jamk.fi

Yritystason varmistusratkaisujen vertailu

Tuomo Miettinen

Opinnäytetyö
Helmikuu 2020
Tekniikan ala
Insinööri (AMK), tieto- ja viestintätekniikka

Jyväskylän ammattikorkeakoulu
JAMK University of Applied Sciences

jamk.fi

Kuvailulehti

Tekijä(t)	Julkaisun laji	Päivämäärä
Miettinen, Tuomo	Opinnäytetyö, AMK	Helmikuu 2020
	Sivumäärä 61	Julkaisun kieli Suomi
		Verkkojulkaisulupa myönnetty: x

Työn nimi
Yritystason varmistusratkaisujen vertailu

Tutkinto-ohjelma
Tieto- ja viestintätekniikka

Työn ohjaaja(t)
Matti Mieskolainen, Esa Salmikangas

Toimeksiantaja(t)
Telia Inmics-Nebula Oy

Tiivistelmä

Telia Inmics-Nebulan jatkuvaa palvelua tuottavan tiimin vastuualueisiin kuuluu sisäisten ja asiakasympäristöjen varmistuspalveluiden tuottaminen ja ylläpito. Varmistuksien sujuvan hallinnan turvaamiseksi yrityksen täytyy pysyä ajan tasalla erilaisista varmistusratkaisuista ja teknologioista. Tavoitteena oli saada tietoa suosituimmista varmuuskopiointiin ja palautukseen tarkoitetuista eri valmistajien tuotteista sekä käsitellä niiden tärkeimpiä ominaisuuksia. Varmistusratkaisujen vertailun pohjalta yrityksen varmistustiimi voisi myöhemmin valita testattavaksi uusia tuotteita varmistuksien toteuttamiseen.

Työtä varten kerättiin tietoa neljästä eri varmistussovelluksesta ja kolmesta julkista pilvipalvelua tarjoavasta organisaatiosta sekä niiden alustaratkaisuista. Yrityksen jo käytössä olevien varmistusratkaisujen lisäksi työssä tutustuttiin ennestään melko tuntemattomiin varmistusratkaisuihin. Käsittelyyn valitut varmistusratkaisut valittiin yhdessä toimeksiantajan kanssa.

Tuloksena saatiin vertailutaulukoita varmistusratkaisujen ominaisuuksista ja julkipilven hinnoittelusta. Työn teoreettisen tiedon pohjalta lukija saa käsitystä varmuuskopioinnista yleisellä tasolla. Saadun tiedon ja taulukoiden avulla voidaan vertailla eri tuotteiden hyviä tai huonoja puolia. Varmistussovelluksia ei voitu suoraan vertailla pilvialustoihin, joten nämä jaoteltiin erillisiksi kokonaisuuksiksi.

Varmistusratkaisun valinnassa on huomioitava vaaditut palautumisajat ja varmistettavan datan määrä sekä sijainti. Varmistusratkaisujen läpikäynnin ohella pohdittiin myös, miten varmuuskopiointia tehdään tulevaisuudessa. Varmistusratkaisun tarve huomataan viimeistään silloin, kun dataa tarvitsisi palauttaa.

Avainsanat (asiasanat)
Varmuuskopiointi, varmistusratkaisu, palautus, julkinen pilvi, tietovarasto

Muut tiedot (Salassa pidettävät liitteet)

jamk.fi

Description

Author(s)	Type of publication	Date
Miettinen, Tuomo	Bachelor's thesis	February 2020
		Language of publication: Finnish
	Number of pages: 61	Permission for web publication: x

Title of publication

Comparison of enterprise level backup solutions

Degree programme

Information and Communication Technology

Supervisor(s)

Matti Mieskolainen, Esa Salmikangas

Assigned by

Telia Inmics-Nebula Oy

Abstract

Thesis was implemented for the company called Telia Inmics-Nebula Oy. One of the company's teams is working on continuous operations services. This particular team is responsible of implementing and maintaining backup and recovery services in variety of environments. To stay up to date of evolving backup solutions team must learn continuously about new technologies.

The main goal was to gather information of different backup solutions and compare features included. Comparison is based on four market leading backup software and three public cloud platforms. Those backup solutions were selected in co-operation with the company. The backup software and public cloud platforms were two main categories to process in thesis. The theory part defines backing up data in general and its methods.

As the result of comparison team would be able to make conclusions which of the backup solutions is best suited on different backup environments. The comparison is implemented by creating tables of features of backup software and public cloud platforms. Alongside the tables the pros and cons of different backup solutions are mentioned.

In addition, the prospects of enterprise level backup solutions were discussed. When choosing a backup solution, the required recovery times must be considered along with the amount of data and location. The importance of backup solution is certainly noticed when data needs to be restored.

Keywords/tags (subjects)

Backup, backup solution, recovery, public cloud, storage

Miscellaneous (Confidential information)

Sisältö

Kuviot

Taulukot

Terminologia

BLOB	Tietotyyppi määrittelemättömän pituisen binääritiedoston tallentamiseen
DEDUPLIKOINTI	Datan pakkaamistekniikka
DELTA-TIEDOSTO	Tiedostotyyppi muuttuneelle datalle
INSTANSSI	Yksittäinen virtuaalinen tai fyysinen palvelin tai työasema
KLUSTERI	Joukko palvelimia, jotka jakavat resursseja keskenään
NODE	Klusterin yksittäinen palvelin
ON-PREMISES	Yrityksen omissa tiloissa sijaitseva palvelinympäristö
OFFSITE	Varmistuksissa tarkoitetaan fyysisesti eri sijaintia verrattuna tuotantoympäristön sijaintiin
PROXY	Välityspalvelin datan siirrossa
REPLIKOINTI	Tiedon monistaminen, esimerkiksi kopio varmuuskopiosta
REST API	Ohjelmointirajapinta web-palveluiden kehityksessä
SCALE-OUT REPOSITORIO	Laajennettava, vähintään kahden tietovaraston yhdistelmä yhdeksi tietovarastoksi

1 Työn lähtökohdat

Tämä opinnäytetyö käsittelee datan varmistamiseen käytettäviä varmistusratkaisuja nykyajan yritysmaailmassa. Datamäärien eksponentiaalinen kasvu vuosituhannen alusta korostaa luotettavien varmuuskopiointiratkaisujen tärkeyttä yrityksille, joiden toiminta on riippuvainen niiden olemassa olevista tiedoista. Kriittisen datan häviäminen voi aiheuttaa huomattavan suuria tappioita yrityksille ja pahimmassa tapauksessa lopettaa liiketoiminnan jatkumisen. Varmuuskopioinnin ajatellaan usein kuuluvan yrityksen IT-toimintaan, mutta todellisuudessa se on tärkeä osa koko yritystoimintaa.

Varmuuskopiointi tarkoittaa tietotekniikassa tiedon kopioimista ja sen tallentamista toiseen sijaintiin. Varmuuskopiointia tehdään, jotta varmistettu data voidaan palauttaa poikkeustilanteen sattuessa. Esimerkiksi kiintolevyn hajoaminen tai muu laitevika on yleisin syy datan menetykseen, mutta seuraavaksi eniten poikkeustilanteita aiheuttaa kuitenkin ihminen. Datan häviämistä voidaan ennaltaehkäistä säännöllisellä ja automatisoidulla varmuuskopioinnilla sekä rajoittamalla varmistusympäristöihin pääsyä. Toimivia varmistusratkaisuja on monia, joista voidaan valita tarpeisiin sopiva ja yrityksen vaatimukset täyttävä vaihtoehto. Tässä opinnäytetyössä keskitytään nykyhetken johtavassa markkina-asemassa olevien yritysten kehittämiin varmistusratkaisuihin ja vertaillaan niiden sisältämiä ominaisuuksia. Vertailun selkeyttämiseksi työssä pyritään luomaan vertailutaulukoita, jotka sisältävät varmistusratkaisujen yhteensopivuuksia eri alustoilla ja keskeisimpiä ominaisuuksia sekä hinnoittelua.

Toimeksiantaja on erikoistunut ICT-ratkaisujen toteuttamiseen yrityksille. Se on perustettu kahden kokeneen IT-talon pohjalta. Inmics perustettiin vuonna 1989 Jyväskylässä Jukka Autereen toimesta vastaamaan kuluttaja-PC:iden kasvavaan kysyntään. Nebula sen sijaan sai syntynsä vuonna 1997 espoolaisen nuoren miehen ideasta toteuttaa IT-palveluja yrityksien tarpeisiin. (Telia Inmics-Nebula 2019.) Vuosina 2017 ja 2018 Telia Company osti molemmat yritykset, jonka jälkeen fuusioituminen Telia Inmics-Nebulaksi alkoi. Yrityksessä on jatkuvaa palvelua tuottava varmistustiimi, jonka toimintaan kuuluu varmuuskopiointi ja palautus.

Opinnäytetyön tavoitteena oli, että vertailun perusteella saadaan tietoa erilaisten yritysten ympäristöihin sopivista varmistusratkaisuista. Työssä pyrittiin tutustumaan varmuuskopiontiin liittyviin alustaratkaisuihin ja sopivan tallennusmenetelmän valintaan sekä tarkastelemaan julkipilven osuutta varmuuskopioinnissa.

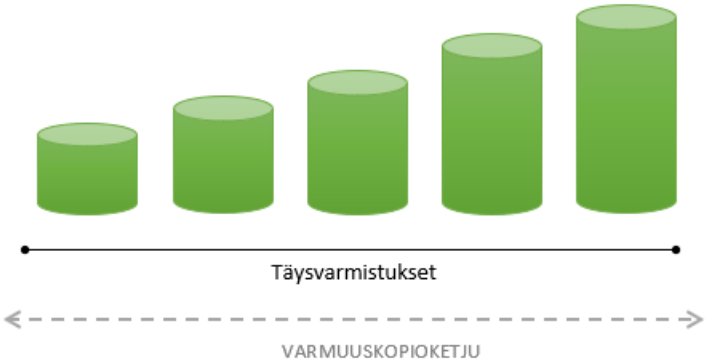
2 Varmuuskopiointi

2.1 Varmuuskopiointimenetelmät

Varmuuskopioinnin suunnittelussa on järkevää ottaa huomioon, minkälaista dataa on tarkoitus varmistaa, ja kuinka usein varmistettava datamäärä muuttuu. Halutaanko datan olevan nopeasti saatavilla, ja minne varmuuskopiot tallennetaan. Erilaisten varmuuskopioiden toteutukseen voidaan valita sopiva menetelmä ja sitä tukeva varmistussovellus. Yleisimpiä varmuuskopiointimenetelmiä ovat täysi varmistus, lisäysvarmistus, ja eroavuusvarmistus.

2.1.1 Täysi varmistus

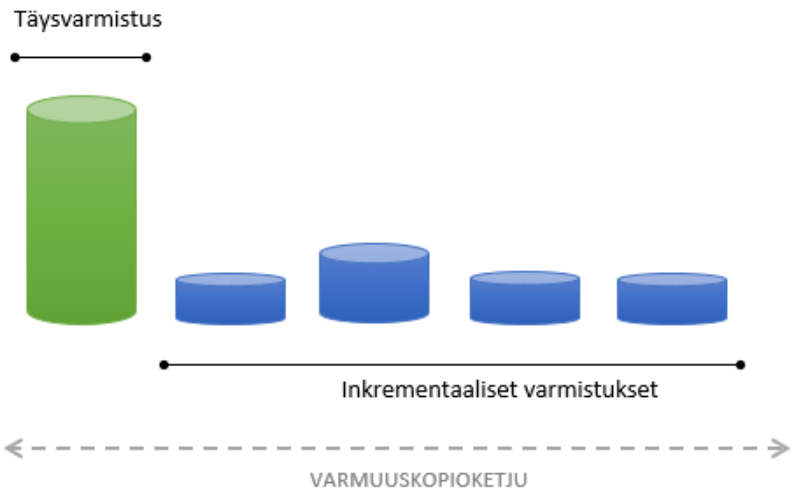
Ensimmäinen kopio varmistettavasta datasta on yleensä täysi varmuuskopio. Täysvarmistus kopioi kaikki varmistukseen valitut tiedostot kansiorakenteineen yhdeksi kokonaiseksi varmuuskopioksi (ks. kuvio 1). Täysvarmistuksen ottaminen vie enemmän aikaa ja tallennustilaa, mutta se on datan varmistamisen ja palautuksen kannalta kattavin menetelmä. (Backup types 2018; Koskinen 2017, 10.)



Kuvio 1. Täysi varmistus

2.1.2 Inkrementaalien varmistus

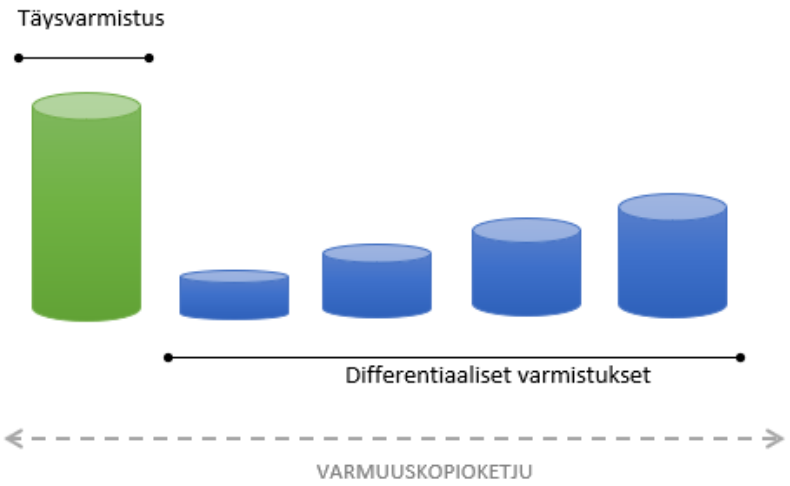
Inkrementaalinen varmistus kopioi viimeisimmän täysvarmistuksen tai inkrementaalin jälkeen vain muuttuneen ja lisätyn datan (ks. kuvio 2). Prosessi muuttaa tiedoston arkistointimäärettä, eli niitä tiedostoja ei varmuuskopioida, jotka on jo merkattu. Inkrementaalisen varmistuksen etuihin kuuluvat varmistuksen ottamisen nopeus ja verrattaen vähäinen tallennustilan tarve, mikäli varmistettava datamäärä ei kasva poikkeavan paljon. Inkrementaalisen varmistusketjun olisi tärkeää pysyä ehjänä, sillä yhdenkin palautuspisteen puuttuminen tai rikkoontuminen tekee varmuuskopiosta palautuskelvottoman. (Backup types 2018.)



Kuvio 2. Inkrementaalinen varmistus

2.1.3 Differentiaalinen varmistus

Differentiaalisessa varmistuksessa (ks. kuvio 3) on sama toimintaperiaate kuin inkrementaalisessa varmistuksessa, mutta muuttuneiden ja lisättyjen tiedostojen arkistointimääre ei muutu. Tällöin myös edellisten differentiaalisten varmistusten tiedostot kopioidaan. Varmuuskopioiden datamäärät voivat kasvaa suuriksi differentiaalisissa varmistuksissa, joten varmistusketjun pituus olisi hyvä määrittää tallennuskapasiteettiin sopivaksi. (Backup types 2018.)



Kuvio 3. Differentiaalinen varmistus

Snapshot

Snapshot (tilannevedos, tilannekuva) tarkoittaa kokonaisen järjestelmän, levyn tai sovelluksen tilan kopiointia (Lyon 2018). Tilannekuvan ottamalla voidaan palautua vaiheeseen, jolloin tilannekuva on otettu. Snapshotin tilaan voidaan palautua nopeasti. Toisaalta tallennustilan tarve on tilannekuvien määrän kasvaessa suurempi verrattuna yhteen syklitettyyn täysvarmistukseen ja inkrementaaliketjuun. Tilannekuvan ottaminen ei ole varsinaista varmuuskopiointia, vaan sen tarkoituksena on tilanteen hetkellinen varmentaminen muutoksia tehdessä.

2.1.4 Menetelmien vertailu

Varmistuksien tarve kuluttajalle on lähtökohtaisesti yhtä tärkeää yritysmaailmaan verrattuna. Varmistuksia toteutetaan vain eri mittakaavassa, eikä kuluttajan tarvitse ottaa huomioon esimerkiksi varmuuskopioinnin ajankohtaa ja verkon kuormitusta. Kuluttajalle voi riittää epäsäännöllisesti tapahtuva lomakuvien manuaalinen kopioiminen puhelimesta tietokoneen kiintolevylle tai pilvipalveluun. Yritystasolla sama toiminta on toteutettu säännöllisesti ajastetulla varmuuskopioinnilla muun muassa tuotantodatasta ja henkilötiedoista. Yhteistä kummallekin taholle on tarve siitä, että tärkeästä datasta on olemassa kopio.

Koskinen (2017) vertailee opinnäytetyössään varmuuskopiointimenetelmiä varmuuskopiointiin ja palautukseen kuluneen ajan sekä datamäärän perusteella. Koska täysvarmistus kattaa varmistettavan datan kokonaisuudessaan, se on menetelmistä helpoin ja nopein palauttaa, ja riski datan häviämiselle on pienempi. Täysvarmistus vie kuitenkin enemmän tallennustilaa, ja varmistuksen ottaminen vie pidempään. Lisäysvarmistuksen etuja on sen ottamisen nopeus ja vähäinen tallennustilan tarve. Toisaalta lisäysvarmistus vaatii enemmän aikaa datan palautukseen. Eroavuusvarmistusta käytettäessä varmistus on hitaampaa, mutta palautusprosessi nopeampi kuin lisäysvarmistuksessa. (Koskinen 2017, 14.)

