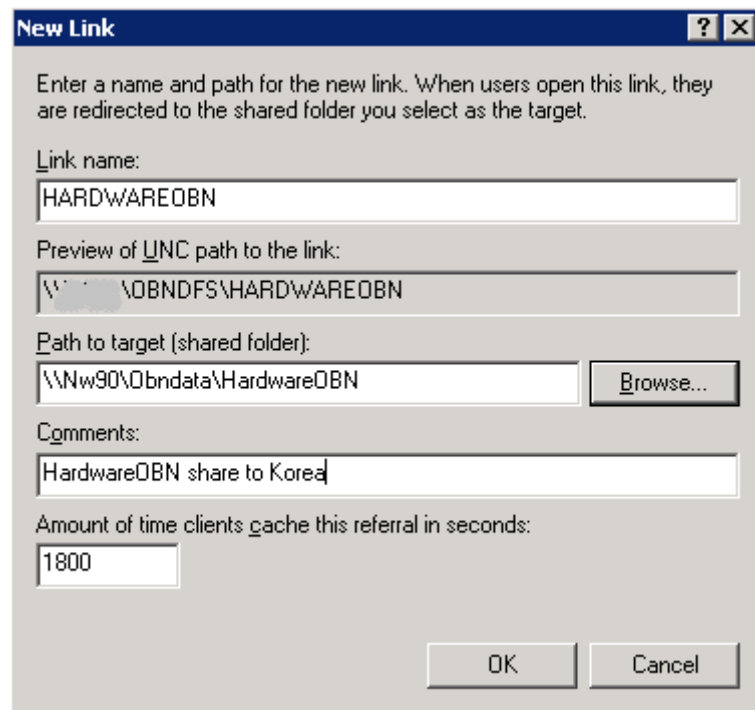
#### 4.7 Linkkitoteutus - Server 2003R2

Tässä ratkaisussa jaetaan suuri määrä dataa, noin 50 GB, Helsingistä Koreaan. Tiedon tulee synkronoitua molempiin suuntiin. Alkuperäisen synkronoinnin jälkeen kopiointimäärä on riippuvainen muuttuneiden tiedostojen määrästä ja muutosten suuruudesta. Yöllisen kopiointin määrä Helsingistä Koreaan vaihtelee siis muutamista sadoista megatavuista useisiin gigatavuihin. Läpikäytävien tiedostojen lukumäärä on noin 170 000, joten jo pelkäästään tiedustelut tiedostojen mahdollisista muutoksista kuluttavat paljon aikaa ja kaistaa. Voidaan olettaa, että näillä latensseilla tiedustelut vievät noin 3 sekuntia per tiedostokohtainen kysely, noin 510 000 sekuntia eli 141 tuntia. Käytössä olevaa xcopya voidaan pitää suhteettoman hitaana ratkaisuna ja kopiointia ei saada yön aikana valmiiksi. Lisäksi Koreasta kopioidaan Helsinkiin muutokset noin 20 GB kansiorakenteesta, joka sisältää noin 30 000 tiedostoa. Tavoitteena on siis saada Helsingissä palvelimella O olevista neljästä kansioista muutokset Koreaan ja päinvastoin. Lisäksi Koreassa olevia tiedostoja tulisi synkronoida Helsingissä palvelimelle T. O-palvelimen synkronointi tulisi olla reaaliaikaista alkuperäisen synkronoinnin jälkeen.

