TAMPEREEN AMMATTIKORKEAKOULU

TUTKINTOTYÖRAPORTTI

VARMUUSKOPIOINTIJÄRJESTELMÄN SUUNNITTELU JA TOTEUTUS

Santtu Putkonen

Tietojenkäsittelyn koulutusohjelma
helmikuu 2007
Työn ohjaaja: Harri Hakonen

TAMPERE 2007

TAMPEREEN
AMMATTIKORKEAKOULU

Tekijä(t)	Santtu Putkonen
Koulutusohjelma(t)	Tietojenkäsittely / Tietoverkkopalvelut
Tutkintotyön nimi **Title in English**	Varmuuskopiointijärjestelmän suunnittelu ja toteutus How to design and implement a backup system.
Työn valmistumis-kuukausi ja -vuosi	02/2007
Työn ohjaaja	Harri Hakonen

Sivumäärä: 37

TIIVISTELMÄ

Tämän tutkintotyön tarkoitus on selittää, mitä tarkoitetaan termillä varmuuskopiointi, mitä vaihtoehtoja sen toteutukselle on tällä hetkellä olemassa ja miten kyseisen kaltainen järjestelmä voitaisiin toteuttaa keskikokoiselle yritykselle. Työssä käydään lävitse varmuuskopioinnin erilaiset tyylit, nauhakiertojen hallintamenetelmät, erilaiset tallennusformaatit sekä luodaan esimerkki ratkaisu varmuuskopiointijärjestelmästä.

Työn pääasiallisina tietolähteinä ovat olleet alan eri julkaisut niin verkossa kuin printtimediassakin, eri valmistajien WWW-sivustot sekä jotkin kirjat. Tämän lisäksi olen soveltanut työhön paljon harjoittelu- ja työelämässäni oppimaani yleistä tietoa eri järjestelmistä ja tekniikoista.

Valmis työ sisältää viimeisimpiä tietoja markkinoilla olevista formaateista sekä tekniikoista sekä yleisiä ja hyväksi havaittuja käytäntöjä varmuuskopiointiin liittyen. Valmiin työn pohjalta on mahdollisuus saada suuntaa antava käsitys varmuuskopiointikentästä ja siellä vallitsevista ratkaisuista sekä niiden mahdollisesta soveltamisesta käytäntöön. Työn lopussa on esimerkki varmuuskopiointijärjestelmän toteutuksesta keskikokoiselle yritykselle, mikä tarjoaa mahdollisuuden nähdä, miten kyseisen kaltainen ratkaisu voitaisiin toteuttaa yritysmaailmassa. Työ tarjoaa varmuuskopiointiin tutustumattomalle helpon tien saada keskitetysti tietoa aihepiiristä sekä tarjoaa jo käyttökokemusta omaavallekin mahdollisuuden palauttaa mieleensä aihepiiriin vaikuttavia tekijöitä ja asioita. Työ voi toimia myös hyvänä suunnanantajana henkilölle/henkilöille, joiden tehtävänä on pohtia varmuuskopiointijärjestelmän uusimista.

Avainsanat Varmuuskopiointi tietojärjestelmät tietojenkäsittely tallennus

1 Johdanto ... 5
1.1 Toimeksiantaja ... 5
1.2 Työn tavoitteet ... 5
1.3 Lähdeaineisto ... 5
2 Varmuuskopiointi ... 6
2.1 Yleistä ... 6
2.2 Varmuuskopioinnin erilaiset tyypit ... 7
2.2.1 Täysi varmistus ... 7
2.2.2 Valikoiva varmistus ... 7
2.2.3 Lisäävä varmistus ... 8
2.3 Nauhakierron Hallinta ... 8
2.3.1 Isoisä–isä–poika–rakenne ... 8
2.3.2 Hanoin torni–rakenne ... 10
2.4 Rakenteiden varioiminen ... 10
3 Varmuuskopiointiin liittyvät laitteistot ja mediat ... 12
3.1 Yleistä ... 12
3.2 Nauha-asemat ... 12
3.3 WORM-tekniikka ... 13
3.4 Nauha-formaatit ... 14
3.4.1 DDS ... 14
3.4.2 DLT ... 15
3.4.3 SDLT ... 15
3.4.4 AIT ... 15
3.4.5 SAIT ... 16
3.4.6. LTO ... 17
4 Varmistusohjelmistot ... 19
4.1 Yleistä ... 19
4.2 Kotikäyttäjän varmuuskopiointiohjelmistot ... 19
4.3 Ammattikäyttöön tarkoitetut varmuuskopiointiohjelmistot ... 20
4.3.1 Hinnoittelu ja vertailu ... 20
4.3.2 Tuotteet ... 21
5 Riskianalyysi ... 22
6 CDP ... 24
6.1 Esittely ... 24
6.2 Ohjelmistot ... 25
6.3 Laitteistot ... 25
6.4 Esimerkki tapauksia CDP:n käytölle ... 25
6.5 Haitat ... 26
7 Varmuuskopiointijärjestelmän toteuttaminen ... 27
7.1 Kolme erityyppistä toteutusta ... 27
7.1.1 Yksinkertainen perustoteutus ... 27
7.1.2 Erillinen tallennuspalvelin yhdistetyllä levy- ja nauhajärjestelmällä ... 27
7.1.3 SAN-järjestelmä ... 28
7.2 Esimerkki yritys ... 28
7.3 Tallennusformaatin valinta ... 28
7.4 Tallennusohjelmisto ... 29
7.5 Laitteisto ... 29
7.6 Nauhakierto ja varmuuskopiointiprosessi ... 30
7.7 Nauhojen säilyttäminen ... 31
7.8 Tulevaisuuden järjestelmämuutokset ... 31
8 Yhteenveto ... 32

8.1 Työn rajaus ja siihen liittyneet ongelmat 32
8.2 Työskentely 32
8.3 Hyödyt 33
8.4 Loppusanat 33
9 Lähteet 35
10 Lyhenteet 36
10.1 Käytettyjen lyhenteiden ja termien selitykset 36

1 Johdanto

1.1 Toimeksiantaja

Päättötyölleni oli aluksi toimeksiantaja, joka oli edellinen työpaikkani. Tarkoituksena oli uusia asiakasyrityksen varmuuskopiointijärjestelmä. Olosuhteiden pakosta projekti kumminkin peruuntui ja työni jäi ilman toimeksiantajaa. Halusin kumminkin tehdä päättötyöni varmuuskopioinnista, koska koin että aihe on mielenkiintoinen ja haastava. Näin ollen tein työni ilman toimeksiantajaa ja työn painopiste siirtyi tutkimaan varmuuskopiointimarkkinoiden tilannetta yleisemmällä näkökulmalla. Tällä näkökulmalla tarkoitan sitä, että työ ensin selvittää markkinoilla olevien medioiden ja tekniikoiden merkityksen varmuuskopioinnissa ja sen jälkeen tarjoaa esimerkin keskikokoisen yrityksen varmuuskopiointiratkaisun uusimisesta.

1.2 Työn tavoitteet

Työn tavoitteena on tarjota katsaus markkinoilla tällä hetkellä oleviin tallennusmedioihin ja ohjelmistoihin sekä valottaa yleisesti varmuuskopiointikentän tilannetta. Työn loppupuolella on lisäksi esimerkkitoteutus kuvitteellisen yrityksen varmuuskopiointijärjestelmän uusimisesta. Tarkoituksena on, että työn luettuaan lukijalla olisi käsitys erilaisten ohjelmien, medioiden ja konfiguraatioiden eroista sekä niiden luomista mahdollisuuksista ja tätä kautta hänellä olisi mahdollista pohjaa valinnoille, jotka tulevat eteen varmuuskopiointijärjestelmää hankittaessa tai uusittaessa. Työssä ei käsitellä erityisemmin eri ohjelmien asentamisprosesseja tai järjestelmien tarkkoja konfigurointeja. Kyseiset asiat ovat yleensä niin ympäristöriippuvaisia.

1.3 Lähdeaineisto

Työssä on käytetty aineistona pääasiallisesti laajaa määrää erinäisiä WWW-sivustoja sekä joitakin alan lehtien artikkeleita ja julkaisuja. Kaikki käytetty materiaali on ollut alun perin englanninkielistä lukuun ottamatta joitakin kotimaisissa alan lehdissä julkaistuja pieniä artikkeleja. Kirjallista materiaalia on käytetty jonkin verran, joskin sen määrä on ollut hyvin rajallista. Kyseisestä aihepiiristä ei ole löydettävissä juurikaan suomenkielistä materiaalia tai mikäli sitä on, ei se ole julkista. Varmuuskopiointijärjestelmät ovat lisäksi tyypillisesti vuosien varrella erinäisistä osista yhteen kasattuja kokonaisuuksia, joten niistä olemassa oleva tieto on lähes pääsääntöisesti ainoastaan kyseisten järjestelmän omaavien organisaatioiden ja niitä ylläpitävien tahojen hallussa.

2 Varmuuskopiointi

2.1 Yleistä

Varmuuskopioinnilla tarkoitetaan prosessia, jossa olemassa olevista tiedostoista otetaan kopio tietojen turvaamista varten. Viimeisen viiden vuoden aikana tallennettavan datan määrässä on tapahtunut pienimuotoinen räjähdys. Yritysten käytössä oleva tiedonmäärä on kasvanut tutkimusten mukaan vuosittain noin 70–100 % 2000-luvun alusta lähtien. Lukema vaihtelee hieman riippuen analyysista. Tämä lukema on monissa paikoissa kuitenkin kyseenalaistettu jopa liian konservatiivisena arviona. Tämä vuosittainen kasvu on seurausta Internetin, sähköpostin, digitaalisen median sekä raskaiden ja mediapitoisten ohjelmistojen määrän massiivisesta lisääntymisestä. Joissakin tapauksissa tallennettavan datan määrän kasvuun vaikuttavana tekijänä on ollut myös lakisääteisten tiedontallennusvaatimusten tarkentuminen. (Toigo 2003)

Yritykset ovat tänä päivänä hyvin riippuvaisia tietojärjestelmissään olevista tiedoista. Yrityksistä niin pienet kuin isotkin käyttävät hyödykseen lukuisia määriä erilaisia ohjelmistoja, jotka taas käyttävät hyväkseen tietokantoja. Näihin tietokantoihin talletettuina on yleensä yrityksen toiminnan kannalta kaikista tärkein tieto. Tänä päivänä esimerkiksi asiakas- ja tilaustietokannat sekä tekniset piirustukset ovat useasti ryhmiteltynä omiin spesifisiin tietokantoihinsa. Tämä tietokantojen lisääntynyt määrä taas on kasvattanut tiedon varmuuskopioimisen merkitystä huomattavasti.

Tietojen menetys vaikuttaa kaiken kokoisissa yrityksissä välittömästi useaan eri liiketoimintaprosessiin. Tästä johtuen yritysten kannalta tietojen varmuuskopioiminen on erityisen tärkeätä, vaikka dataa ei määrällisesti olisikaan olemassa paljon. Tämä johtuu siitä, että useissa organisaatioissa ei välttämättä ole resursseja tai kykyä luoda uudelleen jo kadotettuja tietoja. Vaikka yrityksellä olisikin resurssit luoda kadotetut tiedot uudelleen, maksaa se yritykselle huomattavan määrän aikaa ja taloudellisia resursseja.

Tiukasti kilpailluilla markkinoilla ydintietojen kadottaminen voi johtaa hyvinkin nopeasti markkinaosuuksien kadottamiseen ja yrityksen toiminnan hiipumiseen. Jopa yhden kiintolevyn rikkoutuminen voi pahimmallaan johtaa koko yrityksen toiminnan lamautumiseen, mikäli sillä olleet kallisarvoiset tiedot ovat lopulliset menetetyt. Pahimmassa tapauksessa seurauksena saattaa olla jopa konkurssi. Amerikkalaisen The National Archives and Records Administrationin tutkimuksen mukaan 80 % yrityksistä, joilla ei ole kunnollista varmuuskopiointitoimintoa sekä katastrofista palautumissuunnitelmaa ajautuvat konkurssiin kahden vuoden sisällä isosta onnettomuudesta.