2.2 Varmuuskopioiden tallennus

2.2.1 Yleistä

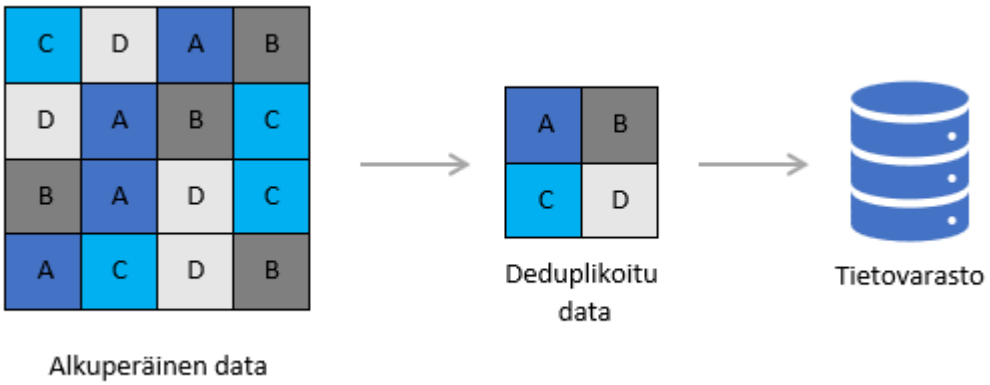
Varmuuskopioiden tallennusmenetelmät eri yrityksien varmistusympäristöissä riippuvat datan käyttötarkoituksesta ja tallennusmediaan investoitavasta summasta. Laitteistot ovat epäluotettavia ja kuluvat ajan myötä.

Hyvänä muistisääntönä varmuuskopioiden tallentamisessa toimii 3-2-1:

- 3 kopiota
- 2 kopiota säilöttynä eri tallennusmedioihin
- 1 kopio maantieteellisesti eri sijainnissa tai pilvipalvelussa

2.2.2 Deduplikoiva tallennusjärjestelmä

Yleisimmin varmuuskopioidaan levylle. Levyvarmistuksen hyötyjä ovat muun muassa nopeus, levyjen jako palvelimien välillä, luotettavuus ja deduplikointi. Deduplikointi on tekniikka, jossa samat datablokit tallennetaan vain kerran (ks. kuvio 4). Tällä saavutetaan huomattava säästö levytilaan, kun tarpeetonta dataa ei kopioida. Deduplikointi voidaan myös tehdä ennen datan siirtoa verkon yli, jolloin yhteys ei kuormitu duplikaattidatasta.



Kuvio 4. Datan deduplikointi

Levyjärjestelmien toimittajat hyödyntävät deduplikointia järjestelmissään ja kilpailu tekniikasta on kovaa. Datan deduplikointi soveltuu varsinkin konesaleihin. Virtuaalipalvelimilla pyörii usein samoja käyttöjärjestelmiä, joissa on deduplikoitavia tiedostoja (Määttä 2013). Deduplikointia voidaan hyödyntää myös virtuaalisissa nauhakirjastoissa. Virtuaalinen nauhakirjasto on tavallisesti kiintolevyllä, johon luodaan perinteisen nauhakirjaston korvike. Tällöin varmistussovellus tunnistaa virtuaalisen nauhakirjaston ja osaa lukea dataa perinteisen nauhakirjaston tavoin. Erona perinteiseen nauhakirjastoon on muun muassa se että, virtuaalinen nauhakirjasto on skaalautuva ja nopeampi siirtonopeuksiltaan (Rouse 2014). Virtuaalisella nauhakirjastolla voidaan korvata perinteinen nauhavarmistusjärjestelmä ja deduplikointia hyödyntämällä saadaan kustannukset minimoitua.

Varmistusjärjestelmien kanssa hyödynnettävistä deduplikoivista tallennusjärjestelmistä tunnetuimpia ovat HPE StoreOnce, Dell EMC PowerProtect DD (aiemmin EMC Data Domain) ja ExaGrid. Tuotteet ovat suorituskykyisiä, mutta kalliita pienempien yritysten käyttöön. Ominaisuudet vaihtelevat valmistajien välillä, mutta esimerkiksi täyteen skaalattu ExaGrid -järjestelmä kykenee täysvarmistukseen 2 petatavun (PB) datamäärälle kerrallaan ja siirtämään sitä jopa 432 teratavun (TB) tuntivauhdilla (ExaGrid 2019).

2.2.3 Tallennusmediat

Kiintolevyjä (hard disk drive, HDD) suositaan datan tallennuksessa hyvän hinta-kapasiteettisuhteen vuoksi. Saatavilla on useiden eri valmistajien levyjä, mutta käytännössä hintavammatkin vaihtoehdot ovat yhtä lailla alttiina korruptoitumaan tai hajoamaan ajan myötä. Suuren tallennuskapasiteetin omaavan kiintolevyn käyttäminen varmistuslevynä voi olla riski. Levykoon kasvaessa yli kymmeneen teratavuun on erityisen tärkeää monitoroida levyä ja varmistaa sen toimivuus säännöllisesti. Tavallisesti suuren kiintolevyn hajotessa myös dataa menetetään enemmän. SSD-levy (Solid State Drive) on kalliimpi mutta nopeampi verrattuna kiintolevyyn. Lisäksi SSD-levy säästää virtaa ja tallennustilaa korkean kaistanleveytensä ansiosta. Vaikka SSD-levyn suorituskyky on parempaa perinteiseen kiintolevyyn verrattuna, sen käyttäminen varmuuskopioinnissa on vielä vähäistä. (Handy 2017.)

Myös NAS-palvelin (Network-attached storage) kuuluu yleisimpiin varmistusmedioihin. Se on tiedostotason verkkotallennuslaite, joka jakaa dataa tietoliikenneverkon kautta. NAS-palvelimeen pääsee käsiksi web-selaimella lähiverkon välityksellä. Tyypillisesti esiasennettu käyttöjärjestelmä on NAS-palvelimen konfigurointia ja hallintaa varten. NAS-palvelin varaa muiden verkkoon kytkettyjen laitteiden tapaan oman IP-osoitteen ja vaatii kirjautumisen käyttäjätunnus-salasanaparilla. NAS-palvelimen hyötyihin kuuluu deduplikoinnin lisäksi muun muassa sen skaalautuvuus. Tallennustilan loppuessa yhdeltä tallennuslaitteelta voidaan hankkia toinen NAS-palvelin ja liittää se samaan lähiverkkoon. (Rouse 2019.)

2.2.4 RAID

RAID (Redundant Array of Independent Disks) on tekniikka, jossa useita levyjä yhdistetään paremman vikasietoisuuden tai nopeuden saavuttamiseksi. Yhdistämällä vähintään kaksi levyä yhdeksi loogiseksi levyksi voidaan vähentää riskiä datan menetykseen levyrikon sattuessa. Eri RAID-tasoista yleisimpiä ovat RAID 0, 1, 5, 6 ja 10. Käytännössä RAID 0 ei käytetä tuotantokäytössä, koska se ei tarjoa minkäänlaista vikasietoisuutta. Uutta RAID-järjestelmää kootessa olisi hyvä miettiä paljonko levytilaa tarvitaan ja kuinka monesta levyrikosta järjestelmän tulisi selvitä kerrallaan. Vikasietoisuuden lisäämiseksi on mahdollista käyttää myös erillistä varalevyä, jotta levyrikon sattuessa palautuminen saadaan aloitettua ilman viivettä. Lisäksi suorituskyky on kannattavaa suhteuttaa siihen, mihin RAID-järjestelmää käytetään. Esimerkiksi RAID 5 -tekniikka vaatii enemmän suorituskykyä toimiakseen moitteettomasti, mutta tarjoaa nopeampaa kirjoitusnopeutta verrattuna RAID 1-tason järjestelmään. (Taylor 2019.)

Suorituskykyisessä RAID 0 -tekniikassa (ks. kuvio 5) datablokit lomitetaan levyille. Tällöin data jakautuu tasaisesti levyjen kesken. Toisin sanoen dataa ei duplikoida ja yhden levyn hajotessa kaikki data menetetään. RAID 1 -tekniikassa puolestaan samaa dataa tallennetaan usealle levylle, jolloin levyn hajotessa on data yhä saatavilla. RAID 1 -tekniikkaa kutsutaan nimellä peilaus. (Mts.)

RAID 0 (lomitus)

Levy 0	Levy 1	Levy 2	Levy 3
0	1	2	3
4	5	6	7
8	9	10	11
12	13	14	15

RAID 1 (peilaus)

Levy 0	Levy 1	Levy 2	Levy 3
0	0	1	1
2	2	3	3
4	4	5	5
6	6	7	7

Kuvio 5. RAID 0 ja RAID 1

RAID 5 -tekniikassa on kyse levyn datablokkien lomituksesta ja pariteetista. Pariteetti on binääritason dataa, jota RAID-järjestelmä lukee ja käyttää datan palautuksessa hajonneelta levyltä. Tekniikka vaatii vähintään kolme fyysistä levyä, joista kaksi on lomitusta varten ja yksi pariteettidatalle. RAID 5 pariteettilevy kuitenkin jaetaan muiden levyjen kesken. Näin ollen jokaiselle lomitetulle RAID 5 -levylle on määritetty oma pariteettiosionsa (P0-P4), jotka esitetään kuviossa 6. (Mts.)

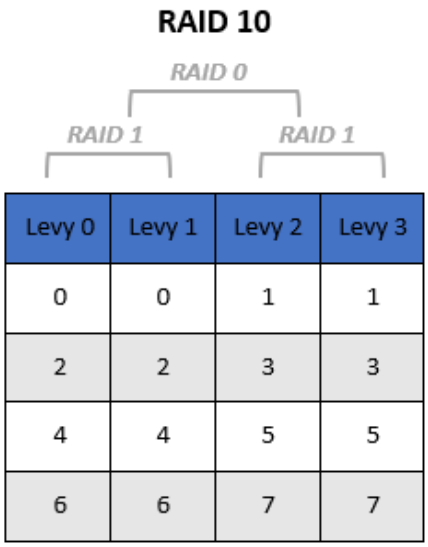
RAID 5 (lomitus + pariteetti)

Levy 0	Levy 1	Levy 2	Levy 3	Levy 4
0	1	2	3	P0
5	6	7	P1	4
10	11	P2	8	9
15	P3	12	13	14
P4	16	17	18	19

Kuvio 6. RAID 5

RAID 6 -tekniikka perustuu kahteen eri pariteettiin. Tällä saavutetaan parempi vikasietoisuus ja RAID 6 kestääkin kahden samanaikaisen levyn hajoamisen. Etenkin levymäärän kasvaessa suureksi RAID 6 on usein paras vaihtoehto varmistuskäyttöön.

RAID 10 (1+0) on sekä vikasietoinen että suorituskykyinen. Siinä yhdistyy lomitus (RAID 0) ja peilaus (RAID 1), jolloin kirjoitus- ja lukuoperaatiot ovat nopeita ja levyn hajotessa dataa ei menetetä. RAID 10 -järjestelmän kokoaminen vaatii vähintään neljä levyä (ks. kuvio 7) ja on RAID -tekniikoista kallein. (Mts.)



Kuvio 7. RAID 10

2.2.5 Nauhavarmistus

Nauhavarmistuksella tarkoitetaan datan varmuuskopiointia fyysiselle magneettinauhalle. Nauhalle varmistaminen on tyypillisin vaihtoehto tietojen pitkäaikaissäilytyksessä tai arkistoinnissa. Varmistusnauhojen heikkouksiin kuuluvat niiden rajallinen tallennuskapasiteetti, hidas kirjoitus- ja lukunopeus, korkeat ylläpitokustannukset sekä palautusten epävarmuus. (Varmuuskopiointipalvelut n.d.)

Varmistusnauhoihin vakiintunut LTO-teknologia (Linear Tape-Open) on tallennusstandardi, joka kertoo nauhan ominaisuuksista, kuten tallennuskapasiteetista ja datan pakkaussuhteesta. Uusimmat käytössä olevat kahdeksannen sukupolven LTO8-nauhat kykenevät säilömään pakkaamatonta dataa 12TB nauhaa kohden ja pakattuna tallennuskapasiteetti yltää 30TB saakka. Teknologian on kehittänyt LTO Ultrium, joka ilmoittaa tulevaisuudessa julkaistavan LTO12-nauhan tallennuskapasiteetin olevan jopa 480TB. (LTO Ultrium 2019.)

Datan kirjoittaminen nauhalle vaatii mekaanisen nauhurin, joka on nauhavarmistuksen elinkaaren kannalta ongelmallista. Nauhurin osat kuluvat ja datan lukemiseksi useita vuosia vanhoilta nauhoilta täytyy olla myös eri nauhateknologioita tukevia

nauhureita saatavilla. Nauhalle varmistaminen on kuitenkin perinteisistä tallennusmenetelmistä kestävin ja luotettavin säilymisen kannalta, sillä magneettinauhat eivät ole alttiita esimerkiksi virtapiikeille ja ne voidaan varastoida vuosikymmeniksi.

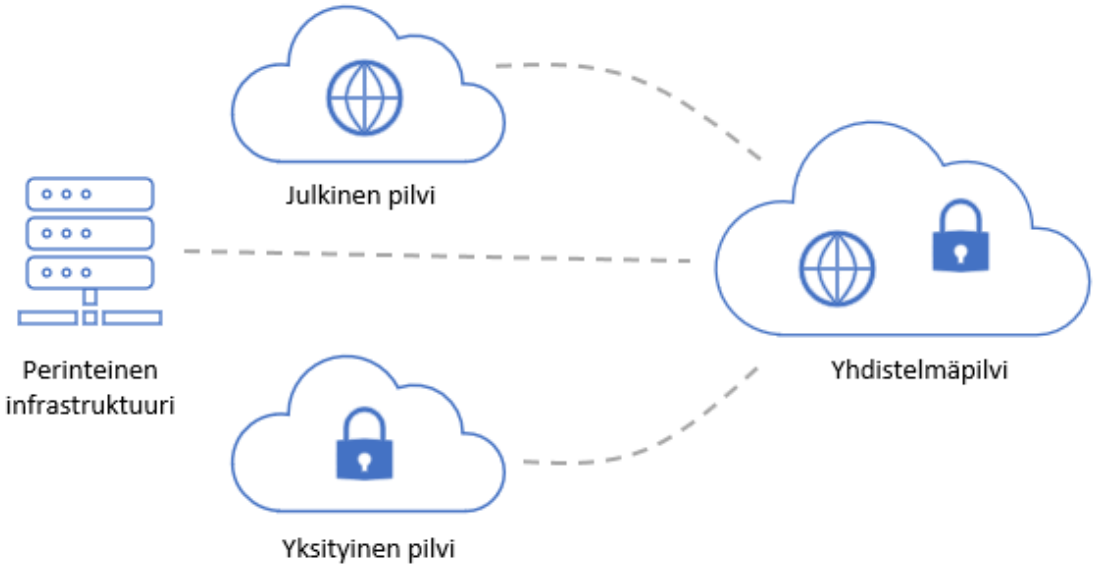
Sähköinen arkistointi

Tietojen arkistoinnilla tarkoitetaan datan pitkäaikaissäilytystä. Arkistoitavia tietoja ei tarvita aktiivisesti, mutta niitä pitää säilyttää esimerkiksi lakiasetusten nojalla tai mahdollista myöhempää käyttöä varten. Arkistoiduille tiedoille määritetään säilytysaika, jonka päättymisen jälkeen tietoja ei tarvitse enää pitää tallessa. Arkistointi sekoitetaan usein varmuuskopiointiin, jossa tiedostosta otetaan kopio säännöllisesti (Poikonen 2012, 10). Arkistoidusta tiedostosta voi olla olemassa vain yksi kopio esimerkiksi magneettinauhalla yrityksen paloturva- tai kassakaapissa.

2.2.6 Etävarmistus

Etävarmistuksessa data tallennetaan muualle kuin yrityksen omissa tiloissa sijaitseviin palvelimiin tai tallennuslaitteisiin. Tällöin varmuuskopiointi ei ole kytköksissä yhteen sijaintiin, ja säilyvyys datalle tuhon sattuessa on paremmin turvattu. Säilyvyyden varmistamiseksi kohdesijainnin tulisi olla paloturvallinen. Etävarmistusta tehdään salattua yhteyttä pitkin, jottei siirron aikana tiedostoja kaapata. Lisäksi etävarmistus vaatii riittävän yhteyden kohde- ja lähdesijainnin välille, jotta siirtoajat varmuuskopioinnissa ja palautumisessa pysyisivät kohtuullisina.

Yleisimmin etävarmistus kohdistetaan pilvipohjaiselle palvelimelle. Pilvi on enemmänkin palvelua, kuin palvelimia. Pilvessä data on säilöttynä kolmannen osapuolen palvelimilla ja sen hakeminen palvelusta ei välttämättä ole suoraviivainen prosessi. Eri pilvityyppeihin (ks. kuvio 8) kuuluu yksityinen pilvi (private cloud), julkinen pilvi (public cloud) ja näiden kahden sekoitus eli yhdistelmäpilvi (hybrid cloud). Tässä työssä keskitytään julkipilven käyttöön varmuuskopioinnissa, jota käsitellään tarkemmin luvussa 4.



Kuvio 8. Pilvityypit (Virtanen 2019, muokattu)

2.3 Datan palautus ja jatkuvuus

Palautustarve datalle ilmenee silloin, kun jotain on mennyt pieleen tai datasta halutaan palauttaa vanhempia versioita. Kadonnut data voidaan palauttaa varmistuksista, jos se on säilytyskäytänteen (Data retention policy) mukaan saatavilla.

2.3.1 BCDR

Poikkeustilanteen tai datan katoamisen sattuessa yrityksen on tärkeää varmistaa liiketoiminnan jatkuminen (BC, business continuity). Liiketoiminnan jatkumiseen liittyvät kaikki ne prosessit, joilla voidaan ennalta ehkäistä poikkeustilanteita. Poikkeustilanteisiin kuuluvat muun muassa sähkökatkot ja laiterikot sekä erilaiset luonnonkatastrofit, kuten tulvat ja tulipalot. Tähän liittyy tärkeänä osana toimiva palautumissuunnitelma (DR, disaster recovery plan, toipumissuunnitelma). Palautumissuunnitelmassa on luotu käytänteet vahinkojen minimointiin ja poikkeustilanteesta palautumiseen. (Aula 2011.)