Taulukko 4. Dataliikenteen suunta ja määrä

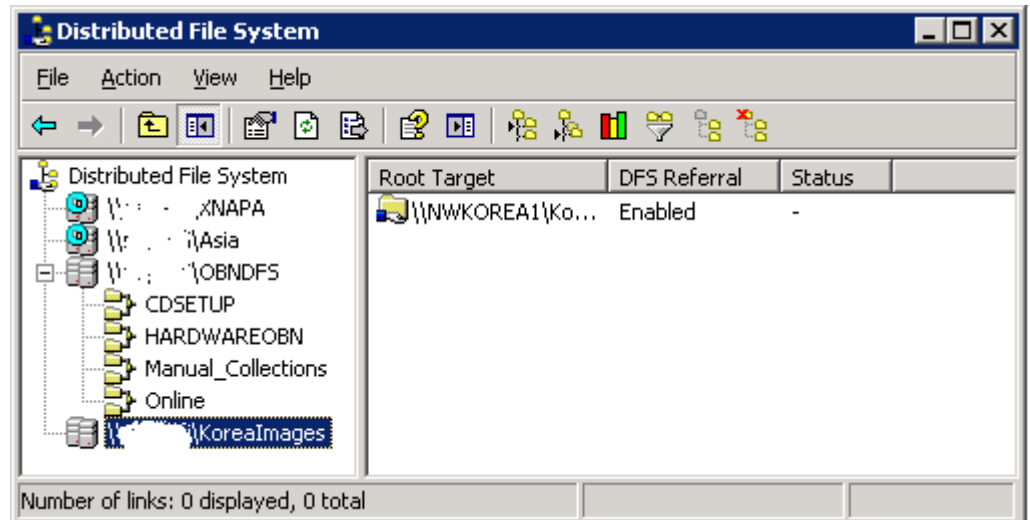
HELSINKI	Koko	Suunta	KOREA
\\O\cdsetup\	28 GB	← →	\\Korea\obn\cdsetup\
\\O\hardwareobn\	11 GB	← →	\\Korea\obn\hardwareobn\
\\O\obn_manual_collections\	10 GB	← →	\\Korea\obn\obn_manual_collections\
\\O\online\	1 GB	← →	\\Korea\obn\online\
\\T\KOREADFS\	20 GB	←	\\Korea\obn\images\

Jokaiselle näistä jaoista muodostetaan oma DFS -linkki jossa määritellään jakokohtaiset asetukset, kuten synkronoinnin ajastus ja jakotopologia, joka on kaikissa näissä tapauksissa point-to-point. Instanssit muodostetaan käyttäen "New Link" -velhoa. Isäntäpalvelimeksi määritetään O-palvelin, koska alkuperäinen data sijaitsee kyseisellä palvelimella. Juurien niminä käytetään selkeyden vuoksi samoja nimiä kuin olemassa olevissa jaoissa, kommenttikentässä selitetään millaiseen käyttöön jako on muodostettu sekaannuksien välttämiseksi. Aluksi luodaan vain uudet linkit mutta niitä ei julkisteta, eikä niihin liitetä vielä kohteita.



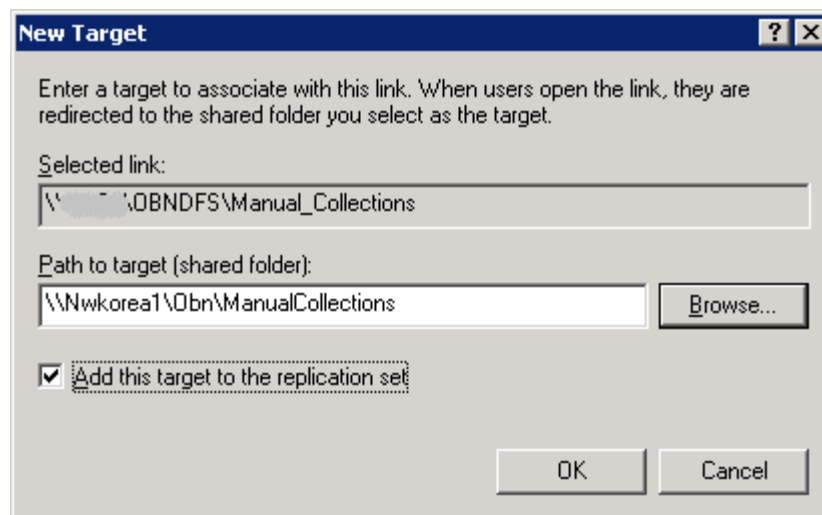
Kuva 31. New Link -velho

Korean suunnasta tulevalla datalla muodostetaan oma juuri, koska kansio käyttäytyy eri tavalla kuin muut.



