Otto Lankia

SUMULASKENNAN TIETOTURVAONGELMAT VERRATTUNA PILVILASKENNAN TIETOTURVAONGELMIIN

JYVÄSKYLÄN YLIOPISTO
INFORMAATIOTEKNOLOGIAN TIEDEKUNTA
2019

TIIVISTELMÄ

Lankia, Otto
Sumulaskennan tietoturvaongelmat verrattuna pilvilaskennan tietoturvaongelmiin
Jyväskylä: Jyväskylän yliopisto, 2019, 34 s.
Tietojärjestelmätiede, kandidaatin tutkielma
Ohjaajat: Luoma, Eetu; Palonen, Teija

Pilvilaskenta on laskentaparadigma, jonka avulla tietojenkäsittelyresurssit voidaan tarjota asiakkaalle verkon välityksellä palveluiden muodossa. Nämä resurssit ovat virtualisoituja, dynaamisesti skaalautuvia sekä nopeasti käyttöönotettavia riippumatta ajasta, sijainnista tai asiakkaan käyttämästä päätelaitteesta. Sumulaskenta on laskentaparadigma, joka laajentaa pilvilaskennan palveluita lähemmäs loppukäyttäjää. Sumulaskennan ydinkomponentteja ovat maantieteellisesti hajautetut sumusolmut, jotka sijaitsevat päätelaitteiden ja keskitettyjen pilviresurssien välillä. Sumulaskenta on kehitetty vastaamaan pilvilaskennan heikkouksiin, kuten ydinverkkojen kuormituksen myötä kasvaviin viiveisiin. Sumulaskennan ei kuitenkaan ole tarkoitus korvata pilvilaskentaa, vaan tukea ja laajentaa sen palveluita. Pilvilaskennan tietoturvaa on tutkittu melko paljon, mutta sumulaskennan osalta tutkimus keskittyy pääasiassa sen hyötyihin. Näin ollen sumulaskennan ongelmat ovat jääneet vähemmälle huomiolle, vaikka tietoturvallisuus on merkittävä tekijä uusien teknologioiden menestyksen kannalta. Tämän kirjallisuuskatsauksen tarkoituksena oli selvittää kuinka sumulaskennan tietoturvaongelmat poikkeavat pilvilaskennan tietoturvaongelmista. Tutkielma osoitti sumulaskennan olevan melko haavoittuvainen laskentaparadigma, vaikka se tarjoaakin tietoturvaa esineiden internetin sovelluksille. Sumulaskenta perii kaikki pilvilaskennan olemassa olevat tietoturvaongelmat ja tämän lisäksi se tuo mukanaan kokonaan uudenlaisia tietoturvauhkia ominaispiirteidensä vuoksi.

Asiasanat: pilvilaskenta, sumulaskenta, sumusolmu, tietoturva, turvallisuusongelma

ABSTRACT

Lankia, Otto
Security issues of fog computing compared to security issues of cloud computing
Jyväskylä: University of Jyväskylä, 2019, 34 pp.
Information Systems, Bachelors Thesis
Supervisors: Luoma, Eetu; Palonen, Teija

Cloud computing is a computing paradigm that can be used to provide computing resources to the customer over the Internet in the form of a service. These resources are virtualized, dynamically scalable and quickly deployable, regardless of time, location or the device used by the end user. Fog computing is a computing paradigm that expands the services of cloud computing closer to the end user. The core components of fog computing are fog nodes that are geographically distributed. Fog nodes are located between end devices and geographically centralized cloud resources. Fog computing is developed to respond to the weaknesses of cloud computing, such as increasing delays due to the load on the core networks. Fog computing is not meant to replace cloud computing, but to support and expand its services. The security of cloud computing is widely researched, but research of fog computing focuses mainly on the benefits of it. Problems of fog computing have received less attention although security is an important factor in the success of new technologies. The purpose of this literature review was to find out how the security issues of fog computing differ from cloud computing security issues. This thesis showed that fog computing is a relatively vulnerable computing paradigm, although it provides security for applications of Internet of Things. Fog computing inherits all the existing security issues of cloud computing. In addition, fog computing also brings new security threats duo to its charasterics.

Keywords: cloud computing, fog computing, fog node, information security, security issue

KUVIOT

SISÄLLYS

1 JOHDANTO

Pilvilaskenta on yksi nopeinten kasvavista segmenteistä informaatioteknologian alalla (Popović & Hocenski, 2010). Pilvilaskenta on laskentaparadigma, jossa tietojenkäsittelyresurssit tarjotaan palveluina asiakkaalle internetin välityksellä ja sillä viitataan sekä palveluina tarjottaviin sovelluksiin että laitteistoihin, jotka mahdollistavat nämä palvelut (Armbrust ym., 2010). Martson, Li, Bandyopadhyay, Zhang ja Ghalsasi (2011) määrittelevät pilvilaskennan palvelumalliksi, jossa sekä laitteisto- että sovelluspalvelut toimitetaan tilausperusteisesti asiakkaan käyttöön verkon välityksellä riippumatta sijainnista tai käytettävästä päätelaitteesta. Tarjottavat resurssit ovat jaettuja, dynaamisesti skaalautuvia, nopeasti käyttöönotettavia, virtualisoituja ja tarjottavissa minimaalisella vuorovaikutuksella asiakkaan ja palveluntarjoajan välillä. Käyttäjät maksavat palveluista käytön määrän perusteella ilman merkittäviä alkuinvestointeja. Lisäksi pilvipalvelut sisältävät järjestelmän, joka jakaa laskentaresurssit sopiviksi lohkoiksi (Marston ym., 2011).

Sumulaskenta on korkeasti virtualisoitu, pilvilaskentaa täydentävä arkkitehtuuri, joka tarjoaa laskenta-, tallennus- ja verkkopalveluja päätelaitteiden ja perinteisten pilvilaskentaresurssien välillä (Bonomi, Milito, Zhu & Addepalli, 2012; Saharan & Kumar, 2015; Stojmenovic, Wen, Huang & Luan, 2016). Perinteisesti pilvilaskennassa käytettyjen datakeskusten laskenta- ja tallennusresurssien lisäksi sumulaskenta hyödyntää sumusolmuja (engl. *fog nodes*), jotka voivat olla esimerkiksi tukiasemia, reitittimiä tai päätelaitteita (Yi, Li & Li, 2015). Sumulaskennan tärkeimpänä tarkoituksena on vähentää pilvilaskennan taakkaa kokoamalla osa sen palveluista, sovelluksista ja datasta lähelle verkon reunaa (Saharan & Kumar, 2015). Sumulaskennan tarkoituksena ei siis ole korvata pilvilaskentaa, vaan täydentää sitä (Dastjerdi & Buyya, 2016; Shropshire, 2014). Saharan ja Kumar (2015) kirjoittavat sumulaskennan syntyneen ratkaisuna pilvilaskennan rajoitteisiin ja haasteisiin, kuten jatkuvasti kasvava verkkojen kuormitus ja sen myötä kasvavat viiveet ja kustannukset sekä datan maantieteellisen sijainnin mukanaan tuomat haasteet. Sumulaskenta voidaan käyttää tukemaan pilvilaskentaa, mutta tulevaisuudessa tämän lisäksi sumulaskennan avulla pystytään

tarjoamaan uusia mahdollisuuksia ja ratkaisuja sekä palveluiden tarjoajille että loppukäyttäjille (Saharan & Kumar, 2015).

Nykyaikaisten teknologioiden menestys riippuu vahvasti tehokkuudesta ja helppokäyttöisyydestä, mutta ennen kaikkea tietoturvasta ja valvonnasta (Ramgovind, Eloff & Smith, 2010). Singhin (2013) mukaan tietoturvalla tarkoitetaan tiedon ja palveluiden suojaamista ulkopuolisilta tahoilta. Tietoturvan keskeisenä tavoitteena on turvata datan eheys, saatavuus ja luottamuksellisuus (Singh, 2013). Zissisin ja Lekkasin (2012) mukaan pilvilaskenta on tuonut mukanaan uusia tietoturvahaasteita, joiden kohdalla perinteiset suojausmekanismit eivät ole enää riittävän tehokkaita. Tämä johtuu siitä, että pilvilaskennan käyttöönottomallit poikkeavat huomattavasti perinteisten arkkitehtuurien ominaispiirteistä (Zissis & Lekkas, 2012). Vuonna 2008 International Data Corporation selvitti tutkimuksessaan, että 74,6% eri yritysten tietohallinnon johto- ja vastuuhenkilöistä pitää tietoturvaa pilvilaskennan keskeisimpänä haasteena (Zhou, Zhang, Xie, Qian & Zhou, 2010). Tietoturva on myös sumulaskennan kannalta mielenkiintoinen näkökulma, sillä sumulaskenta nähdään kirjallisuudessa usein ratkaisuna pilvilaskennan ongelmiin ja näin ollen sumulaskennan haasteet ja ongelmat jäävät helposti vähemmälle huomiolle.