2.2 Varmuuskopioinnin erilaiset tyypit

Varmuuskopion ottaminen voidaan jakaa kolmeen luokkaan sen mukaan millä tavalla ohjelman halutaan käsittelevän tietoja ja niissä tapahtuneita muutoksia. Nämä vaihtoehdot ovat: täysi varmistus, valikoiva varmistus ja lisäävä varmistus.

2.2.1 Täysi varmistus

Täysi varmistus (Full Backup) ottaa nimensä mukaisesti kopiot kaikista tiedostoista. Tässä tallennustavassa etuna on se, että kaikki tiedot ovat helposti löydettävissä viimeisimmältä nauhalta. Yleisimmin täysi varmistus ajetaan viikon viimeisenä varmistuksena. Organisaatioissa, joissa tallennettavan datan määrä on vähäinen verrattuna nauhan kapasiteettiin, on tietenkin myös mahdollista ajaa täysi varmistus vaikka joka päivä.

Haittapuolena täydessä varmistuksessa on se, että monet tiedostot, kuten esimerkiksi systeemitiedostot, muuttuvat harvoin ja tällöin nauhalla on paljon jo aikaisemmilla kerroilla tallennettua tietoa viemässä turhaa tilaa medialta. Harvoin muuttuvat systeemitiedostot, kun löytyvät myös aikaisemmilta täyden varmistuksen nauhoilta täsmälleen samanlaisina. Lisäksi täyden varmistuksen ajaminen joka päivä kuluttaa nauhoja turhan takia. Täyden varmistuksen ajaminen vie myös huomattavasti enemmän aikaa kuin vastaavasti valikoivan tai lisäävän varmistuksen tekeminen.

2.2.2 Valikoiva varmistus

Valikoiva varmistus (Diffrential backup) ottaa kopion kaikista tiedostoista, jotka ovat muuttuneet edellisen täyden varmuuskopioinnin jälkeen. Näin ollen ensimmäisenä päivänä täyden kopion ottamisen jälkeen tallennetaan sen päivän aikana muuttuneet tiedot. Seuraavana päivänä tallennetaan sen päivän aikana muuttuneet tiedostot samoin kuin edellisenäkin päivänä muuttuneiksi merkatut tiedostot. Palautettaessa valikoivalla varmistuksella talletettuja tiedostoja nauhalta tarvitaan täyden varmistuksen-media samoin kuin viimeisin valikoivan varmistuksen-media useimmissa tapauksissa.

Valikoiva varmistus-tavan etu verrattuna täyteen varmistukseen on se, että sen ottaminen kestää huomattavasti lyhyemmän ajan sekä joka päivä otettuna valikoivalla varmistuksella kopioidut tiedostot ovat selkeästi ajantasaisempia kuin täydellä varmistuksella otetut. Etuna valikoivassa varmistuksessa on myös palautusprosessin nopeus verrattuna lisäävän varmistustavan palautukseen.
Haittana voidaan pitää tietojen päällekkäisyyttä medioilla mitä kauemmas edetään viimeisimmästä täydestä varmistuksesta. Toisaalta

jos tiedot mahtuvat esimerkiksi yhdelle medialle ei tästä ole juurikaan merkittävää haittaa.

2.2.3 Lisäävä varmistus

Lisäävä varmistus (Incremental backup)-vaihtoehdossa talteen otetaan ainoastaan edellisestä varmuuskopiointi kerrasta muuttuneet tiedostot. Näin ollen medialle ei talleteta ollenkaan esimerkiksi valikoivassa varmistuksessa mukana tulevia edeltäneiden päivien muuttuneita tietoja. Palautettaessa tiedostoja, jotka on otettu lisäävällä varmistuksella, tarvitaan prosessissa mukana niin täyden varmistuksen-media kuin kaikki lisäävän varmistuksen–mediatkin useimmissa tapauksissa. Lisäävän varmistuksen etuina ovat luonnollisesti tilan säästö, nopeus sekä nopeuden kautta tuleva varmistuslaitteen käyttötuntien säästäminen. Yleisin tallennusvaihtoehto yhdistelmä yrityksissä onkin lisäävä varmistus–täysi varmistus.

2.3 Nauhakierron Hallinta

Nauhakierron hallinnalla tarkoitetaan suunnitelmaa minkä mukaisesti hankittuja nauhoja kierrätetään tietyllä logiikalla. Näin ollen väärät nauhat eivät ajaudu kiertoon liian aikaisin tai myöhään ja halutut tiedostot ovat ennalta määrätyssä nauhassa. Hallintaan on olemassa monia tapoja, mutta kaksi tapaa nousee selvästi muiden ylle käytön helppoudessa, selkeydessä ja yleisyydessä. Ne tavat ovat kohdassa 2.3.1 esiteltävä Isossä-isä-poika-rakenne sekä kohdassa 2.3.2 esiteltävä Hanoin torni-rakenne.

2.3.1 Isoisä–isä–poika–rakenne

Isoisä–isä–poika–rakenne (Grandfather–Father-Son, myöhemmin lyhennetty GFS) on ehkä yleisin nauhakierrossa käytetty hallinnointirakenne. Sen vahvuuksia ovat selkeys sekä helppo muokattavuus omiin tarpeisiin. GFS-rakenteessa käytetään mediasarjoja niin, että viikonpäiville maanantai–torstai on jokaiselle omansa. Kyseisiä nauhoja käytetään siis uudestaan joka viikko sille määrättynä päivänä. Näihin päiväkohtaisiin nauhoihin ajetaan tallennus joko valikoivana tai lisäävänä varmistuksena. Näitä neljää mediasarjaa kutsutaan pojiksi (Sons). Tämän lisäksi joka perjantai ajetaan täysi varmistus erillisille viikkonauhoille. Näitä nauhoja on yhteensä kolme eri sarjaa ja niitä kutsutaan isiksi (Fathers). Niitä käytetään joka kuukausi niille määrätyn viikon perjantaina. Kerran kuussa otetaan lisäksi yksi täysi varmistus erilliselle medialle, jota kutsutaan isoisäksi (Grandfather). Kuukausinauhasarja voidaan tämän jälkeen käyttää uudestaan joko muutaman kuukauden välein (3-6 kk) tai pitää tallessa pysyvästi. Jos kuukausinauhasarja pidetään pysyvästi tallessa, tullaan niitä tarvitsemaan vuodessa 13 kappaletta.

Nauhojen säilytyksen kannalta tarkoituksena tässä kierrossa on, että päivänauhat pidetään tallessa lähellä tallennusasemaa. Viikkonauhat säilytetään hivenen kauempana tallennuspaikan välittömästä läheisyydestä ja kuukausinauhat pidetään tallessa jossakin erillisessä säilytyspaikassa kaukana itse tallennuspaikasta.

Etuna GFS–kierrossa on ehdottomasti sen selkeys ja muokattavuus eri tarpeisiin. Medioiden käyttömäärät saattavat vaihdella paljonkin riippuen siitä mitä mediaa käytetään, kuinka suuri sen tallennuskapasiteetti on, paljonko tallennettavaa tietoa löytyy järjestelmästä ja minkälainen tallennuslaite on käytössä. Joka päivälle voi esimerkiksi olla yksi tai useampia nauhoja. GFS–rakennetta käyttöön otettaessa pitää muistaa sovittaa viikonpäivät sen mukaan kuinka monena päivänä viikossa yrityksen järjestelmään tulee uutta tietoa tallennettavaksi. Kuvassa 1. GFS–rakenne on toteutettu normaalille 5 työpäivän viikolle, mutta se voidaan myös sovittaa esimerkiksi 6 työpäivän viikolle. Haittapuolena GFS–kierrossa voidaan pitää epätasaista nauhojen kulumista, kun päivänauhoja verrattaan viikko– ja kuukausinauhoihin ja tarkastellaan näiden nauhojen käyttötiheyttä.

	Maanantai	Tiistai	Keskiviikko	Torstai	Perjantai
Vk 1	Maanantain päivänauhasarja, Osittainen varmistus	Tiistain päivänauhasarja, Osittainen varmistus	Keskiviikon päivänauha sarja, Osittainen varmistus	Torstain päivänauhasarja, Osittainen varmistus	Viikkonauhasarja 1 Täysi varmistus vk 1
Vk 2	Maanantain päivänauhasarja, Osittainen varmistus	Tiistain päivänauhasarja, Osittainen varmistus	Keskiviikon päivänauha sarja, Osittainen varmistus	Torstain päivänauhasarja, Osittainen varmistus	Viikkonauhasarja 2 Täysi varmistus vk 2
Vk 3	Maanantain päivänauhasarja, Osittainen varmistus	Tiistain päivänauhasarja, Osittainen varmistus	Keskiviikon päivänauha sarja, Osittainen varmistus	Torstain päivänauhasarja, Osittainen varmistus	Viikkonauhasarja 3 Täysi varmistus vk 3
Vk 4	Maanantain päivänauhasarja, Osittainen varmistus	Tiistain päivänauhasarja, Osittainen varmistus	Keskiviikon päivänauha sarja, Osittainen varmistus	Torstain päivänauhasarja, Osittainen varmistus	Kuukausinauha 1 Täysi varmistus kk 1

Kuva 1. Tyypillinen GFS–rakenne.

2.3.2 Hanoin torni–rakenne

Tämä rakenne saa nimensä vanhasta lasten pelistä, jossa niin sanottuja sormuksia liikutellaan kepistä toiseen siten, että ne päätyvät lopussa toiseen keppiin ollen samassa järjestyksessä kuin peli aloitettaessakin. Tavoitteessa onnistuakseen kyseisiä sormuksia on siirrettävä tietyin jaksoin eteenpäin. Hanoin torni – rakenteessa ajatuksena on, että tiedostoista on olemassa useita eri-ikäisiä kopioita talletettuna. Tämä rakenne mahdollistaa useita kopioita viimeisimmistä tiedostoista, useita kopioita useamman viikon vanhoista tiedostoista sekä muutamia pari kuukautta vanhoja tiedostokopioita.

Hanoin torni-rakenteessa nauhojen kierto toteutetaan kuvan 2. mukaisella tavalla. Ajatuksena on, että A-sarjaa käytetään joka toinen viikko, B-sarjaa joka neljäs viikko, C-sarjaa joka kahdeksas viikko ja niin edelleen. Uusia nauhasarjoja lisättäessä uusista nauhoista tehdään A-sarja ja vanhemmat nauhasarjat liikkuvat vastaavasti pykälän eteenpäin rakenteessa eli entisestä A-sarjasta tuleekin B-sarja ja näin edelleen. Etuna Hanoin rakenteessa on se, että tallessa on useita eri-ikäisiä kopioita. Haittapuolena vastaavasti on toisten nauhasarjojen altistuminen kovemmalle kulutukselle. (Seagate Technology LLC 2006)

Hanoin torni-rakenne ei sulje pois GFS-rakenteen käyttöä ja useasti itse asiassa tämän rakenteen sisällä käytetäänkin GFS-rakennetta. Yhdistelmärakennetta luotaessa on tärkeätä, että asiasta tehdään kunnollinen taulukointi ja että itse nauhat nimetään selkeällä ja informatiivisella tavalla.

Viikko	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Tallen-nusme-diasarja	A		A		A		A		A		A		A		A	
		B				B				B				B		
				C								C				
								D								
																E

Kuva 2. Hanoin torni–rakenne.

2.4 Rakenteiden varioiminen

Kaikissa tapauksissa ei ole tarpeen toimia jonkin erityisen rakenteen mukaan. Pienissä yrityksissä tallennettavan datan määrä ei ole välttämättä niin iso, että tarvitaan erillistä nauhakiertoa. Tällöin voi tulla kysymykseen esimerkiksi vain yhden nauhan käyttö, jolla varmistetaan tietojen tallessa pysyminen tallentamalla sille tiedot joka päivä. Tosin yhden nauhan käyttäminen ei ole suositeltavaa, sillä nauhan rikkoutuessa tilanne voi olla toivoton. Toisaalta edellä esiteltyjä rakenteita pystyy myös varioimaan monipuolisempaan suuntaan tar-

peista riippuen. Esimerkiksi GFS–rakenteen voi toteuttaa niin, että jokaiselle päivälle on oma päivänauhasarjansa (tai nauha) välillä maanantai–torstai ja näitä päivänauhasarjoja on sitten viisi kappaletta. Näin kokonaisen kuukauden saa katettua niin, että joka päivälle on käytössä eri nauhasarja.