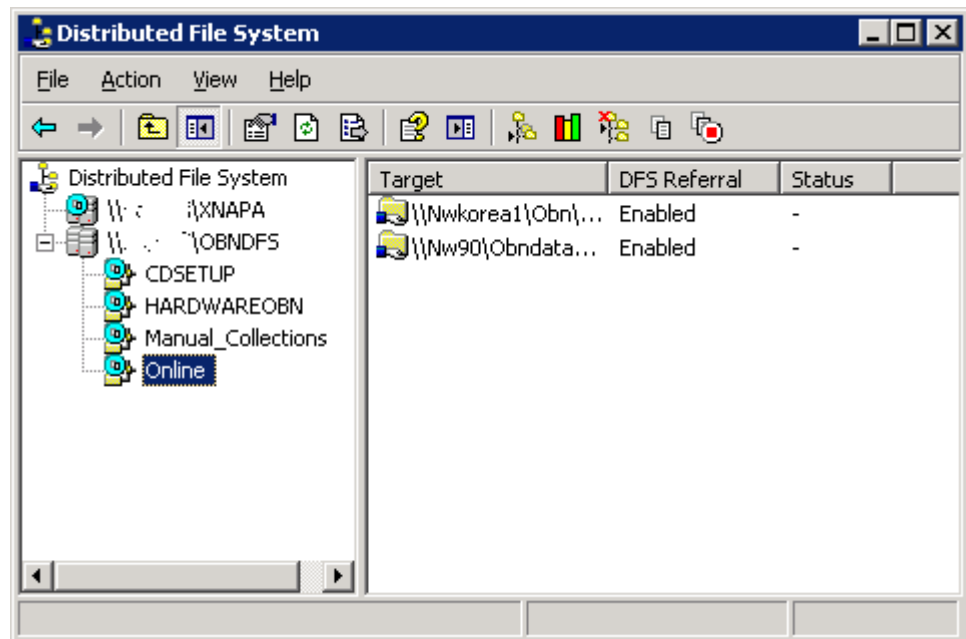
Kuva 32. Uudet linkit ja juuri

Seuraavaksi lisätään jokaiseen linkkiin kansion Korean vastakappale käyttäen "New Target" -velhoa. Velhossa määritellään optio "Add this target to the replication set" mahdollistamaan replikointi kansioden välillä. Lopuksi määritetään että kaikki replikointiasetukset konfiguroidaan vasta, kun kaikki linkit on muodostettu. Linkkien lisääminen on hidas prosessi Korean suuren la-  
tenssin takia.