2.3.2 RTO ja RPO

Palautumissuunnitelmassa asetetaan tavoitteet palautumisajalle (RTO, recovery time objective) ja palautuspisteiden määrälle (RPO, recovery point objective). Palautumisaika määräytyy siten, kuinka kauan palautuminen kestää normaaliin tilaan poikkeustilanteesta. Palautuspistetavoitteella tarkoitetaan palautuspisteiden määrää eli rajaa, kuinka kaukaa voidaan palauttaa dataa varmuuskopioista. RPO määritellään joko päivinä, tunteina, minuutteina tai sekunteina. Esimerkiksi, jos poikkeustilanne ilmenee kalenterikuukauden 24. päivän aamuna ja viimeisin varmistus on edelliseltä yöltä, tilanteesta voidaan palauttaa enintään 10. päivän tilanne RPO:n ollessa 14 päivää. Palautuspistetavoitetta asettaessa arvioidaan, kuinka paljon yritys on valmis menettämään dataa ennen kuin se vaikuttaa liiketoimintaan. (Puricica 2017.)

3 Varmistussovellukset

3.1 Veeam

3.1.1 Yleistä

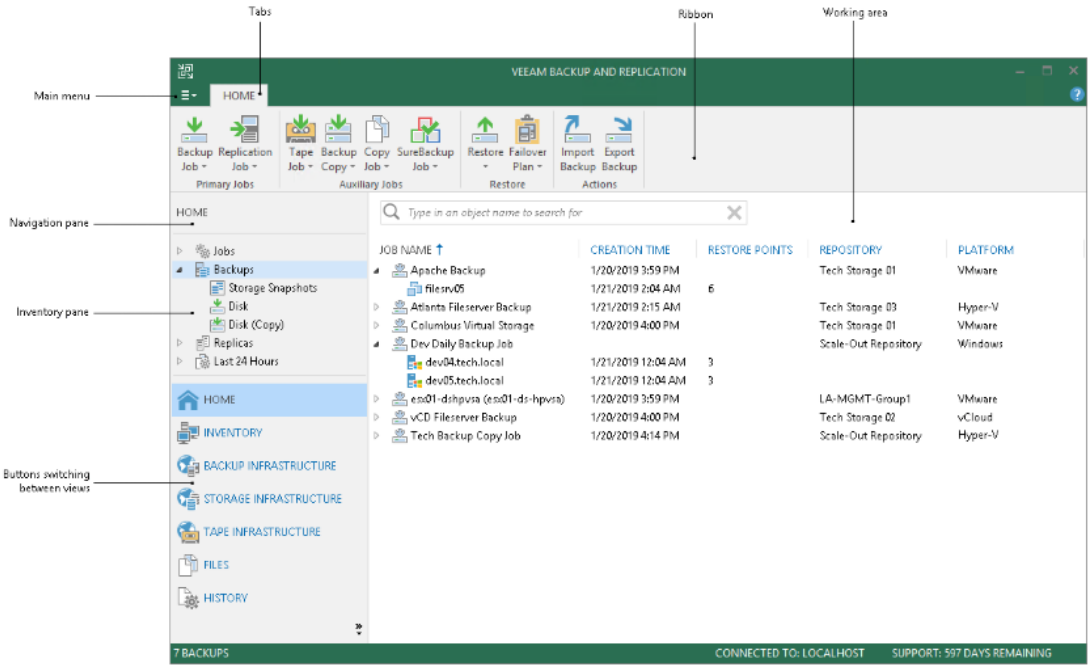
Veeam on kahden venäläisen IT-alan yrittäjän vuonna 2006 perustama varmuuskopiointiin ja palautukseen, datan suojaukseen sekä datanhallintaratkaisuihin erikoistunut ohjelmistoyritys. Veeam kuuluu suurimpiin yksityisen sektorin ohjelmistoalan yrityksiin liikevaihdon yltäessä yli miljardiin dollariin vuonna 2019. Yli 3600 henkilöä työllistävä Veeam onkin vahvassa asemassa varmistusratkaisuja tarjoavien yritysten joukossa. (Veeam 2019.)

Veeam on saanut tunnustusta muun muassa sijoittumalla markkinajohtajien joukkoon kansainvälisen tutkimuslaitos Gartnerin teettämässä varmuuskopiointiratkaisujen vertailussa. Veeamin asiakasmäärät ovat kasvaneet ja toiminta laajentunut, sillä Veeam on investoinut yhteistyöhön konesaliympäristöjä omistavien teknologiayritysten kanssa, joita ovat muun muassa Amazon Web Services, Cisco, HPE, Microsoft, Nutanix, NetApp ja Pure Storage. (Russell 2019.)

Syitä Veeamin menestykseen on monia. Ilman yhteistyötä tietovarastoja tarjoavien yritysten kanssa Veeam ei itsenäisesti olisi voinut saavuttaa markkinajohtajan asemaa (Russell 2019). Veeam kehittää jatkuvasti uusia tuotteita datan suojaukseen ja pyrkii tuottamaan tuotteita, jotka ovat yhteensopivia kuumimpien teknologioiden kanssa. Uusin Veeam-tuoteperheen jäsen on joulukuun 2019 alussa julkaistu Veeam Backup for AWS. Yrityksen 20% vuosikasvun lisäksi sen tuotteiden liikevaihto on kasvanut tasaisesti ja Veeam raportoi sivuillaan, että esimerkiksi luvussa 3.1.3 käsiteltävän Veeam Backup for Microsoft Office 365 -tuotteen suosio on vuodesta toiseen kasvanut 113%. (Veeam 2019.)

3.1.2 Veeam Backup & Replication

Veeamin kehittämä Veeam Backup & Replication -varmistussovellus tarjoaa monipuolisia varmistusratkaisuja eri alustoille, joihin kuuluvat virtuaalisten ja fyysisten ympäristöjen lisäksi pilvivarmistus sekä virtualisoidut konesaliympäristöt. Sovelluksen käyttöliittymä on selkeä ja ohjaava (ks. kuvio 9).



Kuvio 9. Veeam Backup & Replication 9.5 käyttöliittymä (Veeam 2019)

Veeam Backup & Replication vahvuuksia ovat:

- Monipuoliset ja tehokkaat varmuuskopiointivaihtoehdot
- VMware vSphere ja Microsoft Hyper-V
- Kokonaisten virtuaalikoneiden palautus
- Scale-out repositoriot varmuuskopioille
- Replikointi
- Pilvitallennus
- Keskitetty hallinta
- Yhteensopivuus
- Helppokäyttöisyys

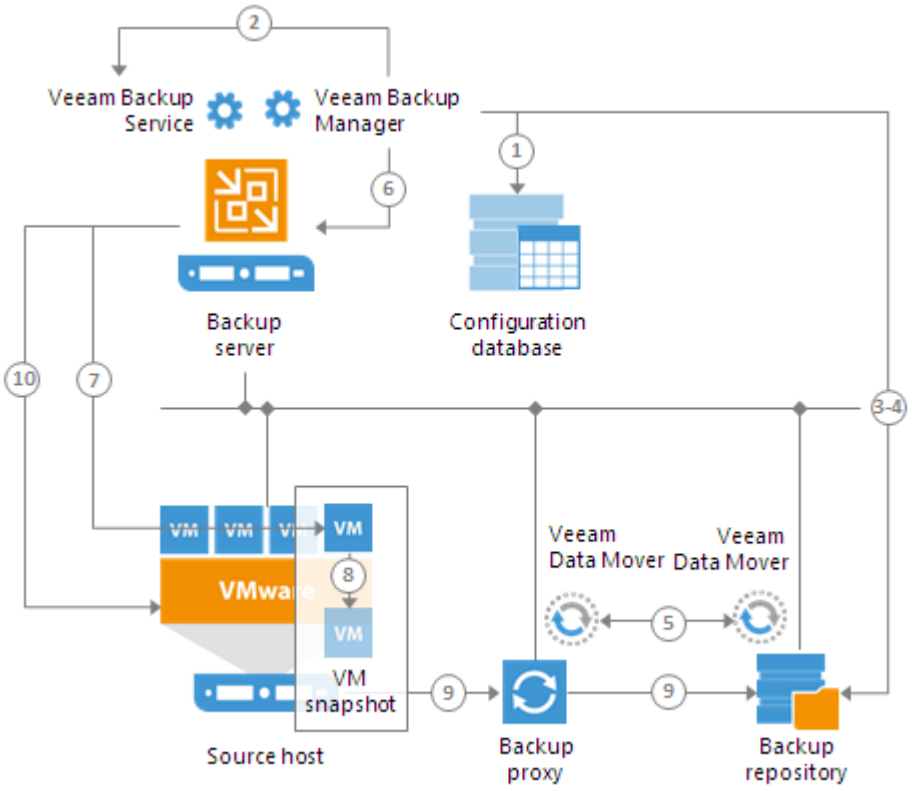
Perinteiseen tapaan varmistuksia organisoidaan varmistussovelluksissa varmistustöillä, eikä myöskään Veeam B&R tee poikkeusta sääntöön. Yksinkertaisuudessaan varmistustyö nimetään kuvaavasti, lisätään varmistettavat virtuaalikoneet tai niiden objektit työhön ja valitaan virtuaalikoneiden varmistusjärjestys sekä tietovarasto, johon varmistukset tallennetaan. Tämän jälkeen varmistustyö ajastetaan suoritettavaksi esimerkiksi yöaikaan tai yrityksen työajan ulkopuolelle, jotta varmistuspalvelimen käyttämän verkon mahdollinen kuormittuminen ei vaikuttaisi suorituskykyyn tai itse varmuuskopiointiprosessiin.

Veeam B&R varmuuskopiointiprosessi vaatii vähintään yhden hostin eli isäntäpalvelimen, backup proxyn ja varmistusrepositorion. Veeam B&R käyttää kahden päätepisteen arkkitehtuuria virtuaalikoneiden datan siirtämiseen, jossa on kaksi Veeam Data Mover -komponenttia. Veeam Data Mover komponentit käsittelevät dataa, pakkaavat sitä ja siirtävät varmistusrepositorioon. Toinen Veeam Data Mover kommunikoi lähdehostin kanssa ja toinen välittää varmistusrepositorion toimintoja. Kummatkin Veeam Data Moverit kommunikoivat keskenään (ks. kuvio 10, kohta 5), jotta yhteys varmistuspalvelimen ja komponenttien välillä olisi mahdollisimman vakaa. (Veeam 2019.)

Varmistustyön alkaessa Veeam Backup Manager prosessi käynnistyy varmistuspalvelimella. Se hakee varmistustyön asetukset konfiguraatiotietokannasta ja määrittää suoritettavat toiminnot jokaiselle työssä olevalle virtuaalikoneelle. Seuraavaksi Veeam Backup Manager muodostaa yhteyden Veeam Backup Serviceen (ks. kuvio 10,

kohta 2), joka hallinnoi koko Veeam varmistusympäristön resursseja varmuuskopiointiprosessin aikana. (Mts.)

Varmuuskopiointiprosessiin kuuluu lisäksi, että Veeam B&R lähettää hostille käskyn ottaa snapshotin (ks. kuvio 10, kohta 8). Prosessoitavan virtuaalikoneen levyt pakotetaan lukutilaan (read-only state) ja jokainen levy saa oman delta -tiedoston. Varmistusprosessin aikana tehdyt muutokset virtuaalikoneilla tallentuvat delta -tiedostoihin, jotka varmistuksen jälkeen yhdistyvät alkuperäisten virtuaalilevyjen tiedostoihin. (Mts.)



Kuvio 10. VMware vSphere varmistusprosessi (Veeam 2019)

Vaikka Veeam B&R varmistussovellus on toimiva ratkaisu monen erilaisen ympäristön varmistamiseen, löytyy siitäkin kehitettävää. Veeam -tuoteperheen lisensointi on ollut monimutkaista ja hämmentävää asiakkaille. Veeam on pitkään käyttänyt tilaus- tai instanssipohjaista lisensointia tuotteilleen perustuen kuormitukseen, eli fyysisten-

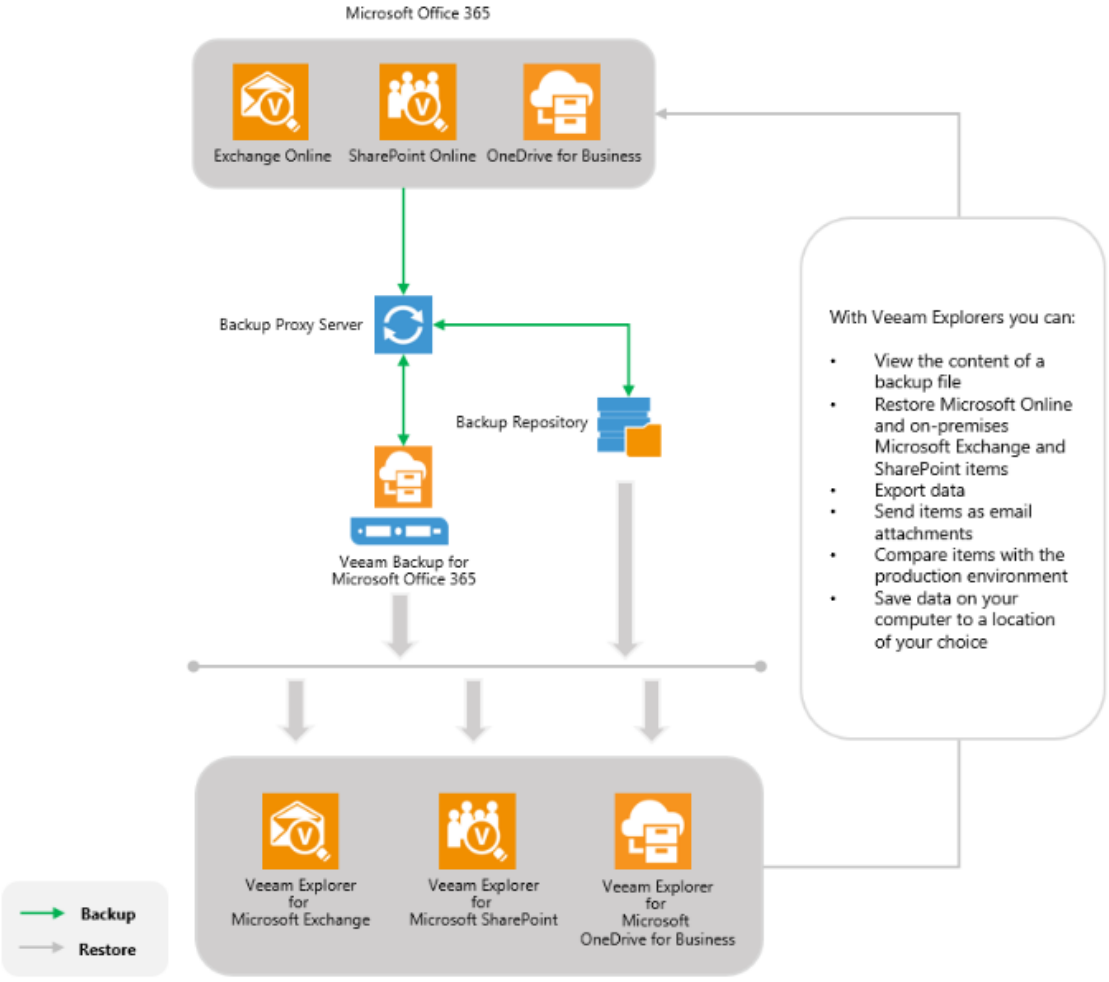
ja virtuaalikoneiden sekä työasemien määrään. Jokaista Veeam-tuotetta varten on tarvinnut oman lisenssin, joten laajemmissa ympäristöissä lisenssikustannukset ovat nousseet korkeiksi. Hinnoittelu on tyypiltään tilauspohjainen eli lisenssejä täytyy uusia säännöllisesti (Veeam 2019). Veeam julkaisi syksyllä 2019 uuden Veeam Universal Licensing -lisensointimallin, joka selkeyttäisi loppukäyttäjiä. Tässä mallissa yksi lisenssi riittää, jolla hallitaan erilaisia ympäristöjä kuormituksiin perustuen (Seget 2019).

3.1.3 Veeam Backup for Microsoft Office 365

Microsoft Office 365 -tuoteperheen suosio on kasvanut etenkin yritysmaailmassa. Jo useamman vuosikymmenen ajan käytössä olleita Microsoftin dokumentinhallintatyökaluja on Microsoft O365 myötä saatavilla pilvipohjaisena tilauspalveluna. Tuoteperheen sovellukset ovat helppokäyttöisiä ja yhteensopivia muiden O365-palveluiden kanssa. Veeam tarjoaa työkalun Microsoft O365 tiedostojen säilytys- ja jakopalveluiden sekä sisällönhallintaratkaisujen varmuuskopiointiin ja palautukseen (ks. kuvio 11). Varmistettavat O365-kohteet on jaoteltu eri tyyppeihin, joita ovat ryhmät, käyttäjät, sivut ja organisaatiot. Ryhmät -objektityyppi nimensä mukaisesti sisältää erilaiset O365 jakeluryhmät, käyttöoikeusryhmät ja dynaamiset käyttöoikeusryhmät. Käyttäjätyyppiin kuuluu jaetut- ja julkiset sähköpostilaatikot sekä käyttäjät. Sivuilla tarkoitetaan Microsoft SharePoint-sivuja ja organisaatiotyyppi kattaa koko yrityksen, eli määritetyn Veeam Backup for Microsoft Office 365 organisaation objektit. (Veeam 2019.)