3 Varmuuskopiointiin liittyvät laitteistot ja mediat

3.1 Yleistä

Varmuuskopiointirintamalla on tällä hetkellä tarjolla laaja valikoima erilaisia formaatteja. Pääasialliset erot formaattien kesken tulevat kapasiteetissa, tiedonsiirtonopeudessa sekä median fyysisessä koossa. Verrattaessa eri formaatteja on hyvä ottaa huomioon, että tallennuskapasiteettilukemat ilmoitetaan monesti kahdella eri luvulla. Toinen luku ilmoittaa natiivin (alkuperäisen tai tässä tapauksessa pakkaamattoman tiedon) tallennuskapasiteetin ja toinen luku pakatun tiedon tallennuskapasiteetin. Monilla valmistajilla on myös tarjottavana laaja skaala erilaisen tallennuskapasiteetin omaavia medioita per formaatti. Samanlaista nauhatyyppiä voi siis saada esimerkiksi 100 gigatavuisena sekä 200 gigatavuisena. (9 to 5 computer.com 2003)

Lisäksi eroja löytyy monesti formaateille annetuista kestävyysluvuista, joilla ilmoitetaan kuinka monta kertaa media voidaan kirjoittaa uudelleen. Näitä lukuja vertaillessa on myös hyvä huomata, että eri valmistajat ilmoittavat lukemiaan toisistaan poikkeavilla laskentatavoilla. Nämä ilmoitettujen lukemien erot johtuvat eri tekniikoiden tavoista kirjoittaa nauhalle. Esimerkiksi AIT-2-nauhalle annettava end-to-end tape passes-luku voisi olla 30 000, kun taas DLT IV-nauhalle annettava luku olisikin 1 000 000. Lukutekniikasta johtuen kumminkin yhden DLT IV-nauhan kirjoittaminen täyteen tarkoittaa 208 raitaa, jonka seurauksena annettu end-to-end luku onkin jaettava 52:lla, koska yhden kirjoituskierroksen kirjoitusmäärä on 4 raitaa kyseisellä tekniikalla. Näin saatava todellinen end-to-end tape passes-luku onkin 19 230.

Formaatin valinnassa tulee muistaa ottaa huomioon se seikka, että tallennettavan tiedon määrä kasvaa koko ajan. Tästä johtuen valinnan tulee olla järkevästi skaalautuva vaihtoehto. Näin ollen on hyvä tarkistaa, että valitulla formaatilla on järkevä tiekartta tulevaisuuteen. Valmistajan laatima tiekartta tarkoittaa, että formaatille on tulossa myös seuraava sukupolvi ja että kyseisen formaatin kehittämiseen panostetaan. Tällä tavoin voi varmistua, että valittu formaatti ei ole niin sanottu end-of-life–tuote tai vain jokin väliaikainen siirtymä-formaatti.

3.2 Nauha-asemat

Formaatin valinnassa olennainen tekijä on sille sopiva nauha–asema. Nauha–aseman valinnassa tärkeitä asioita ovat hinta, skaalautuvuus, taaksepäin yhteensopivuus, nopeus ja toimintavarmuus. Suurin osa nauha-asemien valmistajista tarjoaa tällä hetkellä eri malleistaan useita eri versioita sen mukaan, haluaako asiakas aseman latauksen

toimivan käsikäyttöisesti vai haetaanko esimerkiksi jonkin näköistä nauhakirjastorakennetta. Käsikäyttöisessä mallissa nauhat syötetään itse latauskehikkoon tietyssä järjestyksessä ja tämän jälkeen kehikko laitetaan asemaan. Toisenlaisissa malleissa asemaan taas laitetaan ainoastaan yksi nauha syöttämälle se sisään samoin kuin esimerkiksi VHS-videonauhureissa. Nauhakirjastot taas voivat koostua useista asemista ja kymmenistä nauhoista ja nämä nauhat latautuvat asemiin halutulla tavalla käyttäjän asettamien konfigurointien mukaan.

Nykyisin datamäärien kasvaessa koko ajan huimaa vauhtia, ovat monet tahot ottaneet käyttöönsä myös aikaisemmin mainittuja nauhakirjastoja. Nämä laitteet tarjoavat mahdollisuuden lisätä yhteen kokonaisuuteen useita nauha-asemia, joille sitten määrätään käyttöön tietty määrä halutun formaatin mukaisia nauhoja. Esimerkiksi IBM 3583 Ultrium skaalautuva nauhakirjasto pystyy sisällyttämään itseensä jopa kuusi nauha-asemaa ja 72 nauhaa. Nauhakirjastojen etuna on aina huima tallennuskapasiteetti ja keskitetty laitteiston toiminta, mutta niiden hinnat ovat asemien ja medioiden määrästä riippuen jopa useita kymmeniä tuhansia euroja, joten ne eivät sovi jokaiselle organisaatiolle.

Formaattia ja nauha-asemaa valittaessa onkin monesti viisasta tarkastella taulukkomuodossa eri laitteille ja medioille annettuja lukuarvoja. Kuvassa 3. tarkastellaan neljän eri aseman arvoja niille annettujen tärkeimpien lukuarvojen kautta.

	SAIT-1	SDLT 320	LTO Ultrium II	SDLT 600
KAPASITEETTI JA SUORITUSKYKY				
Natiivi kapasiteetti	500 GB	160 GB	200 GB	300 GB
Pakatun tiedon kapasiteetti	1,3 TB	320 GB	400 GB	600 GB
Natiivi tiedonsiirtokyky	30MB/s	16 MB/s	35 MB/s	36 MB/s
Pakatun tiedon tiedonsiirtokyky	78MB/s	32MB/s	70 MB/s	72 MB/s
Tiedon hakunopeus	70 s	70 s	49 s	79 s
Median latausaika	23 s	12 s	15 s	12 s

Kuva 3. Kooste eri formaattien nauha-asemien numeerisista arvoista.

3.3 WORM-tekniikka

90–luvun loppupuolella markkinoille alkoi ilmestyä koko ajan enenevässä määrin eri valmistajilta WORM–nauhoja (Write Once, Read Many). Tämä tekniikka kehitettiin, koska monilla yrityksillä ja yhteisöillä on paljon lainsäädännöllisistä syistä säilytettäviä tietoja.

Nauhojen erikoisuus on, että niille voi tallentaa tietoja ainoastaan kerran. Tämän seurauksena niille talletetaan tietoja, joidenka säilyminen muuttumattomana on erittäin tärkeätä. Tällaisia tietoja voivat olla esimerkiksi potilaiden terveystiedot, rakennepiirustukset ja tekniset kaaviot sekä tuotantomääritykset. Kyseiselle nauhatyypille on olemassa tiukat määritykset, minkä puitteissa sen pitää toimia. Määritettyjä asioita ovat muun muassa toimintaikä ja varmuus, formaattituki tulevaisuuteen sekä tietojen säilytysvarmuus. Nykyisin lähes jokaisesta formaatista on olemassa myös WORM–versiot ja niitä valmistavat useat eri yritykset.

3.4 Nauha-formaatit

Nauha-formaatteja on markkinoilla kirjava määrä ja niistä osa on poistumassa olevia tuotteita ja osa vasta vähän aikaa sitten lanseerattuja versioita. Formaattien versiot on yleensä helpointa tunnistaa niiden perässä olevasta numerosta. Esimerkiksi AIT-formaatista on olemassa ainakin neljä erilaista ikäversiota. Nauhoja hankittaessa onkin syytä olla tarkkana, että hankkii tietystä formaatista oikeanlaisia versioita, koska tietylle formaatille hankitut nauha-asemat eivät välttämättä aina hyväksy kyseisen formaatin kaikkia ikäversioita.

3.4.1 DDS

DDS-1-formaatti (Digital Data Storage) oli vuonna 1989 ensimmäinen tallennusmedia, joka käytti nykyisin liki kaikissa formaateissa käytettävää metallipartikkeli-teknologiaa. DDS–formaatti tehtiin DAT–formaatin pohjalta. Kyseinen teknologia oli mullistava, sillä se avasi mahdollisuuden käyttää fyysisesti huomattavasti pienempiä nauhoja, joihin kumminkin pystyi tallentamaan merkittävästi isompia datamääriä. Verrattuna aikaisempiin fyysisesti kookkaampiin nauhoihin, DDS-1–nauhoille pystyi tallentamaan neljää kertaa enemmän tietoa, vaikka ne tarjosivat viisi kertaa vähemmän fyysistä tallennusalaa. DDS-formaatista kehittyi halpa, mutta luotettava tallennusformaatti, joka merkitsi pitkän aikaa synonyymia sanalle tallennusnauha. DDS–formaatti oli erittäin suosittu niin sanotussa low-end–segmentissä eli hieman halvempien tallennustuotteiden puolella.

DDS-1–nauhojen natiivi kapasiteetti oli 1,3 gigatavua ja sitä kasvatettiin myöhemmin 2:en gigatavuun. DDS-2–sukupolvessa nauhan rakennetta muutettiin ja tällä tavoin natiivi tallennustila saatiin tuplattua 4 gigatavuun. DDS-3–sukupolvessa muutettiin jälleen nauhan rakennetta sekä määrää ja näin saavutettiin suurin kapasiteetillinen kehitys eri DDS-versioiden välillä 12 gigatavun tallennustilalla. DDS-4–sukupolven aikana päästiin 20 gigatavun natiiviin kapasiteettiin ja tiedonsiirtonopeus nousi lukemaan 3 megatavua sekunnissa (myöhemmin Mb/s).

Vuonna 2001 Sony ja Seagate ilmoittivat, että DDS-4-sukupolvi tulisi olemaan viimeinen versio DDS-formaatista. HP kuitenkin julkaisi vuonna 2003 DAT72–sukupolven (DDS-5–sukupolvi) paikkaamaan markkinoille jäänyttä tyhjiötä. DAT72 esitteli tällöin 32 gigatavun natiivin kapasiteetin, mutta sen tallennusnopeus oli sama kuin DDS4-sukupolvessa. DAT72–formaatti oli kuitenkin yhteensopiva aikaisempien DDS–3 ja DDS-4-formaattien kanssa ja saavutti tämän takia menestystä low–end–segmentissä.

3.4.2 DLT

DLT–formaatin (Digital Linear Tape) ensimmäinen sukupolvi ilmestyi jo vuonna 1989 ja se toi mukanaan silloiseen aikaan huiman 2,6 gigatavun kapasiteetin. Tästä kehitys jatkui tasaisesti ja DLT4–sukupolvi tarjosi natiivia tallennuskapasiteettia 40 gigatavua ja seuraava sukupolvi nimeltään DLTVS1 nosti kapasiteetin jo 80 gigatavuun. DLTVS1–sukupolvi oli Quantumin vastaus DDS–formaatin seuraajaksi. DLT–formaatti oli DDS–formaatin tavoin monien pienten ja keskisuurten yritysten suosituin valinta tallennusformaatikseen melkein 15 vuotta sen tarjoaman halvan, mutta luotettavan toiminnan vuoksi.

