

# DATAN ARKISTOINTI

LAHDEN  
AMMATTIKORKEAKOULU  
Tekniikan ala  
Tietotekniikan koulutusohjelma  
Tietoliikennetekniikka  
Opinnäytetyö  
Syksy 2015  
Eino Kolli

Lahden ammattikorkeakoulu  
Tietotekniikan koulutusohjelma

KOLLI, EINO:

Datan arkistointi

Tietoliikennetekniikan opinnäytetyö, 40 sivua

Syksy 2015

TIIVISTELMÄ

---

Opinnäytetyön tavoitteena oli suunnitella datan arkistointiratkaisu LSK Electrics Oy:lle. LSK Electrics Oy:lle on kertynyt paljon projektikohtaista dataa, jonka tallessa pysyminen on yrityksen toiminnalla erittäin tärkeää. Data sisältää tietoa esimerkiksi yrityksen ylläpitosopimusten alaisista tehtaista ja projekteista. Tiukentuneet vaatimukset laadunvalvonnassa lisäävät yritykselle painetta tallentaa entistä enemmän tiedostoja. LSK Electrics Oy:n palvelimien kovalevypakan tallennustila on loppumassa piakkoin, joten datan arkistointiratkaisu vapauttaisi tilaa kovalevypakalta.

Datan arkistoinnilla on monta toteutustapaa. Arkistointi voidaan toteuttaa paikallisesti, pilvipohjaisesti tai siirrettävien medioiden avulla. Paikallisesti toteutettuna arkistointiratkaisu sisältää joko kovalevypakan tai NAS-kovalevyn. Pilvipohjaisesti toteutettuna arkistointiratkaisu ostetaan ulkoiselta palveluntarjoajalta, kuten Googlelta, ja siitä maksetaan kiinteä kuukausimaksu. Siirrettävillä medioilla toteutettu arkistointiratkaisu toimii hyvin pienillä yrityksillä ja kotitoimistoilla. Siirrettäviin medioihin pohjautuvassa ratkaisussa tieto tallennetaan esimerkiksi ulkoisille kovalevyille.

Kiintolevyjen tallennustekniikoilla on suuri merkitys paikallisen arkistointiratkaisun kehittämisessä. Tallennustekniikoista tärkeimmät ovat RAID ja Hot swap. RAID-tekniikat tarjoavat mahdollisuuden suojautua kovalevyjen hajoamisilta. Yleisin RAID-tekniikka palvelimissa ja levypakoissa on RAID5. Hot swap -tekniikka mahdollistaa kovalevyjen vaihtamisen, ilman että palvelinta tai tietokonetta tarvitsee sammuttaa vaihdon aikana. Hot swap -tekniikka on yleistä nykyisissä kovalevyratkaisuisissa.

LSK Electrics Oy:n datan arkistointiratkaisu tullaan toteuttamaan HP:n laitteita käyttäen paikallisena arkistointiratkaisuna. Yrityksen palvelintiloihin sijoitetaan uusi kovalevypakka. Kovalevypakasta voidaan vapauttaa tilaa myös yrityksen palvelimien käyttöön.

Asiasanat: arkistointi, data, NAS, RAID, tallennus

Lahti University of Applied Sciences  
Degree Programme in Information technology

KOLLI, EINO:

Data archiving

Bachelor's Thesis in Telecommunications technology, 40 pages  
Autumn 2015

ABSTRACT

---

The goal of this thesis was to plan a data archiving solution for LSK Electric Company. LSK Electric Co. have collected huge amounts of data, especially data on their various projects. For the company it is very important to keep that data secure. This data contains information of maintenance services that are sold to for example factories. Tightened quality control requirements have caused the company to save more data than before. LSK Electric Corporation's hard drive rack has used almost all of its storage space, and therefore a new data archive solution would release a lot of space for servers.

There are many ways to create a data archive. The archiving can be implemented locally, as cloud-based, or by using movable media devices. As locally implemented, the archiving solution contains either a hard drive rack or a NAS -hard drive. The cloud-based archiving solution will be purchased from an external service provider, such as Google and it will be paid for with a fixed monthly fee. For small offices and home offices the archiving solution that is made with movable media can be a good choice. When the solution is made with movable media, the data will be saved for example to external hard drives.

When creating a local archiving solution, different data handling techniques have a huge effect on the solution's outcome. The most important techniques are RAID and Hot swap techniques. The RAID- technique offers redundancy for hard drives. The most common RAID technique for hard drive racks and for servers is RAID5. This Hot swap technique offers a chance to switch broken hard drives on the go. Servers and racks do not need to be shut down for this. The Hot swap technique is common in all modern hard drive solutions.

The data archiving solution for LSK Electric Company will be implemented by using the local archiving solution and HP's equipment. There will be a new hard drive rack and it will be located in the server room of the company. The new hard drive rack will provide plenty of space, so some of the free space can be released for the company's servers.

Key words: archive, data, NAS, RAID

## SISÄLLYS

1	JOHDANTO	1
2	DATAN ARKISTOINTI	2
2.1	Data	2
2.1.1	Data tietotekniikassa	3
2.1.2	Data muilla osa-alueilla	4
2.2	Arkistointi	5
2.2.1	Tietojen suojaaminen	5
2.2.2	Datan arkistoinnin avulla saavutettavat säästöt	6
2.2.3	Tietojen palautus ja varmuuskopiointi	6
2.3	Verkkotekniikka	10
2.3.1	OSI-malli	10
2.3.2	TCP/IP-malli ja kerrokset	12
2.3.3	IPv4 avaruudet ja maskit	14
2.4	Kiintolevyjen tallennustekniikat	15
2.4.1	RAID	16
2.4.2	Hot swap	19
2.5	Datan pakkaaminen	21
2.6	Pakkausmenetelmät (kompressio)	21
2.7	Datan arkistoinnin tallennusmediat	24
2.7.1	Paikallinen arkistointi	24
2.7.2	Pilvi tallennustilat	25
2.7.3	Siirrettävät mediat	27
2.8	Datan tallennus lähtökohta ja ongelmat	28
3	DATAN TALLENNUS RATKAISUN TOTEUTUS LSK ELECTRICS OY:LLE	30
3.1	Laitteistoratkaisut	32
3.1.1	Kovalevypakka	32
3.1.2	NAS-palvelin	34
3.2	Ohjelmistoratkaisut	36
3.3	Arkistointiratkaisuehdotus	38
4	YHTEENVETO	39
	LÄHTEET	41

## SYMBOLI- JA LYHENNELUETTELO

CD	Compact Disk. Optinen digitaalisen tiedon tallennusmedia.
DNA	Deoksiribonukleiinihappo. Ihmisen perimä.
DVD	Digital Versatile Disk. Digitaalinen video-/monikäyttölevy, optinen datan tallennusväline.
Gb	Gigabitti. Gigabitti on 1 000 000 000 eli $10^9$ bittiä (miljardi). Informaation määrän mitta.
Gt	Gigatavu. Gigatavu on 1 000 000 000 eli $10^9$ tavua (miljardi). Tietotekniikassa käytettävä mittayksikkö.
IP	Internet Protocol. Internetprotokolla, joka huolehtii IP-tietoliikenne-pakettien toimittamisesta perille pakettikytkentäisessä verkossa.
IPv4	Internet Protocol versio 4. Internetprotokollan neljäs versio.
IT	Informaatioteknologia.
JPEG	Joint Photographic Experts Group. Kuvanpakkaus-formaatti. Tiedostomuoto.
LAN	Local Area Network. Lähiverkko.
Mb	Megabitti. Megabitti on 1 000 000 eli $10^6$ bittiä (miljoona). Informaation määrän mitta.
Mt	Megatavu. Megatavu on 1 000 000 eli $10^6$ tavua (miljoona). Tietotekniikassa käytettävä mittayksikkö.
MPEG	Moving Picture Experts Group. Ryhmä, joka suunnittelee nykyaikaisia videonpakkaustapoja.
NAS	Network-attached Storage tai Network Access Storage, eli verkkotallennusjärjestelmä.
OSI	Open Systems Interconnection Reference model. Kuvaa tiedonsiirtoprotokollien yhdistelmän seitsemässä kerroksessa.
USB	Universal Serial Bus on sarjaväyläarkkitehtuuri ohjeislaitteiden liittämiseksi tietokoneeseen.
RAID	Redudant Array of Independed Disks on Tekniikka, jolla lisätään kovalevyjen vikasietoisuutta tai/ja nopeutta.
SAS	Serial Attached SCSI eli sarjaankytketty SCSI on tietokoneväylä.
SATA	Serial ATA on sarjamuotoinen liitäntä sisäisen tai ulkoisen massamuistilaitteen kytkemiseen tietokoneeseen.

- SSD Solid State Drive, eli puolijohdeasema. Kovalevy, jossa ei ole lainkaan liikkuvia osia.
- TCP Transmission Control Protocol on tietoliikenneprotokolla, jolla luodaan yhteyksiä tietokoneiden välille.
- UDP User Datagram Protocol on yhteydetön protokolla, joka ei vaadi yhteyttä laitteiden välille, mutta mahdollistaa tiedostojen siirron.

# 1 JOHDANTO

Työn tavoitteena on suunnitella LSK Electrics Oy:lle uusi ratkaisu kertyvän datan loppusijoitukseen. Datan loppusijoituksella tarkoitetaan yritykselle paljon kertyvän arkistoitavan datan sijoittamista niin, että kaikki loppusijoitettava data on samassa sijainnissa. Loppusijoitettuna data ei kuormita yrityksen palvelimia levytilan puolesta ja loppusijoitettu data on helposti koko yrityksen henkilöstön saatavilla. Yrityksen tekemistä projekteista ja urakoista jää aina digitaalista informaatiota, jonka tulisi säilyä kymmeniä vuosia. Datan loppusijoittaminen sopii yrityksen projektien ja urakoiden tallennussijainniksi. Datan arkistointi on yritykselle erittäin ajankohtaista, sillä tieto lisääntyy valtavasti. Yritykseen kohdistuvat vaatimukset tallennetun tiedon osalta ovat lisääntyneet: tiedon ei enää vain tule olla saatavilla, vaan tiedon pitää olla todistettavissa laadunvalvontaa varten. Tallennettavien tiedostojen koko on lisääntynyt, sillä laatu tiedostoissa on parantunut viime vuosina huomattavasti.

Opinnäytetyössä käsitellään datan tallennusratkaisuja, verkkotekniikkaa, kiintolevyjen tallennustekniikkaa ja datan pakkaamista teoriassa käytännön esimerkein. Työssä vertaillaan datan arkistointiratkaisujen sopivuutta yrityksen tarpeisiin nähden. Samoin opinnäytetyössä vertaillaan mahdollisia laiteratkaisuja datan arkistoinnin osalta. Yhteenvedona käydään läpi työn onnistuneisuutta sekä tavoitteiden onnistumista.

## 2 DATAN ARKISTOINTI

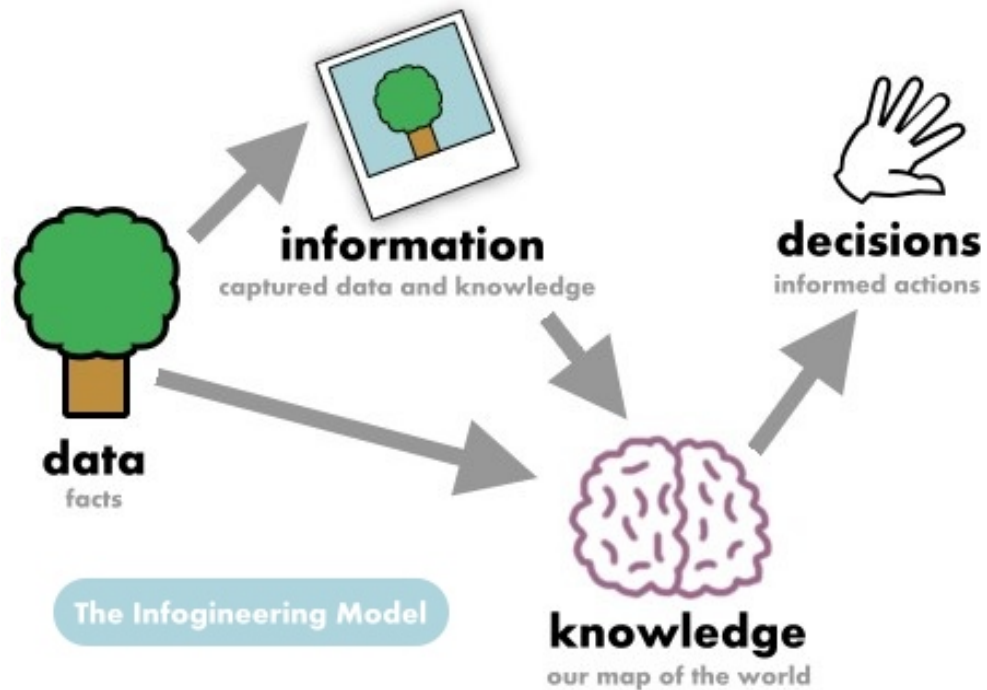
Datan arkistoinnilla tarkoitetaan tiedon, joka ei ole aktiivikäytössä, siirtämistä erilliseen tallennussijaintiin. Arkistot sisältävät tietoa, joka on vanhaa, mutta tulevaisuudessa vielä tärkeää. Yleisesti data arkistoidaan joko sisäisille tai ulkoisille kovalevyille sekä vaihtoehtoisesti muille siirreltäville medioille. Arkistoimalla dataa mahdollistetaan tiedon säilyvyys ja luettavuus helposti. (Rouse 2015.)