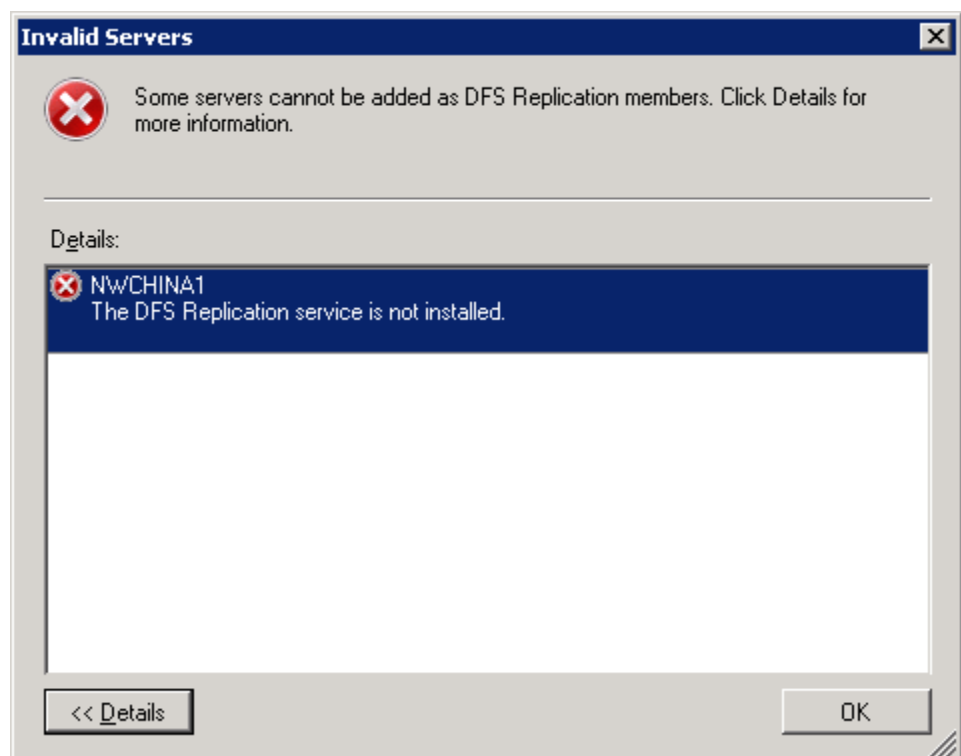
Kuva 33. Linkin lisääminen

Kun kohteet on lisätty konfiguroidaan replikaatiot toimimaan kaksisuuntaisesti neljässä linkkiryhässä.



Kuva 34. Valmiit linkit

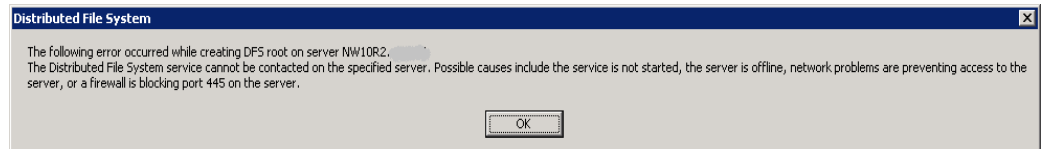
#### 4.8 Toteutuksessa ilmenneitä ongelmia ja virheilmoituksia



Kuva 35. Virhe 2008 hallintaliittymässä

Ongelma ilmeni liitettäessä Kiinan palvelinta replikaatioryhmään. Distributed File System palvelu on asennettu ja palomuri oli kokonaan pois päältä. To-

dennäköisenä syynä on Active Directory:n schema versio. Teoriaa ei kuitenkaan voitu testata scheman tason takia. Yrityksellä on käytössä Server 2003 ympärille rakennettu domain-ympäristö ja vaatimuksena täydellisellä Server 2008 pohjaiselle Distributed File System -ratkaisulle on 2008 palvelinympäristön ympärille rakentuva Active Directory.



*Kuva 36. Yleinen virhe lisättäessä uutta Juurta*

Yllä oleva virheilmoitus esiintyy, kun ollaan hyväksymässä uuden DFS-juuren (root) luontia. 2008 palvelimen ollessa kyseessä ongelma korjautui uudelleenkäynnistämällä DFS Namespace -palvelu kohteena olevasta palvelimesta. Virheilmoitus esiintyi usein käytettäessä Server 2003:n DFS hallintaliittymää. Tämä viittasi useimmiten siihen, että kohdepalvelimella oli unohdettu käynnistää asianmukaiset palvelut. Sama virheilmoitus ilmestyy myös, jos käyttäjällä ei ole tarpeeksi korkeita käyttöoikeuksia, eli jos käyttäjä ei ole Domain Admin tai hänelle ei ole delegoitu DFS-Admin -oikeuksia.