3.4.3 SDLT

Ensimmäisessä SDLT–sukupolvessa (SuperDLT) oli natiivia tallennuskapasiteettia 110 gigatavua ja sen kirjoitusnopeus oli 11 Mb/s. SDLT1–sukupolven natiivi tallennuskapasiteetti nostettiin myöhemmin 160 gigatavuun ja sitä seurannut sukupolvi nimeltään SDLT2 (SDLT 320) tarjosi jo 320 gigatavua tallennustilaa pakkaamattomalle tiedolle ja tiedonsiirtonopeuden 32Mb/s. SDLT–laitteet ovat yhteensopivia DLT–formaatin kanssa, joten DLT-nauhoja pystyy lukemaan myös SDLT–laitteistoilla. Tämä asia onkin tehnyt SDLT suositun formaatin monilla jo ennestään DLT–formaattia käyttäneillä yrityksillä. Etuna tässä tekniikassa on myös sen halpa hinta per tallennettava gigatavu. SDLT–tekniikka on yksi suosituimmista nauhatekniikoista tällä hetkellä ja sillä on selkeän aggressiivinen tiekartta tulevaisuuteen. Se sijoittuu tämän hetken tallennusmarkkinoilla keskisegmenttiin.

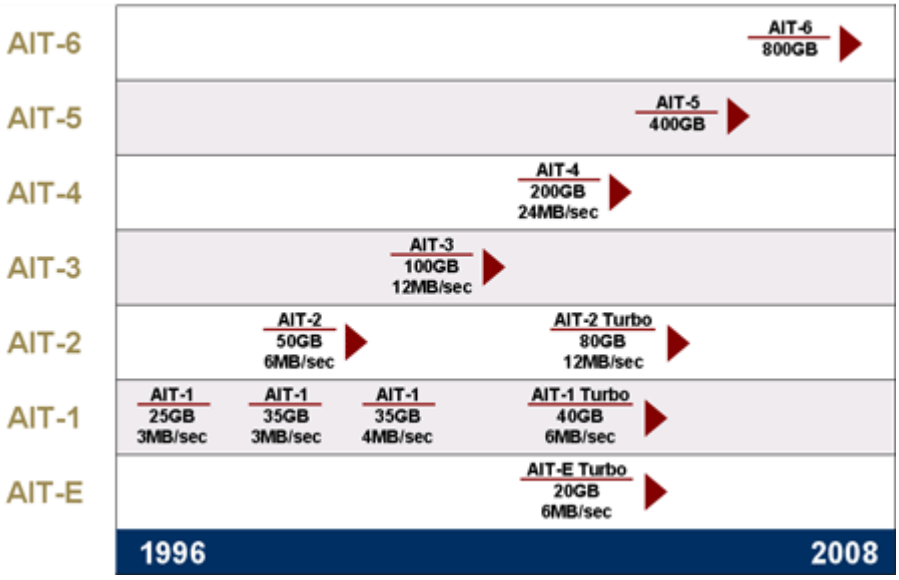
3.4.4 AIT

Sony julkaisi AIT-1–formaatin (Advanced Intelligent Tape) vuonna 1996 ja se oli tarkoitettu päivitysvaihtoehdoksi DDS-3–formaatilla toimivien järjestelmien tilalle. AIT–formaatti erosi muista sen aikaisista ratkaisuista selkeästi, sillä sen koko oli selvästi pienempi kuin muiden vaihtoehtojen. Se vastaa kooltaan käytännössä vanhaa c-kasettia. Tuote hinnoiteltiin alussa varsin maltillisesti, millä pyrittiin vauhdittamaan muutosta DDS-tekniikasta uuteen AIT–tekniikkaan.

Alussa AIT-1–formaatti tarjosi 25 gigatavun verran pakkaamatonta tallennustilaa ja vuoteen 1999 mennessä kapasiteetti saatiin nostettu 35 gigatavuun natiivia tietoa per kasetti. Keväällä 1999 Sony julkaisi AIT-2–formaatin, joka tarjosi 50 gigatavua pakkaamatonta tallennustilaa ja joka oli tarkoitettu DDS-4–tekniikan omaavien yritysten järjestelmien päivitykseen. AIT-2–formaatti oli ensimmäisiä tekniikoita, missä esiteltiin tiedon pitkäaikaisessa varastoinnissa tärkeä WORM-ominaisuus. AIT-3–formaatti esiteltiin loppuvuonna 2001 ja se tarjosi 100 gigatavun verran natiivia tilaa ja 260 gigatavun verran pakattua tallennustilaa. Tiedonsiirtonopeus oli tässä vaiheessa noussut 12Mb/s natiiville tiedolle ja 31Mb/s pakatulle tiedolle. AIT-3 oli myös täysin yhteensopiva aikaisempien AIT–sukupolvien kanssa. Vuoden 2004 aikana Sony esitteli useita AIT–turbo–tuotteita, jotka tehostivat aikaisempia AIT–formaatteja tarjoamalla kasvanutta tiedonsiirtonopeutta ja suurempaa kapasiteettia. Nämä tuotteet olivat AIT-E Turbo, AIT-1 Turbo ja AIT-2 Turbo. AIT-4–formaatin Sony julkaisi alkuvuodesta 2005 ja siinä natiivi tallennuskapasiteetti oli noussut jo 200 gigatavuun ja siirtonopeuskin 24 Mb/s. Kyseinen formaatti tarjoaa tällä hetkellä suurimman kapasiteetin ja tehokkuuden kokoluokassaan, mutta se ei ole täysin yhteensopiva aikaisempien AIT–formaattien kanssa. Tämä siis tarkoittaa käytännössä sitä, että vaihtaminen uuteen formaattiversioon ei onnistu kuin nauha–aseman vaihtamisella.

3.4.5 SAIT

Sony julkaisi vuonna 2004 todella vaativiin tarpeisiin tarkoitetun S-AIT–formaatin. Se eroaa kooltaan aikaisemmasta AIT–formaatista, sillä SAIT–nauha vastaa kooltaan kohdassa 3.4.6 esiteltävää Ultrium–nauhaa. SAIT–nauhan tallennuskapasiteetti on 500 gigatavua natiivina ja yli teratavun verran pakattuna. SAIT oli ensimmäisiä formaatteja, jotka tarjosivat yhdellä kasetilla mahdollisuuden tallentaa yli teratavun verran tietoa. Tiedonsiirtonopeus SAIT–formaatilla on 30Mb/s ja sitä on saatavilla myös WORM-mallisena. Myös SAIT–formaatille on olemassa tiekartta, joka kattaa uusien versioiden ilmestymisen aina SAIT 4–versioon asti, joka lupaa tällöin 4 teratavua natiivia tallennustilaa yhdelle nauhalle. Tällä hetkellä SAIT -1–sukupolvi tarjoaa suurimman natiivin kapasiteetin markkinoilla.

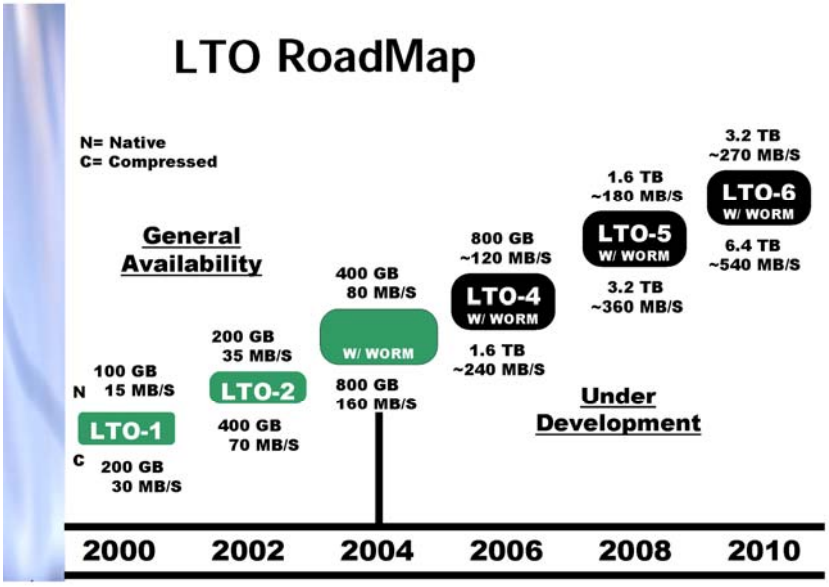


Kuva 4. AIT – formaatin tiekartta (AIT Technology Forum 2006).

3.4.6. LTO

LTO–tekniikan (Linear Tape-Open) kehittivät yhteistyössä HP, IBM ja Quantum vuonna 2000 ja sen tunnetuin avoimeen formaattiin pohjaava versio on Ultrium–formaatti. Ultrium–formaatista on ilmestynyt kolme sukupolvea. Ensimmäinen sukupolvi tarjosi pakkaamattomalle tiedolle tallennuskapasiteettia aluksi 20 gigatavua, mutta nykyisin se tarjoaa parhaimmillaan jo 100 gigatavua. Ultriumin toinen sukupolvi tarjosi 200 gigatavua natiivia tallennustilaa ja kolmas jo 400 gigatavua natiivia tallennuskapasiteettia. Vastaavat tiedonsiirtonopeudet ovat 10–20 Mb/s, 20-40Mb/s ja 40-80Mb/s. Ensimmäinen luku on natiivin tiedon ja jälkimmäinen luku pakatun tiedon tiedonsiirtonopeus. Ultrium on eräs markkinoiden suosituimmista tallennusvaihtoehdoista ja sille on myös julkaistu selkeä tiekartta tulevaisuuteen. (Ultrium.com 2006)

LTO–formaattiin pohjautuvia tallennusmedioita on markkinoilla ollut muitakin, kuten esimerkiksi Accelis, mutta ne eivät ole saavuttaneet merkittävää menestystä ja ovat jo nykyisin käytännössä kadonneet markkinoilta.



Kuva 5. Ultriumin–formaatin tiekartta.

4 Varmistusohjelmistot

4.1 Yleistä

Varmistusohjelmistojen määrä on lisääntynyt nopealla tahdilla jatkuvasti kasvavilla tallennusmarkkinoilla. Nykyisin ne voidaankin jakaa kahteen pääryhmään. Toinen ryhmä on yksityiskäyttäjän ja pienen yrityksen käyttöön tarkoitetut yksinkertaisemmat ja kevyemmät ohjelmat ja toinen ryhmä on isojen tietomassojen tallennukseen tarkoitetut raskaammat ja huomattavasti monipuolisemmat ohjelmat. Monet ammattikäyttöön tarkoitetuista järjestelmistä ovatkin useamman ohjelmapalasen muodostamia ohjelmaperheitä, joihin on mahdollista hankkia useita erilaisia moduuleja varmuuskopiointihallinnan lisäksi.

4.2 Kotikäyttäjän varmuuskopiointiohjelmistot

Kotikäyttäjille tietojen tallennukseen tarkoitettujen ohjelmistojen määrä on kasvanut viime vuosina merkittävästi. Markkinoilta löytyy paljon niin ilmaisia kuin kohtuullisen edullisiakin vaihtoehtoja kotikäyttäjälle. Koti– ja pienkäyttäjälle tarkoitetuissa ohjelmistoissa olevat erot eivät nykyisin ole kovinkaan merkittäviä. Tästä johtuen ohjelmien vertaaminen keskenään ei ole aina aivan yksinkertainen asia. Hyvänä nyrkkisääntönä voisi tässäkin asiassa kumminkin pitää, että tunnetun yrityksen tuotteet ovat varmasti ainakin kohtuullisen yksinkertaisia ja niille on mahdollista saada jonkinlaista tukea. Tietenkin on olemassa varmasti myös poikkeuksia. Ohjelmaa valittaessa kannattaa kumminkin tutustua huolella ohjelman ominaisuuksiin sekä sen tukeen eri laitteille ja ohjelmistoille.

Lähestulkoon kaikki ohjelmat nykyisin sisältävät täyden, valikoivan sekä lisäävän varmistuksen mahdollisuuden, tuet eri Windows–käyttöjärjestelmille sekä Outlook-sähköpostiohjelmalle, mahdollisuuden tallentaa siirtomedialle ja riittävät tietojen palautus-, ajastus– ja tallennusvarmistus vaihtoehdot. Näiden asioiden lisäksi monissa ohjelmissa on vielä paljon ominaisuuksia, jotka ovat ominaisia juuri kyseiselle ohjelmalle. Kaiken kaikkiaan tarjonnan ollessa laaja hyödyn siitä korjaa koti– ja pienkäyttäjä parantuneina ominaisuuksina sekä alentuneina hintoina.