Oikein toteutettuna arkistointi voi olla turvallinen ja varma tapa pitää digitaalista tietoa tallessa useita vuosia. Rakenteeltaan järkevästi toteutetusta arkistosta data on helposti saatavilla.

### 2.1 Data

Data on esitys, joka käsiteltäessä muuttuu merkitykselliseksi. Itsessään datalla ei ole merkitystä. Kun dataa käsitellään, siitä syntyy useimmiten informaatiota ja lopulta tietoa. Suomen kielessä sanaa data käytetään myös tarkoitettaessa tietoa. (Mitra 2012.)





KUVIO 1. Datan muutos tiedoksi (Infogineering 2015)

Datan ja tiedon ero on pieni, suhteellinen. Data muuttuu usein paljon käsiteltäessä, ja data onkin usein aluksi vain ”raakaa” dataa seuraavalle vaiheelle. Lyhyesti: kaikki faktat ovat dataa. Esimerkkinä voidaan pitää vuoren korkeutta. Vuoren korkeus on dataa ja fakta: n-metriä. Tämä data voi olla kirjassa, joka sisältää muuta dataa kyseisestä vuoresta, vaikkapa tutkittaessa parasta keinoa kiivetä vuorelle. Kun tätä dataa ymmärtäen tehdään valinta parhaasta keinosta kiivetä vuorelle, siitä tulee tietoa. (KUVIO 1.) (Sharon 2015).

### 2.1.1 Data tietotekniikassa

Tietotekniikassa data on pääasiassa bittejä, jotka oikeassa järjestyksessä ja oikein luettuina ovat tallennettua tietoa. Dataa esiintyy esimerkiksi kovalevyillä ja muisteissa. Datasta puhuttaessa yleensä tarkoitetaan tallennettua tietoa.

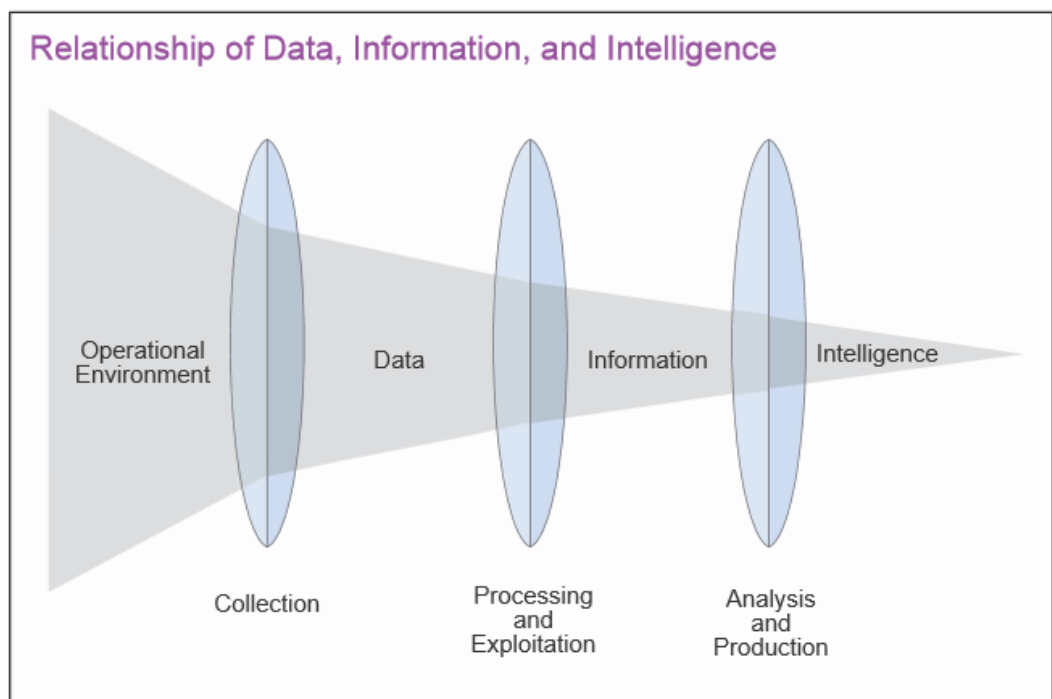
Data voidaan jaotella karkeasti kolmeen luokkaan. Ensimmäinen luokka on niin sanottu perinteinen data: datan rakenne on selkeä ja ennalta määritelty. Toiseen

luokkaan kuuluu löyhemmin määritelty data, kuten internetsivuston keräämät lo-kitiedostot. Toisessa luokassa data sisältää tietoa, joka on merkitty ennalta määrä-tyllä tavalla. Löyheämmin määritellystä datasta pienenkin tiedon etsiminen voi olla vaikeaa. Kolmannessa ja viimeisessä luokassa datalla ei ole lainkaan raken-etta. Tässä luokassa olevasta datasta ei voida tehdä oletuksia siitä, mitä data tulee sisältämään tai missä muodossa data on. Esimerkkinä kolmannesta luokasta on so-siaalisista medioista lähetetyt julkiset päivitykset. (Mitra 2012.)

### 2.1.2 Data muilla osa-alueilla

Dataa on muuallakin kuin tietotekniikassa. Dataa ovat esimerkiksi yksittäiset ään-teet puheessa, kirjaimet painetussa tekstissä ja emäspari DNA:ssa. Puhekielessä harvemmin käytetään edellä mainituista nimitystä data.

Lähes kaikki mitä ihminen tutkii, on dataa. Esimerkiksi rikoksia tutkittaessa kaikki kerätyt näytteet ja tiedot ovat dataa, kunnes data on muunnettu informaati-oksi. (KUVIO 2.) (Sharon 2015.)



KUVIO 2. Datan, informaation ja älykkyyden suhde (Department of Defence 2013)

## 2.2 Arkistointi

Arkistointi tarkoittaa yleisesti jonkinlaista järjestelmää, jonka kohteena ovat kuva-, äänite-, kirjoitus- ja tietokonetuotosten tallennus. Arkistointi kattaa myös edellä mainittujen luettelointia, säilytystä, kuljetusta ja hallinnointia. Sähköisellä arkistoinnilla tarkoitetaan tietokonetuotosten – eli datan – sijoittamista pitkäkestoiseen tallennusmediaan, josta tiedot ovat helposti saatavilla. Sähköisessä arkistoinnissa datan tulee olla hyvin hallittua.

Nykytrendinä on jo pidemmän aikaa ollut lähes kaikesta paperille painetusta tiedosta siirtyminen sähköisiin tietoihin. Laskut, asiakirjat, muistiinpanot ja muut dokumentit tehdään yhä useammin tietokoneilla ja nämä tiedot halutaan pitää tallessa, eli ne tallennetaan. Tälle tallennetulle datalle etsitään turvallista tallennuspaikkaa, joten sähköisten arkistointiratkaisujen kehittäminen on tullut ajankohtaiseksi.

### 2.2.1 Tietojen suojaaminen

Arkistoidun tiedon suojaamiseksi on olemassa useita ratkaisuja. Erilaisia suojausratkaisuja suositellaan käytettävän päällekkäin. Yrityksien arkistoidun tiedon ensimmäinen ja yleisin suojauskeino on rajoittaa käyttäjien oikeuksia. Esimerkkinä käyttäjien oikeuksien rajoittamisesta on pelkkien lukuoikeuksien antaminen arkistoituihin dataan (ee.oulu.fi 2015). Käyttäjät helposti poistavat vahingossa tietoa tai siirtävät tietoa väärin paikkoihin. Tämä pätee myös arkistoidussa datassa, ja tiedon tuhoutuminen on helppo estää ”vain luku” oikeuksilla. Myös oikein toteutettu varmuuskopiointi auttaa käyttäjien poistaman datan palautuksessa. Yleisesti kaikille käyttäjille ei pidä antaa oikeuksia jokaiseen kansioon, vaan oikeudet annetaan vain niille henkilöille, jotka työssään tarvitsevat kyseistä tietoa (ee.oulu.fi 2015).

Toinen lähes itsestäänselvä suojauskeino on varmistaa, ettei järjestelmään pääse ketään ulkopuolisia. Tässä keinossa virusturvat ja palomuurit tulee pitää päällä ja ajantasalla. Valitettavan usein yrityksien järjestelmiin onnistutaan hakkeroimaan takaporttien tai päivittämättömien palomuurien kautta. Hakkerit yleensä kiristävät

uhriaan joko julkaisemalla osan varastetusta tiedosta tai suoraan vaatimalla lunnaita.

### 2.2.2 Datan arkistoinnin avulla saavutettavat säästöt

Arkistoimalla dataa sähköisesti voidaan saada huomattavia säästöjä yrityksille. Työntekijöiden työaika ei mene ”hukkaan”, kun data on tallessa samassa turvallisessa paikassa eikä esimerkiksi jokaisen omalla kovalevyllä tai verkkoasemalla. Kun data on arkistoitu, data pysyy tallessa niin kauan kuin yritys sitä haluaa, pois sulkien vahingot ja laiterikot. Mahdollisten laiterikkojen tapauksissa tiedon tulisi olla tallessa, jos varmuuskopiointi on tarpeeksi tiheätä ja tarkkaa. (Strömberg 2015.)

Käyttöön otettavan arkistointiratkaisun suunnittelu- ja rakennusvaiheessa kulut nousevat hetkellisesti huomattavasti, mutta kun ratkaisu on toiminnassa, kulut laskevat. Tallennettu tieto voi olla hyvinkin arvokasta, joten tiedon säilyminen vuosien ajan on tärkeää. Usein yrityksiä projektien ja muiden aikaa vievien töiden dataa voidaan tarvita vuoden tai useamman jälkeen, tällöin arkistointi on tärkeää ja hyvä keino varmistaa datan säilyvyys (Pehkonen 2013). Kun arkistointi on toteutettu oikein, arkistointi säästää palvelimien ja työasemien resursseja, sillä data on erillisessä säilössä. Palvelimet tai työasemat eivät myöskään ole riippuvaisia tallennusratkaisusta.

### 2.2.3 Tietojen palautus ja varmuuskopiointi

Arkistoitu data voi olla erittäin arvokasta, jolloin sen varmuuskopioinnin tulee olla kunnossa. Datan tulee olla helposti palautettavissa mahdollisen laiterikon tai vahingon tapahtuessa. Yrityksissä yleisesti varmuuskopiointi tapahtuu joko erilliselle kovalevyille tai vaihtoehtoisesti siirrettäville varmuuskopiointinauhoille (Poelker 2013). Kolmantena vaihtoehtona ovat yleistyneet erilaiset pilvipalveluihin perustuvat varmuuskopiointiratkaisut (Rouse 2013), kuten Microsoftin Windows Azure.

Varmuuskopiointiratkaisun toteutuessa ulkoisille kovalevyille ratkaisu sisältää useita kovalevyjä, jotka käyttävät jotain RAID-tekniikkaa (Redundant array of independent disks). Ulkoisia kovalevyjä on yleensä kolme tai useampi. Varmuuskopiointiratkaisu on RAID-tekniikkaa käyttäen myös itsessään vikasietoinen. Yleisimpiä käytettyjä RAID-tekniikoita ovat RAID1- ja RAID5- tilat (Natarajan 2010). Ulkoinen kovalevyratkaisu sijoitetaan tieto- ja paloturvallisista syistä johonkin muuhun tilaan, kuin missä palvelimet tai datasäilö sijaitsevat. Erilliseen tilaan sijoitettu ulkoinen kovalevy on näin turvassa mahdolliselta tulipalolta tai vesivahingolta, joka voi tapahtua palvelintilassa. Kovalevyratkaisu on useimmiten yhdistetty muuhun järjestelmään lähiverkon eli LAN:n välityksellä. Tällöin kovalevyratkaisu käyttää Ethernet-kaapelia tai valokuitukaapelia. Toinen mahdollinen tapa liittää varmuuskopiointiin käytettävät kovalevyt muuhun järjestelmään on käyttää USB-liitäntää (Universal Serial Bus). USB-ratkaisua käytetään erityisesti pienissä ympäristöissä, kuten kotitoimistoissa ja PK-yrityksissä. Tällöin kovalevyt ovat alttiita samassa tilassa tapahtuville vahingoille. (Poelker 2013.)

Varmuuskopioinnin tapahtuessa magneettinauhoille yleensä itse varmuuskopiointinauha-laite sijoitetaan palvelinhuoneeseen. Harvoin varmuuskopiointinauha-laite sijoitetaan erilliseen sijaintiin, sillä liitäntänä käytetään hyvän tiedonsiirtokapasiteetin omaavia kaapeleita, joiden pituus ei voi olla kovin pitkä. Varmuuskopiointinauhoja käytettäessä etuna kovalevyihin verrattuna on nauhojen vaihdettavuus ja siirrettävyys sekä niiden alhainen hinta. (KUVIO 3.)(KUVIO 4.) (Lock 2010). Kun viikoittainen tai kuukausittainen varmuuskopionauha sijoitetaan esimerkiksi pankin holviin, on todella pieni todennäköisyys, että varmuuskopiointi tuhoutuisi.