Veeam Backup for Microsoft Office 365 asennusprosessi on suoraviivainen ja varmuuskopiointia varten O365-sovellukseen lisätään Microsoft Office 365 organisaation käyttäjätunnukset ja konfiguroidaan yhteysasetukset. Asennuksen läpiviemiseksi oletusasetukset yleensä riittävät. Sen jälkeen luodaan varmistustyö, johon yksinkertaisesti määritetään varmistettavat kohteet. Helpointa olisi varmistaa koko organisaatio tai ryhmä, mutta on myös mahdollista valita yksittäisiä objekteja, kuten sähköpostilaatikoita varmistuksen piiriin. Huomioitavaa on, että varmistettavista SharePoint-sivuista voidaan varmistaa joko kokonainen sivurakenne alisivuineen, tai yksittäisiä alisivuja, mutta juurisivun ja sen sisältämien alisivujen valikoivaa valintaa ei

tueta. Veeam Backup for Microsoft Office 365 voidaan asentaa myös AWS- tai Azure -virtuaalikoneelle. Palautus tapahtuu käyttämällä varmistuksien hallintaan ja monitorointiin tarkoitettuja Veeam lisäosia. (Veeam 2019.)



Kuvio 11. Veeam Backup for Microsoft Office 365 toimintaperiaate (Veeam 2019)

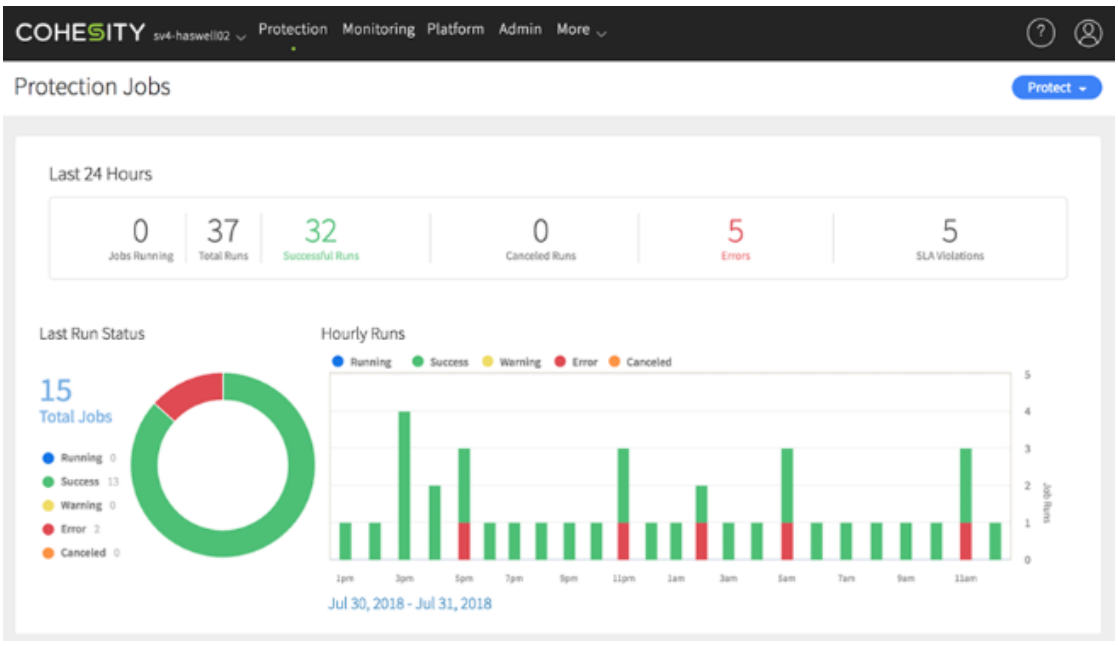
3.2 Cohesity

3.2.1 Yleistä

Cohesityn on perustanut Mohit Aron vuonna 2013. Mohit Aron kuuluu muun muassa Nutanixin perustajiin ja on ollut mukana Googlen tiedostojärjestelmän kehittämisessä. Cohesity on erikoistunut datan hallintaan- ja suojaukseen ja sen päämajan sijaintina on San Jose Kalifornian osavaltiossa. Yrityksen sivuilla kerrotaan, että useissa yrityksissä data on säilöttynä vanhoille laitteille ja eri sijainteihin, jolloin IT-ympäristön dataa on vaikeaa hallita. Ratkaisuna ongelmaan Cohesityn tuotteet muodostavat yhden keskitetyn hallinnan yrityksen datalle. Näistä tuotteista Cohesity Dataplatform on alusta datan hallintaan ja varastointiin, jota täydentää seuraavaksi tarkasteltava yritystason varmuuskopiointiin ja palautukseen kehitetty sovellus nimeltään Cohesity DataProtect. (Cohesity 2019.)

3.2.2 Cohesity DataProtect

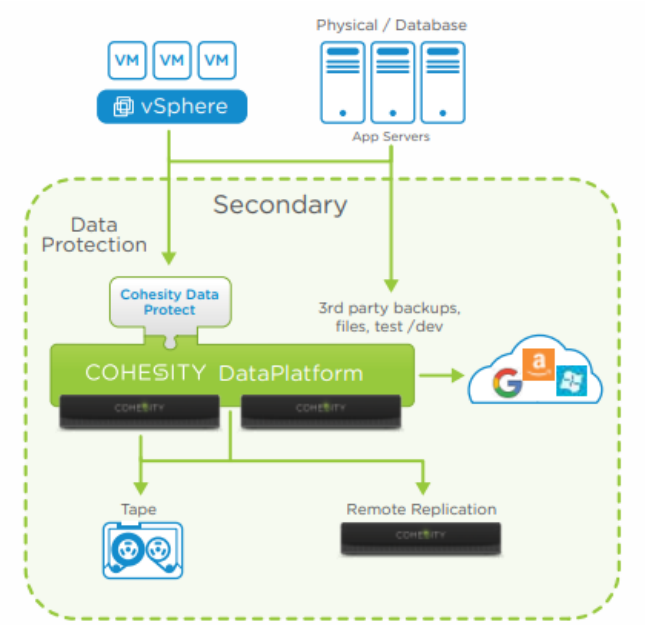
Cohesity DataProtect on varmistussovellus, joka pohjautuu klusteriarkkitehtuuriin. Se toimii scale-out tiedostojärjestelmän päällä, jota käskytetään Cohesity DataPlatform hallintapaneelista (ks. kuvio 12). Hallintapaneeli voidaan asentaa muun muassa konesaliympäristöön, jolloin yritys voi valita käyttääkö Cohesityn omaa tietokonelaitteistoa vai kolmannen osapuolen, kuten Hewlett Packard Enterprisen kanssa yhteensopivaa laitteistoa ympäristön pystyttämiseen. Varsinkin uutta ympäristöä luodessa kätevintä olisi valita Cohesityn toimittama laitteisto, jossa ohjelmisto on esiasennettuna valmiiksi. (Cohesity 2019.)



Kuvio 12. DataProtect näkymä Cohesity DataPlatform hallintapaneelissa (Cohesity 2019)

Cohesity DataProtect sopii fyysisten ja virtuaaliympäristöjen lisäksi erilaisten sovelluksien ja tietokantojen varmistukseen. Cohesityn lisäosien avulla myös yleisimpiin pilvipalveluihin replikointi tai arkistointi on mahdollista (ks. kuvio 13). Varmistettavat kohteen määritetään varmistustöihin, joihin voidaan konfiguroida eri käytänteitä (policy) työkohtaisesti.

Huomionarvoista on, että varmuuskopiot tallennetaan natiivina, eli datan alkuperäismuodossa ja Cohesityn tiedostotason dataan pääsee käsiksi NAS-protokollien välityksellä. Näin ollen dataa ei välttämättä tarvitse palauttaa, kun varmuuskopioita voi tarkastella verkon yli. (Mukhyala, Rao & Simpson 2019.)



Kuvio 13. Cohesity DataPlatform arkkitehtuuri (Cohesity 2019)

Kun tarkastellaan Cohesityn tuotteiden lisensointia voidaan huomata, että vaikka tuotteet on integroitu yhtenäiseen alustaan, jokaista tuotetta varten tarvitsee oman lisenssin. Tällöin on mahdollista hankkia lisenssi esimerkiksi vain tietovarastolle tai varmistussovellukselle, joiden kanssa voidaan käyttää muita yhteensopivia tuotteita. (Mukhyala, Rao & Simpson 2019.)

Cohesity DataProtect sisältää oleellisimmat varmistussovelluksen ominaisuudet ja mahdollisuuden monitoroida tapahtumia, mutta esimerkiksi nauhavarmistuksia se ei itsessään tue, vaan vaatii kolmannen osapuolen QStar archive manager -ohjelman. On myös syytä huomioida, ettei varmistuksien lokien kerääminen sovelluksella ainakaan toistaiseksi ole mahdollista. Toisaalta Cohesityn uusimpiin ominaisuuksiin kuuluvat Microsoft Office 365 Exhange Online varmistus ja GDPR-raporttien (General Data Protection Regulation) luominen sovelluksella. (Mts.)

3.3 Commvault

3.3.1 Yleistä

Commvault on datan suojaukseen ja hallintaan erikoistunut yritys, jonka päämaja sijaitsee Yhdysvalloissa New Jerseyn osavaltiossa. Commvault perustettiin vuonna 1996, ja sen ratkaisut on kohdistettu yritystason asiakkaille. Commvault kuuluu pitkään markkinoilla olleiden varmistussovellusten joukkoon ja onkin vuonna 2019 yksi suurimmista varmistusratkaisuja tarjoavista yrityksistä yli 2500 työntekijällään ja noin 711 miljoonan dollarin liikevaihdollaan. Commvaultin tärkeimpiä kolmannen osapuolen liikekumppaneita ovat muun muassa Cisco, Citrix, Fujitsu, Hewlett Packard Enterprise, Microsoft, Oracle ja VMware. (Commvault Annual Report 2019.)

Commvault tuoteperhe uudistui vuonna 2019. Nykyinen versio muodostuu neljästä eri datanhallintaratkaisusta, jotka on tarkoitettu varmuuskopiointiin, palautukseen, palvelun- ja tiedonhallintaan sekä tietovarastoja varten. Edellä mainitut toiminnallisuudet ovat Commvaultin mukaan oleellisimpia ratkaisuja nopeasti kehittyvien yritysten tarpeisiin. Jotta tuotteisiin liittyvät tukipyynnöt ja käyttöönotot saataisiin tehokkaasti valmiiksi, Commvaultilla on lisäksi saatavilla vuorokauden ympäri oleva maailmanlaajuinen asiakaspalvelu. (Mts.)

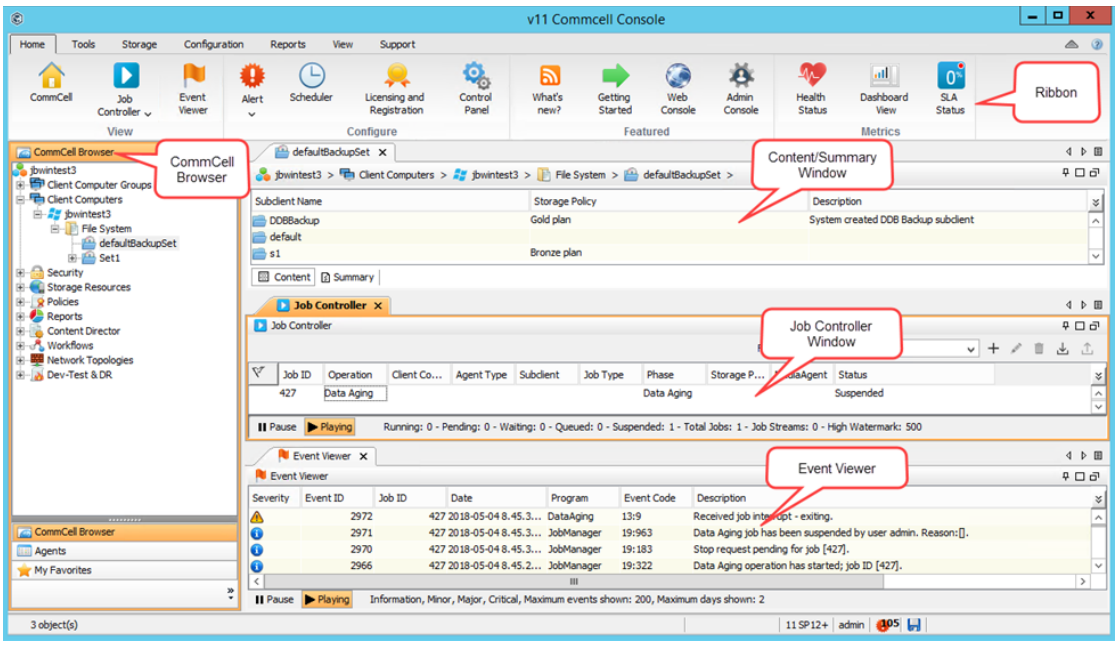
Yksinkertaiseksi mainostetut Commvault tuotteet käyttävät yhden alustan arkkitehtuuria. Se tarkoittaa, että yhden ohjelmistoalustan kautta operoidaan eri tuotteita ja ympäristöjä. Näin ollen varmistusympäristöjen hallinta on helpompaa kuin useaan eri ohjelmistoon hajautettujen kokonaisuuksien ylläpito. (Mts.)

Vuoteen 2018 asti Commvault tuoteperheen lisenssit ovat olleet ikuisesti voimassa olevia ohjelmistolisenssejä, instanssipohjaisia tai tallennuskapasiteettiin perustuvia lisenssejä. Vuosittaisessa raportissaan Commvault on vakuuttunut, että tilauspohjaiset lisensointimallit kasvattavat suosiotaan ja tuottavat tulosta yritykselle tulevina vuosina. (Mts.)

Koska Commvaultilla on muiden varmistussovelluksia tarjoavien organisaatioiden tapaan useampia datanhallintaratkaisuja, käsitellään alempana vain varmuuskopiointiin ja palautukseen tarkoitettua Commvault Complete Backup & Recovery varmistussovellusta. Näin ollen luvussa 3.3.2 Commvault Complete Backup & Recovery -tuotetta käsiteltäessä käytetään yksinkertaisesti nimeä Commvault.

3.3.2 Commvault Complete Backup & Recovery

Commvault sisältää tässä opinnäytetyössä käsiteltäville varmistussovelluksille tyypilliset ominaisuudet. Se tukee levy- ja nauhavarmistuksia, siinä on integroituna datan salaus ja deduplikointiominaisuudet sekä yhteensopivuus pilvitallennukselle. Käyttöliittymä on Commvaultissa yksinkertainen (ks. kuvio 14). Varmistusten hallinta ja monitorointi tapahtuu käytännössä pääikkunasta. Koska Commvault tulkee REST API -rajapintoja, voi pääikkunaa muokata käyttäjän tarpeisiin sopivaksi.



Kuvio 14. Commvault v11 käyttöliittymä (Commvault Documentation 2019)

Commvaultin tärkeimpiä ominaisuuksia ovat:

- Virtuaalikoneiden ja tiedostojen arkistointi

- Kryptaus ja deduplikointi
- Loppukäyttäjän datan suojaus
- Disaster recovery
- Monitorointi
- Tekoälyn ja koneoppimisen hyödyntäminen
- Laaja tuki eri tiedostojärjestelmille ja sovelluksille sekä virtuaalialustoille

Gartnerin tutkimuksessa kerrotaan, että Commvault valitaan kattavien datan suojausominaisuuksien perusteella yleensä yrityksen varmistusratkaisuksi. Muun muassa integroidun deduplikoinnin ja laajan yhteensopivuuden ansiosta Commvaultia suositaan yritysten keskuudessa. (Mukhyala, Rao & Simpson 2019.)

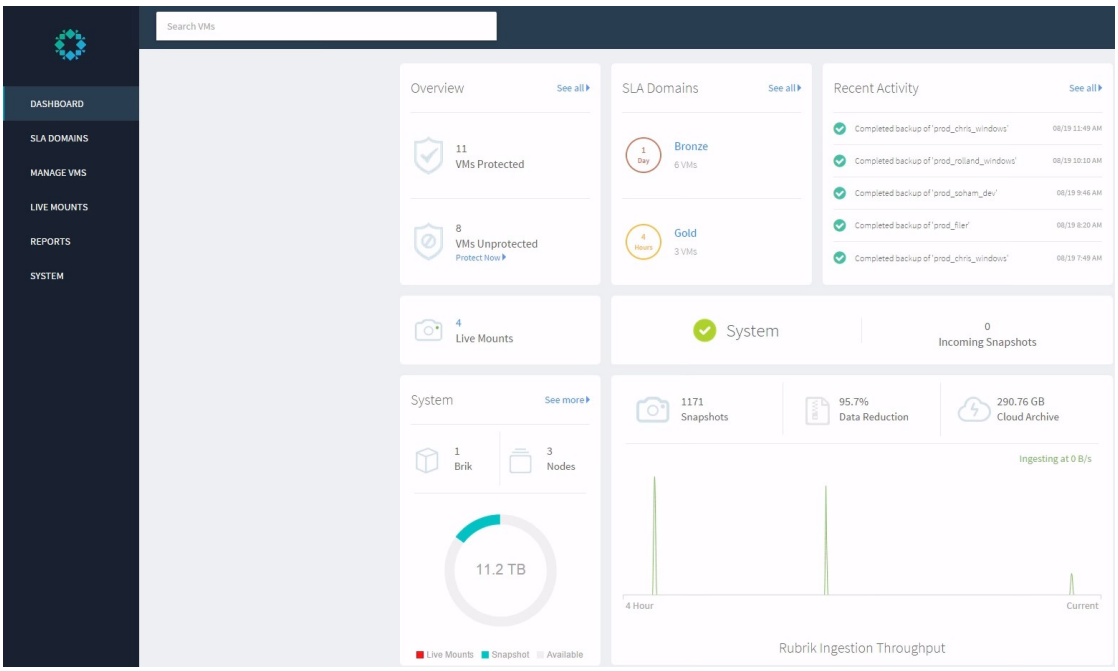
Commvaultin heikkouksiin kuuluu epämääräinen dokumentointi. Commvaultin eri tuotteita ei ole selkeästi erikseen dokumentoitu, vaan jokaisesta julkaisusta versiosta on oma, yhtenäinen dokumentaatio. Tämän seurauksena käyttäjän on vaikea löytää tietoa tuotekohtaisesti ja ymmärtää mihin yleinen dokumentaatio liittyy.