Tällä hetkellä koti– ja pienkäyttöön tarkoitettujen tallennusohjelmistojen hinnat lähtevät noin 40 eurosta. Viimeaikaisissa alan testeissä hyväksi todettuja ohjelmia ovat muun muassa olleet: Genie Backup Manager, NTI Backup NOW, Norton Ghost sekä Handy Backup.

4.3 Ammattikäyttöön tarkoitetut varmuuskopiointiohjelmistot

Ammattikäyttöön tarkoitetut ohjelmistot ovat tavallisesti raskaita erikoisohjelmistoja, jotka yleensä asennetaan omalle palvelimelleen tehokkaan ja luotettavana toiminnan mahdollistamiseksi. Ammattiohjelmistojen ero verrattaessa niitä koti- ja pienkäyttäjien versioihin tulee siitä, että ne ovat alusta alkaen tehty yritysjärjestelmiä ja niiden tarpeita ajatellen. Kyseisiin ohjelmiin on mahdollista lisätä myös uusia ohjelmamoduuleja, jotka tuovat mukanaan erilaisia ominaisuuksia ja laajentavat niiden yhteensopivuutta eri järjestelmien ja laitteiden kanssa. Kyseiset ohjelmistot sisältävät myös valtavan määrän erilaisia ominaisuuksia sekä konfigurointi mahdollisuuksia. Monesti ohjelmistoversiopäivityksen mukana kyseisiin ohjelmistoihin lisätään uusia ominaisuuksia sekä päivitetään yhteensopivien nauhaasemien ja laitteistojen luetteloa.

Ohjelmia ei tietenkään ole pakko asentaa omalle palvelimelleen, mutta tällöin on syytä pohtia palvelimella jo käytössä olevien ohjelmien ja palveluiden resurssitarvetta ja tutkia näiden ohjelmien yhteensopivuutta tallennusohjelmiston kanssa.

4.3.1 Hinnoittelu ja vertailu

Ammattilaisohjelmistojen hintahaitari on varsin laaja. Halvimmat kyseisen luokan ohjelmat maksavat alle tuhat euroa ja kalleimmat jopa useita kymmeniä tuhansia euroja. Ohjelmien hintavertailu ei tietotekniikkatuotteissa ole ikinä ollut helppoa, eikä asiassa poikkeusta tee myöskään varmuuskopiointiala. Hintoihin vaikuttavia asioita ohjelmissa ovat muun muassa ohjelman järeys, sulautettujen ohjelmaosien ja tuotetuen määrä, ohjelman asennusympäristön koko sekä käyttöjärjestelmätuki. Kaiken kaikkiaan vaikuttavia tekijöitä on huomattavasti ja niiden painoarvo riippuu täysin olemassa olevan ympäristön vaatimuksista sekä toivotuista tuloksista.

Kuten edellisessä kappaleessa totesin, ohjelmistojen hintojen suhteellinen vertailu on vaikeata ja sama pätee tuotteiden tarjoamaan sisältöönkin. Parhaiten tietoa eri tuotteiden toimivuudesta ja ohjelmistojen eroista löytääkin alan eri julkaisuista. Tuotteita vertailtaessa on hyvä ottaa huomioon ainakin seuraavia asioita: Asennuksen nopeus ja helppous, lisensointi hinta, pakettiin sisältyvät ohjelmat (tulevatko esimerkiksi agenttiohjelmat mukana), laitteistovaatimukset, ohjelmaosien lisäämismahdollisuus ja päivittäminen, tuki eri tallennusvaihtoehdoille sekä yleinen tuotetuki.

4.3.2 Tuotteet

Suomessa suurimman markkinaosuuden on jo vuosia omannut Symantecin Backup Exec. Ennen kyseinen ohjelma tunnettiin Veritaksen nimellä, mutta Symantec osti yrityksen noin vuosi sitten. Backup Execin suosio Suomessa pohjaa pitkälti ominaisuuksiin nähden halpaan hintaan, kohtalaisen helppoon käytettävyyteen sekä ohjelman laajaan levikkiin, minkä ansioista maassa on paljon osaamista kyseisestä ohjelmasta. Ohjelma on myös hyvin dokumentoitu ja sitä käsitteleviä asiantuntija foorumeita on internetissä useita.

Markkinoilla on olemassa myös useita muita suosittuja vaihtoehtoja, kuten Computer Associatesin BrightStor ARCserve Backup r11.5 for Windows, EMC:n Retrospect 7.5. for Windows Business Editions, ja IBM:n Tivoli Storage Manager 5.3. Kaikki kyseiset tuotteet ovat isojen tai keskisuurten yritysten vaativiin tarpeisiin tarkoitettuja ammattilaisjärjestelmiä. Niistä parhaiten lisätietoa löytää suoraan valmistajien sivuilta.

5 Riskianalyysi

Riskianalyysilla tarkoitetaan yrityksen sisäistä arviointimenettelyä, jossa kerätään tietoa siitä, mitkä seikat voivat aiheuttaa yrityksen toiminnan lamaantumisen tai merkittävästi vaikeuttaa sen toimintaa. Kyseisen analyysin toteuttamiseen on olemassa monia valmiita ohjeita ja selkeitä formaatteja. Ajatuksena on ensin luoda selkeät pääkohdat arvioinnille ja tämän jälkeen jakaa nämä pääkohdat alakohtiin, jolla tarkennetaan kyseisten kohtien toimintaa syvällisemmin.

Varmuuskopiointiin liittyvässä analyysissa voitaisiin paremminkin puhua järkevästä ja johdonmukaisesta tietovarojen ja niiden merkityksen kartoittamisesta sekä kyseisten tietojen vaatimien säilytysolosuhteiden ja metodien sopivuudesta yritykselle. Tärkeimmät pääryhmät tässä pohdiskelussa ovat:

- Paljonko tärkeäksi / talletettavaksi määriteltyä tietoa on ja paljonko kyseisen määrän uskotaan kasvavan?
- Minkälainen laitteisto tarvitaan tiedon tallentamiseen, varastointiin ja palauttamiseen nyt ja tulevaisuudessa?
- Kuka pitää huolta järjestelmän toimivuudesta ja siihen liittyvistä asioista?

Näitä pääkohtia kasattaessa pohdittavaksi tulee varmasti ainakin seuraavia kysymyksiä:

- Mitkä tiedot voivat kadotessaan aiheuttaa yrityksen ajautumisen merkittäviin vaikeuksiin tai jopa konkurssiin?
- Mitkä tiedot ovat ainutkertaisia yritykselle ja mikä on niiden potentiaalinen merkitys esimerkiksi kilpailijalle?
- Missä kyseiset tiedot sijaitsevat ja kenellä on niihin pääsyoikeus?
- Mitkä tiedot tarvitsee pitää tallessa valitsevien säädösten ja lakien vuoksi?
- Pitääkö lakien ja säädösten vuoksi tallennettavien tietojen olla muuttumattomia eli tarvitseeko ne tallettaa WORM–nauhoille?
- Onko yritys kasvamassa lähiaikoina huomattavasti esimerkiksi yritysostojen muodossa tai rekrytoinnin seurauksena?
- Mikä on budjetti hankinnoille? Onko mahdollista hankkia kaikki tarvittavat asiat yhdellä kertaa vai rakennetaanko kokonaisvaltaista ratkaisua pidemmän ajan kuluessa usealla eri hankintakerralla.
- Toteutetaanko tietojen tallennus nauhoilla vai kiintolevyillä?
- Missä nauhat säilytetään?
- Onko nauhojen säilytyspaikka asian mukaisessa kunnossa?
- Kuka kuljettaa nauhoja paikkojen välillä ja vastaa prosessin toiminnasta?

- Onko nauhojen säilytyspaikka varmasti joka paikassa yhtä laadukas?
- Kuka vastaa tallennusprosessista, jos päävastaava on poissa?

Nämä kysymykset esittävät joukkoa niistä asioista joihin on saatava järkevä vastaus, jotta varmuuskopiointiprosessista saadaan varmasti kokonaisvaltaisesti turvallinen ja kattava ratkaisu.

Monissa pienemmissä yrityksissä ei välttämättä ole suurta määrää rahaa sijoitettavana IT–projekteihin puhumattakaan erillisistä varmuuskopiointijärjestelmistä. Tällöin kunnolliseen ratkaisuun päätyminen vaatii monesti kompromisseja. Yleensä pienempi budjetti tarkoittaa käytännössä sitä, että järjestelmiä ja niiden toimintaan ei pystytä kovinkaan pitkälle automatisoimaan. Niiden toimiminen tällöin tarkoittaa suurempaa määrää suoraa käyttäjäkontaktia. Tästä seuraa se, että tapahtuvien virheiden mahdollisuuden määrä kasvaa merkittävästi, kun mukaan tulevat inhimilliset erehdykset. Näin ollen suunnitteluvaiheessa onkin jo syytä keskittyä kartoittamaan ja mahdollisesti eliminoimaan turhia käyttäjistä johtuvia vaaratilanteita ja riskejä. Näitä ovat perinteisesti olleet nauhojen liikuttaminen, vaihtaminen ja asian mukainen varastoiminen. Kattavan ratkaisun aikaansaaminen vaatii monesti aikaa, kärsivällisyyttä sekä huolellisuutta.

6 CDP

6.1 Esittely

CDP tulee sanoista Continuous Data Protection ja se tarkoittaa järjestelmää, missä tiedot sisältäviä kohteita suojataan toimintahäiriöiltä ottamalla niistä kopioita tai seuraamalla niissä tapahtuneita muutoksia reaaliaikaisesti. Ajatus CDP–järjestelmästä lähti tarpeesta saada tiedostopalvelimien ja NAS–laitteiden (Network-Attached Storage) tiedostot turvaan siten, että palautettavat tiedostot olisivat viimeisimmät mahdolliset versiot, jotka olivat olemassa ennen laitteistorikkoa tai käyttöhäiriötä. Perinteisen nauha–tallennuksen avulla tallennetut tiedostot voivat olla pahimmassa tapauksessa melkein vuorokauden vanhoja. Tästä johtuen monissa yrityksissä on siirrytty CDP–järjestelmiin, jotka mahdollistavat käyttökatkosten ajallisen eliminoinnin. Nykyisin CDP–järjestelmän voi hankkia SAN–laitteen tai levyryhmän ominaisuutena, erillisenä ohjelmana tai sulautettuna tallennusohjelmapakettina. Muutama vuosi sitten kyseiset järjestelmät ja järjestelyt olivat kalliita toteutuksia, mutta sen jälkeen markkinoille on tullut useita ohjelmistoja ja laitteita, jotka tukevat kyseistä ominaisuutta.

CDP-järjestelmä ottaa kohteiksi valituista tiedostoista reaaliaikaisesti niin sanottuja valokuvia (snapshot). Näiden valokuvien perusteella mahdollistetaan vahingon sattuessa tietojen palauttaminen juuri vahinkoa edeltäneeseen hetkeen.
CDP–järjestelmästä on selkeätä hyötyä aloilla, joissa liiketoiminta on taukoamatonta. Yhtenä esimerkkinä voisi pitää vaikka sähköistä kaupankäyntiä, jossa järjestelmän kaatuminen ja näin aiheutuva asiakastilausten menettäminen ja toiminnan pysähtyminen voi aiheuttaa huomattavia tappioita. Nykyisin yritysten toiminta saattaa myös lamaantua totaalisesti, jos sähköpostitoiminta jostakin syystä keskeytyy. Tällöin oikeanlaisen CDP–järjestelmän olemassa olo voi tulla korvaamattomaksi, koska tuhoutuneet / kadonneet sähköpostit pystytään palauttamaan takaisin välittömästi.

CDP–järjestelmien määrä on lisääntynyt tasaisesti parin viime vuoden aikana. Verkkosivusto networkingcomputing.com teki artikkeliaan Market Analysis: Continuous Data Protection, joka julkaistiin tammikuun 27. 2006, varten sivustollaan kyselyn, jossa se kysyi yritysten IT-henkilöiden aikomuksia CDP–järjestelmien suhteen. Kyselyssä 40 % vastaajista kertoi, että heillä on käytössään jo CDP–järjestelmä tai että sellainen otetaan käyttöön seuraavan 12 kuukauden kuluessa. Kyselyssä selvisi myös, että yli puolet CDP–järjestelmään investoineista oli suosinut hankinnassa CDP–laitteita. Loput vastaajista olivat jakautuneet hankinnoissaan tasaisesti ohjelmistojen ja levyryhmiin integroitujen järjestelmien kesken.