Tämän kirjallisuuskatsauksen tavoitteena on selvittää koskeeko sumulaskentaa samat tietoturvaongelmat kuin pilvilaskentaa, tuoko sumulaskenta mukanaan kokonaan uusia tietoturvaongelmia, vai onko sumulaskennan avulla mahdollista ratkaista joitakin pilvilaskennan tietoturvaongelmia. Lisäksi tarkoituksena on selvittää vastaako sumu- ja pilvilaskennan tietoturvaratkaisut yhteisten ongelmien osalta toisiaan, vai onko niissä eroavaisuuksia. Tämän kirjallisuuskatsauksen tutkimuskysymykseksi on asetettu:

- Kuinka sumulaskennan tietoturvaongelmat poikkeavat pilvilaskennan tietoturvaongelmista?

Tutkimuskysymys voidaan pilkkoa pienempiin osiin määrittelemällä sen tueksi apukysymyksiä. Tämä tekee laajan kokonaisuuden käsittelystä vaivattomampaa. Tässä kirjallisuuskatsauksessa tutkimuskysymyksen tueksi on määritelty seuraavat apukysymykset:

- Mitä on pilvilaskenta?
- Mitä on sumulaskenta?
- Millaisia tietoturvariskejä pilvilaskentaan liittyy?
- Millaisia tietoturvariskejä sumulaskentaan liittyy?

Tämä kandidaatin tutkielma toteutetaan systemaattisena kirjallisuuskatsauksena. Kirjallisuuskatsaus toteutetaan soveltaen Okolin ja Schabramin (2010) määrittelemää opasta systemaattisen kirjallisuuskatsauksen toteuttamiseksi. Lähdeaineistona käytetään pääasiassa tieteellisissä julkaisuissa julkaistuja artikkeleita. Pääasiallisena hakukoneena toimii Google Scholar -palvelu. Apuna käytetään myös AIS eLibrary -palvelua sekä Scopus-tietokannan dokumenttihakua. Lähdeaineistoa kerätessä kiinnitetään huomiota lähteiden viittausten määrään ja

lähteiksi pyritään valitsemaan mahdollisimman tuoreita julkaisuja. Lisäksi julkaisukanavien tasoluokitukset tarkistetaan julkaisufoorumi-palvelussa. Lähdeaineisto pyritään rajaamaan mahdollisimman laadukkaisiin ja luotettaviin julkaisuihin, jotka sopivat tutkimuskysymysten asettamiin rajoihin. Lähdeaineiston keräämisessä käytetään pääasiallisina hakusanoina termejä ”cloud computing” ja ”fog computing”, joita yhdistetään hakusanapareiksi seuraavien hakusanojen kanssa: security, data security, information security, privacy ja threats.

Kirjallisuuskatsauksen rakenne koostuu kahdesta sisältöluvusta, joita seuraa yhteenveto. Ensimmäisessä sisältöluvussa perehdytään pilvilaskentaan sekä sumulaskentaan. Kirjallisuudessa pilvilaskennalle on esitelty useita vaihtoehtoisia määritelmiä, joita tässä luvussa vertaillaan. Määritelmän lisäksi luvussa tutustutaan pilvilaskennan pääpiirteisiin. Lisäksi luvussa määritellään sumulaskennan käsite, jolle on myös tarjolla vaihtoehtoisia määritelmiä. Tämän jälkeen esitellään sumulaskennan pääpiirteet ja tutustutaan sen hyödyntämismahdollisuuksiin. Toisessa sisältöluvussa pyritään vastaamaan tutkimuskysymykseen paneutumalla sekä pilvi- että sumulaskennan tietoturvaongelmiin ja vertailemalla niitä keskenään. Yhteenvedossa kerrataan tutkielman keskeinen sisältö ja pohditaan mahdollisia jatkotutkimusaiheita.

2 PILVI- JA SUMULASKENTA

Tässä luvussa tutustutaan kahteen laskentaparadigmaan, pilvi- ja sumulaskentaan. Ensimmäisenä vuorossa on pilvilaskentaan tutustuminen, sillä sumulaskentaa on helpompi lähestyä, kun pilvilaskenta on pääpiirteittäin jo tuttu. Pilvilaskentaan tutustutaan määrittelemällä aluksi sen käsite. Tämän jälkeen perehdytään pilvilaskennan pääpiirteisiin ja sen eri malleihin. Sumulaskenan osalta esittely aloitetaan käsitteen määrittelyllä, jonka jälkeen tutustutaan sumulaskennan ominaispiirteisiin, toimintaperiaatteeseen sekä hyötyihin. Luvun lopuksi esitellään sumulaskennan tarjoamia hyödyntämiskohteita ja -mahdollisuuksia.

2.1 Pilvilaskenta

Zhang, Cheng ja Boutaba (2010) toteavat artikkelissaan, ettei ajatus pilvilaskennan taustalla ole kovinkaan uusi. Jo 1960-luvulla tietojenkäsittelytieteilijä John McCarthy visioi, että laskentaresursseja pystyttäisiin tarjoamaan kuluttajille julkisen hyödykkeen tavoin. Vuonna 2006 pilvi-termin käyttö alkoi yleistyä, kun Googlen silloinen toimitusjohtaja Eric Schmidt käytti termiä kuvaamaan liiketoimintamallia, jossa palvelut tarjotaan asiakkaalle verkon välityksellä (Zhang ym., 2010). Pilvilaskennassa asiakas maksaa palveluista käyttön määrän mukaisesti, eikä asiakkaan tarvitse itse huolehtia käyttämiensä resurssien ja niiden taustalla olevan infrastruktuurin hallinnasta (Aazam & Huh, 2014). Shropshiren (2014) mukaan pilvilaskennasta on tullut suosituin malli tietojenkäsittelytarpeiden täyttämiseksi. Pilvilaskenta perustuu resurssien yhdistämiseen (engl. *resource pooling*) ja se tarjoaa joustavan alustan, joka skaalautuu tarvittaessa vaihtelevan käyttötason mukaan. Lisäksi pilvilaskenta tarjoaa sijaintiriippumattomuutta, sillä resurssit ovat käytettävissä verkon välityksellä (Shropshire, 2014).

2.1.1 Pilvilaskennan määritelmä

Vaikka pilvilaskentaa on pyritty määrittelemään useiden tutkijoiden toimesta, sille ei ole muodostunut yhtä yleisesti hyväksyttyä määritelmää (Takabi, Joshi & Ahn, 2010). Foster, Zhao, Raicu ja Lu (2008) ovat määritelleet pilvilaskennan seuraavanlaisesti:

> Laajamittainen hajautettu laskentaparadigma, jota ohjaa mittakaavaedut, (engl. *economies of scale*) jossa jaetun sijainnin (engl. *pool*) abrtraktoidut, virtualisoidut, dynaamisesti skaalautuvat, hallinnoidut laskentakapasiteetit, tallennustilat, alustat ja palvelut toimitetaan kysynnän perusteella ulkoisille asiakkaille internetin välityksellä. (Foster ym., 2008)

Määritelmässään Foster ym. (2008) tuovat hyvin esille pilvilaskennan luonteenomaisen piirteen, jossa skaalautuvat palvelut tarjotaan asiakkaalle internetin välityksellä perustuen asiakkaan henkilökohtaisiin tarpeisiin. Määritelmä ei kuitenkaan huomioi pilvilaskentaresurssien käyttöönoton ominaispiirteitä tai tuo esille palveluiden käytettävyyttä ajasta, paikasta tai laitteesta riippumatta.