3.4 Rubrik

3.4.1 Yleistä

Rubrik on vuonna 2014 perustettu datanhallintaan- ja suojaukseen erikoistunut yksityinen yritys, jonka pääkonttori sijaitsee Kalifornian Piilaaksossa kaupungissa nimeltä Palo Alto. Rubrik on verrattaen uusi varmistusratkaisu, mutta sen toiminta on kasvanut nopeasti vuoden 2014 alusta 3,3 miljardin dollarin markkina-arvoon vuoteen 2019 mennessä. Suurista yrityksistä Cohesity on Rubrikin varteenotettavin kilpailija, sillä varmistusratkaisu on kummallakin hyvin samanlainen. (Chhabra, Flug, Nagel & O'Donnell 2019.)

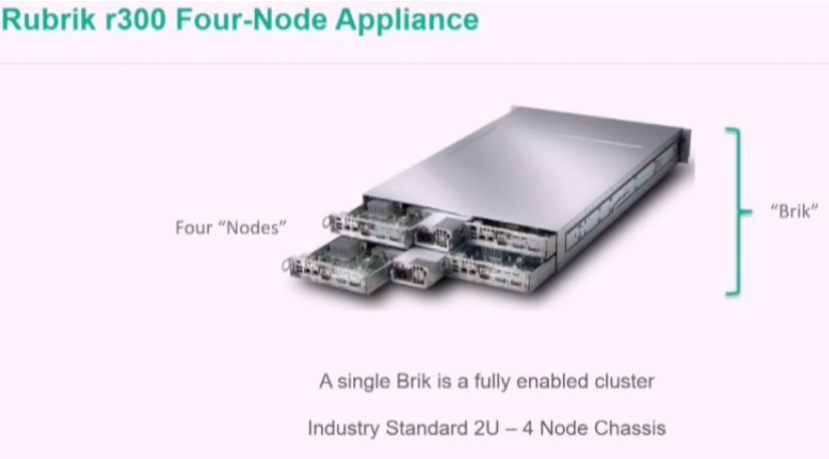
Forresterin kolmannen vuosineljänneksen raportissa kuvaillaan Rubrikia selkeäksi, yksinkertaiseksi ja moderniksi ratkaisuksi varmuuskopiointiin ja palautukseen (ks. kuvio 15). Rubrikin päätuote on Rubrik Cloud Data Management (CDM), joka on varmistussovelluksen keskitetty alusta. Se tukee erilaisia virtualisointialustoja, joita ovat VMware vSphere, Microsoft Hyper-V ja Nutanix AHV.



Kuvio 15. Rubrikin hallintapaneeli

3.4.2 Rubrik Cloud Data Management

Rubrikin käyttöönotto on suunniteltu vaivattomaksi. Asennus vaatii klusteriarkkitehtuuria noudattavan ”Brik”-laitteen, jossa on neljä palvelinta sisällä (ks. kuvio 16). Brik tuodaan verkkoon ja sille annetaan IP-osoite. Rubrikin hallintapaneeliin pääsee selaimella laitteen IP-osoitteella ja hallintapaneeliin kirjaudutaan esimerkiksi VMware ympäristön tunnuksilla. Kirjautumisen jälkeen Rubrik CDM hakee VMwaressa olevat koneet alustalle, joilloin käyttöönotto on valmis. Seuraavaksi käyttäjä voi konfiguroida varmistuksia eri palvelutasosopimuksien mukaisiksi.



Kuvio 16. Rubrik appliance (Rubrik 2019)

Rubrik CDM rakentuu seuraavista komponenteista (ks. kuvio 11):

- Atlas (Pilvipohjainen tiedostojärjestelmä)
- Callisto (Hajautettu metadatajärjestelmä)
- Cerebro (Datanhallintakerros)
- Infinity (Rajapinta ja sovellusyhteydet)

Atlas on hajautettu pilvipohjainen tiedostojärjestelmä datan käsittelyyn ja varastointin. Rubrikin kotisivuilla sen kerrotaan olevan vikasietoinen, itseohjautuva ja loputtomasti skaalautuva. Vikasietoisuus perustuu yksinkertaiseen poistamiskoodiin (erasure coding, EC), joka muuntaa ja osittaa dataa. Sen ansiosta noden tai levyn hajotessa alkuperäinen data voidaan palauttaa. Itseohjautuvuudella Rubrik tarkoittaa jatkuvuutta vikatilanteen sattuessa eli automaattista reagointia ja tilanteen korjaamista. Lisäksi Atlas optimoi dataa ja kaistanleveyttä parantaakseen suorituskykyä. (The Definitive Guide to Rubrik Cloud Data Management 2019.)

Arkkitehtuurin kokoonpanoa täydentää Callisto. Se on hajautettu metadatajärjestelmä, joka toimii hakukoneena Atlas tiedostojärjestelmälle. Callisto suorittaa nopeat tiedostotason haut ja huolehtii, että järjestelmä on jatkuvasti saatavilla. (Mts.)

Käyttäjä voi valita, varastoiko datan Rubrik-laitteelle vai julkipilveen. Cerebron tehtävänä on ohjata datavirtaa pilvivarastoihin ja konesaleihin. Sen toiminta perustuu kahteen pääkomponenttiin, joiden tehtävänä on varmistaa, että data on saatavilla ja ehjänä palautettavaksi. Toiselle näistä komponenteista ei löytynyt vielä suomen kieleen vakiintunutta selkeää käännöstä, mutta kyse on Rubrikin hajautetusta versionhallintajärjestelmästä nimeltään Blob Engine. Se käsittelee dataa palvelutasosopimuksen mukaisesti. Palvelutasosopimukseen määritetään, kuinka usein varmistus suoritetaan ja säilytysajat varmuuskopioille, replikoille sekä arkistoidulle datalle. Blob Enginen tehtäviin kuuluu datanhallinnan lisäksi datan suojaus ja deduplikointi. Cerebron toinen komponentti on hajautettu tehtäväjärjestelmä, joka ylläpitää korkeaa suorituskykyä. Se huolehtii tehtävien aikataulutuksesta ja ylläpidosta palvelutasosopimuksen mukaisesti, eli priorisoi tehtäviä ja jakaa kuormaa tasaisesti järjestelmässä. (Mts.)

Ohjelmointirajapintojen komennot kulkevat Rubrik CDM -järjestelmässä Infinityn kautta. Se toimii rajapintana Cerebron ja sovellusyhteyksien välillä. Kaikki ohjelmointirajapintoihin vaikuttavat toiminnot kulkevat Infinityn kautta aina palvelutasosopimuksen mukaisesti. (Mts.)

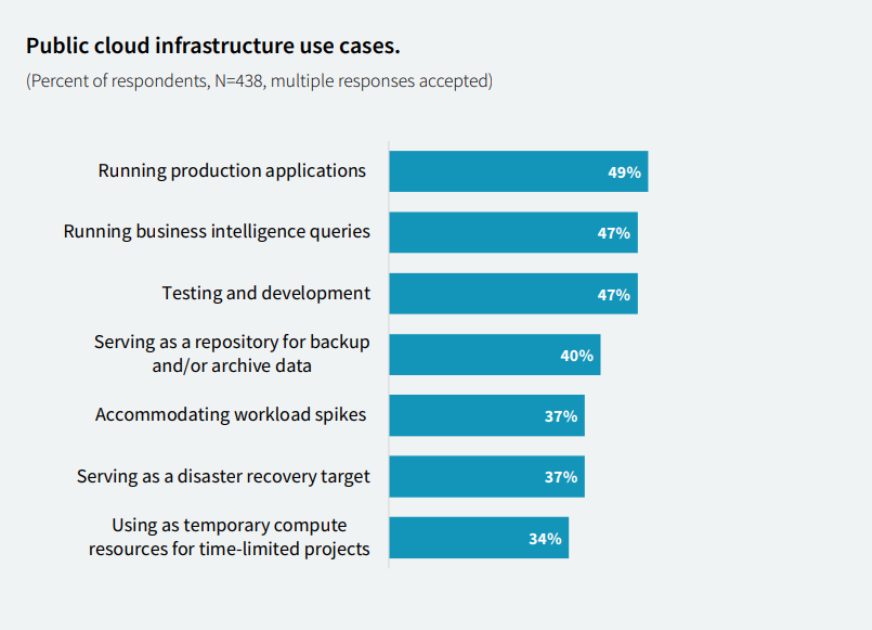
Rubrikin toiminnallisuuksista mainittavin asia on varmuuskopioidun datan hyödyntäminen. Perinteinen tyyli on ollut varmuuskopioida data ja jättää se passiivisena säilöön, kun taas Rubrik laittaa datan töihin, jolloin järjestelmä analysoi dataa mahdollisten haittaohjelmien löytämiseksi ja etsii poikkeamia varmistuksista. Tällöin varmistetaan myös nopeat palautukset, kun data on tarkastettu ehjäksi.

4 Julkiset pilvialustat

4.1 Yleistä

Nykyään on olemassa monia pilvitallennuspalveluja ja niistä sopivimman valitseminen yrityksen tarpeisiin ei ole aina helppoa, sillä jokainen projekti on omanlaatuinen.

Pilveen varmuuskopiointi on tunnettu suhteellisen edullisena verrattuna vaihtoehtoisiin varmistusratkaisuihin, mutta kustannusarvio kannattaa kuitenkin tehdä eri palvelumalleille. Hinnoittelu perustuu usein käytettyyn tallennuskapasiteettiin, datan siirtoon pilvipalvelussa tai pilveen säilöttävään kiinteään datamäärään sekä ylläpitoon, jolla varmistetaan palvelun suorituskyky ja datan eheys. Julkista pilvipalvelua suositaan sen saatavuuden ja helppouden takia sekä vaatimukset palvelun käyttöönottoon ovat matalat (ks. kuvio 17). Pilvipalvelun etuina ovat muun muassa, että omiin tietokonelaitteistoihin ei tarvitse investoida, eikä ylläpitää erillistä tietokantaa datasta.



Kuvio 17. Julkisen pilven käyttötapaukset (Lundell 2019)

Pilvipalveluita tarjoavien kolmansien osapuolien toteutukset ovat samankaltaisia, sillä niistä jokaisella on

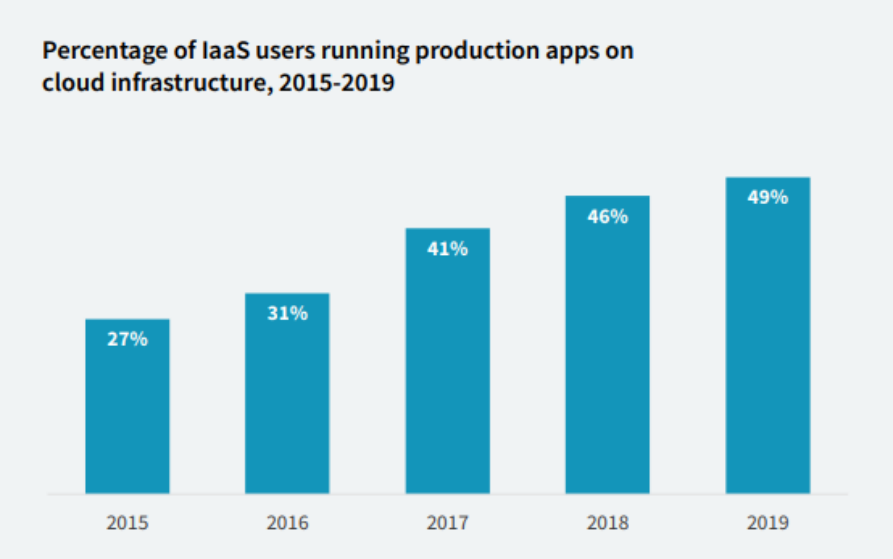
- Infrastruktuuri kattavalla laskentateholla, tietovarastoilla ja verkkoyhteyksillä ympäri maailmaa
- Joustavia IaaS-palveluita asiakkaan tarpeisiin
- Käytännössä rajoittamaton tallennuskapasiteetti tietovarastointijärjestelmissä

Julkisen pilvipalvelun infrastruktuuri on fyysisesti palveluntarjoajan tiloissa ja siitä myydään resursseja asiakkaan käyttöön palveluna. Eri pilvialustojen toiminnallisuuksissa on kuitenkin eroavaisuuksia. Seuraavaksi opinnäytetyössä tutustutaan Googlen, Amazonin, Microsoftin pilvialustoihin varmistusten näkökulmasta.

4.2 Palvelumallit

4.2.1 IaaS

IaaS (Infrastructure as a service) palvelumalli muodostaa pohjan IT-infrastruktuurille. Siihen sisältyy palvelimien, virtualisoinnin, laskentatehon, verkkoyhteyksien ja tietovarastojen tarjoaminen palveluna. Tuotantoympäristön siirtäminen IaaS palvelumallin pilvipalveluun on yleistynyt viime vuosina (ks. kuvio 18). IaaS-palvelun tarjoaja voi liittää lisäksi palveluun esimerkiksi komponenttien monitorointia ja tietoturvaa. (Eronen 2016.)



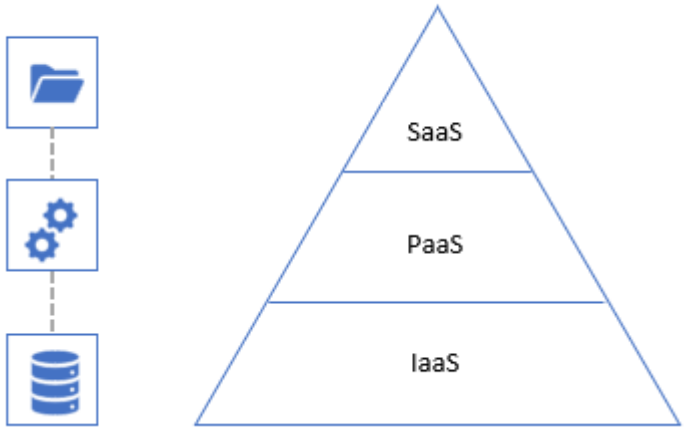
Kuvio 18. Prosenttiosuus IaaS käyttäjistä, joilla on tuotantosovelluksia pilvipalveluissa (Lundell 2019).

4.2.2 PaaS

PaaS (Platform as a service) mallissa esimerkiksi ohjelmistokehitykseen tarvittavat työkalut asennetaan valmiiksi virtuaalipalvelimelle, joka toimii palvelun tarjoajan ympäristössä. Tässä palvelumallissa asiakas tekee itse tarvittavat ohjelmistopäivitykset ja on vastuussa tietoturvasta. PaaS toimii alustana SaaS-palveluille (ks. kuvio 19). PaaS nopeuttaa projektien aloittamista yrityksissä, joissa resursseja ei haluta käyttää ympäristön pystyttämiseen ja tarvittavien sovelluksien asentamiseen. (Eronen 2016.)

4.2.3 SaaS

SaaS (Software as a service) eli ohjelmiston tarjoaminen palveluna. SaaS-ohjelmisto pyörii palvelun tarjoajan ympäristössä ja asiakas käyttää ohjelmistoa web-selaimella. Ohjelmistoa ja sen alustana toimivaa infrastruktuuria ylläpidetään niin, ettei asiakkaan tarvitse huolehtia muusta, kuin ohjelmiston käytöstä. (Mts.)



Kuvio 19. Pilvipalvelumallit

4.3 Amazon Web Services

4.3.1 Yleistä

Vuonna 2006 perustettu Amazon Web Services (AWS) tarjoaa erilaisia IaaS ja PaaS -pilvipalvelumalleja yksityiselle sekä julkiselle sektorille. Gartnerin teettämän toisen vuosineljänneksen raportin mukaan AWS on pysynyt viimeisen yhdeksän vuoden ajan markkinajohtajan asemassa pilvilaskennan jättinä ja jatkaa kasvuaan laajentamalla konesaliympäristöjään sekä palveluidensa saatavuutta globaalisti. (Mukhyala, Rao & Simpson 2019.)

AWS palveluissa on useita vaihtoehtoja datan tallentamiseen ja säilyttämiseen. AWS tietovarastopalveluita on kätevintä käyttää yhdessä muiden AWS palveluiden kanssa. Varmistuksien ohjaaminen niille sopivaan tietovarastoon voi kuitenkin aiheuttaa hämmennystä. Tässä luvussa käsitellään tietovarastopalveluja ja Amazonin omaa varmuuskopiointialustaa nimeltään AWS Backup.

4.3.2 Elastic Block Storage

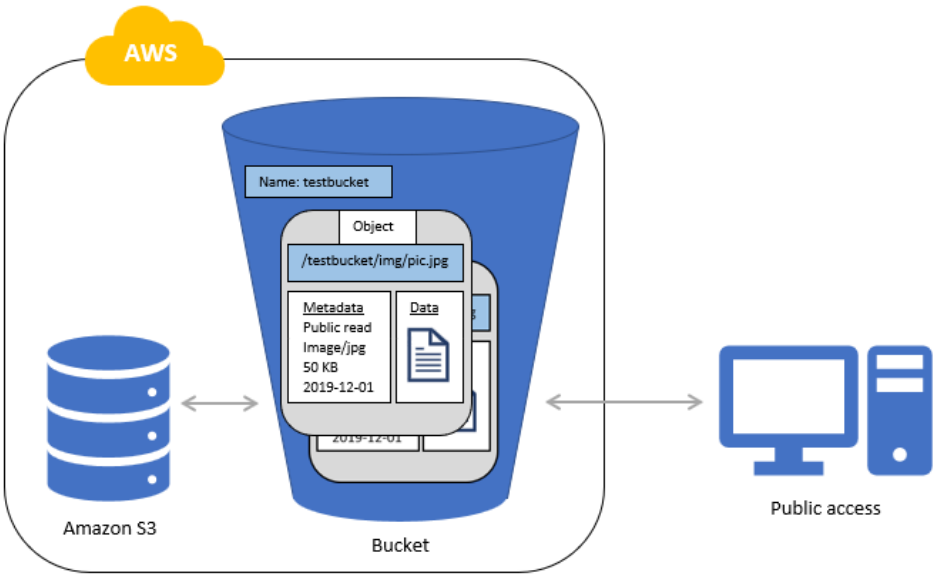
Amazon Elastic Block Storage (EBS) on blokkitason tietovarastopalvelu, jonka tehtävänä on toimia levyjärjestelmänä Amazon Elastic Compute Cloud (EC2) virtuaalikoneissa. EBS-levyt soveltuvat parhaiten datan pitkäaikaissäilytykseen, jossa data on kuitenkin nopeasti saatavilla. (AWS Documentation 2019.)