6.2 Ohjelmistot

Tällä hetkellä markkinoilla on jo useita ohjelmistotaloja, jotka tarjoavat CDP–järjestelmiä yrityksille. Kyseisiä yrityksiä ovat muun muassa IBM, Symantec, Microsoft, Availl, LiveVault, TimeSpring ja XOSoft. CDP–ohjelmistoissa on paljon eroja keskenään ja niiden hinnat vaihtelevat suuresti. Edellä mainittujen ohjelmistotalojen tuotteiden hinnat esimerkiksi vaihtelevat välillä 950 $ -25 000 $.

Ohjelmistoissa on isoja eroja keskenään, joten niitä valittaessa on niihin syytä syventyä kunnolla. Availlin ja LiveVaultin tuotteet tarjoavat tukea esimerkiksi SQL Serverille ja Microsoft Exchange Serverille, kun taas XOSoftin ja TimeSpringin tuotteet antavat jopa mahdollisuuden yksittäisen sähköpostin tai sähköpostilaatikon hakemiselle tietomassasta. Tuotteista taas XOSoftin ja LiveVaultin ohjelmat tukevat joitakin Linuxeja ja Unixeja, mihin muut ohjelmistot eivät anna mahdollisuutta. Lisäksi ohjelmissa on paljon eroja liittyen valokuvien määriin sekä niiden hallintaan ja automatisointiin. Ohjelmista halvin on tällä hetkellä Symantecin Backup Exec 10D, mikä myös voitti networkingcomputing.com:n testin eri ohjelmistojen välillä. (Marks 2006)

6.3 Laitteistot

CDP–laitteiden etu verrattaessa niitä CDP-ohjelmistoihin on niiden nopeampi asennus, prosessoritehon säästäminen ja käyttöjärjestelmä riippumattomuus. Laitteistovalmistaja pystyy lisäksi panostamaan tutkimus– ja kehitystyö rahansa tehokkaammin laitteen ominaisuuksiin, toimintoihin ja testaukseen, koska varoja ei joudu käyttämään käyttöjärjestelmämuunnoksiin. Tämän lisäksi erillinen laite on yleensä vakaampi, vähemmän altis virhetoiminnoille sekä helpommin eristettävissä oleva piste virheenkorjausta tehtäessä kuin ohjelmisto. Laitteiden heikkona puolena voidaan pitää korkeampaa hankintahintaa kuin ohjelmistoilla. CDP–laitteita valmistavat tällä hetkellä muun muassa Revivio, Sonicwall, Storactive, Lucid 8´s ja Mimosa Systems.

6.4 Esimerkki tapauksia CDP:n käytölle

CDP on erityisen hyödyllinen seuraavassa kolmessa tapauksessa:

1) Tiedostopalvelimien/ NAS–laitteiden tiedostot on tarkoitus suojata palvelinhuoneessa reaaliaikaisesti.
2) Haarakonttorin tiedostojen kopiointi aiotaan keskittää päätoimipaikalle.
3) Yrityksen kannettavien tietokoneiden tiedot päätetään varmuuskopioida silloin, kun koneet liitetään yrityksen verkkoon.

Ensimmäisessä tapauksessa korvataan kyseisten laitteiden yöllinen varmuuskopiointi nauhoille CDP–järjestelmällä. CDP-palvelimelta ajetaan yöllinen varmistus etäpisteen nauhoille tietojen varmistamiseksi. (Marks 2006)

Tapauksessa kaksi CDP–järjestelmällä replikoidaan haarakonttorin tiedostot päätoimipaikalle. Tällä tavoin saadaan kaikki sivutoimipisteet keskitetyn varmuuskopioinnin piiriin ja eliminoidaan mahdolliset käyttäjien virheet liittyen kyseisten pisteiden varmuuskopiointiprosessin seurantaan. Ideaalisessa tilanteessa sivukonttorille hankitaan erillinen varastopalvelin, jota käytetään valokuvien ja tiedostojen paikalliseen varastoimiseen. Tällöin tiedostojen palauttaminen on huomattavasti nopeampaa, koska tiedot löytyvät myös paikallisesta verkosta. (Marks 2006)

Tapauksessa kolme CDP–tekniikalla tallennetaan kannettavan tietokoneen tiedostojen muutokset kyseisen koneen kiintolevylle. Tekniikka, jolla muutokset taltioidaan, tuntee nimen journaling. Kun kone liitetään yrityksen verkkoon, informaatio tiedostojen muutoksesta liikutetaan palvelimelle, jossa verrataan vanhaa ja uutta tiedostoa keskenään ja tämän jälkeen viimeisimmät versiot tiedostoista ladataan palvelimelle talteen. (Marks 2006)

6.5 Haitat

CDP–järjestelmissä on olemassa myös selviä heikkouksia, jotka tekevät niistä vain osittaisia ratkaisuja tietojen varmistukseen. Jotkin järjestelmät eivät mahdollista kuin rajallisen määrän valokuvia. Tämä taas vaikuttaa järjestelmän tehokkuuteen, koska palautuspisteiden aika saattaa tulee liian lähelle kuluvaa hetkeä. Tästä johtuen järjestelmällä ei pystytä mahdollisesti palauttamaan tietoja kuin muutaman päivän takaa. Toisaalta runsas valokuvien määrä voi aiheuttaa huomattavia vaikeuksia palauttamistyöhön ryhdyttäessä, jos järjestelmässä ei ole kunnollista valokuvien (snapshot) hallintaa. (Marks 2005)

CDP–järjestelmät eivät ole vielä tässä vaiheessa siinä asti, että niihin kannattaisi luottaa ainoastaan ainoana tallennusmenetelmänä. Ne ovat erittäin toimivia sähköpostin ja tietokantojen palautuksissa, koska valokuvat ja tiedostot voidaan tallentaa näppärästi kiintolevylle ja palauttaa sieltä tarvittaessa nopeasti. CDP–järjestelmän rinnalla on kuitenkin järkevää pitää jonkin muukin tallennusmenetelmä. Nykyaikaisten nauhatekniikoiden ansioista nauhoille tallentaminen on edelleenkin erittäin suositeltava vaihtoehto tähän työhön. Nauhoilla järjestelmän toiminta voidaan turvata isommankin vahingon sattuessa, kunhan niiden asian mukaisesta säilyttämisestä ja sijoittamisesta on vain huolehdittu oikein. (Schuchart Jr. 2003)

7 Varmuuskopiointijärjestelmän toteuttaminen

7.1 Kolme erityyppistä toteutusta

Tallennusjärjestelmien toteuttamisen voi jakaa kolmeen tasoon. Ensimmäisessä tasossa yhteen palvelimeen hankitaan tallennusasema, mille data ohjataan. Tämä ratkaisu on yksinkertainen ja selkeä pienessä ympäristössä. Seuraavassa ratkaisumallissa yrityksellä on hallussaan erillinen tallennuspalvelin ja sen yhteydessä NAS-järjestelmä. Tämä yhdistelmä on monipuolinen ja helposti muokattavissa, mutta se ei vielä sisällä erityisen kehittyneitä tallennustekniikoita. Viimeisessä tasossa yrityksellä on hallussaan erillisiä SAN-laitteita, jotka ovat suunniteltu varta vasten datan tallennukseen ja siirtoon. Nämä järjestelmät ovat toiminta varmoja ja selkeästi kalliimpia kuin edellisen tason ratkaisut.

7.1.1 Yksinkertainen perustoteutus

Yrityksellä on muutamia palvelimia ja yhdessä niistä on nauha-asema. Data liikutetaan nauha-asemalle joko käyttämällä hyväksi Windowsin-käyttöjärjestelmässä olevia metodeja, kuten xcopy:a tai Windowsin backup:a tai sitten tämän toiminnon hoitaa jokin erillinen varmistusohjelma kuten Veritaksen Backup Exec.

7.1.2 Erillinen tallennuspalvelin yhdistetyllä levy- ja nauhajärjestelmällä

Yrityksellä on käytössään useita palvelimia ja yksi niistä on erityisesti varmuuskopiointia ja datan tallennusta varten. Tässä palvelimessa on käyttöjärjestelmänä Microsoftin Storage Server 2003 R2. Itse palvelin on niin sanottu keskiraskas palvelin ja sen yhteyteen on lisäksi kiinnitetty erillinen kiintolevy rypäs tai NAS-laite. Tässä laitteessa on kahden tyylisiä levyjä sisällään; SAS-levyjä sekä SATA-levyjä. Budjetista ja tarpeiden määrittelystä riippuen levyiksi voitaisiin mahdollisesti myös valita kuitukanavalevyjä. Käyttöjärjestelmä pitää huolen, että tuoreeksi ja tärkeäksi määritelty tieto (maksimissaan kaksi kuukautta vanhaa) on sijoitettu nopeammille SAS-levyille ja vanhempi, vähemmän käytetty tieto, on majoitettu halvemmille ja hieman hitaammille SATA-levyille. Lisäksi käyttöjärjestelmä pitää huolen, että samasta tiedosta ei ole turhia kopioita sijoittuneena ympäriinsä levylle. Käyttäjä ei tätä asiaa huomaa mistään sillä kaikessa on kyse järjestelmän tiedon sisäisestä indeksoinnista ja linkittämisestä. Palvelimeen on lisäksi liitettynä nauha-asema tai nauha-kirjasto, mihin liikutetaan jo hieman vanhentunut, mutta yhä tärkeäksi määriteltävä tieto. Näistä nauhoista osa liikutetaan paloturvakaappiin suojaan katastrofin varalta.

NAS-järjestelmä yhdistettynä tallennuspalvelimeen on toimiva ratkaisu, kun pääasiallisena tarkoituksena on tukea tiedostojen jakoa

sekä pitää huolta niiden helposta varmistuksesta. Kyseinen laitteistokokoonpano on helposti laajennettavissa ja sen hinta on maltillinen verrattuna SAN-toteutuksiin. Tämän tyylisen järjestelmän hinta on noin kolminkertainen verrattuna edellisen kohdan toteutukseen.

7.1.3 SAN-järjestelmä

SAN-järjestelmä muodostetaan käyttämällä hyväksi erillistä laitteistoa, jossa on niin sanottu äly-yksikkö ja tämä laite yhdistetään verkkoympäristöön ja palvelimiin omilla nopeilla liitännöillään. Näiden liitäntöjen muodostamien kanavien kautta on mahdollista muodostaa erillinen SAN-verkko, mikäli käytössä on lisäksi kuitukanavakytkin. Äly-yksikkönä voi palvella esimerkiksi Dell:n Dell|EMC CX-300. Tässä yksikössä käytetään nopeita ja äärimmäisen varmoja kuitukanavalevyjä, gigabitin verkkokortteja sekä erillisiä tallennus- ja varmennusohjelmistoja. Tämä laite on fyysiseltä rakenteeltaan täysin kahdennettu lähtien prosessoreista ja päätyen virtalähteisiin. Myös datan liikutuksessa käytettävät kuitukanavatekniikalla toteutetut tietoväylät ovat kahdennetut virhetoiminnallisuuksien varalta. Itse laite kytketään kuitukanavaliitännöillä keskeisiin palvelimiin joko suoraan tai sitten hankkimalla nopeaan datansiirtoon tarkoitettu kuitukanavakytkin, joka puolestaan

7.2 Esimerkki yritys

Otetaan kuvitteellinen yritys, jolle kehitetään kokonaisuudessaan toimiva tallennusratkaisurunko. Yrityksellä on 50 työntekijää, 45 työasemaa, 10 kannettavaa ja kaksi IBM:n palvelinta. Toinen palvelin on kolme vuotta vanha Windows 2003 Server ja toinen hieman uudempi, mutta samanlainen 2003 Server. Vanhempi palvelin toimii yrityksen tiedosto- ja ohjainpalvelimena ja uudemman tehtävänä on olla yrityksen Exchange–palvelin. Aikaisempi varmuuskopiointiratkaisu on tullut nyt tiensä päähän ja sen kapasiteetti on muuttunut hyvin rajoittuneeksi. Tarkoituksena on kohtuullisella budjetilla kehittää tallennusratkaisu kokonaan uusiksi niin, että se on ajantasainen sekä tulevaisuuden tarpeisiin skaalautuva. Yrityksen tallennettavan datan määrä on kohonnut kuluvan vuoden aikana noin 100 gigatavuun ja sen odotetaan kasvavan vähintään kolminkertaiseksi seuraavan kolmen vuoden kuluessa.