Martson ym. (2011) ovat ottaneet artikkelissaan myös nämä seikat huomioon ja tarjoavat pilvilaskennalle liiketoimintalähtöistä määritelmää, joka kuuluu seuraavasti:

> Pilvilaskenta on informaatioteknologian palvelumalli, jossa laskentapalvelut (sekä laitteisto että ohjelmisto) toimitetaan tarpeen vaatiessa asiakkaalle verkon välityksellä itsepalvelu tyylisesti, riippumatta laitteesta ja sijainnista. Resurssit, jotka tarvitaan vaaditun palvelunlaatutason saavuttamiseksi, ovat jaettuja, dynaamisesti skaalautuvia, nopeasti saavavilla olevia, virtualisoituja ja vapautettavissa minimaalisella vuorovaikutuksella palvelun tarjoajan kanssa. Käyttäjät maksavat palvelusta käytön mukaan ilman merkittäviä alkuinvestointeja, ja pilvipalvelut hyödyntävät mittausjärjestelmää, joka jakaa laskentaresurssit sopiviksi lohkoiksi (engl. *blocks*). (Marston ym., 2011)

Mertsonin ym. (2011) määritelmä on kattava ja se ottaa huomioon useimmat pilvilaskennan ominaispiirteet. Kirjallisuudessa eniten viitattu pilvilaskennan määritelmä on kuitenkin esitelty National Institute of Standards and Technologyn (NIST) julkaisussa. Tässä julkaisussa Mell ja Grance (2011) ovat määritelleet pilvilaskennan seuraavanlaisesti:

> Pilvilaskenta on malli, jolla mahdollistetaan tarpeen vaatiessa ajasta ja paikasta riippumaton kätevä pääsy varantoon (engl. *pool*) konfiguroitavissa olevia resursseja (esimerkiksi verkkoyhteyksiä, servereitä, tallennustilaa, sovelluksia ja palveluita), jotka voidaan nopeasti valjastaaa käyttöön tai vapauttaa vähäisillä hallintotoimilla ja minimaalisella vuorovaikutuksella asiakkaan ja palveluntarjoajan välillä. Pilvilaskennan malli muodostuu viidestä olenaisesta ominaisuudesta, kolmesta palvelumallista ja neljästä käyttöönottomallista. (Mell & Grance, 2011)

Kyseinen määritelmä tuo ominaispiirteiden lisäksi esille myös pilvilaskennan palvelu- ja käyttöönottomallit. Useat pilvilaskentaa käsittelevät artikkelit käyttävät pilvilaskennalle juuri tätä määritelmää, ja se vaikuttaakin olevan kirjallisuudessa yleisesti hyväksytty määritelmä.

2.1.2 Pilvilaskennan pääpiirteet

Mell ja Grance (2011) määrittelevät pilvilaskennan mallin muodostuvan viidestä olennaisesta ominaisuudesta, kolmesta palvelumallista ja neljästä käyttöönottomallista. He luettelevat pilvilaskennan ominaisuuksiksi seuraavat piirteet:

Tarvepohjainen itsepalvelu (engl. *on-demand self-service*). Kuluttaja voi itse varata pilvilaskentaresursseja tarpeensa mukaan ilman ihmisten välistä vuorovaikutusta palvelun tarjoajan kanssa (Mell & Grance, 2011).

Laaja saatavuus verkossa (engl. *broad network access*). Resurssit ovat saatavilla verkon välityksellä ja käytettävissä eri alustoilla, kuten tietokoneilla ja mobiililaitteilla (Mell & Grance, 2011).

Resurssien yhdistäminen (engl. *resource pooling*). Palvelun toimittaja tarjoaa joukon tietojenkäsittelyresursseja, jotka on mahdollista jakaa dynaamisesti kulutuskysynnän mukaan käyttäen moniasiakkuus-mallia (engl. *multi-tenant model*) (Mell & Grance, 2011; Zhang ym., 2010). Lisäksi paikan riippumattomuus ilmenee siten, ettei asiakas pysty kontrolloimaan datan tarkkaa sijaintia, mutta pystyy mahdollisesti määrittämään sen esimerkiksi maan tai osavaltion tasolla (Mell & Grance, 2011).

Nopea joustavuus (engl. *rapid elasticity*). Resurssit ovat joustavasti ja osittain automaattisesti skaalautuvia siten, että kuluttajalle resurssit näyttäytyvät käytännössä rajattomina (Mell & Grance, 2011).

Mitattu palvelu (engl. *measured service*). Pilvijärjestelmät kontrolloivat ja optimoivat automaattisesti resursseja ja niiden käyttöä voidaan tarkkailla sekä asiakkaan että palveluntarjoajan toimesta (Mell & Grance, 2011).

2.1.3 Pilvilaskennan palvelumallit

Zhangin ym. (2010) mukaan pilvilaskennan arkkitehtuuri voidaan yleisesti ottaen jakaa neljään erilliseen kerrokseen, jotka ovat laitteistokerros (engl. *hardware layer*), infrastruktuurikerros (engl. *infrastructure layer*), alustakerros (engl. *platform layer*) ja sovelluskerros (engl. *application layer*). Laitteistokerros on pilvilaskennan arkkitehtuurin alin kerros, joka tarjoaa fyysiset laitteistoresurssit pilvilaskennan muiden kerrosten palveluille. Nämä laitteistoresurssit sijaitsevat datakeskuksissa, jotka tyypillisesti sisältävät tuhansia palvelimia, jotka ovat kytketty toisiinsa kytkinten, reitittimien ja muiden rakenteiden avulla. Laitteistokerros mahdollistaa pilvilaskennan ylempien kerrosten palvelumallien mukaiset palvelut (Zhang ym., 2010).

Mell ja Grance (2011) tunnistavat pilvilaskennalle kolme palvelumallia, jotka ovat Infrastructure as a Service (IaaS), Platform as a Service (PaaS) ja Software as a Service (SaaS). Nämä palvelumallit rakentuvat pilvilaskennan arkkitehtuurin mukaan kerroksittain ja ovat toisistaan riippumattomia (Ali, Khan & Vasilakos, 2015).

Infrastructure as a Service. Käyttäjälle tarjotaan laskentaresursseja, tallennuskapasiteettia, verkkoyhteiksyä tai muita tavanomaisia tietojenkäsittelyresursseja (Mell & Grance, 2011). Käyttäjä ei voi kontrolloida tai hallita infrastruktuuria palvelun alla, mutta käyttöjärjestelmä, muisti ja sovellukset ovat käyttäjän hallittavissa (Mell & Grance, 2011; Saharan & Kumar, 2015). Lisäksi käyttäjällä voi olla mahdollisuus hallita verkkokomponentteja, kuten palomuuria (Mell & Grance, 2011).

Platform as a Service. Käyttäjällä ei ole mahdollisuutta hallita palvelun taustalla olevaa infrastruktuuria, käyttöjärjestelmää tai tallennustilaa (Mell & Grance, 2011). Mallissa käyttäjä maksaa pääsystä alustaan, jossa hän voi ajaa haluamiaan alustan tarjoajan tukemia sovelluksia ja ohjelmistoja (Mell & Grance, 2011; Saharan & Kumar, 2015).

Software as a Service. Mallissa käyttäjälle tarjotaan mahdollisuus käyttää palveluntarjoajan sovelluksia, jotka toimivat pilvi-infrastruktuurissa (Mell & Grance, 2011). Sovelluksia voi käyttää ajasta ja paikasta riippumatta erilaisilla laitteilla jonkin käyttöliittymän, kuten web-selaimen kautta (Mell & Grance, 2011; Saharan & Kumar, 2015). Käyttäjä ei voi hallita sovelluksen taustalla olevia tekijöitä, vaan käyttäjällä on pääsy vain palveluna tarjottavaan sovellukseen (Mell & Grance, 2011).

2.1.4 Pilvilaskennan käyttöönottomallit

Mell ja Grance (2011) määrittelevät pilvilaskennalle neljä käyttöönottomallia, jotka ovat yksityinen pilvi, julkinen pilvi, yhteisöllinen pilvi ja hybridipilvi. Käyttöönottomallien erot perustuvat siihen, kuinka yksinomaisesti tietojenkäsittelyresurssit tarjotaan kuluttajalle (Liu ym., 2011).

Yksityinen pilvi (engl. *private cloud*). Pilvilvipalvelu tarjotaan yksinomaiseen käyttöön tietylle organisaatiolle, joka voi koostua useista kuluttajista, kuten liiketoimintayksiköistä (Mell & Grance, 2011). Yksityistä pilvieä voi hallita joko itse asiakas, palveluntarjoaja tai kolmas osapuoli (Liu ym., 2011; Mell & Grance, 2011). Lisäksi yksityisen pilven infrastruktuuri voidaan järjestää joko kuluttajaorganisaation tiloissa, tai se voidaan ulkoistaa pilven isännöintiyritykselle (Liu ym., 2011).