EBS-levyjä voidaan käyttää fyysisten kiintolevyjen tavoin ja ne voidaan varmuuskopioida snapshot -tekniikalla Amazon S3 palveluun. Ensimmäisen täyden varmistuksen jälkeen snapshotit varmistetaan aina inkrementaaleina, jotta EBS-levyjen kokoon ja käytettyyn tallennuskapasiteettiin perustuva hinnoittelu olisi S3-palvelussa kohtuullista. (Mts.)

4.3.3 Simple Storage Service

Amazon Simple Storage Service (S3) on web-pohjainen objektitason tietovarastopalvelu. Objekti S3-palvelussa on erillinen yksikkö, joka koostuu tiedostosta ja sen metadatasta. Objektit tallennetaan erilleen datasta, joten perinteisiä kansiorakenteita ei tarvita. Tämä nopeuttaa datan käsittelyä ja mahdollistaa REST API -sovellusrajapinnan käytön. Jokainen tiedosto saa ID-tunnisteen, joita sovellukset käyttävät hakiessaan objektien tietorakenteita. Lisäksi objekteille on olemassa S3-palvelussa versiointi ominaisuus datan häviämistä varten. Tällöin objekteista luodaan uusi versio aina, kun sen sisältö muuttuu, jolloin palautuminen edelliseen versioon helpottuu. Amazon ei kuitenkaan ilmoita versioiden tarkkaa määrää vaan, kertoo niitä olevan muutamia jokaisesta objektista versioinnin ollessa käytössä. (Alapati 2019.)

S3-palvelussa toinen tärkeä käsite on bucket. Bucket toimii konttitekniikan tavoin, siihen voidaan tallentaa objekteja (ks. kuvio 20), joille voidaan määrittää käyttöoikeuksia. Bucketin sisältämään dataan pääsee käsiksi vain ne, joilla on oikeat käyttöoikeudet objekteihin. Vaikka Amazon S3 käyttöliittymää hallitaan web-selaimella URL protokollan välityksellä, käyttöoikeuksien hallinnan ansiosta data ei ole julkisesti saatavilla ja bucketiin oikeutetut käyttäjät tarvitsevat AWS-tunnukset sen sisällön katseluun. (Mts.)



Kuvio 20. Amazon S3 rakenne (Alapati 2019, muokattu)

Amazon S3 tietovaraston vahvuuksia ovat

- Loputon skaalautuvuus
- Saatavuus, AWS konesaleja sijaitsee ympäri maailmaa
- Yhteensopivuus muiden AWS palveluiden kanssa
- 99,999999999% varmuus datan säilyvyydestä
- Kattava dokumentaatio

4.3.4 S3 Glacier

Amazon S3 Glacier on tietovarastopalvelu, joka sopii varmuuskopioiden varastointiin ja datan arkistointiin. Se on tarkoitettu harvoin tarvitun datan pitkäaikaissäilytykseen. S3 Glacieriin varastoitu data ei ole muiden S3 tietovarastoluokkien tavoin jatkuvasti saatavilla, vaan sen hakeminen palvelusta kestää muutamia tunteja. (Alapati 2019.)

4.3.5 Elastic File System

Elastic File System (EFS) on Amazonin hallinnoima tietovarasto EC2 instansseille. Se on palvelu, joka hoitaa verkkolevyjärjestelmän (network file system, NFS) roolin. EFS ansiosta yrityksen ei tarvitse ylläpitää omaa verkkolevyjärjestelmää. (Alapati 2019.)

4.3.6 Snowball

Amazon Snowball on fyysinen laite, joka on tarkoitettu todella suurien datamäärien siirtämiseen yrityksen omasta konesalista Amazonin pilveen ilman verkkoa. Se aktivoidaan AWS tilillä ja liitetään oman konesalin verkkoon, jonka jälkeen käyttöliittymän kautta voidaan määritellä toimenpiteitä datan siirrolle Snowballiin. Kryptaus on määritetty pakolliseksi asetukseksi (256-bit) ennen datan siirtoa laitteelle. Kun haluttu data on siirretty, täytetään laitteen mukana tullut rahtikirja ja toimitetaan Snowball postitettavaksi AWS tietovarastoon. AWS siirtää toimitetun datan Snowballista S3-palveluun. (Alapati 2019.)

4.3.7 Storage Gateway

AWS Storage Gateway on tietovarastopalvelu, joka toimii yhdistelmäpilven tavoin. Sen avulla voidaan esimerkiksi yhdistää yrityksen omaa laitteistoa tai sovelluksia Amazon S3 kanssa. Yhteensopivuuden johdosta yrityksen ympäristö saadaan yhdistettyä AWS pilvipalvelun tietovarastoon. (Alapati 2019.)

4.3.8 AWS Backup

Vuoden 2019 alussa julkaistu AWS Backup on varmuuskopiointipalvelu, joka on yhteensopiva edellä käsiteltyjen AWS tietovarastopalveluiden kanssa. AWS Backupin kautta voi hallinnoida tietovarastojen varmuuskopioita ja luoda varmistussuunnitelmia. (AWS Backup Documentation 2019.)

Jotta varmuuskopiointi toimisi järkevästi, täytyy sitä varten tehdä varmistussuunnitelma (ks. kuvio 21). Varmistussuunnitelma on AWS Backupissa joukko asetuksia (policy), joilla määritetään varmuuskopioinnin ajastus, milloin varmuuskopiointi on sallittua ja varmuuskopioiden säilytyspolitiikka. Varmistussuunnitelmia voi luoda useampia eri datamääriä ja ympäristöjä varten. (Mts.)

Kuvio 21. AWS Backup toimintaperiaate

AWS Backupissa varmuuskopioita voi lajitella eri säiliöihin (backup vault). Varmistussuunnitelmassa määritellään, mihin säiliöön varmistukset ohjataan. Säiliöön voidaan asettaa käyttöoikeuksia, joilla pystyy rajoittamaan säiliöön ja AWS Backup -sovellusrajapintaan pääsyä sekä estämään varmuuskopioiden poistamisen käyttäjältä. (Mts.)

Datan palautusta varten AWS Backup luo uuden palautustyön, joka saa yksilöllisen ID-tunnisteen. Varmistuksista valitaan haluttu palautuspiste ja määritetään palautusparametrit, jotka AWS Backup osaa luoda automaattisesti riippuen palautuksen kohteena olevasta AWS tietovarastosta. AWS-palveluiden välinen yhteensopivuus mahdollistaa datan palautuksen eri tietovarastopalveluihin. Palautus aloitetaan AWS käyttöliittymän kautta. Palautustyöstä voi seurata palautuksen tilaa ja onnistumista. (Mts.)

AWS Backup muistuttaa varmistussovelluksia toiminnoiltaan, mutta sen käyttäminen ei vaadi erillisen ohjelmiston asennusta, vaan hallinta tapahtuu täysin AWS konsolin kautta. Palvelu on saatavilla toistaiseksi kahdellatoista eri AWS alueella, mutta AWS-sivustolla kerrotaan saatavuuden laajenevan loppuvuoden aikana.

4.4 Microsoft Azure

4.4.1 Yleistä

Microsoft Azure tarjoaa kaikkia pilvipalvelumalleja. Azure on Microsoftin vuonna 2010 julkaisema pilvipalvelu, joka tunnettiin aiemmin nimellä Windows Azure. Se perustettiin alun perin vastaamaan Amazonin pilvialustan kilpailuun projektina nimeltään *Red Dog* (Harvey 2017). Azure tukee Microsoftin tuotteiden lisäksi kolmannen osapuolen ohjelmistoja ja järjestelmiä sekä erilaisia ohjelmointikieliä. Azure sopii joustavuutensa ansiosta pilvialustaksi sekaympäristöihin tai täysin pilvipohjaiseksi ympäristöksi. Se on saatavilla maailmanlaajuisesti 140 maassa ja Microsoft Azuren sivuilta löytyy kartta, joka havainnollistaa konesalien sijainteja eri alueilla (Microsoft Azure 2019). Tämän kappaleen alaluvuissa tutustutaan Microsoft Azuren eri tietovarastoihin ja Azure Backup varmistusratkaisuun.

4.4.2 Blob Storage

Azure Blob Storagen voisi ajatella olevan Microsoftin versio Amazon S3 -palvelusta. Se on myös objektitason loputtomasti skaalautuva tietovarasto jäsentämättömälle datalle. Kuten Amazon S3 -palvelussa, myös Azuressa on erilaisia varastointiluokkia. ”Kuuma” varastotaso on tarkoitettu aktiiviselle datalle, jota tarvitaan usein. ”Kylmä”

varastotaso sopii epäsäännöllisesti tarvittavan datan varastointiin, jota säilytetään vähintään 30 päivää. Kolmas varastotaso on arkistointia ja datan pitkäaikaissäilytystä varten. Arkistossa vähimmäisaika datan säilytykseen on 180 päivää ja se on näistä kolmesta varastotasoista halvin. Kuuma varastotaso muodostaa Azure Blob Storage -palvelussa eniten varastointikustannuksia, mutta vähiten datan käsittelykustannuksia. Kylmän varastotason puolestaan veloittaa enemmän datan käsittelystä, kuin varastoinnista. Arkisto on halvin datan säilytykseen, mutta datan palauttaminen arkistosta on hinnaltaan korkeinta. (Introduction to Azure Blob storage 2019.)

4.4.3 File Storage

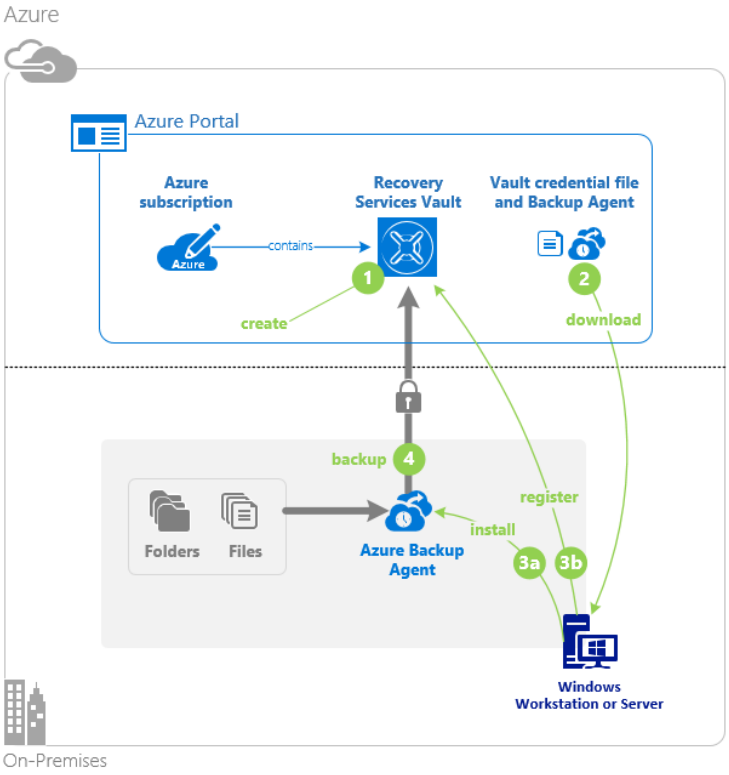
Azure Files on tiedostojakopalvelu, joka toimii SMB-protokollan (Server message block) avulla. Se voi toimia erillisen tiedostopalvelimen ja NAS-laitteen rinnalla tai vaihtoehtoisesti korvata molemmat keskitettynä tiedostojakona pilvessä. Näin ollen Azuren tiedostojakoa käytettäessä ei tarvitse tietokonelaitteistoa tai käyttöjärjestelmää. Tiedostojako on muiden pilvipalveluiden tavoin jatkuvasti saatavilla ja se voidaan liittää Azuren virtuaalikoneeseen tai on-premises ympäristöihin esimerkiksi paikallisen levyn rinnalle. Liittäminen vaatii Azuren tunnuksien lisäksi SAS-suojausavaimen (Shared access signature), joka luodaan Azuren hallintapaneelissa. Azure Files tukee Windows, Linux, tai macOS käyttöjärjestelmiä ja ohjelmointirajapintoja. (Microsoft Azure 2019.)

4.4.4 Disk Storage

Levyllä tarkoitetaan Azuressa virtuaalista kiintolevyä (VHD, virtual hard disk). Levyt jaotellaan kahteen pääluokkaan; hallittu ja hallitsematon. Hallitsematonta levyä varten täytyy luoda tallennustili ja määrittää muun muassa IOPS-rajoitus (input/output operations per second). Hallittua levyä käyttäessä ei tarvitse luoda erillistä tiliä, vaan uusi levy voidaan yksinkertaisesti luoda ja lisätä virtuaalikoneeseen loogisena yksikkönä. Tällöin Azure huolehtii taustalla olevista tehtävistä, kuten levyn salauksesta ja palautumisvalmiudesta. Koska Azure ylläpitää hallittua levyä, käyttäjällä ei tällöin ole täysiä käyttöoikeuksia levyyn. Seuraavassa alaluvussa käsiteltävä Azure Backup tukee hallittuja levyjä. (Microsoft Azure 2019.)

4.4.5 Azure Backup

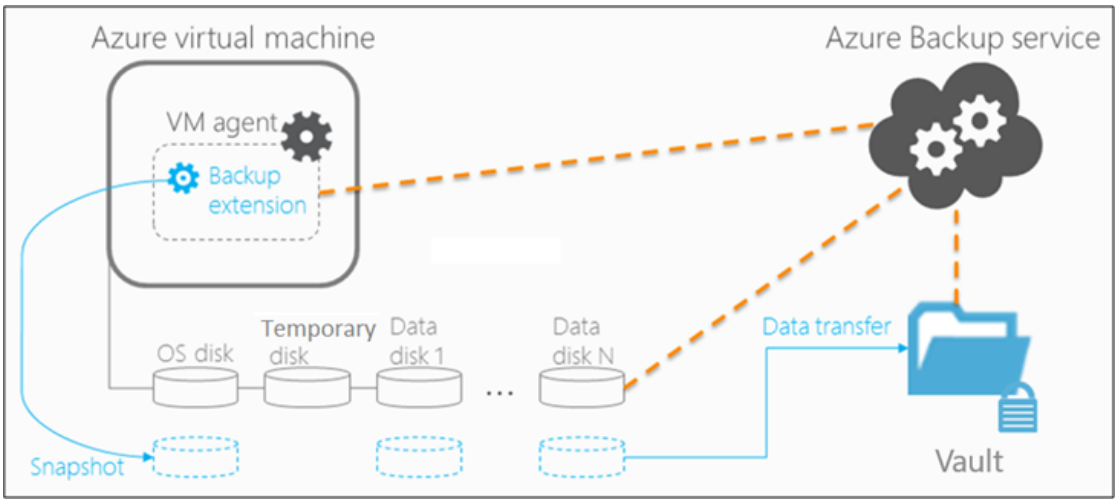
Azure Backup on varmuuskopiointipalvelu Microsoft Azuressa. Sillä voi varmistaa fyysisiä ja virtuaalisia palvelimia, työasemia sekä Azuren omia virtuaalikoneita. Se sopii parhaiten Windows-pohjaisten koneiden varmistamiseen (ks. kuvio 22) Microsoft-yhteensopivuutensa takia, mutta Linux-yhteensopivuutta ei kuitenkaan ole rajattu pois. Azure Backup tarjoaa muiden pilvivarmistuspalveluiden tavoin automatisoitua varmistusta ja datan suojausta sekä on-premises että pilvialustoille. (Microsoft Azure 2019.)



Kuvio 22. Windows koneen varmuuskopiointi Azure Backup agentilla (Microsoft Azure 2019)

Virtuaalikoneen varmuuskopiointi tapahtuu Azure Backup palvelussa siten, että varmistustyön alkaessa laajennusosa ottaa tilannevedoksen Windowsin VSS-ominaisuutta (Volume Shadow Copy Service) käyttäen ja siirtää sen säilöön, jonka jälkeen vain muuttunut data varmuuskopioidaan (ks. kuvio 23). On-premise varmistuksissa data suojataan AES256-salauksella ennen säilöön siirtoa HTTPS-yhteydellä, kun

taas Azure virtuaalikone salataan SSE-salauksella (Storage service encryption). Tilannevedos poistuu varmuuskopioinnin jälkeen säilöstä. Varmistamisen tiheys, eli kuinka usein voidaan suorittaa varmistus, on Azuressa rajoitettu. Windows kone voidaan varmistaa kolmesti- ja Azure virtuaalikone kerran päivässä. (Documentation for the Azure Backup service 2019.)



Kuvio 23. Azure virtuaalikoneen varmistusprosessi (Microsoft Azure 2019)

Onnistuneen varmuuskopioinnin jälkeen datasta luodaan palautuspiste. Azure Backupissa on kahdenlaisia palautuspistetyyppejä virtuaalikoneille, joista nopeampi tapa palauttaa on tilannevedos. Tilannevedos voidaan palauttaa ennen sen siirtämistä säilöön. Se on välitön palautustapa, eikä dataa tarvitse hakea tietovarastosta palautusta varten. Tilannevedos on tässä palautustyypissä inkrementaalinen ja tilannevedosta säilytetään Azure Backupissa enintään seitsemän päivää määritetyn varmistustyön mukaisesti. Näin ollen välitön palautus yhdestä tilannevedoksesta on mahdollista vain viikon ajan varmistuksen ottamisesta. Toisaalta vain yhden tilannevedoksen säilyttäminen välitöntä palautusta varten säästää tallennustilaa ja samalla kustannuksia. Toinen palautuspistetyyppi on tilannevedos säilössä. Kun tilannevedos on jo siirtynyt säilöön, voidaan se palauttaa Azuren hallintaportaalista. Palautus voidaan tehdä virtuaalikoneesta, levystä tai tiedostotasolla. (Mts.)