7.3 Tallennusformaatin valinta

Tällä hetkellä paras hinta / tallennettava gigatavu–suhde on tarjolla Ultrium2–sukupolven nauhoissa. Tosin tuotteiden hinnat elävät koko ajan, joten tilanne voi tältä osin olla erilainen jo vaikka seuraavalla viikolla. Ultrium2–nauhaan mahtuu pakattuna 400 gigatavua per nauha ja se on tällä hetkellä varmasti riittävä määrä tälle yritykselle. Kyseisten nauhojen kapasiteetti takaa myös sen, että tallennusvara ei

lopu aivan pian. Muutaman nauhan lisäyksellä kapasiteetin kasvu on huima ja markkinoilla on jo nyt olemassa Ultrium3–sukupolvi, missä kapasiteetti on tuplattu lukemiin 400 / 800 gigatavua. Järkevin vaihtoehto on hankkia Ultrium3–sukupolven nauha-asema ja ottaa siinä sitten käyttöön joko Ultrium1 tai Ultrium2-nauhat. Näin laitteisto antaa myöten tuleville skaalauksille ja hankinnoissa tarvitsee vain vaihtaa nauhan tyyppiä.

7.4 Tallennusohjelmisto

Keskisuuren yrityksen valinta Suomessa tallennusohjelmistokseen on yleisimmin Symantecin Backup Exec. Sen viimeisin versio on tällä hetkellä 10D. Symantecin suosio johtuu sen kohtuullisesta hankinta hinnasta ja toimivuudesta sekä hyvästä tuesta Windows–käyttöjärjestelmille. Ohjelmasta ilmestyy lisäksi tasaiseen tahtiin uusia versioita, sille löytyy paljon Suomen kielistä tukea ja se on hyvin monipuolisesti konfiguroitavissa erilaisiin ympäristöihin. Viimeisin 10D-versio sisältää myös CDP–ominaisuuden, mikä tekee siitä ideaalisen tallennusohjelmisto vaihtoehdon. Backup Exec tarvitsee toimiakseen keskivahvan palvelimen ja on suositeltavaa, että palvelimella, jolla tallennusohjelmisto suoritetaan, ei ajeta mitään muuta palvelinta erityisesti rasittavaa ohjelmistoa. Keskivahvalla palvelimella tarkoitetaan tässä tapauksessa yrityskäyttöön tarkoitettua palvelinta, joka ei ole valmistajan hintahaitarin ala- eikä yläpäässä vaan sijoittuu mallistossa keskivaiheille. Tällä tavoin voidaan varmistaa, että tallennusohjelmiston toimintaa ei ainakaan haittaa minkään toisen ohjelman resurssien varaus tai yleinen ohjelmallinen yhteensopimattomuus. Sinällään Backup Exec ei ole kuitenkaan erityisen herkkä häiriö tekijöille. (Symantec Corporation 2006)

7.5 Laitteisto

Toimiakseen Backup Exec tarvitsee palvelimen, joka voisi olla esimerkiksi IBM xSeries 226 seuraavalla kokoonpanolla:

Käyttöjärjestelmä:	Windows Server 2003
Prosessori:	Intel Xeon 2,8 GHz
Keskusmuistia:	2 Gigatavua DDR2
Kiintolevyt:	5 x 200 Gigatavua, RAID 5
DVD - asema:	8x

Palvelimeen hankitaan hot swap–kehikko, mikä mahdollistaa kiintolevyjen irrottamisen ja liittämisen lennossa. Lisäksi palvelimeen hankitaan lisäkorttina Adaptecilta Ultra 320 SCSI–ohjainkortti, mihin liitetään kiinni nauha–asema. Nauha–asemaksi tuli Ultrium–nauhojen valinnan myötä uusinta Ultrium3–sukupolvea tukeva HP StorageWorks Ultrium 960 External Tape Drive. Aluksi asemassa otetaan käyttöön Ultrium2–sukupolven nauhat, koska niiden kapasi-

teetti on varsin riittävä olemassa olevalle datamassalle. Palvelimeksi Backup Execille sopii selkeästi tehottomampikin palvelin, mutta vanhaa laitteistoa on turha hankkia uutta ratkaisua luotaessa.

7.6 Nauhakierto ja varmuuskopiointiprosessi

Nauhakierron hallintaan sovelletaan yleisesti käytössä olevaa ja yksinkertaista GFS- rakennetta ja tämän rakenteen mukaisesti nauhoja hankitaan 22 kappaletta. Tarkoituksena on, että kuun jokaiselle arkipäivälle on olemassa oma nauhansa. Nauhoja on siis 16 arkipäivänauhaa (kuun maanantait–torstait), 3 viikkonauhaa (kuun kolme ensimmäistä perjantaita) ja 3 kuukausinauhaa (kuun viimeinen perjantai). Yhtenä mahdollisuutena on tietenkin hankkia ainoastaan yhden viikon mittaiselle ajanjaksolle omat arkipäivänauhansa, joita sitten kierrätetään joka viikko. Tällä tavalla nauhahankinnoissa on mahdollista säästää rahaa.

Nauhat nimetään loogisesti käyttöpäivien mukaan. Esimerkiksi kuun ensimmäisen maanantain nauha on ”set 1 Monday” ja seuraavan viikon tiistain nauha on ”set 2 Tuesday” ja näin edelleen. Kuukausinauha viedään kerran kuukaudessa turvaan pankin tallelokeroon. Tämä tehdään joka kuukauden ensimmäisenä maanantaina. Nauhoista tehdään lista ja listaan päivitetään tiedot siitä, mitä kyseisellä nauhalla sijaitsee. Tätä listaa säilyttää IT–henkilöstö ja päivitetty

kopio siitä toimitetaan aina pankin tallelokeroon talteen, kun nauhoja viedään sinne turvaan. Joka kolmas kuukausi kuukausinauhoista vanhin haetaan tallelokerosta ja se otetaan käyttöön uudelleen seuraavan kuukauden kuukausinauhaksi.

Backup Exec konfiguroidaan ottamaan lisäävä varmistus yrityksen tiedostoista maanantaista torstaihin ja perjantaisin ajetaan täysi varmistus. Yrityksellä ei ole toimintaa viikonloppuisin, joten viikonloppuna ei ole tarvetta ajaa erillisiä varmistuksia. Tosin on huomioitava, että varmistukset ajetaan aina öisin, joten perjantain varmistuskin sinällään otetaan talteen vasta lauantain puolella. CDP–ominaisuus konfiguroidaan toimintaan Backup Execistä ja sen avulla otetaan valokuvat sekä tiedostopalvelimesta että Exchange–palvelimesta ja kyseisten palvelimien tiedostot ja valokuvat talletetaan uudelle palvelimelle. Kannettaviin tietokoneisiin asennettava agenttiohjelma varmistaa, että kannettavien liittyessä työyhteisön sisäverkkoon, otetaan niistäkin varmistus. Varmistukset ajetaan aina ensiksi palvelimen kiintolevylle, josta ne sitten siirretään yön aikana nauhalle.

Nauhojen vaihdosta vastaa sille määritelty henkilö. Kyseisen kokoisessa yrityksessä tehtävään määrätään IT-henkilöstöstä yksi henkilö. Hänen tehtävänään on joka aamu hakea kyseisen päivän nauha palo-

kaapista ja vaihtaa se nauha- asemaan edellisen päivän nauhan tilalle. Tehtävään kuuluu myös tarkistaa, että edellisen yön ajo on mennyt oikein lävitse ja että mitään poikkeuksia tai virheitä ei ole sattunut ajon aikana. Vastaavan tehtäviin kuuluu myös varmistaa, että nauhoja kierrätetään oikein ja ajallaan ja että uusia nauhoja hankitaan sopivin aikavälein vanhojen tilalle.

7.7 Nauhojen säilyttäminen

Yritykselle hankitaan uusi palokaappi nauhojen säilyttämistä varten. Palokaapin pitää olla paloturvallinen ja sen tulee kestää perinteiset ryöstöyritykset. Kaapin valmistajalle ei ole merkitystä, mutta ainakin Kasola-konsernin palokaapit ovat VTT:n testaamia. Kaappi sijoitetaan yrityksen toimitiloihin, mutta sellaiseen paikkaan, jossa se ei ole ensimmäisenä vahingoittumassa tai varkaiden ulottuvilla. Optimaalisessa tilanteessa kaapin sijoituspaikka olisi tila, missä on jo etukäteen otettu rakenteiden osalta huomioon tulipalot ja vesivahingot. Kaapin avainta säilytetään sille määrätyssä turvallisessa paikassa ja sitä ei luovuteta kuin asianomaisille henkilöille, jotka ovat vastuussa tallennusprosessin valvomisesta. Avaimesta on optimaalisessa tilanteessa olemassa kolme kopiota, joista yksi on vastaavalla IT-henkilöllä, yksi toimitusjohtajalla ja yksi pankin tallelokerossa.

7.8 Tulevaisuuden järjestelmämuutokset

Kyseinen järjestelmä on toimintakuntoon konfiguroinnin jälkeen tehokas ja toimiva ratkaisu kyseisen kokoiselle yritykselle. Tulevaisuuden kasvavia tilantarpeita varten nauha – asema mahdollistaa jo nyt seuraavan sukupolven nauhojen käyttöönottamisen nauhakierrossa. Tämän lisäksi itse palvelimeen pystytään helposti lisäämään uusia kiintolevyjä tilantarpeen niin vaatiessa. Tulevia muutoksia itse tallennusratkaisussa voisi olla esimerkiksi CDP–järjestelmän käytön monipuolistaminen siten, että yritys perustaa itselleen etätietokeskuksen, missä sijaitsevalle palvelimelle tiedot replikoidaan reaaliaikaisesti. Tämän kaltainen järjestely olisi järkevintä rakentaa tällöin reaaliaikaisen levyryhmien peilauksen kautta niin, että replikointitekniikalla siirretään tiedot reaaliaikaisesti etäpisteen palvelimen vastaavan kaltaiseen levyryhmään. Kyseisen järjestelmän toteuttaminen ei ole halpaa, mutta sen kunnollinen käyttöön saattaminen mahdollistaa katastrofitilanteissa nopeamman ja tehokkaamman palauttamisen onnistumisen. Tämänkaltaisen järjestelmän toteuttamisen yhteydessä kysymykseen voisi tulla jo SAN-järjestelmä ratkaisut.

8 Yhteenveto

8.1 Työn rajaus ja siihen liittyneet ongelmat

Työn rajausta vaikeutti huomattavasti IT-järjestelmille tyypillinen ”kaikki liittyy samaan kokonaisuuteen” -ongelma. Kun järjestelmiä ja laitteistoja verkotetaan, ne liitetään yhteen, jonka jälkeen niiden eriyttäminen kokonaisuudesta ei ole helposti tehtävä asia. Tästä johtuen varmuuskopiointijärjestelmät ovat aina uniikkeja toteutuksia, jotka ovat räätälöityneet kyseisen yrityksen tarpeisiin. Näiden järjestelmien toiminnasta ja toimivuudesta paras tieto onkin tällöin yleensä kyseisissä yrityksissä itsessään eivätkä ne luonnollisesti anna tehdyistä ratkaisuistaan ilmoituksia ulkomaailmaan.