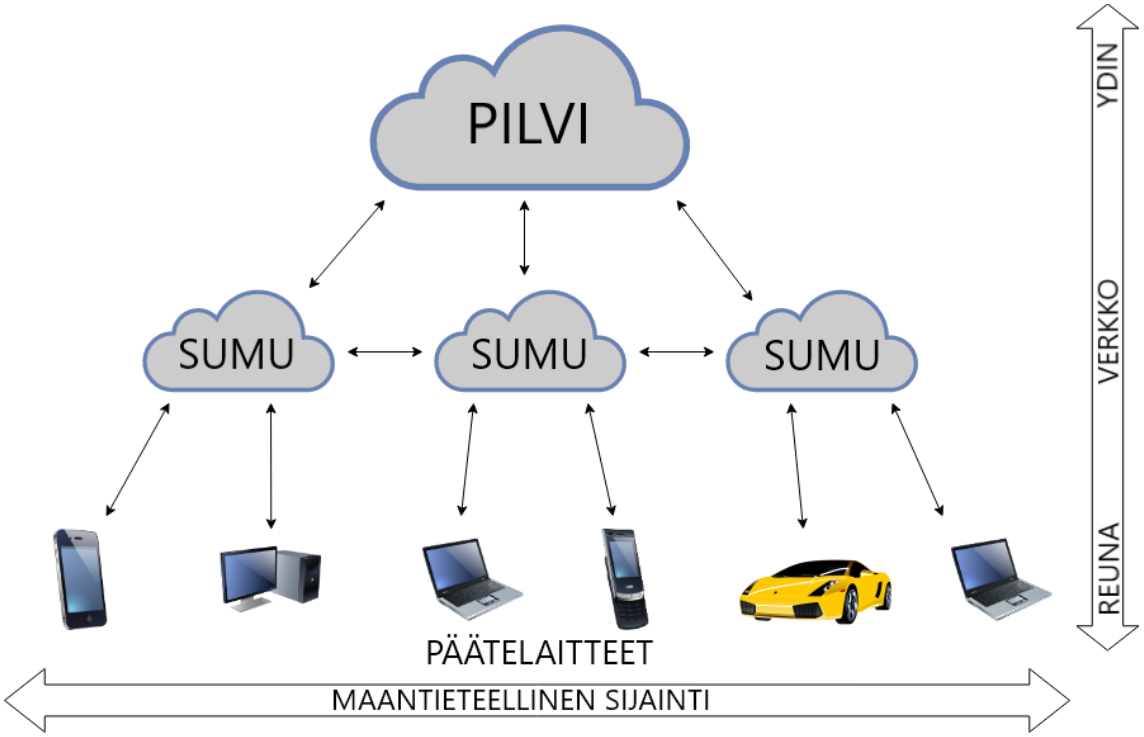
Julkinen pilvi (engl. *public cloud*). Julkinen pilvi tarjotaan avoimesti suurelle yleisölle palveluntarjoajan toimesta (Mell & Grance, 2011; Saharan & Kumar, 2015). Sen infrastruktuuri on järjestetty palveluntarjoajan tiloihin, joka myös hallinnoi julkista pilveä (Mell & Grance, 2011).

Yhteisöllinen pilvi (engl. *community cloud*). Yhteisöllisen pilven infrastruktuuri jaetaan useiden organisaatioiden kesken, joilla on samankaltaisia vaatimuksia ja intressejä (Mell & Grance, 2011; Saharan & Kumar, 2015). Kuten yksityinen pilvi, yhteisöllinen pilvi voi olla hallinnoitu asiakkaan, palveluntarjoajan tai kolmannen osapuolen toimesta ja infrastruktuuri voidaan toteuttaa asiakkaan tai ulkoisen toimijan tiloihin (Liu ym., 2011).

Hybridipilvi (engl. *hybrid cloud*). Hybridipilvi on yhdistelmä kahdesta tai useammasta edellämainitusta pilvilaskennan infrastruktuurista, jotka pysyvät erillisinä kokonaisuuksina, mutta mahdollistavat datan ja sovellusten siirrettävyyden eri infrastruktuurien välillä (Liu ym., 2011; Mell & Grance, 2011).

2.2 Sumulaskenta

Sumulaskenta on alunperin tietoliikenne- ja elektronikkateollisuuden alan yrityksen Ciscon lanseeraama termi (Tordera ym., 2016). Se on laskentaparadigma, joka laajentaa pilvilaskennan palveluita lähemmäs loppukäyttäjää ja verkon reunaa (Bonomi, Milito, Natarajan & Zhu, 2014; Shropshire, 2014). Sumulaskennan nimi onkin syntynyt pilvilaskennan kautta, sillä sumu on kuin pilvi, joka on laskeutunut lähemmäs maan pintaa (Bonomi ym., 2012). Sumulaskennan ei ole tarkoitus korvata pilvilaskentaa, vaan toimia sen tukena ja tuoda tuoda esille molempien laskentaparadigmojen parhaat puolet (Shropshire, 2014). Seuraavassa kuviossa esitetään sumulaskennan arkkitehtuuri.



KUVIO 1 Sumulaskennan arkkitehtuuri (muokattu: Stojmenovic, Wen, Huang & Luan, 2016)

2.2.1 Sumulaskennan määritelmä

Pilvilaskennan tavoin myös sumulaskennalle on tarjolla vaihtoehtoisia määritelmiä. Sumulaskenta on melko tuore käsite, joten sille ei ole muodostunut yhtä täsmällistä määritelmää. Seuraavaksi tutustutaan kahteen kirjallisuudessa yleisesti esille nousevaan määritelmään ja vertaillaan niitä keskenään.

Vaquero ja Rodero-Merino (2014) esittävät sumulaskennalle määritelmän, jossa he pyrkivät korostamaan sumulaskennan tärkeimpiä piirteitä. Heidän tarjoama määritelmä on seuraavanlainen:

> Sumulaskenta on skenaario, jossa suuri määrä monimuotoisia, (langattomia ja joskus autonomisia) ajasta ja paikasta riippumattomia ja hajautettuja laitteita kommunikoivat

> keskenään, toimivat yhteistyössä ja suorittavat tiedon tallennus- ja käsittelytehtäviä verkon välityksellä ilman kolmansien osapuolten toimia. Nämä tehtävät voivat tukea tavallisia verkkotoimintoja tai kokonaan uusia palveluita ja sovelluksia, joita ajetaan eristetyssä ympäristössä (engl. *sandboxed environment*). Käyttäjät saavat kannustimia, jotta he vuokraavat osan laitteistaan isännöimään näitä palveluja. (Vaquero & Rodero-Merino, 2014)

Määritelmä tuo esille monimuotoisten laitteiden välisen kommunikoinnin ja se korostaa ajasta ja paikasta riippumattomuutta sekä näiden laitteiden välistä yhteistyötä ilman kolmansien osapuolten toimia. Määritelmässä mainitaan myös, että sumulaskenta mahdollistaa kokonaan uudenlaisia palveluja. Tämä määritelmä jää kuitenkin hieman puutteelliseksi, sillä Vaquero ja Rodero-Merino (2014) eivät huomioi määritelmässään sumulaskennan suhdetta pilvilaskentaan, mikä on yksi sumulaskennan keskeisimmistä periaatteista. National Institute of Standards and Technologies (NIST) tarjoaa määritelmän, joka huomioi myös pilvilaskennan roolin. Tämä määritelmä kuuluu seuraavasti:

> Sumulaskenta on kerrostettu malli, joka tarjoaa ajasta ja paikasta riippumattoman pääsyn skaalautuvien laskentaresurssien joukkoon. Malli tekee helpommaksi hajautettujen sekä viiveherkkien sovellusten ja palveluiden käyttöönoton, ja se muodostuu fyysisistä tai virtuaalisista sumusolmuista (engl. *fog nodes*), jotka sijaistevat älykkäiden päätelaitteiden ja keskitettyjen (pilvi) palveluiden välillä. Sumusolmut ovat kontekstitietoisia ja ne tukevat sekä yleistä datan hallinointia että viestintäjärjestelmää. Ne voidaan järjestää klustereissa – joko vertikaalisesti (tukemaan eristämistä), horisontaalisesti (tukemaan yhteen liittämistä), tai suhteessa sumusolmujen ja älykkäiden päätelaitteiden väliseen viiveeseen. Sumulaskenta minimoi pyyntöihin vastausajan tuetuille sovelluksille ja tarjoaa päätelaitteille paikallisia laskentaresursseja sekä tarvittaessa myös verkkoyhteyden keskitettyihin palveluihin. (Iorga ym., 2018)

Iorgan ym. (2018) hyvin tuore määritelä vaikuttaa kirjallisuuden kattavimmalta määritelmältä sumulaskennalle. Määritelmän mukaan sumulaskenta muodostuu joko fyysisistä tai virtuaalisista sumusolmuista, jotka sijaitsevat päätelaitteiden ja keskitetyn pilviarkkitehtuurin välillä. Sumusolmut ovat sumulaskenta-arkkitehtuurin keskeisiä rakennuspalikoita, joita käsitellään enemmän luvussa 2.2.3. Iorga ym. (2018) huomioivat myös sen, että sumusolmujen kautta on mahdollista muodostaa verkkoyhteys keskitettyihin pilvilaskentaresursseihin. Lisäksi määritelmä tuo esille sumulaskennan hyödyt hajautettujen ja viiveherkkien sovellusten toteutuksessa.

2.2.2 Sumulaskennan ominaispiirteet

Iorgan ym. (2018) mukaan sumulaskennalla on kuusi ominaispiirrettä, jotka ovat olennaisia sumulaskennan erottamiseksi muista laskentaparadigmoista. Nämä kuusi ominaispiirrettä ovat kontekstuaalinen sijaintitietoisuus ja matala viive, maantieteellinen jakautuminen, heterogeenisyys, yhteentoimivuus ja yhdistäminen, reaaliaikainen vuorovaikutus sekä yhdistettyjen sumusolmuklustereiden skaalautuvuus ja keteryys. On kuitenkin hyvä

huomioida, ettei käyttäjä välttämättä hyödynnä kaikkia näitä ominaispiirteitä käyttäessään sumulaskennan palveluita (Iorga ym., 2018). Seuraavaksi avataan mitä näillä ominaispiirteillä tarkoitetaan.