KUVIO 3. Kovalevypohjaisen tallennusratkaisun hinta verrattuna varmuuskopiointinauhojen hintaan (Poelker 2013)



KUVIO 4. Varmuuskopiointinauhan hinta verrattuna kovalevypohjaisen varmuuskopiointin hintaan (Poelker 2013)

Yritykset käyttävät järjestelmänsä varmuuskopintiin myös yleisesti nauha- ja kovalevyvarmuuskopiointin yhdistelmää (Snyder 2013). Esimerkkinä tästä voitaisiin pitää ratkaisua, jossa joka yö varmuuskopiointi menee kovalevyille talteen ja lisäksi viikonloppuisin koko järjestelmästä otetaan levykuva nauhoille, jotka sijoitetaan varmaan paloturvalliseen paikkaan (Spector 2013). Näin pieniä palautuksia

tehtäessä voitaisiin käyttää kovalevyjä. Suuremman tuhon, kuten tulipalon, sattuessa nauhavarmuuskopioilta voitaisiin palauttaa koko järjestelmän varmuuskopio kokonaan uuteen järjestelmään nopeasti ja varmasti.

Pilvipalvelut ovat tehneet tuloaan erilaisten palvelujen tarjoajana perinteisten ratkaisujen ohelle. Pienehköjen tiedostojen tai järjestelmien varmuuskopiointi pilveen on mahdollistunut nopeiden internetyhteyksien ansiosta. Monelle yritykselle pilvipalvelut ovat jokapäiväisiä. Pilvipalveluiden käyttö varmuuskopiointiratkaisuna vaatii yritykseltä erittäin nopean internetyhteyden, kuten valokuituyhteyden. Etuna pilvipalveluissa on se, että tieto on sijoitettuna ympäri maailmaa eikä esimerkiksi naapurirakennukseen. Dataan on myös mahdollista päästä käsiksi mistä vain. Tietysti haittana ja suurena kysymyksenä yrityksille on suojauksen taso: Onko tieto varmasti tallessa turvallisessa sijainnissa, jonne kukaan ulkopuolinen ei pääse? Esimerkkinä pilvipalvelupohjaisesta varmuuskopiointiratkaisusta on Microsoftin Windows Azure. Azuressa on mahdollista tehdä kuukausimaksua vastaan oma varmuuskopiointisäilö, johon saa ladata talteen tietoa niin paljon, kuin tilaus sallii. Suuremmalla summalla rahaa saa pilvipalveluista enemmän tallennustilaa.

## 2.3 Verkkotekniikka

Arkistointiratkaisut perustuvat usein kovalevyihin, jotka ovat tietoliikenne-ve-kossa. Tiedostoihin pääsy tapahtuu usein myös lähiverkon kautta, jolloin lähiverk-kotekniikka tulee tuntea.

NAS-levypalvelimet (Network-attached storage) toimivat vain verkossa. NAS-levypalvelimet käyttävät pääasiassa TCP/IP-yhteyksiä (Transmission Control Protocol / Internet Protocol). Samoin levypalvelimien hallinnointi onnistuu hel-pommin verkon kautta, kuin esimerkiksi paikallisesti. Tällöin myös itse palveli-mien sijainnilla ei ole suurta merkitystä, kunhan palvelimet ovat hyvien verkkoyh-teyksien päässä.

### 2.3.1 OSI-malli

OSI-malli, eli Open Systems Interconnection Reference Model koostuu seitse-mästä kerroksesta, joiden avulla tiedonsiirtojärjestelmät tulisi suunnitella (KUVIO 5). OSI-malli on kehitetty jo 1980-luvun alussa, ja OSI-malli on vieläkin ajankoh-tainen. OSI-malli on kansainvälinen standardi. Tiedonsiirto perustuu OSI-mallin kerrokseen. (Krimaka.net 2012.)





## KUVIO 5. OSI-malli (Wikipedia 2005)

OSI-mallin kerroksia on seitsemän. Kerrostuksen ideana on pyramidimainen malli. Kerroksissa aina ylempi kerros käyttää hyväkseen alempia kerroksia, jotka eivät ole niin kehittyneitä, kuin ylemmät kerrokset. (Raisamo 2014.)

Ensimmäinen niin sanottu fyysinen kerros määrittelee tiedonsiirron fyysisen rakenteen, eli siirretäänkö dataa langattomasti, valokuidun kautta, vai peräti sähköverkon yli (Krimaka.net 2012). Toinen kerros on siirtoyhteyskerros. Siirtoyhteyskerros hoitaa yhteyden luomisen, virheiden korjaamisen ja yhteyden katkaisemisen. Siirtoyhteyskerros huolehtii siitä, että dataa ei lähetetä liian nopeasti, vaan sitä tahtia, että vastaanottaja pystyy sitä käsittelemään. (Raisamo 2014.)

Kolmantena kerroksena OSI-mallissa on verkkokerros. Verkkokerros tarjoaa verkon rakenteesta riippumattoman tiedonsiirron. Kolmannen kerroksen tarkoituksena on lähinnä piilottaa tiedonsiirron fyysiseen toteutukseen liittyvät piirteet. Verkkokerroksen yhtenä tehtävänä on valita, mitä reittiä pitkin data lähetetään monihaarisessa tietokone- ja muissa verkoissa. (Krimaka.net 2012.)

Kuljetuskerros on taas neljäs kerros OSI-mallissa. Kuljetuskerroksen tehtävänä on taata luotettava päästä-päähän yhteys. Kuljetuskerros huolehtii myös vaihtoehtoisista reiteistä häiriöiden sattuessa. Viides kerros on yhteysjaksokerros. Yhteysjaksokerroksen tehtävänä on pitää huolta, että tiedonsiirto ei sekoja esimerkiksi fyysisten yhteyksien katketessa. Yhteysjaksokerros huolehtii lisäksi tiedon salaamisesta. Kuudes kerros on esitystapakerros. Esitystapakerros huolehtii tiedon esitysmuodon oikeaksi tiedonsiirron yhteydessä. Esitystapakerroksen tehtäviin kuuluu myös päättää, missä muodossa erinäinen data (kokonaisluvut, teksti, kuva, ääni) esitetään tiedonsiirron yhteydessä, eli sen tarkoituksena on se, että lähetetty tieto on oikeassa muodossa niin, että vastaanottaja sen ymmärtää. (Krimaka.net 2012.)

Viimeinen ja seitsemäs kerros on sovelluskerros. Sovelluskerroksen tehtävänä on lähinnä toimia linkkinä siihen ohjelmaan, joka tiedonsiirtoa tarvitsee. (Krimaka.net 2012.)

OSI-mallin kerrokset toimivat hierarkkisesti niin, että ylempi kerros on tekemisissä yhtä alemman kerroksen kanssa, eli esimerkiksi kolmas kerros on tekemisissä neljännen ja toisen kerroksen kanssa. Kolmas kerros vaikuttaa toiseen kerrokseen enemmän kuin neljänteen. (Krimaka.net 2012.)

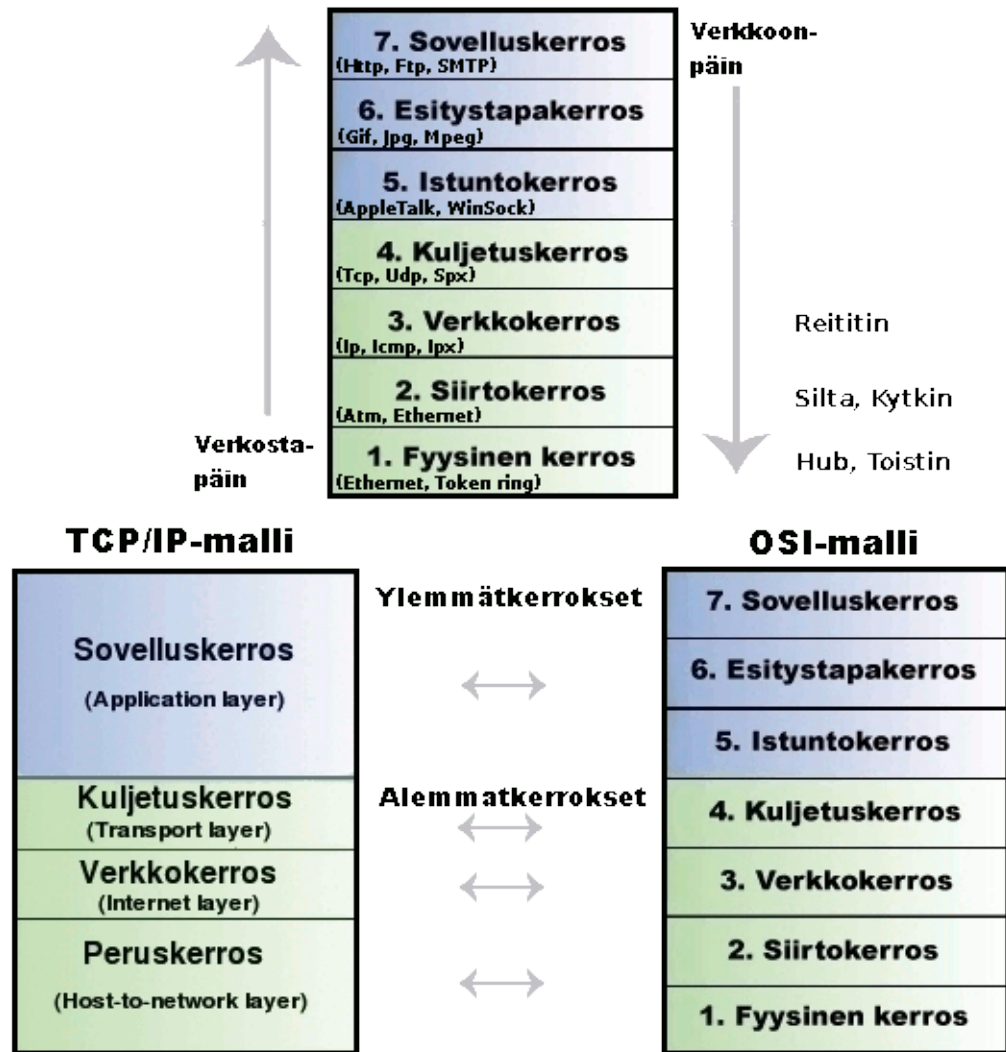
OSI-malli on ISO:n (International Standard Organization) määrittelemä malli verkko-ohjelmistojen kerroksittaisesta rakenteesta. OSI-standardia käytetään kaikessa tietoliikenneverkon pohjana juuri sen standardoinnin vuoksi. (Krimaka.net 2012.)

### 2.3.2 TCP/IP-malli ja kerrokset

TCP/IP-mallilla on paljon yhteistä OSI-mallin kanssa. Molemmat koostuvat kerroksista, ja näiden kerroksien välillä on paljon samanlaisuutta. (KUVIO 6.)

OSI-mallissa rajapinnat, palvelut ja protokollat on pidetty erillään, toisin kuin TCP/IP-mallissa. TCP/IP-malli kehitettiin vasta TCP- ja IP-standardien jälkeen. (Krimaka.net 2012.)

## OSI- TCP/IP-mallit



KUVIO 6. TCP/IP-malli ja OSI-malli (Krimaka.net 2012)

TCP/IP-malli koostuu neljästä kerroksesta, joista kolmella on suorat vastineet OSI-mallista. Ensimmäinen kerros on sovelluskerros. Sovelluskerroksessa ei ole määritelty, mitä protokollaa käyttäen sovellus lähettää tai vastaanottaa dataa. Sovellukset huolehtivat siitä itse. (Krimaka.net 2012.)

Toinen kerros on kuljetuskerros. Kuljetuskerroksen tarkoituksena on sallia tietojenvälitys lähde- ja kohdekoneiden välillä. TCP/IP-mallin kuljetuskerros on sama kuin OSI-mallin kuljetuskerros. Yleisimmin TCP/IP-mallin kuljetuskerros käyttää TCP:tä ja UDP:tä. (Krimaka.net 2012.)

Kolmas kerros on verkkokerros. Verkkokerros ei ole tietoinen ylempien kerroksien tapahtumista, jokainen IP-paketti käsitellään erillisenä tapahtumana, eikä verkkokerros ole tietoinen sovelluksia yhdistävästä TCP-yhteydestä. TCP/IP-mallin verkkokerros vastaa varsin hyvin OSI-mallin verkkokerrosta. (Krimaka.net 2012.)

Neljäs ja alin kerros on peruskerros. TCP/IP-malli ei määritä tarkkaan, mitä tapahtuu verkkokerroksen alapuolella. Peruskerros olettaa, että IP:n alla on jokin protokolla, joka pystyy välittämään IP-paketteja tietokoneen muistista toiselle tietokoneelle. (Krimaka.net 2012) IP-protokolla ei ota mitenkään kantaa fyysiseen mediaan tai pakettien kehystykseen, vaan olettaa saavansa käyttöön alemman kerroksen palveluita jonkin rajapinnan kautta, niin että IP-protokolla voi lähettää paketteja reitittimille. Tyypillisiä peruskerroksia ovat lähiverkkojen Ethernet, matkapuhelimien GPRS ja 3G ja ADSL. (Mrin 2001.)