## 5 YHTEENVETO

Tekniikkana Distributed File System -ratkaisun paras puoli on sen ilmaisuus, joka kerää sille varmasti paljon asiakkaita. Käyttöönotto on suhteellisen helppoa kunhan vaivautuu perehtymään Microsoftin TechNetin tarjoamaan materiaaliin. Osittain vanhan laitekannan takia jouduin rajoittamaan toteutuksesta pois kaikkein uusimmat Distributed File Systemin tarjoamat ominaisuudet, mutta ratkaisu näyttää kehittyvän hyvin Microsoftin palvelinperheen mukana, joten odotettavissa on kiinnostavia uusia ratkaisuja. Hallintaominaisuudet ovat vielä Server 2003 -pohjaisessa DFS-ratkaisuissa hiukan rajoittuneet mutta kehityssuunta on oikea, koska Server 2008 -ratkaisussa toteutus on lähes miellyttävä käyttää

Toteutus yrityksessä onnistui hyvin pienten alkukangertelujen jälkeen ja vastaanotto käyttäjien keskuudessa oli positiivinen. Microsoft lupaa esitteissään jopa 60% parannusta tiedostojen siirroissa. Tällaisiin lukuihin ei päästy, mutta etäkonttoreiden tiedostosynkronoinneista on tullut toimintavarmempia ja huomattavasti nopeampia. Tätä työtä varten toteutimme kolme erilaista



ratkaisua, mutta tarkoituksena on tulevaisuudessa laajentaa olemassa olevaa kolmea toteutusta ja luoda myös uusia toteutuksia.

Ratkaisua, varsinkin uudempaan Server 2008R2 pohjautuvaa, voi suositella varauksettomasti kaikille Microsoft AD -ympäristössä työskenteleville, jotka painivat erilaisten tiedonsiirto- ja synkronointiongelmien parissa.

**VIITELUETTELO**

- [1] Domain Name System, Wikipedia [verkkodokumentti, viitattu 5.4.2010] Saatavissa: [http://en.wikipedia.org/wiki/Domain\\_Name\\_System](http://en.wikipedia.org/wiki/Domain_Name_System).
- [2] Albitz, Liu *DNS and BIND (4. painos)*. USA: O'REILLY. 2001.
- [3] Langfeldt, *The Concise Guide to DNS and BIND*. USA: QUE. 2001.
- [4] Wikipedia artikkeli Active Directory:stä [verkkodokumentti, viitattu 11.4.2010] Saatavissa: [http://en.wikipedia.org/wiki/Active\\_directory](http://en.wikipedia.org/wiki/Active_directory).
- [5] Kivimäki, *Active Directory Verkonhallinta*. Helsinki: Edita. 2003.
- [6] Kouti, Seitsonen, *Inside Active Directory - A System Administrator's Guide*. USA: Addison & Wesley. 2002.
- [7] MSDN verkkokirjasto, RDC:stä [verkkodokumentti, viitattu 15.5.2010] Saatavissa: [http://msdn.microsoft.com/en-us/library/8d6431d5-a66e-4faf-904e-67c47b9d6efc\(v=PROT.13\)#rdc](http://msdn.microsoft.com/en-us/library/8d6431d5-a66e-4faf-904e-67c47b9d6efc(v=PROT.13)#rdc).
- [8] Wikipedia artikkeli Distributed File System:istä [verkkodokumentti, viitattu 5.4.2010] Saatavissa: [http://en.wikipedia.org/wiki/Distributed\\_File\\_System\\_\(Microsoft\)](http://en.wikipedia.org/wiki/Distributed_File_System_(Microsoft)).
- [9] Microsoft Technet artikkeli Distributed File System:istä [verkkodokumentti, viitattu 5.4.2010] Saatavissa: [http://technet.microsoft.com/en-us/library/cc782417\(WS.10\).aspx](http://technet.microsoft.com/en-us/library/cc782417(WS.10).aspx).
- [10] Microsoft Technology Center sivusto Distributed File System:istä [verkkodokumentti, viitattu 5.4.2010] Saatavissa: <http://www.microsoft.com/windowsserversystem/dfs/default.aspx>.

- [11] Microsoft Technet artikkeli Distributed File System:istä [verkkodokumentti, viitattu 5.4.2010] Saatavissa: <http://technet.microsoft.com/en-us/library/cc787066%28WS.10%29.aspx>.
- [12] Microsoft Technet Blog Distributed File System:istä [verkkodokumentti, viitattu 5.4.2010] Saatavissa: <http://blogs.technet.com/filecab/archive/tags/DFS+Replication/default.aspx>.