Kontekstuaalinen sijaintitietoisuus ja matala viive (engl. *contextual location awareness, and low latency*). Iorgan ym. (2018) mukaan sumulaskenta tarjoaa alhaisimman mahdollisen viiveen johtuen sumusolmujen tietoisuudesta niiden loogisesta sijainnista koko järjestelmän kontekstissa sekä viivekustannuksista (engl. *latency costs*) muiden sumusolmujen kanssa kommunikoidessa. Sumulaskennan alkuperänä voidaan pitää ehdotuksia verkon reunalla olevista päätepisteistä, jotka tukevat monipuolisia palveluita, kuten sovelluksia, jotka vaativat alhaisia viiveitä. Sumusolmut sijaitsevat usein älykkäiden päätelaitteiden yhteydessä, joten datan analysointi ja käsittely on paljon nopeampaa näissä laitteissa, kuin keskitetyissä pilvipalveluissa tai datakeskuksissa (Iorga ym., 2018).

Maantieteelinen jakautuminen (engl. *geographical distribution*). Pilvipalveluiden ollessa keskitettyjä sumulaskennan palvelut ovat maantieteellisesti hajautettuja (Iorga ym., 2018). Sumulaskennan avulla voidaan esimerkiksi tarjota korkealaatuisia suoratoistopalveluita liikkuville kulkuneivoille teiden varsille sijoitettujen välityspalvelimien kautta (Bonomi ym., 2012; Iorga ym., 2018).

Heterogeenisyys (engl. *heterogeneity*). Sumulaskenta tukee datan keräämistä ja käsittelyä toisistaan poikkeavien teknologioiden ja verkkoon kytkettyjen laitteiden välillä (Iorga ym., 2018).

Yhteentoimivuus ja yhdistäminen (engl. *interoperability and federation*). Tiettyjen palveluiden, kuten reaaliaikaisten suoratoistopalveluiden saumaton tuki edellyttää eri toimijoiden yhteistyötä (Bonomi ym., 2012; Iorga ym., 2018). Tällöin sumulaskennan komponenttien tulee olla yhteentoimivia ja palvelut on yhdistettävä eri alueiden välillä (Iorga ym., 2018).

Reaaliaikainen vuorovaikutus (engl. *real-time interactions*). Sumulaskennan sovellukset hyödyntävät reaaliaikaista vuorovaikutusta eräkäsittelyjen sijaan (Iorga ym., 2018).

Yhdistettyjen sumusolmu-klustereiden skaalautuvuus ja ketteryys (engl. *scalability and agility of federated, fog-node clusters*). Sumulaskenta on klusteritasolla luonteeltaan mukautuvaa ja joustavaa, joka tukee muun muassa joustavaa laskentaa, resurssien yhdistämiestä, kuormituksen muuroksia ja verkon tilan vaihteluita (Iorga ym., 2018).

Iorga ym. (2018) määrittelevät julkaisussaan näiden ominaispiirteiden lisäksi sumulaskennan lisäominaisuudeksi tuen liikkuvuudelle. Vaikka sumulaskentaa käytetään myös langallisessa ympäristössä, niin monet sumulaskennan keskeiset sovellukset tarvitsevat langattomia yhteyksiä (Bonomi ym., 2012; Iorga ym., 2018). Näiden sumulaskennan piirteiden lisäksi Bonomi ym. (2012) mainitsevat sumusolmujen maantieteellisen jakautumisen ohessa yhdeksi ominaispiirteeksi sumusolmujen todella suuren määrän.

2.2.3 Sumusolmut

Sumusolmut ovat sumulaskenta-arkkitehtuurin ydinkomponentteja, jotka sijaitsevat päätelaitteiden ja keskitettyjen pilvilaskentaresurssien välillä (Iorga ym., 2018; Tordera ym., 2016; Yi ym., 2015). Sumusolmut voivat olla joko fyysisiä- (esim. kytkimiä, reitittimiä tai palvelimia) tai virtuaalisia (esim. virtuaalisia kytkimiä tai virtuaalikoneita) komponentteja (Iorga ym., 2018). Tietoliikenne- ja elektronikkateollisuuden alan yrityksen Ciscon mukaan käytännön tasolla katsoen mikä tahansa laite, joka tarjoaa laskentatehoa, tallennuskapasiteettia ja mahdollisuuden verkkoyhteyteen voi mahdollisesti toimia sumusolmuna (Matt, 2018). Tordera ym. (2016) ovat määritelleet sumusolmun käsitteen seuraavanlaisesti:

> Sumusolmut ovat sumulaskennan hajautettuja laskentayksiköitä, jotka mahdollistavat sumupalveluiden käyttöönoton ja jotka muodostuvat ainakin yhdestä fyysisestä laitteesta, jolla on prosessointi- ja tunnistusonminaisuuksia (engl. *sensing capabilities*) (esim. tietokone, matkapuhelin, verkon älykäs reunalaite (engl. *smart edge device*), auto, lämpötila-anturi jne.). Kaikki sumusolmun laitteet kytketään yhteen erilaisilla verkkoteknologioilla (langallinen ja langaton) ja yhdistetään yhdeksi loogiseksi kokoneisuudeksi, eli sumusolmuksi, joka pystyy suorittamaan saumattomasti hajautettuja palveluita, kuten ne olisivat yhdellä laitteella. (Tordera ym., 2016)

Iorgan ym. (2018) mukaan sumusolmut ovat tietoisia sekä maantieteellisestä sijainnistaan että loogisesta sijainnistaan klusterin sisällä. Lisäksi sumusolmut tarjoavat tiedonhallinta- ja viestintäpalvelun verkon reunakerroksen ja keskitettyjen resurssien välille (Iorga ym., 2018; Luan ym., 2015). Sumusolmut voivat verkkoon kytkettyinä toimia missä tahansa, esimerkiksi teiden varsilla tai tehtaiden lattioissa hyödyntämistarkoituksestaan riippuen (Cisco, 2015). Jotta sumusolmut mahdollistavat sumulaskennan käyttöönoton, niin niiden on tuettava yhtä tai useampaa seuraavista Iorgan ym. (2018) määrittelemistä ominaispiirteistä:

Autonomia (engl. *autonomy*). Sumusolmut voivat toimia itsenäisesti tekemällä päätöksiä yksittäisen sumusolmun tai klusterin tasolla (Iorga ym., 2018).

Heterogeenisyys (engl. *heterogeneity*). Sumusolmut voivat olla monimuotoisia laitteita ja ne voivat toimia monissa erilaisissa ympäristöissä (Iorga ym., 2018).

Hierarkinen klusterointi (engl. *hierarchial clustering*). Sumusolmut tukevat hierarkisia rakenteita, joissa eri kerrokset tarjoavat erilaisia palvelutoimintojen osa-alueita yhteisenä jatkumona (Iorga ym., 2018).

Hallittavuus (engl. *manageability*). Sumusolmuja hallitaan ja organisoidaan monimutkaisilla järjestelmillä, jotka pystyvät suorittamaan useimmat rutiininomaiset toiminnot automaattisesti (Iorga ym., 2018).

Ohjelmoitavuus (engl. *programmability*). Sumusolmut ovat luonnostaan ohjelmoitavissa useilla eri tasoilla sekä useiden eri sidosryhmien, kuten esimerkiksi verkko-operaattoreiden, laitteiden tarjoajien ja loppukäyttäjien toimesta (Iorga ym., 2018).

2.2.4 Sumulaskennan edut ja hyödyntämismahdollisuudet

Shropshiren (2014) mukaan sumulaskenta on kehitetty tukemaan pilvilaskentaa ja sen heikkouksia. Pilvilaskennan resursseja isännöidään usein suurissa datakeskuksissa, jotka sijaitsevat tyypillisesti kaukana suurista kaupungeista. Tämä johtaa väistämättä suuriin viiveisiin, mikä onkin pääasiallinen syy sumulaskennan kehittämiselle (Shropshire, 2014). Kattepur, Dohare, Mushunuri, Rath ja Simha (2016) onnistuivatkin tutkimuksessaan laskemaan sumulaskennan avulla viiveitä 77% verrattuna pilvilaskentaan. Tutkimuksessa havaittiin, että sumulaskenta paransi myös energiatehokkuutta, mutta Dastjerdi ja Buyya (2016) huomauttavat artikkelissaan, ettei sumulaskenta ole välttämättä yhtä energiatehokas laskentaparadigma kuin pilvilaskenta lukuisten hajautettujen sumusolmujen vuoksi.