### 2.3.3 IPv4 avaruudet ja maskit

IPv4-osoitteet (Internet Protocol version 4) ovat 32 bittiä (4 tavua) pitkiä. Yleensä IPv4-osoite esitetään neljänä desimaalilukuna, jotka on eroteltu toisistaan pistellä. Arvot, joita desimaaliluvut voivat saada, ovat 0 - 255. Esimerkkinä perinteisestä IP-osoitteesta voidaan pitää osoitetta 192.168.0.1. IP-osoitteet jaoteltiin aikaisemmin viiteen luokkaan, mutta nykyään käytetään luokatonta järjestelmää. Tällöin jäljellä olevat IP-osoitteet saadaan paremmin hyödynnettyä. Luokattomissa osoitteissa verkko- ja laiteosat erotellaan aliverkon maskin avulla. (Rintala 2001.)

Aliverkoituksessa IP-avaruuksien maskit ovat tärkeässä osassa. Maskeilla saadaan jaoteltua verkkoa pienempiin osiin, aliverkkoihin, tai vastaavasti yhdistää aliverkkoja suuremmiksi verkoiksi (yliverkotus). (Rintala 2001.)

Aliverkoituksessa käytetään laiteosan bittejä jakamaan verkkoa useammaksi osaksi, jolloin jokaisen syntyneen verkon laitemäärä pienenee (Raisamo 2014). Nykyisin yleisesti aliverkoitus tehdään maskin (Subnet Mask) avulla. Kuten IP-osoitteenkin, maskinkin pituus on 32 bittiä. Esimerkkinä usein käytetyistä mas-

keista on 255.255.255.0. Verkot on lajiteltu kolmeen ryhmään suuruutensa mukaan: A-luokan verkot ovat suuria, B-luokan verkot keskikokoisia ja C-luokan verkot ovat pienehköjä. (Rintala 2001.)

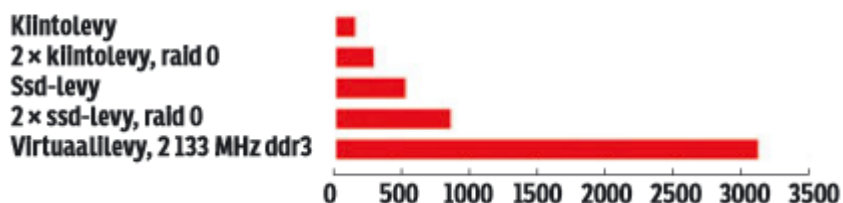
#### 2.4 Kiintolevyjen tallennustekniikat

Tallennustekniikat ovat kehittyneet valtavasti läpi tietokoneiden historian. Viimeisin suuri harppaus on tapahtunut SSD (Solid State Drive) -kiintolevyjen tullessa markkinoille. Perinteinen kovalevy sisältää liikkuvia osia, kuten lukupään ja kovalevykiekon. Perinteinen kovalevy pystyy parhaillaan vain muutamiin satoihin luku- ja kirjoitustapahtumiin sekunnissa. Tämä on johtunut pääasiassa liikkuvista osista: lukupään täytyy liikkua edestakaisin kovalevyn kiekolla, ja tämä on hidastanut toimintaa. Myös tallennetun datan sijainti kovalevyllä on vaikuttanut nopeuksiin: mitä ulompana kiekolla tieto on, sitä hitaammin tieto on saatavilla.

SSD-puolijohdeasemat eivät sisällä liikkuvia osia ja ovatkin siksi paljon perinteisiä kovalevyjä nopeampia. SSD-kovalevyt suorittavat useita kymmeniä tuhansia luku- ja kirjoitustapahtumia sekunnissa. Tämänhetkisissä kovalevyissä, eritoten SSD-asemissa, pullonkaulana on uusin SATA-liitäntä (Serial ATA). Sen teoreettinen luku ja kirjoitusnopeus on rajoitettu 600 megatavuun sekunnissa. (KUVIO 7.) SSD-kovalevyjen käyttö levypalvelimissa on harvinaista, sillä niiden hinnat ovat moninkertaiset verrattuna perinteisiin kovalevyihin. Tulevaisuudessa niiden hintojen ennustetaan laskevan, joten niitä tullaan luultavasti näkemään myös palvelinympäristöissä (Lal Shrimpi 2010).

#### Jatkuvat lukunopeudet

Mt/s. Pidempi parempi



KUVIO 7. Vertailu jatkuvista lukunopeuksista (Jääskeläinen 2013)

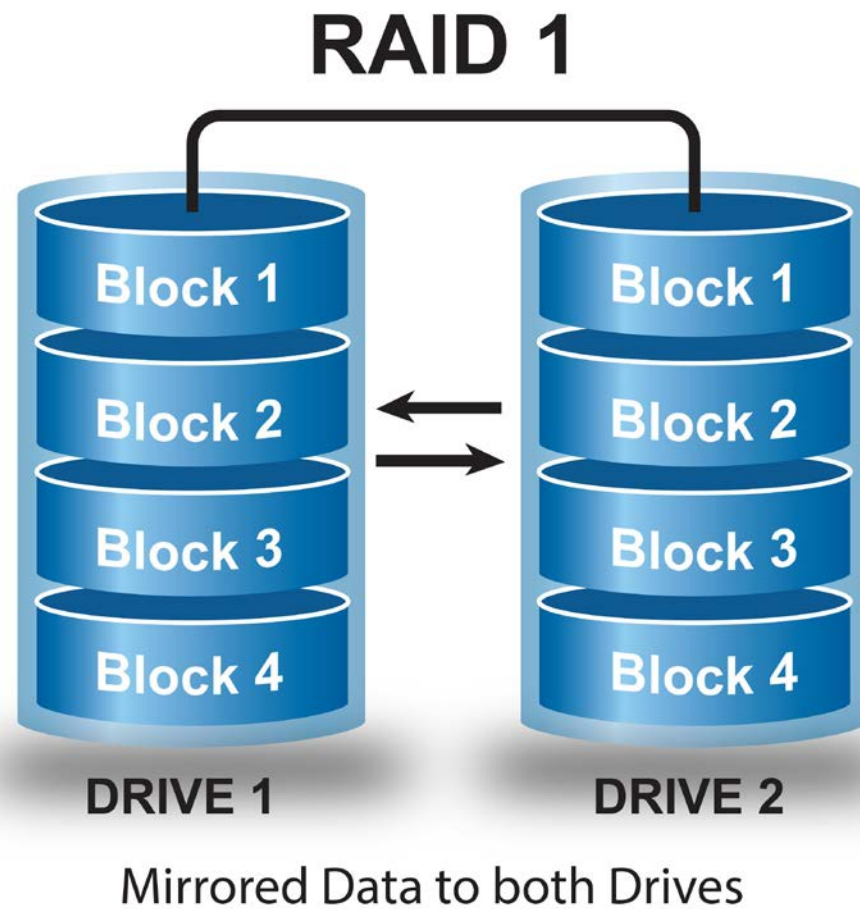
### 2.4.1 RAID

RAID (Redudant Array of Independent Disks) on tekniikka, jolla on mahdollista kasvattaa luku- ja kirjoitusoperaatioiden määrää kovalevyillä tai vastaisesti parantaa vikasietoisuutta. Yksinkertaistettuna RAID-tekniikalla yhdistetään kovalevyjä yhdeksi virtuaaliseksi muistiksi. Yleisimpiä RAID-tekniikoita ovat RAID0, RAID1 ja RAID5 (Natarajan 2010). RAID-tekniikoissa voidaan käyttää merkintöjä: N-kappaletta ja C-kokoisia levyjä. RAID-tekniikkaa käytettäessä pyritään käyttämään samankokoisia levyjä. Jos levyjä on erikokoisia, niin koko määräytyy pienimmän mukaan, eli kaikki levyt toimivat samankokoisina, kuin pienin levyistä.

RAID-tekniikka vaatii tietokoneen tai palvelimen suorittimelta laskentatehoa. Tämän takia RAID-tekniikkaa tehdään usein laiteetasolla, eli laitteistoon asennetaan erillinen RAID-kortti. RAID-kortti huolehtii laskennasta ja kovalevyjen loimituksesta ja peilaamisesta. RAID on mahdollista toteuttaa myös ohjelmistotasolla. Ohjelmistotasolla käytetään laitteiston omaa prosessoria. Ohjelmistotason RAID-ratkaisut ovat huomattavasti halvempia, mutta kuormittavat laitteistoa enemmän, kuin laitteistotason RAID (Natarajan 2010).

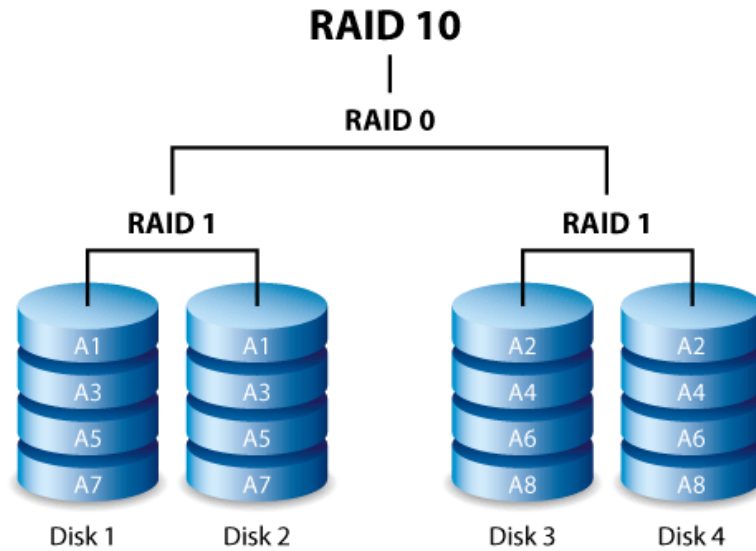
RAID0 on loimitusta. Loimituksessa kaksi tai useampi levy on yhdistetty yhdeksi loogiseksi tallennustilaksi. Loimituksen kapasiteetti ja nopeus on laskettavissa:  $N \times C$  (Natarajan 2010). Loimituksessa jos yksikin levy hajoaa, menetetään kaikki data. Kun data kirjoitetaan loimitettuna, data jakautuu tasaisesti kaikille levyille. Tällöin kirjoitus- ja lukunopeus kasvavat (Segate 2015).

RAID1 on peilausta. Peilauksessa sama data kirjoitetaan kahdelle tai useammalle levyille. Jos jokin levyistä hajoaa, niin data on yhä tallessa muilla levyillä. (KUVIO 8.) Peilauksessa siis data on tallessa kahdella tai useammalla levyllä, jolloin periaatteessa datan lukunopeuden kuuluisi olla kaksinkertainen, verrattuna yhteen levyyn (Natarajan 2010). RAID1 on yleinen kuluttajille suunnatuissa levypalvelimissä, erityisesti NAS-asemissa.



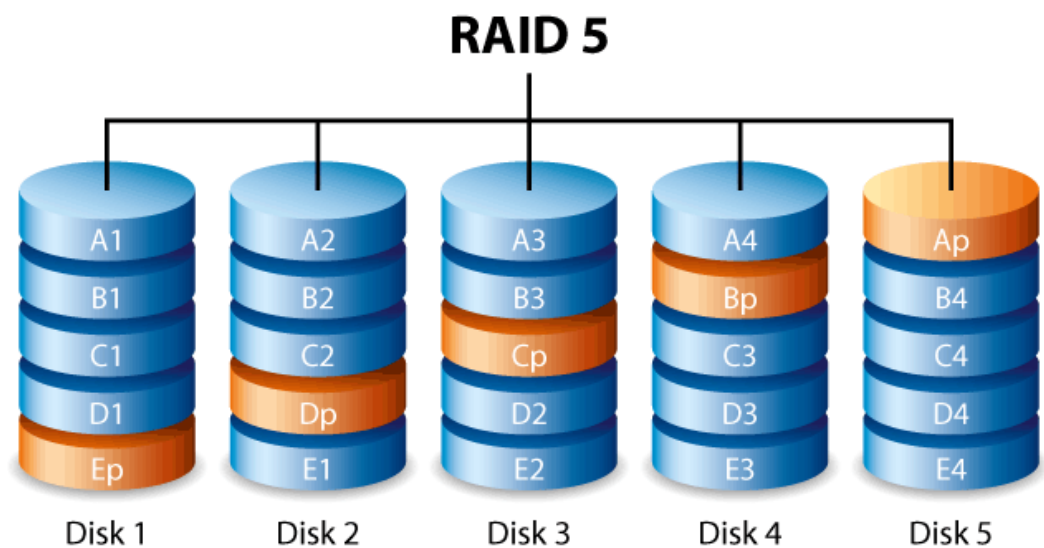
KUVIO 8. RAID 1 (Segate 2015)

RAID1+0 (kutsutaan myös RAID10) on loimitusta ja peilausta. Levystä puolet loimittavat keskenään RAID0:n tapaan ja toinen puoli peilaa näiden levyjen kanssa RAID1:n lailla. (KUVIO 9.) Tällä RAID-tekniikalla saavutetaan nopeus ja luotettavuus, eli data on palautettavissa, jos peilaavat levyt ovat ehjiä. Samalla lukunopeus on kaksinkertainen yhteen levyyn verrattuna. Levyjä RAID1+0 vaatii vähintään neljä (Natarajan 2010).