Varmuuskopioinnin voi jakaa kolmeen aihealueeseen, jotka ovat vapaasti englannista käännettynä: arkistointi, tallentaminen ja palautus. Kaikki nämä aihealueet ovat kiinteästi kytköksissä toisiinsa. Niihin liittyvät laitteistot, ohjelmat ja konfiguroinnit ovat näin ollen luonnollisesti riippuvaisia paljolti toisistaan. Jokainen kolmesta aihealueesta on jo itsessään varmasti tarpeeksi laaja, jotta siitä voisi tehdä helposti yhden päättötyön.

Päädyin lopulta ratkaisuun, jossa käsittelen pelkästään varmuuskopioinnin alkuvaihetta. Eli kerron mitä tulisi ottaa huomioon varmuuskopiointia suunniteltaessa, mitä vaihtoehtoja on valittavana ja miten järjestelmän voisi sitten toteuttaa keskikokoiselle yritykselle. Tein siis tietoisen valinnan ja jätin työstäni pois useita siihen liittyviä aihealueita kuten esimerkiksi: katastrofista palautumisen ja SAN-verkon sekä erilaiset peilaustoteutukset.

Aiheenrajauksen tekeminen oli minulle vaikea asia ja en ole vieläkään varma, että se onnistui parhaalla mahdollisella tavalla. Rajauksesta voi olla varmasti montaa mieltä, mutta kuitenkin kaiken kaikkiaan olen ratkaisuuni kohtuullisen tyytyväinen.

8.2 Työskentely

Työn tekeminen kesti kokonaisuudessaan noin neljä kuukautta vaikkakin sen olisi varmasti saanut valmiiksi puolta lyhyemmässä ajassa pakottavan tarpeen niin vaatiessa. Suurin osa työn tekemiseen liittyvästä ajasta menikin asioiden selvittämiseen ja kokonaiskuvan luomiseen varmuuskopiointikentän tilanteesta. Pääasiallisesti suurin osa materiaalista työhön liittyen on haettu ja koostettu eri puolilta Internetiä. Tämän kaltainen tiedon hakeminen ja koostaminen on aikaa vievää, koska löydetyn materiaalin sisältöön sokeasti luottaminen ei luonnollisesti järkevää. Näin ollen monen tiedon kohdalla asioita joutui selvittämään useasta eri lähteestä luotettavamman kuvan saamiseksi asiasta.

8.3 Hyödyt

Tämän työn tärkein hyöty on siinä, että se tarjoaa selkeän erittelyn varmuuskopioinnin perusteista sekä varmuuskopiointiin vaikuttavista tekijöistä. Työn rajauksen takia työ luonnollisesti keskittyy vain tiettyyn osaan varmuuskopioinnissa, mutta sen avulla on varmasti mahdollista luoda toimiva suunnitelma pk-yrityksen varmuuskopiointiratkaisuksi. Työn lukemalla on mahdollista määrittää niitä tekijöitä, jotka merkitsevät kunnollisen varmuuskopiointiratkaisun luomisessa sekä saada apua omien ajatustensa selvittämiseksi asian tiimoilta. Työ antaakin ehkä eniten niille henkilöille, jotka eivät ole vielä kunnolla tutustuneet varmuuskopiointiin. Mikäli tämä työ pystyy auttamaan edes yhtä ihmistä, joka on suunnittelemassa varmuuskopiointiratkaisua, on se jo sinällään onnistunut loistavasti. Itselleni työn tekeminen antoi paljon uutta näkemystä varmuuskopiointiin sekä niihin asioihin, jotka on otettava huomioon, kun järjestelmää lähtee miettimään. Työni aihepiirin valinnan alkuperäisenä ajatuksena olikin, että se hyödyttäisi lähinnä minua itseäni ja antaisi minulle mahdollisuuden ymmärtää aihealuetta syvällisemmin. Omasta mielestäni tässä asiassa onnistuinkin varsin hyvin, joten olen työhöni ihan tyytyväinen.

8.4 Loppusanat

Jo vuosia eri ammatilliset tahot ovat ennustaneet, että nauhalle tallentamisen loppu on häämöttämässä ja että markkinoita tulee hallitsemaan selvästi kiintolevytallennus. Nauhalle tallentamisen loppua on aavisteltu sen takia, että kriitikot kuvittelivat, että nauhavarmistuksessa on tultu kapasiteetin kasvun määränpäähän. Markkinoilla olevien kiintolevypohjaisten ratkaisujen määrä onkin lisääntynyt selvästi, mikä on paljolti seurausta NAS–laitteiden määrän lisäyksestä sekä CDP-järjestelmien käyttöönotosta. Kuitenkin samaan aikaan nauhatekniikka on ottanut omalla tahollaan niin merkittäviä harppauksia eteenpäin kapasiteetin kehityksessä, että sen merkitys tallennusmenetelmänä varmuuskopioinnissa ei ole kadonnut mihinkään. Kiintolevyjärjestelmät tulevat kuitenkin tulevaisuudessa varmasti kasvattamaan osuuttaan markkinoista, mutta nauhan edes osittaista kuolemaa ei tulla näkemään vielä useisiin vuosiin.

Tulevaisuuden suuntauksina varmuuskopioinnissa näyttäisi olevan tällä hetkellä erilaiset virtualisointitekniikat SAN–järjestelmien yhteydessä sekä replikointijärjestelmien muunnelmat. Erityisesti replikointijärjestelmät, jotka siirtävät tietoa kahden pisteen välillä levyryhmistä toisiin, ovat yleistymään päin. Tällä hetkellä kuitenkin kyseisten järjestelmien kustannukset ovat vielä sen verran merkittäviä, että niitä ei ole käytössä kuin isoilla ja kohtuullisen hyvin rahoitetuilla organisaatioilla. Vasta tulossa olevista tekniikoista, holografinen tallennus vaikuttaisi olevan seuraava todella merkittävä askel tallennuksessa. Tällä hetkellä ainakin Hitachi Maxell on suunnitellut

julkaisevansa ensimmäisiä tätä tekniikkaa käyttäviä tuotteita tämän vuoden loppupuolella.

Aloittaessa työtäni olin pitkän aikaa tuuliajolla materiaalin kanssa. Vasta kun sain aikaiseksi kuvitteellisen sisällysluettelon ja aihealueen rajauksen niin tunsin, että ryhdyin saamaan aikaiseksi jotakin konkreettista. Pitäisinkin kyseisiä seikkoja tärkeimpinä työhöni vaikuttaneina raja-arvoina.

Viimeisenä haluan kiittää kaikkia työn lukeneita ja toivon, että tulevaisuudessa työstäni voi olla apua myös muille varmuuskopioinnista kiinnostuneille ihmisille.

9 Lähteet

Toigo Jon William, 2003. Disaster Recovery Planning preparing for the unthinkable third edition. New Jersey: Prentice Hall

Marks Howard, 2005. Product Analysis: Continuous Data Protection Products http://www.networkcomputing.com/showArticle.jhtml?articleID=177103369&pgno=5

Schuchart Jr Steven, 2003. Making the Case for Both Disk and Tape http://www.networkcomputing.com/showitem.jhtml?docid=1404ws1

9to5computer.com http://www.9to5computer.com/howtobuyatapedrive.htm

AIT Technology Forum http://www.aittape.com/ait-tape-roadmap.html

Seagate Technology LLC http://www.seagate.com/products/tapesales/backup/A2g1.html

Symantec Corporation http://www.symantec.com/small_business/products/overview.jsp?pcid=br&pvid=bewin_svr

Ultrium.com http://www.ultrium.com/newsite/html/format.html

10 Lyhenteet

10.1 Käytettyjen lyhenteiden ja termien selitykset

DDR2

DDR2 (double data rate SDRAM) on toisen sukupolven DDR–muistia. Se eroaa edellisestä DDR-sukupolvesta korkeampien väylänopeuksien ja latenssien puolesta.

Exchange-palvelin

Exchange on Microsoftin tekemä palvelinohjelmisto yritysten sähköpostin vastaanottamiseen ja lähettämiseen.

Journaling

Journaling tarkoittaa vikasietoista tiedostojärjestelmää, missä datan eheys varmistetaan kirjoittamalla tiedostoihin tulevat muutokset ensin sarjalokitiedostoon. Vasta tämän tapahtuman jälkeen alkuperäinen kiintolevyn lokitiedosto päivitetään. Näin ollen järjestelmävirheen sattuessa lokitiedoston tietojen avulla järjestelmä pystytään palauttamaan takaisin virhettä edeltäneeseen tilaan. Tällä tavalla pystytään myös palauttamaan tiedot, joita ei ollut vielä ehditty tallentaa ja nämä tiedot on mahdollista palauttaa siihen paikkaan mihin ne olivat menossa.

NAS

NAS tulee sanoista Network-Attached Storage ja sillä tarkoitetaan kiintolevyillä varustettu erillistä laitetta, jota voidaan käyttää tarjoamaan erinäisiä resursseja verkon käyttäjille. Laitteissa on yleensä useita kiintolevyjä, jotka ovat yhdistetty toisiinsa RAID–tekniikalla nopeuden ja/tai vikasietoisuuden saavuttamiseksi. Tämän lisäksi kyseisissä laitteissa on erillinen verkkoliitäntä, joten sitä voidaan käsitellä yksittäisenä laitteena verkon sisällä.

RAID 5

RAID 5 on vähintään kolmen kiintolevyn levyohjaimella muodostama joukko, jossa tieto kirjoitetaan pieninä palasina tasaisesti kaikille levyille. RAID 5 mahdollistaa kirjoitus– ja lukuprosessit eri levyiltä yhtäaikaisesti. Jokaisen kirjoittamisen yhteydessä levyille lisätään lisäksi pariteetti tieto, jonka avulla on mahdollista myöhemmin uudelleen rakentaa kadonnut tieto yhden kiintolevyn rikkoutuessa.

SAN

SAN on lyhenne sanoista Storage Area Network ja se tarkoittaa suomennettuna tallennusverkkoa. Tämä tarkoittaa käytännössä tekniikkaa, jossa palvelimien massamuistit yhdistetään toisiinsa. SAN –verkossa tallennuskapasiteetti on erillään käyttäjien verkosta ja yhdistävänä tekijänä käytetään jonkinlaista käyttöön tarkoitettua kuitukanava–kytkintä.

Snapshot

Snapshot on kuvaus kiintolevyllä olevista tiedoista, jonka perusteella on mahdollista palauttaa tiedot, niin tarvittaessa. Näitä niin sanottuja valokuvia on olemassa kaksi päätyyppiä. Toinen on copy-on-write tai low-capacity–valokuva ja toinen tyyppi on split-mirror–valokuva. Ensimmäisessä vaihtoehdossa tuloksena on aina uusi valokuva, kun tiedostoon tulee muutoksia tai uutta tietoa luodaan. Jälkimmäisessä vaihtoehdossa luodaan aina uusi valokuva koko osiosta eikä vain muuttuneesta tai uudesta datasta. Tätä tekniikkaa käytetään peilattujen kiintolevyjärjestelmien yhteydessä.

SQL Server

SQL Server on Microsoftin tekemä relaatiotietokantojen hallintajärjestelmä, joka on tarkoitettu yritysten käyttöön. Se lisää laajennuksia ja moninaisia ominaisuuksia normaaliin SQL:n. Ensimmäinen versio kyseisestä ohjelmasta kehitettiin 80 – luvun lopulla (SQL Server 4.2 for OS/2) ja uusin versio kyseisestä ohjelmasta kantaa nimeä SQL Server 2005.

Xeon

Xeon on nimitys Intelin palvelinprosessoriperheelle. Xeon–mallin prosessorit ovat tarkoitettu nimenomaan yrityskäyttöön tuleviin palvelimiin ja niiden arkkitehtuuri on selvästi eroava työasema prosessorien arkkitehtuurista. Xeon-prosessorit ovat Intelin korkeimman teknologian prosessoreita ja malliperhe onkin ollut jatkuvan kehityksen kohteena koko sen olemassa olo ajan, joka on noin kymmenen vuotta. Ensimmäinen Xeon–prosessori oli nopeudeltaan 400 MHz.