Chiang ja Zhang (2016) kuitenkin toteavat artikkelissaan, ettei matalat viiveet ole ainoa sumulaskennan tarjoama etu, vaan sen lisäksi keskeisimmät edut ovat kognitio, tehokkuus ja ketteryys. Kognitiolla tarkoitetaan sitä, että lähellä loppukäyttäjää olevat sumusolmut voidaan rakentaa siten, että ne ovat tietoisia käyttäjän tarpeista ja vaatimuksista. Tämän ansioista sumusolmut pystyvät määrittelemään, missä laskenta-, tallennus- ja muut tietojenkäsittelytoimenpiteet on paras suorittaa. Tämä näkyy myös tehokkuutena, sillä osa tietojenkäsittelyn toimenpiteistä voidaan suorittaa lähellä loppukäyttäjää. Tällöin kaikki käytettävissä olevat resurssit tulee hyödynnettyä, eikä kaikkea dataa tarvitse viedä pilvilaskennan keskitettyjen resurssien käsiteltäväksi. Sumulaskennan ketteryys mahdollistaa resurssien edullisen ja nopean käyttöönoton (Chiang & Zhang, 2016). Näiden etujen lisäksi Dastjerdi, Gupta, Calheiros, Ghosh ja Buyya (2016) mainitsevat sumulaskennan eduiksi myös verkkoliikenteen vähenemisen päätelaitteiden ja keskitettyjen pilviresurssien välillä sekä sumulaskennan soveltuvuuden IoT (*Internet of Things*) -sovelluksiin.

Iorgan ym. (2018) mukaan älykkäiden, toisiinsa liittettyjen laitteiden määrän oletetaan kasvavan 50 miljardiin vuoteen 2020 mennessä. Laitteiden määrän nopea kasvu johtuu mobiililaitteiden, kuten matkapuhelinten ja tablettien sekä älykkäiden sensorien ja anturien lisääntymisestä (Iorga ym., 2018). Vaquero ja Rodero-Merino (2014) kuitenkin ennustavat laitteiden määrän kasvavan paljon suuremmaksi älykkäiden sensorien ja anturien lisääntymisen vuoksi. Näin suuri määrä monimuotoisia ja älykkäitä laitteita tarjoaa hyvät puitteet sumulaskennan laajamittaiselle hyödyntämiselle.

Monet sumulaskennan hyödyntämiskohteet vaikuttavat kirjallisuuden perusteella liittyvän esineiden internetin (engl. *Internet of Things*) palveluihin. Datta, Bonnet ja Haerri (2015) kirjoittavat artikkelissaan tämän johtuvan siitä, että monet esineiden internetin palveluista edellyttävät tekijöitä, kuten reaaliaikaisia yhteyksiä, sijaintitietoisuutta, tukea liikkuvuudelle ja maantieteellistä hajautuneisuutta. Nämä tekijät on mahdollista tarjota sumulaskennan avulla, mikä tekee siitä ihanteellisen alustan esineiden internetin sovelluksille (Datta ym., 2015).

Kirjallisuudessa käytetään sumulaskennan käytännön tason esimerkkinä hyvin usein älykästä liikennevalojärjestelmää (engl. *Smart Traffic Light System*),

joka tuo hyvin esille sumulaskennan edut sekä kommunikoinnin sumu- ja pilviresurssien välillä. Bonomi ym. (2014) kirjoittavat artikkelissaan älykkään liikennevalojärjestelmän olevan sumulaskennan avulla toteutettavissa oleva järjestelmä, jonka toiminta perustuu sensoreihin, jotka mittaavat ajoneuvojen nopeuksia ja etäisyyksiä sekä tunnistavat jalankulkijoiden ja pyöräilijöiden sijainnin ja kulkusuunnan. Järjestelmä mukauttaa liikennevalojen toimintaa liikenteen kulun sujuvuuden lisäämiseksi sekä onnettomuuksien välttämiseksi. Järjestelmä voi myös lähettää verkkoon liitettyihin ajoneuvoihin varoituksia mahdollisesti lähestyvistä vaaratilanteista. Älykkään liikennevalojärjestelmän kolme päätarkoitusta ovat onnettomuuksien estäminen, tasaisen liikennevirran ylläpitäminen sekä datan kerääminen järjestelmän kehittämiseksi (Bonomi ym., 2014).

Älykäs liikennevalojärjestelmä täyttää useita sumulaskennan ominaispiirteitä. Järjestelmässä sensorit ovat maantieteellisesti hajautettuja, sillä niiden on sijaittava laajalla alueella pystyäkseen seuraamaan kokonaisvaltaisesti liikenteen kulkua, mikä edellyttää myös sumusolmujen kontekstuaalista sijaintitietoisuutta. Lisäksi järjestelmä edellyttää laitteiden välistä yhteentoimivuutta, sillä se muodostuu monimuotoisista laitteista, kuten sensoreista, liikennevaloista ja pilviresursseista. Jotta järjestelmän avulla voidaan välttää onnettomuuksia, siltä vaaditaan reaaliaikaista vuorovaikutusta sekä matalia viiveitä sumusolmujen välillä. Datan kerääminen ja pitkäaikainen säilyttäminen järjestelmän kehittämiseksi tuo puolestaan esille vuorovaikutuksen sumu- ja pilviresurssien välillä.

3 TIETOTURVAONGELMAT

Tässä luvussa tarkastellaan sekä pilvi- että sumulaskennan tietoturvaongelmia. Aluksi tutustutaan pilvilaskennan tietoturvaongelmiin yleisellä tasolla, jonka jälkeen tarkastellaan dataan kohdistuvia tietoturvaongelmia luotettamuksellisuuden, eheyden ja saatavuuden näkökulmista. Tämän jälkeen perehdytään pilvilaskennan eri palvelumallien turvallisuusongelmiin ja niiden välisiin yhteyksiin. Seuraavaksi siirrytään tarkastelemaan sumulaskennan tietoturvaongelmia verraten niitä pilvilaskennan tietoturvaongelmiin. Tämän jäljeen tarkastellaan sumulaskennan tietoturvaa sen pääasiallisen käyttökohteen, eli IoT-sovellusten näkökulmasta. Tämän luvun tarkoituksena on vastata tutkimuskysymykseen, eli kuinka sumulaskennan tietoturvaongelmat poikkeavat pilvilaskennan tietoturvaongelmista.

3.1 Tietoturvaongelmat pilvilaskennassa

Ali ym. (2015) toteavat artikkelissaan turvallisuuden olevan yksi suurimmista pilvilaskennan ongelmista. Pilvilaskennan tietoturvasta ongelmallisen tekee palveluiden ulkoistaminen, jolloin käyttäjä menettää datan fyysisen hallinnan ja on riippuvainen palvelun tarjoajasta myös turvallisuuden suhteen (Ali ym., 2015; Saharan & Kumar, 2015). Pilvilaskennan turvallisuutta onkin tutkittu paljon, ja sitä voidaan tarkastella useista eri näkökulmista (Chen & Zhao, 2012).

Pilviympäristöt kohtaavat perinteisestä IT-infrastruktuurista poikkeavia tietoturvaongelmia erityispiirteidensä ja käyttämiensä teknologioiden vuoksi, mutta myös perinteisen infrastruktuurin tietoturvaongelmat ovat läsnä pilvilaskennassa (Ali ym., 2015; Chen & Zhao, 2012; Rong, Nguyen & Jaatun, 2013). Tämän vuoksi Rong ym. (2013) jakavat pilvilaskennan kohtaamat tietoturvaongelmat tavanomaisiin tietoturvaongelmiin sekä pilvilaskennan tietoturvaongelmiin. Tavanomaisilla tietoturvaongelmilla tarkoitetaan ongelmia, jotka ovat yleisesti liittettävissä perinteiseen viestintäteknologiaan, ja pilvilaskennan tietoturvaongelmilla tarkoitetaan ongelmia, jotka aiheutuvat pilvilaskennan ominaispiirteistä (Rong ym., 2013). Aiheen rajauksen vuoksi tässä tutkielmassa keskitytään tarkastelemaan pilvilaskenta-arkkitehtuurin erityispiirteiden mukanaan tuomia tietoturvaongelmia, eikä tavanomaisiin tietoturvaongelmiin kiinnitetä huomiota.

Ryan (2013) kuitenkin huomauttaa artikkelissaan, että vaikka pilvilaskenta altistaa käyttäjän datan erilaisille tietoturvauhille, niin pilviympäristö voi osittain myös parantaa tietoturvallisuutta. Pilvipalvelun tarjoaja voi pystyä investoimaan ajantasaisempaan tietoturvateknologiaan ja takaamaan täsmällisempiä tietoturvakäytäntöjä kuin asiakas itse pystyisi (Ryan, 2013).