KUVIO 9. RAID 1+0 (Segate 2015)

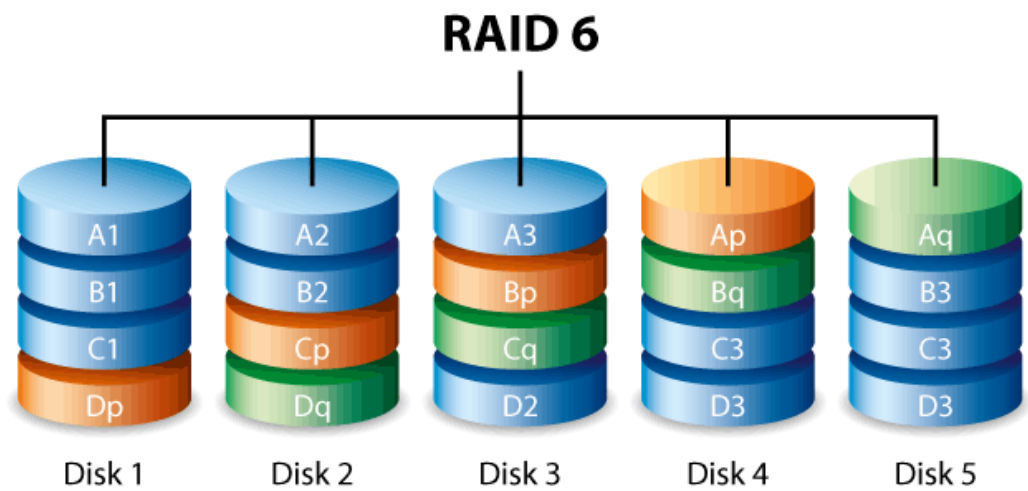
RAID5 sallii yhden minkä tahansa levyn hajoamisen. RAID5:n kapasiteetti on laskettavissa:  $C \times (N-1)$ . Eli RAID5:ssä vähintään kolme kovalevyä toimivat keskenään. Tässä tekniikassa yksi levy toimii muiden levyjen pariteettidatan tallennussijaintina, jolloin mikä tahansa kovalevyistä voi hajota ja data on yhä palautettavissa. (KUVIO 10.) Tätä tekniikkaa käytetään yleisesti yritysten levypalvelimissa, koska tekniikka mahdollistaa suuren tallennuskapasiteetin ja silti vikasietoisuuden (Murelli 2007).



KUVIO 10. RAID5 (Segate 2015)



RAID6 on lähes kuin RAID5. Erona RAID5-tekniikkaan on, että RAID6 mahdollistaa kahden minkä tahansa levyn hajoamisen ilman että dataa menetettäisiin. RAID6:ssa on toinenkin levy, jolle pariteettidata tallentuu. (KUVIO 11.) Tämä kuitenkin tarkoittaa sitä, että levyjen määrästä menetetään vähintään kahden levyn verran tallennustilaa. Tämä RAID-tekniikka ei ole vielä lyönyt läpi, vaan RAID6:n sanotaan olevan liian tehoton kustannussyistä (Murelli 2007).



KUVIO 11. RAID 6 (Segate 2015)

#### 2.4.2 Hot swap

Kovalevyistä on tarjolla palvelinkäyttöön niin sanottuja “hot swappable” -kovalevyjä. Tällä tarkoitetaan kovalevyjä, jotka voidaan vaihtaa ilman, että palvelinta tai tietokonetta tarvitsee sammuttaa. Suurin osa nykypalvelimista käyttää pelkästään hot swappable -levyjä yhdistettynä johonkin RAID-tekniikoista. Hot swap -kovalevyjä käytettäessä yrityksen palvelin jatkaa toimintaansa normaalisti, silloin kun vioittunutta levyä vaihdetaan. Tämä auttaa ylläpidossa, sillä turhia alasajoja ei tarvitse tehdä.

Hot swap on enemmän tuttua jokapäiväisistä USB (Universal Serial Bus) -tikuista. Ne voidaan irrottaa tietokoneesta lähes milloin vain ja kytkeä samoin. Hot swap -ominaisuus on kovalevyn SATA-liitäntöissä vakio-ominaisuus, joten hot swap on toteutettavissa periaatteessa myös kotikoneille. Hot Swap -tekniikkaa

käytetään todella paljon palvelimissa, erityisesti RAID5-tekniikan kanssa. Tällöin kovalevyn vikaantuessa kovalevyn vaihtaminen on erittäin nopeata ja vaivatonta.

## 2.5 Datan pakkaaminen

Datan pakkaamisella tarkoitetaan datan viemän fyysisen tilan pienentämistä käyttämällä pakkausalgoritmeja. Pakkaus voi olla häviöllistä tai häviötöntä riippuen siitä, muuttuuko data käsittelyssä vai ei.

Kuluttajalle yleisin ja tutuin pakkausmenetelmä on häviötön. Lähes jokainen on törmännyt ".zip"- tai ".rar"-tiedostoihin, jotka on täytynyt purkaa jollain tietokoneen käyttöjärjestelmään sopivalla pakkausohjelmalla.

Tiedon pakkaaminen kehitettiin, jotta tallennus tulisi halvemmaksi; kun data menee pienempään tilaan, kustannukset laskevat. Tiedonpakkaamisen teoreettisen pohjan loi matemaatikko Claude Shannon. Hän loi määritelmän, että tiedon pakkaamisessa on rajansa: tietoa ei voida pakata loputtomiin, ilman tiedon menettämistä. Claude Shannon loi myös teorian häviöllisestä pakkaamisesta. Hänestä käytetäänkin nimitystä "informaatioteorian isä". Shannon loi teorian myös häviöllisestä datasta. (Data-Compression.com 2015.)

Tiedonpakkaamisen teoriassa pääosissa oli turhien merkkien poistaminen tiedosta. Turhien merkkien ja usein toistuvien merkkien lyhentämiseen, tai poistamiseen, perustuvat kaikki nykyiset pakkausmenetelmät. Pakkauksesta käydään kiivasta keskustelua alan asiantuntijoiden keskuudessa koskien pakattujen tietojen turvallisuutta. Pakkauksen purkamisessa voi moni asia epäonnistua, jolloin tieto voi päätyä lukukelvottomaksi tai ainakin laatu voi huonontua radikaalisti. (Cover&Thomas 2006, 103-105, 149-150.)

## 2.6 Pakkausmenetelmät (kompresio)

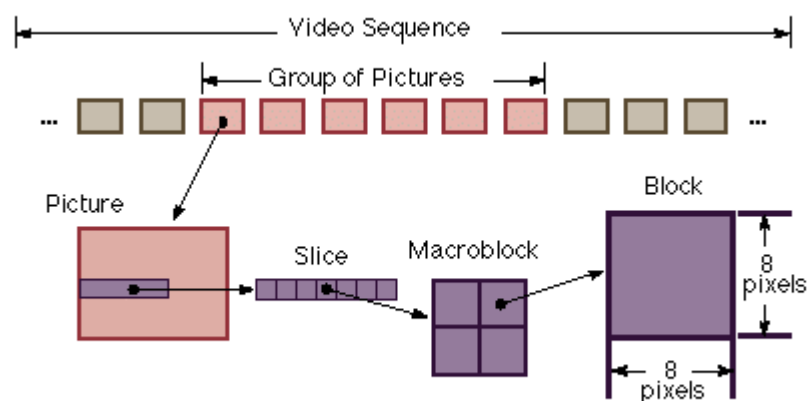
Yksinkertaistettuna pakkausmenetelmiä on kaksi: häviöllinen- ja häviötön pakkausmenetelmä. Näistä menetelmistä valitaan sen perusteella, halutaanko tieto alkuperäisessä muodossaan ja rakenteessaan vai ei (Cover&Thomas 2006, 105).

Häviötön pakkausmenetelmä tarkoittaa, että tieto ei muutu lopullisesti vaan tieto kuvataan toisella tapaa pakkauksen yhteydessä. Häviöttömässä pakkauksessa tieto palautuu alkuperäiseen muotoonsa ja rakenteeseensa, kun pakattu tieto puretaan.

Esimerkkinä häviöttömästä pakkauksesta on Huffmanin koodaus, jossa useammin tiedossa esiintyviä merkkejä korvataan lyhemmillä symboleilla. Huffmanin koodauksessa tutkitaan siis sitä, montako kertaa esimerkiksi sana ”kolme” esiintyy. Jos sana ”kolme” esiintyy tarpeeksi monta kertaa, niin sana ”kolme” korvataan jollain lyhemmällä symbolilla, kuten symbolilla ”3”, ja korvatuista sanoista pakkauksen yhteydessä tehdään oma lyhyt ”sanakirja”, jotta pakattu tieto osattaisiin kääntää takaisin. Purkuvaiheessa pakattu tieto muutetaan takaisin sanaksi ”kolme”. (Cover&Thomas. 2006, 121–127.)

Häviötön pakkaus etsii myös toistuvuuksia, kuten usean tietyn symbolin jälkeen toistuvaa symbolia. Esimerkkinä toistuvuuksista symboleissa lause: ”kissa ja koira näkivät ankan”. Lauseessa usein symbolin ”a” jälkeen esiintyy välilyönti. Pakattaessa siis voitaisiin korvata symboli ”a” ja välilyönti vaikkapa symbolilla ”\_”. Tällöin lause olisi pakattuna ”kiss\_j\_koir\_näkivät ankan”. (Harris 2015.)

Häviöllisiä pakkausmenetelmiä käytetään pääasiassa kuvien, videoiden ja äänen pakkaamiseen. Tässä pakkausmenetelmässä käsitelty tieto muuttuu huonompilataiseksi, jotta tieto veisi vähemmän tilaa. Esimerkiksi internetsivuille upotetut videot ja kuvat ovat usein pakattuja. Häviöllisillä pakkausalgoritmeilla pakattua dataa ei voida palauttaa täysin samaan muotoon, kuin alkuperäistä pakkaamatonta.



KUVIO 12. Videon pakkaaminen MPEG-2 formaatilla (Julaakso 2015)

Häviöllisessä pakkauksessa tiedosta poistetaan epäolennainen tieto. Käytännössä tämä tarkoittaa sitä, että videon ja kuvan resoluutiota eli tarkkuutta pienennetään.

Lähes kaikki internetistä löytyvät kuvat ja videot ovat koodattuja jollakin pakkausmenetelmällä, kuten JPEG-pakkauksella (Joint Photographic Experts Group). MPEG-animaatioformaatti (Moving Picture Experts Group) (KUVIO 12.) on hyvä esimerkki videon ja äänen pakkauksesta häviöllisesti. (Ala-Mutka 2002.)

## 2.7 Datan arkistoinnin tallennusmediat

Dataa arkistoidessa on tärkeää, että tallennusmedia ja sijainti ovat luotettavia. Datan tulee olla helposti saatavilla, mutta silti sijainnin tulee olla tietoturvallisesti vahva. Mahdollisten levypalvelimien ja muiden tallennusmedioiden tulee sijaita tilassa, johon vain yrityksen IT-tuella ja henkilöillä, jotka huoltavat tallennusmedioita on pääsy. (Strömberg 2015.)

Yrityksen koosta ja tallennettavan datan määrästä riippuen, arkistointi tapahtuu yleisesti joko paikallisesti, siirrettäville medioille tai nykytrendin mukaisesti ”pilveen”. Yleisin tapa arkistoida dataa on yhä paikallisesti levypalvelimille, joilla vikasietoisuus on kohtuullisen hyvä.

### 2.7.1 Paikallinen arkistointi

Paikallinen arkistointi tarkoittaa yrityksen tiloihin sijoitettavaa levypalvelintä tai yleistä palvelintä, jolla tietyt kovalevyt tai niiden osiot on määritetty arkistointikäyttöön. Paikallisen arkistoinnin mediana levypalvelimet ovat pitkällä aikavälillä huokein ratkaisu suurille yrityksille.

Paikallisessa arkistoinnissa pitää aina huomioida varmuuskopiointi ja vikasietoisuus. Usein levypalvelimet on kofiguroitu niin, että levy peilaavat toisiinsa. Tällöin jos vikaa ilmenee jossain kovalevyistä, niin muut kovalevyt jatkavat toimintaa normaalisti ja dataa ei menetä. Tätä kutsutaan RAID-tekniikaksi. Paikallisen arkistoinnin vahvuus on tiedon käsittelyn helppous: Tieto sijaitsee yrityksen tiloissa ja tieto on helposti käsiteltävissä. Suurienkin tiedostojen siirtäminen ja varmuuskopiointi on nopeaa paikallisessa arkistoinnissa, sillä tieto liikkuu lähiverkon välityksellä (Pehkonen 2013). Paikallinen arkistointi on mahdollista toteuttaa myös kovalevyakoilla, jotka voidaan liittää palvelimiin. Kovalevypaketit toimivat joko lähiverkossa tai suoraan SAS-kaapelilla (Serial attached SCSI) palvelimissa. Kovalevypakkojen tallennustilaa on helppo käyttää myös virtuaalisien palvelimien tai muiden vastaavien ratkaisujen tallennussijaintina.

## 2.7.2 Pilvi tallennustilat

Arkistointiin on mahdollista käyttää pilvitalennustiloja, jotka sijaitsevat palveluntarjoajan palvelimilla. Tällöin yrityksen ei tarvitse huolehtia tiedon varmuuskopiointista, sillä varmuuskopiointi tapahtuu palveluntarjoajan omilla järjestelmillä. Pilvipohjaisessa ratkaisussa joko itse luodaan vuokrattuun tallennustilaan oma palvelin, joka huolehtii tiedostoista ja niiden jakelusta, tai vaihtoehtoisesti palveluntarjoajalta vuokrataan vain levytilaa. Levytilaan päästään käsiksi yleisesti kirjautumalla internetselaimella palveluntarjoajan kotisivuille. Googlen ja Microsoftin pilvitalennusratkaisut sisältävät omat sovellukset, joilla yleisimmät tiedostot on mahdollista avata suoraan palvelusta. Tällöin tiedostoja ei tarvitse välttämättä edes ladata palvelimelta työasemalle (Mitroff 2014). Microsoftilla, Googlella ja Dropboxilla on mahdollista jakaa tiedostoja myös ulkopuolisille. Tämä onnistuu lähettämällä linkin sähköpostilla tai jakamalla halutun henkilön tilille tiedosto tai kansio (TAULUKKO 1.) (Parker 2015.)