4.5 Google Cloud

4.5.1 Yleistä

Googlella on vahva tausta teknologiaosaamisessa ja se tunnetaan luotettavana tietoteknisten palveluiden ja työkalujen tarjoajana ympäri maailmaa. Koska kyseessä on massiivinen IT-alan yritys, joka panostaa pilvipalveluiden lisäksi eri teknologiasektoreihin omaa infrastruktuuriaan hyödyntäen, myös Google Cloud Platformin (GCP) kehitys on jatkuvaa (Geewax 2018). GCP rakentuu erilaisista pilvipalveluista, jonka pääkomponentit ovat vuonna 2008 kehitetty palvelualusta nimeltään Google App Engine ja vuonna 2012 markkinoille tullut virtualisointiin tarkoitettu Google Compute Engine (Rouse 2017).

4.5.2 Google Cloud Storage

Google Cloud Storage on tietovarastopalvelu pilveen säilöttävälle jäsentämättömälle datalle. Palvelua hallitaan selaimen välityksellä tai komentoriviltä. Dataobjektit varastoidaan säiliöihin (bucket), jotka voidaan jaotella datan käyttötarpeen mukaan (Google Cloud Storage 2019). Huomioitavaa on, että Googlen pilvialustassa on ollut viime vuosina eri hinnoittelumalleja datan lokerointiin, joten opinnäytetyön tekovaiheessa käsitellyt mallit voivat vanhentua lyhyelläkin aikavälillä muutoksien myötä. Jaottelu tapahtuu kolmeen eri varastotasoon (ks. kuvio 20), joita ovat ”Standard”, ”Nearline” ja ”Coldline”. Ensiksi mainittu on korkean käyttöasteen objektivarasto usein tarvittavalle datalle ja myös suosituin loppukäyttäjien keskuudessa. Nearline ja coldline sopivat harvemmin tarvitulle datalle ja arkistointiin. Jokainen varastotaso tarjoaa ominaisuuksia, joita ovat:

- Loputon tallennuskapasiteetti
- Saatavuus maailmanlaajuisesti
- 99,999999999% kestävyys datalle
- Varmuus datan säilyvyydestä
- Yhteensopivuus eri ohjelmointirajapintoihin

Kun varastotasoon luodaan säiliö, sille määritetään lisäksi maantieteellinen sijainti. Googlen dokumentaatiossa kerrotaan, ettei sijainnin valitsemista ole rajoitettu erityisemmin. Kannattavinta on kuitenkin valita sijainti, joka on lähimpänä datan käyttäjiä tehokkaan saatavuuden varmistamiseksi. Lisäksi järkevästi valittu sijainti optimoi suorituskykyä ja viivettä sekä minimoi kustannuksia. Sijaintityyppejä on kolmenlaisia. *Regional* on jokin yksittäinen maantieteellinen sijainti, kuten Suomi. Jos laskentatehosta vastaavat pilvialustan komponentit ja tietovarasto ovat samassa sijainnissa, tämä sijaintityyppi sopii käytettäväksi parhaiten. *Dual-region* puolestaan replikoi datan kahteen eri sijaintiin, jolloin saatavuus datalle on korkeampi. *Multi-region* on tarkoitettu maantieteellisesti hajautetuille data-alueille, kun tarve saatavuudelle on suurin. Kun säiliö tallennetaan useampaan sijaintiin, se on parhaiten turvassa esimerkiksi luonnonkatastrofeilta. (Google Cloud Storage 2019.)

Varmuuskopiointi Google Cloud Storage -palveluun perustuu pääasiassa yhteensopivuuteen kolmansien osapuolien palveluiden ja varmistussovelluksien kanssa. Varmuuskopiointia varten täytyy luoda ensiksi Google-tili, jolla kirjaudutaan pilvialustaan. Seuraavaksi GCP-palveluun luodaan vähintään yksi projekti, johon lisätään säiliö datalle. Säiliöitä lisätään tarpeen mukaan useampia ja ne toimivat varastoina dataobjekteille. Säiliö nimetään ja sille määritetään varastotaso ja sijainti, jonka piiriin säiliö kuuluu. Nimeämisessä on tiettyjä rajoituksia, joita ei työssä käydä läpi tarkemmin, mutta yksinkertaisuudessaan on käytettävä pieniä kirjaimia ja yhdysmerkkejä sekä jokainen säiliö täytyy nimetä yksilöllisesti.

Viimeinen vaihe on muodostaa yhteys pilvivarastoon. Jotta yhteys Googlen pilvivarastoon toimii, täytyy luoda ID eri palveluita ja ohjelmointirajapintoja varten. ID-tunnistetta käytetään säiliöiden tunnistamiseen Google-tilillä. Luotu tunniste syötetään esimerkiksi varmistussovellukseen, jonka halutaan kopioivan Google Cloud Storage -palveluun. (Google Cloud Storage 2019).

Kun datalle ilmenee palautustarve, voidaan palautukseen käyttää joko varmistussovelluksen käyttöliittymää tai Googlen pilvialustaa. Objekteja varten on olemassa versiointi ominaisuus. Sen ollessa käytössä objekteista luodaan ajastetusti uusia versioita, joita voidaan palauttaa tarvittaessa. Jos palautus tehdään varmistussovelluksen

kautta, silloin käytetään ensisijaisesti varmistussovellukseen määriteltyjä palautuspisteitä. (Mts.)

4.5.3 Google Cloud Filestore

Jokaisen käyttöjärjestelmän perustana on tiedostojärjestelmä. Tiedostojärjestelmä tekee tiedostojen käsittelystä helpompaa ja sen avulla dataa voidaan varastoida massamuistiin tai pilvipalveluun. Google Cloud Filestore on palvelu sovelluksille, jotka tarvitsevat datalle tiedostojärjestelmän. Se tarjoaa hallinnoidun pilvipohjaisen tiedostojärjestelmän käyttöliittymällä ja verkkojaolla, jolloin yritys ei tarvitse omaa erillistä NAS-laitetta. Google Cloud Filestore sopii esimerkiksi nopeasti muuttuviin ympäristöihin sen skaalautuvuutensa ansiosta. (Google Cloud Platform 2019.)

5 Vertailu

5.1 Yleistä

Tässä kappaleessa käydään läpi yleisesti tuloksia työssä tehdyn tutkimuksen pohjalta ja vertaillaan eri varmistussovelluksia ja pilvialustoja. Neljää eri työssä käsiteltyä varmistussovellusta vertaillaan oleellisimpien varmistus- ja palautuskyvykkyyksien kannalta taulukoihin kootuilla ominaisuuksilla. Varmistussovellukset esitetään taulukoissa yksinkertaisesti numeroituina tuotteina, joiden selitteet ovat listattu kappaleen loppuun (ks. taulukko 3). Tämä toteutustapa vertailulle tuottaa taulukoita, joita voi tarkastella esimerkiksi uutta varmistusratkaisua mietittäessä ilman painotusta tiettyyn tuotteeseen. Taulukoihin on pyritty keräämään tyypillisimpiä nykyhetken varmistusratkaisun ominaisuuksia, joita on merkattu eri tuotteille kyvykkyyksien mukaan.

Pilvialustoja analysoidaan erikseen omassa kappaleessaan. Ensin vertaillaan tärkeimpien tietovarastojen hinnoittelua ja vertailu eri tuotteista yritetään pitää puolueettomana. Myöhemmin vertailu rajataan yleiselle tasolle pilvipalveluiden laajuuden

vuoksi. Vertailun tarkoituksena ei ole nimetä parhainta tuotetta, vaan kerätä oleellisia huomioita eri yritysten tarjoamista pilvipalveluista varmistuksien tueksi.

5.2 Varmistussovellukset

Jokainen käsitelty tuote osaa varmistaa virtuaalikoneen, joten kyseistä ominaisuutta ei varmuuskopiointitaulukkoon (ks. taulukko 1) kirjattu. Taulukosta nähdään, että suurin osa tuotteista kykenee pääpiirteissään listattujen ominaisuuksien vaatimiin toimintoihin. Varmuuskopiointi scale-out repositorioon tuottaa yhdelle tuotteista vaikeuksia ja tuki nauhavarmistukselle ei ole natiivia kahdelle neljästä tuotteesta, joten ominaisuus on tällöin merkattu epäkäytännölliseksi.

Taulukko 1. Varmuuskopiointikyvykkyydet

Ominaisuus	Tuote 1	Tuote 2	Tuote 3	Tuote 4
Tiedostotason varmuuskopiointi	K	K	K	K
Varmuuskopiointi scale-out repositorioon	K	EK	K	K
Fyysisen koneen varmistus (Windows, Linux)	K	K	K	K
Tuki nauhavarmistukselle	EK	K	EK	K
Offsite-varmistus	K	K	K	K
Microsoft Office 365 datan varmistus (Exchange, SharePoint, OneDrive)	R	K	R	K
Arkistointi julkipilveen (AWS, Azure, GCP)	K	R	K	K
Vieraskäyttöjärjestelmän tiedostojen haku varmistuksen jälkeen	K	K	K	K
Varmistuksien suojaus salauksella	R	K	R	K
Jatkuva varmuuskopiointi	K	K	K	K
Varmistuksien deduplikointi ja pakkaaminen	K	K	K	K

Palautuskyvykkyyksistä (ks. taulukko 2) oleellisin on kokonaisen fyysisen tai virtuaalisen koneen palautus, jotta tiedot saadaan palautettua kokonaisvaltaisesti. Koska tahattomia käyttäjävirheitä tapahtuu usein, on tiedostotason palautus palautustyypeistä yleisin ja taulukon ensimmäiseltä ominaisuusriviltä nähdään, että jokainen tuote siihen myös kykenee. Puolet tuotteista pystyvät palautukseen varastotason tilannekuvista. Tämä ominaisuus on tärkeä etenkin tuotannossa olevien koneiden palautuksessa ja katastrofista palautumisessa, jotta palautus saadaan nopeasti vietyä läpi ja liiketoiminta jatkumaan normaalisti.

Taulukko 2. Palautuskyvykkyydet

Ominaisuus	Tuote 1	Tuote 2	Tuote 3	Tuote 4
Tiedostotason palautus	K	K	K	K
Virtuaalikoneen välitön palautus	K	K	K	K
Virtuaalilevyn palautus	K	K	E	K
Fyysisen koneen palautus	R	K	K	K
Palautuminen varastotason tilannekuvista	E	K	E	K
Microsoft Office 365 datan palautus (Exchange, SharePoint, OneDrive)	R	K	R	K
Palautus julkipilveen (AWS, Azure, GCP)	R	K	K	R
Skannaus tietoturvauhkien varalta ennen palautusta	R	E	R	K
Vieraskäyttöjärjestelmän tiedostojen palautus	K	K	K	K
Sovellustason palautus	EK	EK	EK	K

Helpointa olisi asettaa varmistussovellukset paremmuusjärjestykseen ominaisuuksien kattavuuden mukaan. Pelkästään ominaisuuksien määrä mittarina ei kuitenkaan

olisi järkevää, sillä erilaisiin käyttötapauksiin ei kaikkia ominaisuuksia tarvita. Jos yritykselle riittää esimerkiksi sen virtuaalikoneiden varmistus ja tiedostotason palautus niin lisäominaisuudet olisivat siinä tapauksessa turhia. Vaikka taulukoiden perusteella yksittäinen tuote vaikuttaisi ominaisuuksien perusteella paremmalta, tuotteen valinnassa kyse on myös makuasioista esimerkiksi käyttöliittymän, hinnoittelun ja tukipalveluiden saatavuuden suhteen.

Taulukko 3. Selitteet

Otsikko	Selite
Tuote 1	Cohesity
Tuote 2	Commvault
Tuote 3	Rubrik
Tuote 4	Veeam
K	Kyllä
R	Rajoitettu kyvykkyys
EK	Epäkäytännöllinen kyvykkyys
E	Ei

5.3 Pilvialustat

5.3.1 Pilvivarastojen hinnoittelu

Pilvivarastoille on kehittynyt yleinen luokittelu lämpötilan mukaan. Data voidaan ohjata kuumaan, viileään tai kylmään varastotasoon ja hinnoitella käytön perusteella (ks. taulukko 4). Kuuma varastotaso on usein käytettävälle ja muuttuvalle datalle. Näin ollen kuuma varastotaso sopii myös tuoreimpien varmuuskopioiden varastoksi. Viileää varastotasoa käytetään harvemmin tarvitulle, mutta yhä kriittiselle datalle, kuten DR-tiedostoille. Kylmä varastotaso on arkistointia ja passiivisen datan säilytystä varten. Data on varastoituna kylmään varastotasoon säilytyskäytänteen mukaisesti tavallisesti vähintään 90 päivää ja datan hakeminen ennen säilytysajan loppumista

muodostaa lisäkustannuksia. Käytännössä järkevintä olisi siirtää varmistukset asteittain kuumasta kylmimpään varastotasoon säilytyskäytänteen mukaisesti. (Cloudberrylab 2019.)

Taulukko 4. Varastotasojen hinnoittelu yhtä gigatavua kohden kuukaudessa

Palvelu	Amazon Web Services	Microsoft Azure	Google Cloud Platform
Tietovarasto	Amazon S3	Azure Blob Storage	Google Cloud Storage
Kuuma	0,021 €/kk	0,018 €/kk	0,023 €/kk
Viileä	0,0090-0,011 €/kk	0,0090 €/kk	0,0090 €/kk
Kylmä	0,0036-0,00089 €/kk	0,00089 €/kk	0,0036-0,0063 €/kk

Pilvipalveluja voidaan hinnoitella käytön mukaan jopa sekuntien tarkkuudella. Googlella on jo saatavilla kyseistä hinnoittelumallia pilvipalveluilleen, jossa käyttäjä säästää enemmän verrattuna minuuttipohjaiseen hinnoitteluun (Detailed Cloud Comparison 2019). Pilvipalvelutarjoajien sivustoilla on saatavilla laskureita, joilla pystyy laskemaan hinta-arvioita eri palveluille. Tällöin palvelua harkitseva voi itsenäisesti vertailla muun muassa instanssien hintoja tarvitsemilleen virtuaalikoneilleen ennen palveluntarjoajan valitsemista. AWS ja Azure tarjoavat käyttöön perustuvaa laskutusta minuutin tarkkuudella, mutta datan varastointipalveluiden hintoja tarkasteltaessa kuukausitasolla (ks. taulukko 5) voidaan huomata, ettei samantyyppisten palveluiden hinnoissa kovin suuria eroja kuitenkaan ole. Perustason tietovarasto hinnoitellaan yksinkertaisuudessaan kapasiteetin ja kirjoitus- ja lukuoperaatioiden määrän perusteella. Käyttäjä maksaa hakuoperaatioista, datan lataamisesta ja siirtämisestä sekä joissain tapauksissa jokaisesta autentikoinnista palvelua käyttäessään (Cloudberrylab 2019).

Taulukko 5. Tietovarastojen hinnoittelu

Palvelu	Amazon Web Services	Microsoft Azure	Google Cloud Platform
Tietovarasto	Amazon S3 Standard Storage	Azure Blob Standard Storage	Google Cloud Standard Storage
Kapasiteetti	10 TB	10 TB	10 TB
Kirjoitus- ja lukuoperaatiot	10 000	10 000	10 000
Arvioitu hinta	212 €/kk	191 €/kk	184 €/kk

Palveluna tarjottavien tiedostojärjestelmien hinnat ovat lähes samat teratavun datamäärälle Amazonin ja Googlen ratkaisuissa, kun kirjoitus- ja lukuoperaatioita ei ole rajoitettu (ks. taulukko 6). Azure Files on kolmesta tiedostojärjestelmästä nykyhetkellä halvin. Käyttäjän kannalta oleellista on, että jokaista tiedostojärjestelmää varten on saatavilla asiakastukipalvelua, joka sisältyy hintaan.

Taulukko 6. Tiedostojärjestelmien hinnoittelu

Palvelu	Amazon Web Services	Microsoft Azure	Google Cloud Platform
Tiedostojärjestelmä	Amazon EFS	Azure Files	Google Cloud Filestore
Kapasiteetti	1 TB	1 TB	1 TB
Arvioitu hinta	274 €/kk	215 €/kk	275 €/kk

5.3.2 Pilvipalvelun tarjoajan valinta

Usein ajatellaan, että pilvi on jokin aineeton teknologian ihme. Taustalla on kuitenkin tavallista tietokonelaitteistoa ja palvelimia, kuten luvussa 4.1 kerrottiin. Suosituimmista palveluntarjoajista sopivimman valitseminen on hankalaa, sillä saatavilla olevat palvelut ovat hyvin samanlaisia. Pisimpään markkinoilla ollut AWS tarjoaa eniten eri-

laisia palveluita, mutta Azure ja Google Cloud kehittyvät jatkuvasti vastatakseen kilpailuun. Kaikille alustoille yhteistä on loputon tallennuskapasiteetti ja saatavuus maailmanlaajuisesti.

Kannattavaa olisi ensiksi kartoittaa yrityksen tarve pilvipalvelun käytölle ja mitä haasteita sillä voidaan ratkaista. Seuraavaksi on huomioitava pilvipalvelua tarvitsevan yrityksen data- ja käyttäjämäärät sekä pilvipalvelun käyttötarkoitus. Mikäli säilöttävää dataa tai käyttäjiä on vähän, ei erillisen pilvi-infrastruktuurin pystyttäminen ole välttämättä tehokkain ratkaisu. Oleellista on myös huomioida, että pilvipalvelussa olevalla datalla ei ole takuuta. Lähtökohtaisesti asiakas on täysin vastuussa pilveen säilötystä datasta ja datan sijainnista riippuen olisi hyvä huomioida myös tietosuojalakien rajoitteet. Toimittajan vastuu on varmistaa palvelun saatavuus palvelutasosopimuksen mukaisesti, johon on syytä tutustua huolella ennen palvelun valintaa.