3.1.1 Dataan kohdistuvat ongelmat

Kun käyttäjä ulkoistaa datansa pilvipalvelun tarjoajalle, hän menettää datansa täyden kontrollin (Ali ym., 2015). Datan siirtäminen pois omistajansa hallinnasta tuo väistämättä mukanaan erinäisiä tietoturvaongelmia (Ali ym., 2015; Ryan, 2013). Pilviresursseihin tallennettu data on paljon alttiimpaa riskeille luottamuksellisuuden, eheyden ja saatavuuden suhteen, kuin paikalliseen tallennustilaan tallennettu data (Wang, Wang, Ren, Cao & Lou, 2012).

Datan luottamuksellisuuden tavoitteena on tarjota pääsy dataan vain valtuutetuille osapuolille (Zhou ym., 2010; Zissis & Lekkas, 2012). Luottamuksellisuus tarjotaan käyttäjälle erilaisten suojausmekanismien avulla, joilla estetään tietojen luovuttaminen luvattomille osapuolille (Farooq, Waseem, Khairi & Mazhar, 2015). Pilvilaskenta kuitenkin aiheuttaa merkittäviä uhkia datan luottamuksellisuudelle (Subashini & Kavitha, 2011).

Zissis ja Lekkas (2012) pitävät laitteiden, osapuolten ja käyttäjäoikeuksien määrän kasvua pilviympäristössä datan luottamuksellisuuden uhkana. Kun dataan pääsy on mahdollista useiden laitteiden ja identiteettien kautta, niin mahdollisen tietomurron riski kasvaa (Zissis & Lekkas, 2012). Lisäksi pilvilaskennassa käytetty moniasiakkuusmalli altistaa datan luottamuksellisuuden säilymisen (Chen & Zhao, 2012). Kun asiakkaiden prosesseja ajetaan fyysisesti samalla palvelimella ja käyttäjien resurssit erotellaan toisistaan virtuaalisesti, on haasteellista varmistaa, että vain valtuutetut osapuolet pääsevät käsiksi dataan (Ali ym., 2015; Rong ym., 2013; Takabi ym., 2010; Zissis & Lekkas, 2012). Laitteiden, muistin ja ohjelmistojen kontrollointi ja valvonta onkin käyttäjän tilin suojausta suurempi ongelma (Zissis & Lekkas, 2012).

Rong ym. (2013) huomauttavat, että datan luottamuksellisuutta voi myös uhata datan maantieteellinen sijainti, sillä eri alueiden lainsäädännöt voivat poiketa huomattavasti toisistaan. Esimerkiksi joissakin maissa lainsäädäntö antaa valtiolle pääsyn dataan melko kevein perustein. Tähän on kuitenkin kiinnitetty huomiota esimerkiksi Euroopan Unionin toimesta, joka on säätänyt direktiivin, mikä estää tietojen siirtämisen maahan, jonka lainsäädäntö ei tarjoa riittävää suojaa datan yksityisyydelle. Loppukäyttäjän kannattaa kuitenkin ottaa huomioon datan mahdollinen sijainti palveluntarjoajaa valitessaan (Rong ym., 2013).

Zissis ja Lekkas (2012) määrittelevät datan eheyden tarkoittavan sitä, että vain valtuutetut osapuolet voivat käsitellä dataa, sovelluksia ja laitteistoa. Se suojaa dataa ulkopuolisilta muutoksilta ja väärinkäytöltä (Zhou ym., 2010; Zissis & Lekkas, 2012). Luottamuksellisuuden ohella eheys on yksi keskeisimpiä datan turvallisuuden tekijöitä (Chen & Zhao, 2012; Subashini & Kavitha, 2011).

Datan eheys voi vaarantua useista eri syistä, kuten esimerkiksi ohjelmisto- tai laitteistovirheistä (Zissis & Lekkas, 2012). Rong ym. (2013) kuitenkin pitävät datan eheyden vaarantumisen pääasiallisena syynä datan ja sen hallinnan ulkoistamista. Kun data tallennetaan kolmannen osapuolen infrastruktuuriin, niin datan omistaja menettää suuren osan datansa kontrollista. Tällöin on

mahdollista, että dataan voidaan tehdä muutoksia datan omistajan tietämättä (Rong ym., 2013). Zissisin ja Lekkasin (2012) mukaan datan eheyden säilymistä voidaan kuitenkin edesauttaa estämällä ulkopuolinen pääsy dataan esimerkiksi indentiteettien todentamismenetelmien ja kulunvalvonnan avulla. Nämä keinot voivat myös edesauttaa löytämään datan eheyden vaarantaneen osapuolen, mikäli data joutuu suojauskeinoista huolimatta väärinkäytön kohteeksi (Zissis & Lekkas, 2012).

Datan saatavuus takaa valtuutetulle osapuolelle jatkuvan, ajasta ja paikasta riippumattoman pääsyn tietojenkäsittelyresursseihin sekä dataan (Zhou ym., 2010; Zissis & Lekkas, 2012). Pilvilaskennassa tavoitteena on ylläpitää datan saatavuutta olosuhteista ja mahdollisista ongelmatilanteista, kuten tietomurrosta riippumatta (Farooq ym., 2015; Zissis & Lekkas, 2012). Tietomurron tapauksessa on kuitenkin haastavaa eristää tietty fyysinen resurssi, sillä dynaamisesti skaalautuvilla resursseilla ei ole kiinteitä suojarajoja (Chen & Zhao, 2012).

Pilvilaskennassa datan saatavuus on täysin palveluntarjoajan vastuulla (Chen & Zhao, 2012). Ongelmat palvelun saatavuudessa koskevat samanaikaisesti suuria käyttäjämääriä (Rong ym., 2013). Armbrustin ym. (2010) mukaan saatavuus voi olla uhattuna useista eri syistä. Laitteiston ongelmien ja ulkoisten hyökkäyksien lisäksi saatavuus voi vaarantua puutteellisten varmuuskopioiden ja varotoimenpiteiden vuoksi (Armbrust ym., 2010; Chen & Zhao, 2012). Lisäksi on mahdollista, että palveluntarjoaja lopettaa toimintansa kokonaan (Armbrust ym., 2010; Fox ym., 2009). Luottamuksellisuudesta ja eheydestä poiketen saatavuus on tietoturvatekijä, joka usein määritellään palvelutasosopimuksessa (engl. *Service Level Agreement*) (Dillon, Wu & Chang, 2010). Palvelutasosopimus ei ole kuitenkaan keino saatavuuden takaamiseksi, vaan palveluntarjoajan ja asiakkaan välinen sopimus, jossa määritellään palvelun vaadittu laatu (Wu & Buyya, 2012).

3.1.2 Palvelumallien turvallisuusongelmat

Kirjallisuudessa hyvin yleinen tapa tarkastella pilvilaskennan turvallisuutta on eri palvelumallien näkökulmat. Ali ym. (2015) kirjoittavat artikkelissaan, että palvelumallit ovat rakennettu kerroksittain, jolloin ne ovat riippuvaisia taustalla olevasta palvelumallista. IaaS on alimman tason palvelumalli, jonka päälle PaaS on rakennettu, joka taas tarjoaa alustan SaaS-palvelumallille (Subashini & Kavitha, 2011). Palvelumallien välinen riippuvuus koskee myös turvallisuusnäkökulmaa, sillä esimerkiksi SaaS on osin riippuvainen alempien palvelumallien turvallisuudesta (Ali ym., 2015). Tässä tutkielmassa eri palvelumallien turvallisuusongelmia tarkastellaan kerroksittain alkaen alimmasta kerroksesta.

IaaS-palvelumallissa käyttäjät pystyvät itse kontrolloimaan turvallisuutta laajemmin, kuin muissa palvelumalleissa (Subashini & Kavitha, 2011). Käyttäjät ovat itse vastuussa käyttämiensä järjestelmien turvallisuudesta, joita he ajavat palveluntarjoajan tarjoaman alustan päällä (Jaeger & Schiffman, 2010). Alustan taustalla olevat laskenta-, verkko- ja varastointi-infrastruktuurit ovat kuitenkin

palveluntarjoajan hallinnoimia (Hashizume, Rosado, Fernández-Medina & Fernandez, 2013). Näin ollen palveluntarjoaja on vastuussa tarjoamansa alustan turvallisuudesta (Dawoud, Takouna & Meinel, 2010).

Virtualisointi on teknologia, joka mahdollistaa skaalautuvien resurssien tarjoamisen pilviympäristössä (Garfinkel & Rosenblum, 2005; Lombardi & Di Pietro, 2011). Se on myös pääasiallinen tietoturvariskien aiheuttaja IaaS-palvelumallissa (Dawoud ym., 2010). Reubenin (2007) mukaan virtualisoidut ympäristöt ovat alttiita samankaltaisille hyökkäyksille, kuin tavanomaiset IT-infrastruktuurit. Virtualisoiduissa ympäristöissä tietoturvauhkien välttäminen on kuitenkin huomattavasti haasteellisempaa johtuen virtualisoinnin monimutkaisuudesta ja yhteyspisteiden suuresta määrästä (Reuben, 2007).