TAULUKKO 1. Pilvitalennusratkaisujen vertailua (Mitroff 2014)

### Cloud storage comparison

	OneDrive	Dropbox	Google Drive	Box	Amazon Cloud Drive	Copy
File size restrictions?	2GB	None with Dropbox apps	5TB	250MB for free plan, 5GB for paid plan	2GB*	None
Free storage?	15GB	2GB	15GB	10GB	No**	15GB
Can I earn extra free storage?	Yes	Yes	No	No	No	Yes
Paid plans	\$2/month for 100GB, \$4/month for 200GB, \$7/month for 1TB	\$10/month for 1TB	\$2/month 100GB, \$10/month for 1TB	\$10/month for 100GB	\$12/month for unlimited photos, \$60/year for unlimited files	\$10/month for 250GB
OSes supported	Windows, Mac, Android, iOS, Blackberry	Windows, Mac, Linux, Android, iOS, Blackberry, Kindle Fire	Windows, Mac, Android, iOS	Windows, Mac, Android, Blackberry, iOS	Windows, Mac, Android, iOS, Kindle Fire	Windows, Mac, Linux, Android, iOS

Pienet arkistointiratkaisut on helppo toteuttaa pilven avulla. Esimerkiksi Dropbox, Google ja Microsoft tarjoavat omia ratkaisujaan. (TAULUKKO 2.) Microsoftin OneDriven tallennustila on ilmaista 15 gigatavuun asti, jonka jälkeen lisätilasta veloitetaan kuukausimaksu periaatteella (Mitroff 2014). Microsoftilla on myös yrityksille suunnattu OneDrive-palvelu. Yrityksille suunnatussa palvelussa on mahdollista käyttää tallennussijaintia usean käyttäjän yhtä aikaa omilla tunnuksillaan ja hallita käyttöoikeuksia (Mitroff 2014). Microsoft tarjoaa myös Office 365-palvelua, jossa Office-tuoteperheen tuotteet ovat palvelua haluavan yrityksen käytettävissä. Samaan pakettiin kuuluu myös lisätallennustila. Microsoftin palvelu toimii myös vertailtavista palveluista parhaiten PC-koneilla (Archambault 2015).

TAULUKKO 2. Paikallisen arkistointiratkaisun ominaisuudet verrattuna pilvipohjaiseen arkistointiratkaisuun (Abate 2011; Archambault 2015)

	<b>Paikallinen arkistointiratkaisu</b>	<b>Pilvipohjainen arkistointiratkaisu</b>
<b>Tiedonsiirtonopeus</b>	122-204Mt/s	Riippuen yrityksen ulkoverkon nopeudesta
<b>Yhteensopivuus</b>	Windows, OSX, Unix	Windows, OSX, Unix, Android, Windows Phone, iOS
<b>Kustannukset / GB</b>	n.0,8€	n.0,05€1€

Microsoftilla on toinenkin palvelu, joka toimii arkistointiratkaisuna: Microsoft Windows Azure-palvelu. Azure-pilveen on mahdollista tehdä niin omia virtuaalikoneita, kuin levypalvelimia. Azuren on enimmäkseen suunnattu yrityksille, joilla on mahdollisuus rakentaa omia ratkaisuja verkkoon (Microsoft 2015a). Microsoftin Azure-palvelu eroaa muista sillä, että Azure-palvelusta maksetaan sen mukaan, miten kulutetaan resursseja. Esimerkiksi palvelin, jolla on käytössä kahdeksan prosessorin ydintä ja kaksi kovalevyä, maksaa enemmän, kuin palvelin, jossa on



vain kaksi prosessorin ydintä ja yksi kovalevy. Veloitettava summa kasvaa prosessori- ja kovalevykuorman perusteella. Samoin internetliikenne palveluun lisää veloitettavaa summaa (Microsoft 2015b).

Googllella on kokonaan yrityksille tehty palvelu, jossa Google Apps for Work-palveluun luodaan oma tallennusympäristö pelkästään kohdeyritykselle. Google kauppaa tallennustilaan liitettäväksi sovelluksia, joilla tiedostojen käyttöoikeuksia voidaan rajoittaa. Samoin kuin Microsoftin OneDrive, niin Googlenkin Apps for Work sisältää mahdollisuuden käyttää tallennussijaintia useilla käyttäjätunnuksilla (Mitroff 2014). Google Apps for Work sisältää 30 gigatavua ilmaista tallennustilaa Google Drive-palveluun. Jos 30 gigatavun tallennustila ei riitä, niin on mahdollista ostaa rajaton tallennustila, jossa viidelle käyttäjälle annetaan teratavun verran tallennustilaa. Tästäkin on mahdollista ostaa suurempia tallennuspaketteja (Parker 2015).

Dropbox on kolmas yleinen vaihtoehto pilvitallennus ratkaisuna. Dropboxia mainostetaan helpolla tiedostojen jakamisella, mutta käytännössä Dropbox suuntautuu yksityisiin käyttäjiin. Yrityksille Dropboxilla ei ole omaa palveluaan. Dropbox on helppo valinta silloin, kun tarvitsee jakaa tiedostoja arkistosta usealle eri henkilöille oman organisaation ulkopuolella. Lähes kaikilla yrityksillä ja yksityishenkilöillä on jo oma Dropbox tilinsä. Dropboxin laitetuki on kattavin vertailtavista palveluista. Dropboxin oma synkronointi sovellus löytyy lähes kaikille alustoille jo. Dropbox on yleisesti helpoin palvelu käyttää, koska Dropboxin käyttöliittymästä on tehty niin yksinkertainen.

### 2.7.3 Siirrettävät mediat

Pienehköt siirrettävät arkistointiratkaisut on mahdollista toteuttaa erilaisin siirrettävin medioin. Siirrettäviksi medioiksi luokitellaan CD/DVD-levyt (Compact Disk / Digital Versatile Disc), ulkoiset kovalevyt, magneettinauhat ja muistikortit (pcworkld.co.uk 2015). Ulkoiset kovalevyt ovat yleisesti USB- tai Thunderbolt-liitännällä varustettuja. Kotitoimistoissa tai todella pienissä yrityksissä arkistointi on mahdollista toteuttaa siirrettäville medioille.

Musiikin ja videoiden arkistointi toteutetaan usein CD/DVD-levyille. Tällöin sisällön toistaminenkin onnistuu esimerkiksi autoissa ja muissa soittimissa.

CD/DVD- pohjaisissa ratkaisussa etuna on hinta. Levyt ovat edullisia, ja niitä saa lähes jokaisesta kaupasta. Haittana CD/DVD-pohjaisessa ratkaisussa on ratkaisun vanheneminen. Tietokoneet ovat nykyisin niin sanottuja Ultrabook-tyyppisiä, eli niissä ei ole mitään tarpeetonta, kuten levyasemaa. Ulkoiset kovalevyt ovat yleistyneet viimeisen viiden vuoden aikana. Niitäkin saa lähes jokaisesta suuremmasta kaupasta. Ulkoisten kovalevyjen etuna siirrettävien medioiden ratkaisussa on niiden suuri tallennuskapasiteetti. DVD-levyille mahtuu normaalisti 4,7 gigatavua dataa, kun taas ulkoisille kovalevyille dataa saadaan mahtumaan jopa neljän teratavun verran. Tällöin tosin hintakin nousee huomattavan suureksi. Muistikortteja käytetään myös paljon tallennussijaintina kameroissa ja puhelimissa. Muistikortit ovat pieniä kooltaan, ja niihin mahtuu paljon tietoa. Yleisesti laitteet tukevat jo 64 gigatavun kokoisia muistikortteja. Muistikortit toimivat monissa erilaisissa laitteissa eivätkä ne ole kehittyneet paljoa viimeisen parin vuoden aikana, joten ne tulevat toimimaan myös tulevaisuudessa.

## 2.8 Datan tallennus lähtökohta ja ongelmat

LSK Electrics Oy:n tiedot tallentuvat yrityksen omassa palvelinhuoneessa olevalle kovalevypakalle, jota myös palvelimet käyttävät. Kovalevypakka on jo kolme vuotta vanha, joten sen uusiminen on ajankohtaista. Kovalevypakasta jaotellaan tallennustilaa tarpeen mukaan yrityksen palvelimille, joilla käyttäjien tallennussijainnit sijaitsevat.

LSK- konserniin kuuluu useita tytäryrityksiä, joilla on tallennettua dataa vuosikymmenien ajalta. Nämä tytäryritykset ovat fuusioituneet yksi kerrallaan takaisin emoyhtiöön, joten datan tulee olla laajemmalti saatavilla. Yrityksen kovalevyillä ei ole tilaa loputtomiin, joten arkistointiratkaisun tulee olla helposti laajennettavissa ja sinne tulee olla helppo pääsy.

Suurimpana ongelmana yrityksen kannalta on tallennustilan loppuminen. Varmuskopiointinauhoille mahtuu vain rajallinen määrä dataa, jolloin harvemmin

tarvittavaa tietoa pitäisi saada sijoitettua muualle, kuin yrityksen omalle kovalevy-pakalle. Koko kovalevy pakka menee joka yö varmuuskopiontina magneettinauhoille. Nauhat sijoitetaan paloturvakaappiin, jotta mahdollisen tulipalon tai rakennuksen muuten tuhoutumisen sattuessa ne olisivat helposti palautettavissa. Paloturvakaappi sijaitsee yrityksen tiloissa eri paloalueella. Data menee myös muuttuneiden tiedostojen osalta erilliselle NAS-kovalevy palvelimelle talteen joka yö. Tieto on nopeammin palautettavissa eri tiloissa olevalta kovalevy palvelimelta, kuin magneettinauhoilta.

LSK-yhtiöiden erinäisten projektien dataa on todella paljon. Vaikka data on vanhaa, sitä ei voida hävittää, sillä projektin kohteena on voinut olla esimerkiksi tehdas, joka on vielä kunnossapitosopimuksien alaisena. Tällöin projektin tietoja voidaan tarvita useiden vuosienkin kuluttua.

### 3 DATAN TALLENNUS RATKAISUN TOTEUTUS LSK ELECTRICS OY:LLE

LSK Electrics Oy:llä on ajanmukainen lähiverkko ja ulkoinen internetyhteyskin on nopea, joten yrityksen tiloihin on mahdollista toteuttaa joko pilvipohjainen arkistointi ratkaisu tai kovalevypohjainen arkistointiratkaisu. (TAULUKKO 3.) Arkistoitavat tiedostot ovat suuria. Tiedostot sisältävät projekteihin liittyviä valokuvia, PDF- dokumentteja (Portable Document Format) ja CAD-kuvia (Computer-aided Design). Erityisesti PDF-dokumentit ja CAD-kuvat ovat todella suuria. Valokuvista on usein jälkikäteen pienennetty laatua, joten ne vievät vähemmän tilaa. Tallennettua dataa on useita satoja gigatavuja, joten pilvipalveluun perustuva ratkaisu voisi tulla todella kalliiksi ja hitaaksi. Taulukossa 3 on esitetty eri arkistointiratkaisujen vahvuuksia ja heikkouksia.

Tiedostoja tarvitaan harvoin, mutta silloin kun niitä tarvitaan, on tärkeää, että ne ovat heti saatavilla. Tämän perusteella verkkokovalevyihin perustuva tallennusratkaisu olisi paras vaihtoehto yritykselle.

TAULUKKO 3. Datan tallennusratkaisujen vertailu LSK Electrics Oy:n kannalta

<b>Datan tallennusratkaisun vertailu</b>			
<b>LSK Electrics Oy:n kannalta</b>	<b>Paikallinen arkistointiratkaisu</b>	<b>Pilvi tallennustilat</b>	<b>Siirrettävät mediat</b>
<b>Tietoturva</b>	Erittäin vahva	Vahva, tieto ulkopuolisilla palvelimilla.	Lukitussa huoneessa. Kuljetuksissa voi kadota.
<b>Siirtonopeus</b>	SAS-portin / lähiverkon maksimi nopeus: 12Gb/s / 1Gb/s	Yrityksen internet yhteyden nopeus: 100Mb/s	CD/DVD: 11Mb/s Kovalevy, USB: 640Mb/s
<b>Vikasetoisuus</b>	RAID5-tekniikalla erinomainen	Internet yhteyksien katketessa, arkistoon ei pääse käsiksi	Laitteen rikkoutuessa on tieto usein menetty
<b>Käyttöönotto</b>	Aikaa vievä prosessi. LSK Electrics Oy:n henkilöstölle tuttua	Kohtuullisen vaikea. Tieto siirrettävä etäpalvelimille	Nopea

### 3.1 Laitteistoratkaisut

Yrityksellä on nyt oma kovalevypakka, joten vastaavanlainen ratkaisu olisi jo tuttu entuudestaan. Yritys käyttää Hewlett-Packardin laitteita palvelin- ja työasemapuolella, joten Hewlett-Packardin laitteet tulevat luultavasti kysymykseen arkistoinnissakin. Yrityksen toiveena olisi toteuttaa arkistointi mahdollisimman edullisesti. Tämä on toteutettavissa, sillä kovalevyjen ei tarvitse olla nopeita.