Palvelun valinnan jälkeen suositellaan huomioimaan muutamia seikkoja. Lopuksi kannattaa varmistaa palvelun toimivuus ja yhteensopivuus määritettyihin järjestelmiin sekä hinnoittelumalli. Lisäksi tietoturvakäytänteistä sopiminen ja mahdolliset lisäkustannukset olisi hyvä ottaa huomioon. Jotta datanhallinta pilvipalvelussa olisi mutkatonta, täytyy yrityksen IT-osaston osata palvelun käyttäminen ja hallinta. Mikäli varmuuskopiointi on epäonnistunut pilvipalvelussa eikä dataa voida palauttaa, on yritys itse siitä vastuussa.

6 Pohdinta

Työn tavoitteena oli löytää vaihtoehtoisia varmistusratkaisuja, joita toimeksiantaja voisi hyödyntää yrityksen palvelutuotannossa toimivien varmistusratkaisujen rinnalla. Tätä varten pinnalla olevista yrityksille suunnatuista varmistusratkaisuista valittiin käsittelyyn ne, jotka olivat suosituimpia maailmanlaajuisesti ja toimitusmalliltaan samankaltaisia.

Työn alkuvaiheessa onnistuttiin luomaan lukijalle yleiskatsaus siitä, mitä kaikkea varmuuskopiointiin liittyy, jonka pohjalta varmistusratkaisujen käsittely työssä oli

luontevampaa. Varmistussovelluksien kohdalla ominaisuuksien vertailu toteutui yleisimpien kyvykkyyksien kannalta, jotka valittiin tekijän oman näkemyksen pohjalta. Pilvialustat ovat oma lukunsa, mutta niiden hyödyntäminen varmistuksien suunnittelussa vaikuttaisi olevan kannattavaa varsinkin silloin, jos dataa on jo pilvessä. Huomion arvoista on, että yhteen pilvipalveluntarjoajaan ei tarvitse välttämättä sitoutua. Esimerkiksi virtuaalikoneiden siirtoihin eri pilviympäristöjen välillä on olemassa palveluita, joiden avulla koneen formaatti saadaan muutettua ympäristöön sopivaksi (Wallenius 2018). Työssä läpikäytyjen varmistusratkaisujen ominaisuuksien tarkastelu laajemmin vaatisi enemmän tutkimustyötä tai aiheen rajauksen johonkin tiettyyn ratkaisuun.

Työn tekovaiheessa aineiston kerääminen varmistusratkaisujen ja palveluiden ominaisuuksista ja toiminnallisuuksista vei ajallisesti eniten resursseja. Tietoa kerättiin usean eri lähteen avulla, ja monessa tapauksessa se oli täysin uutta tai ollut saatavilla vasta vähän aikaa. Tutkimuslaitos Gartnerin mukaan vuoteen 2022 mennessä 40% yrityksistä korvaavat vuonna 2018 hankkimansa varmistusratkaisunsa (Mukhyala, Rao & Simpson 2019). Tämän oletuksen perusteella varmistusratkaisujen tulisi kehittyä jatkuvasti ja vastata yrityksen tarpeisiin mahdollisimman hyvin. Vertailua muihin varmistusratkaisuihin tulisi säännöllisesti tehdä, sillä saatavilla on useita vaihtoehtoja jatkossakin. Tiedon hankinnan lisäksi haasteena oli se, että lähteiden luotettavuus täytyi joissain tapauksissa varmistaa, sillä esimerkiksi uusimmissa tutkimuslaitoksien tuottamissa raporteissa saattoi olla virheitä, joita ei oltu vielä oikaistu.

Jatkokehityksen kannalta seuraava vaihe olisi varmistusratkaisujen testaaminen. Ennen sopivan varmistusratkaisun valintaa tulisi suorittaa testausta yrityksen infrastruktuuria vastaavassa testiympäristössä. Ensiksi pitäisi selvittää testattavan tuotteen mahdolliset haasteet asennuksessa ja käyttöönotossa. Testiympäristön varmistamista voisi testata eri tuotteilla ja kerätä tuloksia esimerkiksi varmuuskopiointi- ja palautusnopeuksista sekä ympäristön kuormituksesta toimenpiteiden aikana. Varmistusympäristöjen uudistaminen ei ole aina suoraviivainen prosessi ja olisikin mielenkiintoista tietää, suoriutuuko jokin varmistusratkaisu paremmin esimerkiksi vanhentuneen ympäristön varmistamisesta

tai sen siirtämisestä uudelle alustalle. Myöhemmin varmistusratkaisujen toimintaa tulisi testata yhdessä muiden ympäristön komponenttien kanssa, jotta testaaminen saataisiin tuotantoympäristöä vastaavalle tasolle.

Onko varmuuskopiointi siirtymässä kokonaan pilvipohjaisille alustoille ja tarvitaanko nauhavarmistuksia vielä jatkossa? Perinteisen varmistusratkaisun korvaaminen uudella voi olla kannattavaa ja jossain vaiheessa jopa välttämätöntä. Usein varmistussovelluksen asennusta varten tarvitaan yhteensopiva käyttöjärjestelmä. Viimeistään silloin, kun varmistettavassa palvelinympäristössä olevasta käyttöjärjestelmästä loppuu tuki ja päivitykset, tulisi uusia vaihtoehtoja harkita. Ennen kaikkea varmistusratkaisun tulisi vastata yrityksen tarpeisiin. Liiketoiminnallisesti määritettyjä säilytyskäytänteitä datalle tulisi noudattaa huolimatta siitä, minkälaista ympäristöä varmistetaan.

Ihanteellinen tilanne varmistuksien näkökulmasta olisi aika-ikkunasta riippumaton varmuuskopiointi, johon jatkuva tietojen suojaaminen (continuous data protection, CDP) voisi olla varteenotettava vaihtoehto. Tällöin varmuuskopio luodaan aina, kun varmistettava data muuttuu. Tekniikka vaatii suorituskykyä järjestelmältä, mutta parhaimmillaan palautuminen olisi mahdollista tarkasti haluttuun ajankohtaan.

Toimiva varmistusratkaisu tuo turvaa ja mielenrauhaa yritykselle. Tietojen varmistaminen ei kuitenkaan ole aina itsestäänselvyys ja varmistuksien tarve huomataan viimeistään silloin, kun dataa tarvitsisi palauttaa. Varmistusratkaisu tulisi nähdä strategisena osana liiketoiminnan jatkuvuutta ja valintaan olisi hyvä käyttää aikaa, sillä jokaiseen tarpeeseen löytyy varmasti sopiva ratkaisu.

Lähteet

Airinen, P. 2011. Pilvilaskenta ja pilvipalvelut: pilvialustojen vertailu. Opinnäytetyö, AMK. Turun ammattikorkeakoulu, tietojärjestelmät, tietojenkäsittelyn koulutusohjelma. Viitattu 12.10.2019. http://urn.fi/URN:NBN:fi:amk-201103213428

Alapati S. R. 2019. AWS Certified SysOps Administrator Associate All-in-One-Exam Guide (Exam SOA-C01). Ohio: McGraw-Hill Education.

Aula, J. 2011. Varmista liiketoimintasi jatkuvuus. Verkkoartikkeli Tivi-sivustolla. Viitattu 25.10.2019. https://www.tivi.fi/uutiset/varmista-liiketoimintasi-jatkuvuus/a6daf903-ddbd-324b-b8ac-3b0ef0004fe3

Amazon S3, MS Azure and Google Cloud Storage Pricing Comparison. 2019. Verkkoartikkeli Cloudberry-sivustolla. Viitattu 10.12.2019. https://www.cloudberrylab.com/resources/blog/amazon-s3-azure-and-google-cloud-prices-compare/

AWS Documentation. 2019. Tekniset dokumentaatiot eri AWS-palveluista. Viitattu 2.11.2019. https://docs.aws.amazon.com/

AWS Backup Documentation. 2019. Tekninen dokumentaatio AWS-varmistuspalvelusta. Viitattu 28.10.2019. https://docs.aws.amazon.com/aws-backup/latest/devguide/whatisbackup.html

AWS vs Azure vs Google – Detailed Cloud Comparison. 2019. Verkkoartikkeli Intellipaat-sivustolla. Viitattu 10.12.2019. https://intellipaat.com/blog/aws-vs-azure-vs-google-cloud/

Backup types. 2018. Artikkeli Backup4all-sivustolla. Viitattu 1.10.2019. https://www.backup4all.com/backup-types-kb.html

Chhabra, N., Flug, M., Nagel, B. & O'Donnell, G. 2019. The Forrester Wave™: Data Resiliency Solutions, Q3 2019. Forresterin raportti varmistusratkaisuista. Viitattu 25.10.2019. https://reprints.forrester.com/?kui=jeCCEMCMwWgsGJ_O1gIHIA#/assets/2/917/RES146735/reports

Cohesity. 2019. Yrityksen kotisivut. Viitattu 10.12.2019. https://www.cohesity.com/

Commvault Annual Report. 2019. Commvaultin vuosittainen raportti sijoittajille. Viitattu 24.11.2019. https://commvault.gcs-web.com/investor-overview

Commvault Documentation. 2019. Dokumentaatio Commvault-sivustolla. Viitattu 15.11.2019. https://documentation.commvault.com/commvault/v11/article?p=3759.htm

Documentation for the Azure Backup service. 2019. Tekninen dokumentaatio Azure Backup -palvelusta Microsoft-sivustolla. Viitattu 9.11.2019. https://docs.microsoft.com/en-us/azure/backup/

Eronen, H. 2016. IaaS, PaaS, SaaS? Mikä pilvipalvelu sopii yrityksellesi. Verkkoartikkeli. Viitattu 12.10.2019. https://www.planeetta.fi/2016/03/15/iaas-paas-saas-mika-pilvipalvelu-sopii-yrityksellesi/

ExaGrid Product Line. 2019. ExaGrid -tuotesivusto. Viitattu 22.12.2019 https://exagrid.com/exagrid-products/product-line/

Geewax. J. 2018. Google Cloud Platform in Action. New York: Manning Publications.

Google Cloud Platform documentation. 2019. Tekninen dokumentaatio Googlen pilvialustasta. Viitattu 7.11.2019. https://cloud.google.com/docs/

Google Cloud Storage. 2019. Tekninen dokumentaatio Googlen pilvivarastosta. Viitattu 12.11.2019. https://cloud.google.com/storage/docs/

Handy, J. 2017. HDD vs. SSD: Is there room for disk in a solid-state world? Verkkoartikkeli. Viitattu 3.10.2019. https://searchstorage.techtarget.com/feature/HDD-vs-SSD-Is-there-room-for-disk-in-a-solid-state-world

Harvey, C. 2017. Microsoft Azure. Verkkoartikkeli Datamation-sivustolla. Viitattu 20.11.2019. https://www.datamation.com/cloud-computing/microsoft-azure.html

Hewlett Packard Enterprise, IBM and Quantum. 2019. What is LTO technology? LTO Ultrium verkkosivut. Viitattu 22.12.2019. https://www.lto.org/technology/what-is-lto-technology/

Introduction to Azure Blob storage. 2019. Tekninen dokumentaatio Azure Blob -tietovarastosta Microsoft-sivustolla. Viitattu 8.11.2019. https://docs.microsoft.com/en-us/azure/storage/blobs/storage-blobs-introduction

Koskinen, M. 2017. Varmuuskopiointipalvelun suunnittelu ja toteutus. Opinnäytetyö, AMK. Kaakkois-Suomen ammattikorkeakoulu, tekniikan ala, tieto- ja viestintätekniikan koulutusohjelma. Viitattu 2.10.2019. http://urn.fi/URN:NBN:fi:amk-2017121220804

Lundell, B. 2019. 2019 Public Cloud Computing Trends. Tutkimusraportti julkipilven suunnasta Enterprise Strategy Group-sivustolla. Viitattu 6.12.2019. https://www.esg-global.com/hubfs/pdf/ESG-Research-Report-2019-Public-Cloud-Trends-Apr-2019.pdf?hsCtaTracking=cb4a5b98-896d-42b4-8da2-4cc325170ba3%7C2e2d925a-9721-4a66-8dbd-9bb8dd6bdd30

Lyon, D. F. 2018. What is snapshot and how is it different than backup? Verkkoartikkeli. Viitattu 2.10.2019. https://blog.qnap.com/snapshot-different-backup/

Mayer, A. 2017. The 3-2-1 Backup Rule – An Efficient Data Protection Strategy. Verkkoartikkeli. Viitattu 5.10.2019. https://www.nakivo.com/blog/3-2-1-backup-rule-efficient-data-protection-strategy/

Microsoft Azure. 2019. Dokumentaatio Azure pilvialustasta ja sen palveluista. Viitattu 17.11.2019. https://azure.microsoft.com/

Microsoft Office 365 is Being Adopted and Used at an Enormous Rate. 2019. Verkkoartikkeli. Viitattu 3.12.2019. https://blog.goptg.com/microsoft-office-365-statistics

Mukhyala, C. & Rao, S. 2019. Critical Capabilities for Data Center Backup and Recovery Solutions. Raportti varmistusratkaisuista Gartner-sivustolla. Viitattu 14.11.2019. https://www.gartner.com/doc/reprints?id=1-6PV91PF&ct=190521&st=sb

Mukhyala, C., Rao, S. & Simpson, N. 2019. Magic Quadrant for Data Center Backup and Recovery Solutions. Raportti varmistusratkaisuista Gartner-sivustolla. Viitattu 7.10.2019. https://www.gartner.com/doc/reprints?id=1-KNKFYAC&ct=190620&st=sb

Määttä, T. 2013. Palvelinten varmistusjärjestelmät. Tampereen teknillinen yliopisto, tietotekniikan diplomi-insinöörin tutkinto-ohjelma, tietoliikenneverkot ja protokollat. Viitattu 21.11.2019. https://trepo.tuni.fi/bitstream/handle/123456789/22690/maatta.pdf?sequence=3&isAllowed=y

Poikonen, S. 2012. Sähköinen arkistointi ja pitkäaikaissäilytys Suomessa valmisohjelmistolla. Jyväskylän yliopisto, tietotekniikan pro gradu -tutkielma, ohjelmistotekniikan linja. Viitattu 16.10.2019. http://urn.fi/URN:NBN:fi:jyu-201201241061

Puricica, C. A. 2017. Demystifying Recovery Objectives. Verkkoartikkeli. Viitattu 25.10.2019. https://www.veeam.com/blog/rto-rpo-definitions-values-common-practice.html

Requests Prices in Azure, Google and AWS Compared. 2016. Verkkoartikkeli Cloudberry-sivustolla. Viitattu 10.12.2019. https://www.cloudberrylab.com/resources/blog/requests-and-data-transfer-prices-in-azure-google-and-aws-compared/

Rouse, M. 2014. What is virtual tape library (VTL). Verkkoartikkeli. Viitattu 22.12.2019. https://searchdatabackup.techtarget.com/definition/virtual-tape-library-VTL

Rouse, M. 2017. What is Google Cloud Platform (GCP)? Verkkoartikkeli. Viitattu 15.11.2019. https://searchcloudcomputing.techtarget.com/definition/Google-Cloud-Platform

Rouse, M. 2018. Microsoft Azure (Windows Azure). Verkkoartikkeli. Viitattu 8.11.2019. https://searchcloudcomputing.techtarget.com/definition/Windows-Azure

Rouse, M. 2019. What is Network Attached Storage? Verkkoartikkeli. Viitattu 5.10.2019. https://searchstorage.techtarget.com/definition/network-attached-storage

Russell, D. 2019. Veeam again in the Leader quadrant! Verkkoartikkeli Veeam-sivustolla. Viitattu 9.12.2019. https://www.veeam.com/executive-blog/gartner-magic-quadrant-2019-leader.html

Seget, V. 2019. What is Veeam Universal License? (VUL). Verkkoartikkeli. Viitattu 18.10.2019. https://www.vladan.fr/what-is-veeam-universal-license/

Taylor, C. 2019. RAID Levels Explained. Verkkoartikkeli. Viitattu 16.12.2019. https://www.enterprisestorageforum.com/storage-management/raid-levels.html

Telia Inmics-Nebula. 2019. Tietoa yrityksestä Telia Inmics-Nebula-sivustolla. Viitattu 4.10.2019. https://www.inmicsnebula.fi/fi/tietoa-yrityksesta

The Definitive Guide to Rubrik Cloud Data Management. 2019. Tekninen käyttöopas Rubrik-sivustolla. Viitattu 5.11.2019. https://pages.rubrik.com/rs/794-OHF-673/images/WTPR_CDM_Digital_Letter_20190430_v3.pdf

Varmuuskopiointipalvelut. N.d. Verkkoartikkeli WMHost-sivustolla. Viitattu 6.10.2019. https://www.wmhost.com/palvelut/sovellukset/varmuuskopiointi

Veeam Backup for Microsoft Office 365 User Guide. 2019. Dokumentaatio Veeam O365-varmistussovelluksesta. Viitattu 16.11.2019. https://helpcenter.veeam.com/docs/vbo365/guide/vbo_introduction.html?ver=30

Veeam Backup & Replication 9.5 Update 4 User Guide for VMware vSphere. 2019. Tekninen dokumentaatio Veeam-sivustolla. Viitattu 10.10.2019. https://helpcenter.veeam.com/docs/backup/vsphere/overview.html?ver=95u4

Veeam Continues to Dominate Cloud Data Management with Q3’19 Double-Digit Growth. 2019. Verkkoartikkeli Veeam-sivustolla. Viitattu 7.12.2019. https://www.veeam.com/news/veeam-continues-to-dominate-cloud-data-management-with-q3-19-double-digit-growth.html

Veeam Leadership Team. 2019. Johtoportaan esittely Veeam-sivustolla. Viitattu 2.10.2019. https://www.veeam.com/management-team.html

Wallenius, N. 2018. Onko hybrid cloud kuollut idea? Verkkoartikkeli. Viitattu 21.12.2019. https://niklaswallenius.fi/teknologiat/onko-hybrid-cloud-kuollut-idea/