Yrityksen kovalevypakka on niin vanha, että arkistoinnissa ei voida käyttää samaa mallia. Kovalevypakasta on julkaistu uusi päivitetty malli, joten uusi päivitetty malli voisi olla sopiva vaihtoehto. Yrityksellä on muutama QNap-valmistajan NAS-kovalevy, joten sekään ei ole poissuljettu vaihtoehto. Yrityksen kovalevypakka ja palvelimet ovat palvelinräkissä, joten suositeltavaa olisi, että ratkaisu saisi palvelinhuoneeseen ja räkkiin. Näin myös nopeat verkkoyhteydet ovat helposti toteutettavissa.

Kovalevypohjaisessa ratkaisussa vaatimuksena ovat mahdollisuudet laajentamiseen eli paikkoja useille kovalevyille ja vähintään kaksi ethernet-liitäntää. Tämä pätee niin NAS-kovalevypalvelin ratkaisulle, kuin uudelle kovalevypakallekin.

#### 3.1.1 Kovalevypakka

Yrityksen tiloihin helpoin ja toimivin ratkaisu olisi uuden kovalevypakan sijoittaminen palvelinhuoneeseen. Yrityksen palvelimet olisi helppo laittaa kommunikoimaan uuden kovalevypakan kanssa, sillä yrityksellä on jo paljon kokemusta erillisistä kovalevypakoista. Koska palvelimet toimivat virtuaalisina palvelimina, uuden tallennustilan lisääminen onnistuisi helposti käytössä olevasta VMwaren Vsphere clientistä. Kovalevypakka-ratkaisuun olisi mahdollista käyttää yrityksen tämänhetkisestä kovalevypakasta päivitettyä versiota HP MSA 2040 SAN Storage-laitteesta. Uusi versio olisi ajantasainen vielä useiksi vuosiksi, ja sen suurta kovalevykapasiteettia olisi mahdollista käyttää tarvittaessa jopa muuhunkin, kuin vain arkistointiin.



KUVIO 13. HP MSA 2040 SAN Storage (hp.com 2015b)

HP MSA 2040 SAN Storage poikkeaa yrityksen nykyisestä kovalevypakasta: nykyisessä ratkaisussa on kovalevy vaakaasennossa, jolloin niitä ei mahdu niin monta kovalevypakkaan. 2040 SAN Storage mahdollistaa 576 teratavun tallennustilan, joka on riittävä paljon suuremmallekin yritykselle. Tämä mahdollistaa myös erilaisien RAID-asetusten käyttöönoton, sillä kovalevyt eivät ainakaan lopu kesken. Laitteessa on siis yhteensä 24 Hot Swappable kovalevyn paikkaa. Laitteessa on neljä 12 gigatavun SAS-liitäntäporttia, jotka käyvät suoraan palvelimien liitäntöihin mahdollistaen suuren tiedonsiirto nopeuden. Yritys ei luultavasti käytä heti kaikkia mahdollisia kovalevypaikkoja, vaan niitä jää runsaasti tyhjiksi. Laitteesta on myös mahdollista saada versio neljällä valokuituportilla. Nämä ovat hyödyllisiä, jos laite päätetään sijoittaa muualle, kuin palvelin huoneeseen, sillä yrityksen kerroksien välissä menee valokuitukaapelia. Laitteessa on myös kaksi tallennuslaitteiden kontrolleria. Perinteisiin ratkaisuihin laitteessa on myös portteja 10 Gigabit-ethernetille (KUVIO 13.) (hp.com 2015a).

Toinen vastaava vaihtoehto kovalevypakaksi olisi Dell PowerVault MD3420. Dellin kovalevypakassa on myös paikat 24 Hot Swappable- kovalevylle. HP:n kovalevypakasta poiketen Dellin versiossa on kaksi 12 gigatavun SAS-

liitäntäporttia. Kovalevy pakka ei sisällä ollenkaan ethernet- tai valokuituportteja. Ne on mahdollista ostaa erikseen (KUVIO 14.) (dell.com 2015b).



KUVIO 14. Dell PowerVault MD3420 (dell.com 2015a)

### 3.1.2 NAS-palvelin

LSK Electrics Oy:n tarpeisiin myös NAS-kovalevy palvelin olisi vartenotettava ratkaisu. Yrityksellä on jo nyt käytössä yksi QNAP:n NAS-palvelin. NAS-kovalevy palvelin toimii palvelimien yhtenä varmuuskopiointisijaintina. Tämä olemassa oleva NAS-palvelin sijaitsee eri tiloissa, kuin palvelinhuone. Olisi siis todennäköistä, että NAS-palvelinratkaisu sijoitettaisiin myös eri tiloihin, kuin missä palvelinhuone on. Tämä onnistuu helposti yrityksen lähiverkon hyvän nopeuden puolesta. NAS-palvelimen etuna olisi se, että NAS-palvelin tekee itse laskutoimitukset ja sisältää oman prosessorin, muistit ja kaiken-mitä normaalisti palvelin sisältää. Tällöin NAS-palvelin ei kuormita yrityksen omia palvelimia.

QNAP TVS-EC1080+ on erittäin vartenotettava vaihtoehto yrityksen tarpeisiin nähden. EC1080+ sisältää kymmenen kovalevyn paikat. Samoin laite sisältää kattavat liitäntämahdollisuudet: kaksi kappaletta 10 Gigabit ethernet- porttia, neljä kappaletta Gigabit ethernet- portteja ja yhteensä yhdeksän USB-väylää. Näistä USB-väylistä kolme on USB3-väyliä. Laite sisältää myös Intel Xeon E3-1245 v3 - prosessorin, joka toimii 3,4 gigahertsin kellotaajuudella. Välimuistia laitteella on vakiona 32 gigatavua. Nämä levy palvelimen tehot ovat enemmän kuin tarpeeksi yritykselle vielä useaksi vuodeksi eteenpäin. Laite kustantaa tällä hetkellä noin



3000 euroa, ilman kovalevyjä. Laitetta ei voida sijoittaa räkkiin, niin kuin yrityksellä oli toiveena (QNAP.com 2015). Laitteen hyvänä ominaisuutena on se, että laite ei ole riippuvainen muista palvelimista. Tällöin laiterikon sattuessa NAS-palvelin jatkaisi toimintaansa. Laitteen kovalevyapaikat ovat niin sanottuja Hot Swappable -kovalevyapaikkoja, eli ne voidaan kovalevyn vaurioitua vaihtaa ilman laitteen sammuttamista.

QNAP SS-EC1279U-SAS-RP on palvelinräkkiin sijoitettava NAS-palvelin. Laitteen ominaisuudet ovat hieman heikommät suorituskyvyltään, kuin mitä TVS-EC1080+ -laitteella. Laitteessa on 12 kovalevyapaikkaa. Kovalevyt on ostettava laitteeseen erikseen. Kovalevyjen kapasiteetissa ei ole mitään rajoituksia, eli mitkä tahansa SAS- tai SATA-kovalevyt toimivat. Laitteen kaikki kovalevyapaikat ovat myös Hot Swappable -paikkoja. Liitännöinä laitteessa on neljä Gigabit Ethernet -porttia. Vakiona muistia laitteella on kahdeksan gigatavuuta, mutta muisti on laajennettavissa 32 gigatavuun asti. Prosessorina toimii sama prosessori kuin TVS-EC1080+-laitteessa: Intel Xeon E3-1245 v3-prosessori. (KUVIO 15.) (QNAP.com 2015b).



KUVIO 15. QNAP SS-EC1279U-SAS-RP- NAS palvelin (QNAP.com 2015b)

### 3.2 Ohjelmistoratkaisut

Yritys käyttää Microsoftin Windows -käyttöjärjestelmiä kaikissa laitteissa, paitsi roskapostin suodatuspalvelimena toimii UNIX-pohjainen ratkaisu. Koko palvelinympäristö on rakennettu VMwaren vSphere Clientin ympärille. Käytännössä tämä tarkoittaa sitä, että yksi palvelin huolehtii siitä, että virtuaaliset palvelimet toimivat ja samalla palvelin jakaa virtuaalisille palvelimille resursseja fyysisiltä palvelimilta.

Yrityksen varmuuskopiointi suoritetaan Acronis VM protect-ohjelmistolla ja Arcserv-ohjelmistolla. Tämä ohjelmisto on helposti muokattavissa uuteen arkistointiratkaisuun. Koska yrityksen varmuuskopiot menevät joka yö niin magneettinauhoillekin kuin levypakallekin talteen, on varmuuskopiointiin otettava huomioon uuden arkistointiratkaisun varmuuskopiot. Kun data siirretään arkistointikovalevyille, niin itse palvelimien varmuuskopiot pienenevät ja näin varmuuskopiointi suoriutuu nopeammin. Tällöin olisi mahdollista toteuttaa erillinen varmuuskopiointityö vain arkistoidusta datasta. Varmuuskopiointityö on mahdollista konfiguroida Acronis VMprotect-sovellukseen. Oma arkistointi menisi omalle varmuuskopiointinauhalleen vain muuttuneiden tietojen osalta. Tämä suorituisi ensimmäisten varmuuskopioiden jälkeen nopeasti, sillä arkistoitu data ei muutu usein.

Jos ratkaisuksi valitaan NAS-pohjainen ratkaisu, on NAS-palvelin liitettävä yrityksen domainiin ja asennettava tarvittavat sovellukset, kuten virusturva ja keskitettyyn hallintaan tarkoitettavat sovellukset. Jos taas ratkaisuna on kovalevypakka, niin kovalevypakka liitetään yrityksen vcenter-palvelimeen, joka jakaa resursseja eteenpäin. Näin jokainen tytäryritys voisi pitää harvemmin käytetyt tiedostonsa omilla kovalevyillään kovalevypakassa ja Windows-ympäristössä ne kovalevyt voidaan liittää oikeille käyttäjille, oikein oikeuksin.

Yritys käyttää Pydio-sovellusta tiedostojen jakamiseen yrityksen ulkopuolisille henkilöille. Tämä sovellus on helppo ottaa käyttöön niin NAS- palvelin pohjaisessa ratkaisussa, kuin kovalevypakka pohjaisessa ratkaisussa. Sovelluksella henkilö jolla, on oikeudet tiedostoihin, voi jakaa tiedoston kenelle vain. Jaetusta tiedostosta lähetetään halutulle henkilölle linkki sähköpostitse ja määritellään,

kuinka kauan tiedosto on henkilöllä käytettävissä. Samalla määritellään käyttöoikeudet, eli saako henkilö tehdä muutoksia tiedostoon vai ei.

### 3.3 Arkistointiratkaisuehdotus

Yrityksen toiveiden ja tulevaisuuden suunnitelmien pohjalta varteenotettavin ratkaisu olisi ottaa käyttöön Hewlett-Packardin kovalevypakka. Yritykselle Hewlett-Packardin laitteet ovat tuttuja, joten ylläpitoa varten ei tarvitse järjestää koulutuksia IT-osastolle, sillä HP:n ohjelmistot eivät ole muuttuneet paljoa vuosien aikana.

Ratkaisussa HP MSA 2040 SAN Storage-kovalevypakka sijoitetaan yrityksen palvelinhuoneeseen palvelinräkkiin. Räkistä kovalevypakka yhdistetään yrityksen fyysisiin palvelimiin SAS-porteilla, jolloin resurssit tulevat palvelimien käyttöön. Tämän jälkeen yrityksen IT-osaajat määrittelevät, paljonko arkistointitilaa tarvitaan ja mitä RAID-tekniikkaa halutaan käyttää. Tässä tapauksessa luultavin RAID-tekniikka on RAID5. RAID5-tekniikka mahdollistaa yhden kovalevyn hajoamisen, ilman että tiedostoja menetetään. Toinen vaihtoehto on RAID6, joka mahdollistaa kahden kovalevyn rikkoutumisen, ilman että dataa menetetään. Kovalevykapasiteettia tulee olla reilusti tämänhetkisiin tarpeisiin nähden, sillä sitä voidaan ottaa pieni erä kerrallaan käyttöön. Kokemus on osoittanut, että jos tilaa on todella paljon käytettävissä, yrityksen henkilöstö ei kiinnitä huomiota, mitä kovalevyille ja verkkosijainteihin tallentavat. Tällöin kertyy todella paljon turhaa tietoa. Yritykselle laitteeseen sopivat halvemmat hitaat kovalevyt, koska tietoa ei käytetä koko aikaa, kuten esimerkiksi palvelimien käytössä olevilla levyillä. Ne sijaitsevat samalla kovalevypakalla kuin nykyiset tallennetut tiedot. Palvelimet tarvitsevat tietoa koko ajan, ja tietoa muokataan koko ajan. Tällöin kovalevyjen tulee toimia nopeasti, jotta luku- ja kirjoitustehtävät eivät vaikuttaisi käyttäjien toimiin. Kovalevypakassa on niin paljon kovalevypaikkoja, että tarpeen tullen voidaan sijoittaa muutama paikkaan nopeammat kovalevyt. Näitä nopeita kovalevyjä voidaan käyttää mahdollisille tuleville virtuaalipalvelimille.

Yrityksen palvelimiin lisättiin juuri muistia ja prosessoreja, joten ne eivät rasitu liikaa uudesta kovalevypakasta. Vuoden sisään yritys tulee päivittämään palvelimensa fyysisen raudan, joten palvelimet eivät tule olemaan ylikuormitettuja. Ratkaisuna arkistointiongelmaan toimii myös vaihtoehto, jossa nykyinen kovalevypakka siirretään arkistointikäyttöön. Näin uusi kovalevypakka voisi toimia palvelimien tallennussijaintina ja muutenkin nopean tiedon tallennuspaikkana.

#### 4 YHTEENVETO

Toimivalla arkistointiratkaisulla mahdollistetaan LSK Electrics Oy:lle pitkäikäinen tiedontallennussijainti, joka on myös turvallinen. Yrityksen henkilöstöllä on vaivaton pääsy tietoon tarvittaessa ja tieto on myös helposti hallittavaa. Tieto on myös turvassa ulkopuolisilta, sillä ratkaisu sijoitetaan yrityksen omaan palvelinhuoneeseen. Tallennuskapasiteetti tulee nousemaan arkistointiratkaisun myötä huomattavasti, mikä on vastaus yrityksen kovalevytilan vähenemiselle. Arkistointiratkaisun tuoma lisätallennustila toimii myös nykyisien palvelimien ja ratkaisujen kanssa ja mahdollistaa resurssien tehokkaamman käyttämisen. Yritykselle arkistoinnin käyttöönotto on tuttua, sillä ratkaisu sisältää yrityksen IT-henkilöstölle tuttuja laitteita ja ohjelmistoja. Ratkaisu toimii suoraan nykyisten ohjelmistojen kanssa.

Mahdollisia ongelmia voi tulla varmuuskopioinnin toteutuksessa, sillä magneettinauhoille mahtuu vain rajallinen määrä dataa. Vaikka data onkin pakattua, niin sitä on nyt jo todella paljon. Varmuuskopioinnit on perinteisesti suoritettu yö aikana, kun kukaan ei tarvitse palvelimien resursseja. Nykyinen ratkaisu ehtii juuri ja juuri toteuttaa varmuuskopioinnin yön aikana. Mahdollisesti yritys joutuu sijoittamaan osan arkistoidun datan varmuuskopioinnista tapahtumaan yön aikana. Varmuuskopiointiratkaisulle voidaan joutua ostamaan erikseen uusia lisenssejä, jos varmuuskopioitavat sijainnit on rajoitettu sovelluksen lisenssissä.

Ongelmaksi voi koitua myös palvelimien SAS-porttien määrä. Tämä ongelma on helposti korjattavissa, jos yritys sijoittaa lisä SAS-portteihin. Niitä voidaan ostaa palvelimien laajennuskorttipaikkoihin.

Tulevaisuudessa uusi kovalevypakka mahdollistaa palvelimien määrän nostamisen, sillä kovalevytilaa on runsaasti. Kovalevypakka mahdollistaa myös testiympäristöjen rakentamisen virtuaali palvelimitse. Lisä tallennustila mahdollistaa myös uusien ohjelmistoratkaisuiden testaamisen ja käyttöönoton.

Tiedon tallennus on tärkeää lähes jokaiselle yritykselle, sillä oikein toteutettuna tiedontallennus lisää tehokkuutta henkilöstön ja laitteiden toiminnassa. Tiedontal-

lennuksella ja arkistoinnilla yritykset saavat varmuutta tärkeän tiedon tallessa pysymiseen. Ilman oikein toteutettua tiedontallennusta ja arkistointiratkaisua yrityksen tärkeät sähköiset dokumentit ovat vaarassa tuhoutua esimerkiksi laiterikon sattuessa.

## LÄHTEET

Abate, L. 2011. Storage costs: What is average cost savings for a GB of data stored on premise vs. in the cloud? [viitattu 20.9.2015]. Saatavissa: <https://www.quora.com/Storage-costs-What-is-the-average-cost-savings-for-a-GB-of-data-stored-on-premise-vs-in-the-cloud>

Ala-Mutka, K. 2002. Tiedon pakkaus (kompresio) [viitattu 26.3.2015]. Saatavissa: <http://www.cs.tut.fi/etaopetus/titepk/luku20/pakkaus.html>

Archambault, M. 2015. Battle of the cloud storage: OneDrive vs. iCloud vs. Dropbox vs. Google Drive [viitattu 13.5.2015]. Saatavissa: <http://www.win-beta.org/news/battle-cloud-storage-onedrive-vs-icloud-vs-dropbox-vs-google-drive-0>

Cover, T.& Tothomas, J. 2006. Elements of Information Theory. uud. p. Kanada: Wiley-Interscience.

Data-Compression.com. 2015. Theory of Data Compression [viitattu 21.4.2015]. Saatavissa: <http://www.data-compression.com/theory.html#rd>

dell.com. 2015b. PowerVault MD3 Serial Attached SCSI Storage Series [viitattu 14.5.2015]. Saatavissa: <http://www.dell.com/us/business/p/powervault-md32x0-series/fs>

dell.com. 2015a. Dell PoverVault MD3420 [viitattu 14.5.2015]. Saatavissa: <http://www.dell.com/us/business/p/powervault-md32x0-series/fs>

Department of Defence. 2013. Joint Publication 2-0 [viitattu 10.3.2015]. Saatavissa: [http://www.dtic.mil/doctrine/new\\_pubs/jp2\\_0.pdf](http://www.dtic.mil/doctrine/new_pubs/jp2_0.pdf)

ee oulu.fi. 2015. Tiedostojen suojaaminen [viitattu 12.5.2015]. Saatavissa: [http://www.ee oulu.fi/~olli/Courses/unixc/suoj\\_tied.html](http://www.ee oulu.fi/~olli/Courses/unixc/suoj_tied.html)

Harris, T. 2015. How File Compression Works [viitattu 13.3.2015]. Saatavissa: <http://computer.howstuffworks.com/file-compression.htm/printable>

hp.com. 2015b. HP MSA 2040 SAN Storage [viitattu 14.5.2015]. Saatavissa: <http://www8.hp.com/us/en/products/disk-storage/product-detail.html?oid=5386548#!tab=features>

hp.com. 2015a. HP MSA 2040 SAN Dual Controller SFF Storage/S-Buy [viitattu 14.5.2015]. Saatavissa: [http://store.hp.com/webapp/wcs/stores/servlet/us/en/pdp/hp-msa-2040-san-dual-controller-sff-storage-s-buy?Redirect\\_SMB\\_ETR=Yes&CatalogCategoryID=#!&TabName=specs](http://store.hp.com/webapp/wcs/stores/servlet/us/en/pdp/hp-msa-2040-san-dual-controller-sff-storage-s-buy?Redirect_SMB_ETR=Yes&CatalogCategoryID=#!&TabName=specs)

Infogineering. 2015. The Differences Between Data, Information and Knowledge [viitattu 11.3.2015]. Saatavissa: <http://www.infogineering.net/data-information-knowledge.htm>

Julaakso. 2015. Pakkaaminen [viitattu 26.3.2015]. Saatavissa: <http://koti.mbnet.fi/julaakso/pakaaminen.htm>

Jääskeläinen. 2013. Tallennustekniikat [viitattu 21.4.2015]. Saatavissa: <http://mpc.fi/nettilehti/pdf/1402201336.pdf>

Krimaka.net. 2012. OSI-malli [viitattu 12.2.2014]. Saatavissa: <http://www.krimaka.net/tietotekniikka/verkko-ja-ethernet/osi-ja-tcp-ip-mallit.html>

Lal Shimpi, A. 2010. Intel & Micron Announce 25nm NAND Flash Production, SSDs to get Bigger/Cheaper in Q4 [viitattu 21.4.2015]. Saatavissa: <http://www.anandtech.com/show/2928>

Lock, I. 2010. Tape backup vs disk backup [viitattu 11.5.2015]. Saatavissa: <http://www.computerweekly.com/podcast/Tape-backup-vs-disk-backup>

Mitra, A. 2012. Classifying data for successful modeling [viitattu 13.3.2015]. Saatavissa: <http://www.dwbiconcepts.com/data-warehousing/12-data-modelling/101-classifying-data-for-successful-modeling.html>

Mitroff, S. 2014. Cloud Storage Comparison [viitattu 14.5.2015]. Saatavissa: <http://www.cnet.com/how-to/onedrive-dropbox-google-drive-and-box-which-cloud-storage-service-is-right-for-you/>



Mitroff, S. 2014. OneDrive, Dropbox, Google Drive and Box: Which cloud storage service is right for you? [viitattu 14.5.2015]. Saatavissa:

<http://www.cnet.com/how-to/onedrive-dropbox-google-drive-and-box-which-cloud-storage-service-is-right-for-you/>

Murelli, G. 2007. RAID6 – Do you really want it? [viitattu 21.4.2015]. Saatavissa:

<http://www.techrepublic.com/blog/the-enterprise-cloud/raid-6-do-you-really-want-it/>

Natarajan, R. 2010. RAID 0, RAID 1, RAID 5, RAID 10 Explained with Dia-

grams [viitattu 21.4.2015]. Saatavissa: <http://www.thegeekstuff.com/2010/08/raid-levels-tutorial/>

Parker, M. 2015. What's the best cloud storage service to use? [viitattu

14.5.2015]. Saatavissa: <http://www.trustedreviews.com/opinions/google-drive-vs-icloud-drive-vs-dropbox-vs-onedrive>

pcworld.co.uk. 2015. Data Storage [viitattu 13.5.2015]. Saatavissa:

<http://www.pcworld.co.uk/gbuk/data-storage-76-u.html>

Pehkonen, J. 2013. Sähköinen arkistointi tarvitsee kokonaisarkkitehtuuria [viitattu

12.5.2015]. Saatavissa: <https://gofore.com/it-johtaminen/sahkoinen-arkistointi-tarvitsee-kokonaisarkkitehtuuria/>

Poelker, C. 2013. Varmuuskopio tekniikoiden hintavertailu: kovalevy pohjainen

ratkaisu [viitattu 11.5.2013]. Saatavissa: <http://www.computerworld.com/article/2475237/data-center/tape-versus-disk--the-backup-war-exposed.html>

Qnap.com. 2015b. SS-EC1279U-SAS-RP [viitattu 14.5.2015]. Saatavissa:

<https://www.qnap.com/i/useng/product/model.php?II=111&event=2>

QNAP.com. 2015a. TVS-EC1080+ [viitattu 14.5.2015]. Saatavissa:

<https://www.qnap.com/i/useng/product/model.php?II=184&event=2>

Raisamo, R. 2014. OSI-malli [viitattu 12.2.2014]. Saatavissa:

<http://www.cs.uta.fi/~rr/tvpk/tvpk2.pdf>

- Rouse, M. 2015. Data archieving [viitattu 26.3.2015]. Saatavissa: <http://searchdatabackup.techtarget.com/definition/data-archiving>
- Rouse, M. 2013. Cloud backup (online backup) [viitattu 11.5.2015]. Saatavissa: <http://searchdatabackup.techtarget.com/definition/cloud-backup>
- Seagate.com. 2015. RAID Modes [viitattu 16.11.2015]. Saatavissa: <http://www.seagate.com/gb/en/manuals/network-storage/business-storage-nas-os/raid-modes/>
- Sharon, D. 2015. The Database: An Aesthetics of Dignity [viitattu 12.3.2015]. Saatavissa: [http://artsites.ucsc.edu/faculty/sdaniel/publications\\_all/DBA\\_essay.pdf](http://artsites.ucsc.edu/faculty/sdaniel/publications_all/DBA_essay.pdf)
- Spector, L. 2013. Backing up your entire drive: Cloning vs. imaging [viitattu 12.5.2015]. Saatavissa: <http://www.pcworld.com/article/2029832/backing-up-your-entire-drive-cloning-vs-imaging.html>
- Snyder, S. 2013. Tape Backup vs Hard Disk Backup: What Does the Future Hold? [viitattu 12.5.2015]. Saatavissa: <http://www.storagecraft.com/blog/tape-backup-vs-hard-disk-backup-what-does-the-future-hold/>
- Strömberg, J. 2015. Sähköisen tiedon arkistointi yrityksen näkökulmasta [viitattu 11.5.2015]. Saatavissa: [http://www.liikearkistoyhdistys.fi/files/sahkoinen-arkistointi-nytko-totta!\[4\].pdf](http://www.liikearkistoyhdistys.fi/files/sahkoinen-arkistointi-nytko-totta![4].pdf)
- Microsoft. 2015b. What is Microsoft Azure? [viitattu 13.5.2015]. Saatavissa: <http://azure.microsoft.com/fi-fi/overview/what-is-azure/>
- Microsoft. 2015a. Microsoft Azure Calculator [viitattu 13.5.2015]. Saatavissa: <http://azure.microsoft.com/fi-fi/pricing/calculator/>
- Wikipedia. 2005. OSI-malli [viitattu 12.2.2014]. Saatavissa: <http://fi.wikipedia.org/wiki/K%C3%A4ytt%C3%A4j%C3%A4:MiaK/Kuvat>
- Wikipedia. 2014. Data [viitattu 13.3.2015]. Saatavissa: <http://en.wikipedia.org/wiki/Data>

Wikipedia. 2015. Big Data [viitattu 13.3.2015]. Saatavissa: [http://fi.wikipedia.org/wiki/Big\\_data](http://fi.wikipedia.org/wiki/Big_data)