Dawoud ym. (2010) toteavat, että virtualisoinnin näkökulmasta suurin riski on isäntätietokoneen vaarantuminen, sillä virtuaalikoneita hallinnoidaan isäntätietokoneen kautta. Mahdollinen hyökkääjä voi käyttää isäntätietokoneen ominaisuuksia väärin sekä seurata virtuaalikoneiden verkkoliikennettä, joka kulkee isäntätietokoneiden kautta (Dawoud ym., 2010). Virtuaalikoneiden turvallisuutta voidaan kuitenkin parantaa suojaamalla ne hypervisorilla. Hypervisor, eli Virtual Machine Monitor (VMM) on matalan tason ohjelmisto, joka vastaa virtuaalikoneiden eristämisestä (Dawoud ym., 2010; Hashizume ym., 2013; Takabi ym., 2010). Hypervisor sijaitsee virtuaalikoneiden ja isäntätietokoneen välissä ja se suojaa virtuaalikoneita ulkopuolisilta tietorvariskeiltä (Reuben, 2007; Subashini & Kavitha, 2011). IaaS-palvelumallia voidaan pitää kokonaisuudessaan melko turvallisena palvelumallina, mikäli käytetyssä virtualisointisovelluksessa ei ole tietoturva-aukkoja (Subashini & Kavitha, 2011).

Mather, Kumaraswamy ja Latif (2009) kirjoittavat PaaS-palvelumallin tarjoavan ympäristön, jossa voidaan suunnitella, kehittää, testata, ottaa käyttöön ja tukea mukautettuja sovelluksia, jotka on kehitetty alustan tukemalla ohjelmointikielellä. Heidän mukaansa PaaS-palvelumallin turvallisuus koostuu kahdesta kokonaisuudesta, jotka ovat itse PaaS-alustan turvallisuus sekä käyttäjän ohjelmistojen turvallisuus. Yleisesti ottaen PaaS-palveluntarjoajat ovat vastuussa alustan ohjelmistojen turvallisuudesta, jotka mahdollistavat asiakkaan omien sovellusten ajamisen (Mather ym., 2009). Asiakkaan on taas oltava itse tietoinen omien sovellustensa turvallisuudesta (Mather ym., 2009; Subashini & Kavitha, 2011).

PaaS-palvelumalli mahdollistaa myös kolmannen osapuolen sovellusten, verkkotapveluiden ja komponenttien käytön, mikä tuo mukanaan uusia tietoturvariskejä (Hashizume ym., 2013; Mather ym., 2009). Asiakkaiden tulisikin tästä syystä ymmärtää sovelluksien ja niiden turvallisuuden riippuvuus kaikista käytetyistä palveluista sekä osata arvioida kolmannen osapuolen palveluntarjoajia koskevia riskejä (Mather ym., 2009).

Hashizume ym. (Hashizume ym., 2013) toteavat, että myös PaaS-palvelumallissa palveluntarjoaja on vastuussa ohjelmistojen alla olevan infrastruktuurin turvallisuudesta. Vaikka palvelun käyttäjät hallitsevat omien sovellustensa turvallisuutta, heillä ei ole täyttä varmuutta palveluntarjoajan

tarjoaman ympäristön turvallisuudesta. PaaS-käyttöönottomallin turvallisuutta ei ole tutkittu yhtä paljoa, kuin muiden käyttöönottomallien osalta, mutta siinä on selkeitä yhtymäkohtia IaaS- että SaaS-palvelumallien turvallisuuteen (Hashizume ym., 2013).

SaaS-palvelumalli tarjoaa sovelluspalveluita, kuten sähköposti-, toiminnanohjaus- ja asiakkuudenhallintajärjestelmiä (Ju, Wang, Fu, Wu & Lin, 2010). SaaS-käyttäjillä on vähäisemmät mahdollisuudet vaikuttaa pilvipalvelun turvallisuuteen, kuin muiden palvelumallien käyttäjillä (Hashizume ym., 2013). SaaS-palvelumallissa käyttäjä onkin täysin riippuvainen palveluntarjoajan tietoturvatoimenpiteistä (Subashini & Kavitha, 2011). Kirjallisuuden perusteella vaikuttaa siltä, että suuri osa SaaS-mallin tietoturvaongelmista liittyvät verkkoon ja ovat hyvin samankaltaisia perinteisen viestintäteknologian ongelmien kanssa. Pilvilaskenta-arkkitehtuurin ominaispiirteet tekevät ongelmien ratkaisusta kuitenkin ongelmallisempaa.

Subashinin ja Kavithan (2011) mukaan SaaS-palveluntarjoaja voi isännöidä sovelluksia omalla yksityisellä palvelimellaan tai ulkoistaa sovelluksen käyttöönoton kolmannen osapuolen palveluntarjoajan infrastruktuuriin. Vaikka tämä voi monessa tapauksessa parantaa SaaS-palvelun tehokkuutta, niin se monimutkaistaa turvallisuuden hallintaa entisestään. Tämän seurauksena myös asiakkaan on hyvin vaikea selvittää, onko SaaS-palvelun turvallisuusseikat otettu huomioon (Subashini & Kavitha, 2011). Sen lisäksi myös palveluntarjoajan voi olla tällaisessa tilanteessa vaikeaa tai jopa mahdotonta osoittaa missä dataa maantieteellisesti tarkalleen säilytetään (Zhou ym., 2010). Tämä korostaa väistämättä datan maantieteelliseen sijaintiin liittyviä ongelmia ja uhkaa datan luottamuksellisuutta.

Subashini ja Kavitha (2011) pitävät web-sovelluksien haavoittuvuutta yhtenä keskeisimmistä SaaS-mallin tietoturvariskeistä. Verizon Business on selvittänyt tutkimuksessaan, että kaikista pilvilaskennan järjestelmiä, sovelluksia ja palveluita koskevista hyökkäyksistä 39% kohdistuu ohjelmistokerrokseen. SaaS-palvelumallin osalta ongelma on samanlainen, kuin tavanomaistenkin viestintäteknologioiden osalta. SaaS-sovellukset eivät myöskään juuri poikkea muista web-sovelluksista (Subashini & Kavitha, 2011). Pilviarkkitehtuurin ominaispiirteiden vuoksi perinteiset suojausmenetelmät eivät kuitenkaan ole SaaS-mallin kohdalla riittäviä, sillä SaaS-sovellusten haavoittuminen voi osoittautua huomattavasti tuhoisemmasi kuin perinteisten web-sovellusten haavoittuminen (Ali ym., 2015; Subashini & Kavitha, 2011). Zissis ja Lekkas (2012) pitävät sovellusten luottamuksellisuutta yhtä tärkeänä kuin datan luottamuksellisuutta järjestelmän yleisen turvallisuuden kannalta.

3.2 Tietoturvaongelmat sumulaskennassa

Sumulaskennan turvallisuutta ei ole vielä tutkittu kovinkaan paljoa, mutta olemassa olevasta kirjallisuudesta nousee esille muutamia keskeisiä tekijöitä sumulaskennan turvallisuuteen liittyen. Sumulaskenta on pilvilaskentaa

täydentävä paradigma, joten pilvilaskennan tietoturvaongelmat koskevat myös sumulaskentaa (Alrawais, Alhothaily, Hu & Cheng, 2017; Yi, Qin & Li, 2015). Tämän lisäksi sumulaskenta tuo erityispiirteidensä, vuoksi mukanaan myös uusia tietoturvahaasteista (Alrawais ym., 2017; Chiang & Zhang, 2016). Tällaisia erityispiirteitä ovat esimerkiksi sumusolmujen heterogeenisyys, sijaintitietoisuus ja suuri määrä, hajautetut resurssit sekä liikkuvuustukea koskevat vaatimukset (Yi ym., 2015).

Chiang ja Zhang (2016) kirjoittavat artikkelissaan sumulaskennan keskeisimmän ominaispiirteen turvallisuuden kannalta olevan resurssien hajauttaminen, sillä hajautetut järjestelmät ovat yleisesti ottaen haavoittuvaisempia, kuin keskitetyt järjestelmät. Pilvilaskenta toimii keskitetysti vahvasti suojatussa ympäristössä, kun taas sumulaskennassa resurssit ovat hajautettuja ja sijaitsevat haavoittuvaisemmissa olosuhteissa (Chiang & Zhang, 2016; Stojmenovic ym., 2016). Lisäksi sumulaskennan hajautetut sumusolmut ovat hyvin pieniä verrattuna pilviresursseihin, joten sumusolmuilla ei ole käytettävissä yhtä paljoa kapasiteettia tietoturvauhkilta suojautumiseen, kuin pilviresursseilla (Chiang & Zhang, 2016). Monet pilvilaskennan tietoturvaratkaisuista eivät sovellu sumulaskentaan, ja tämän seurauksena sumulaskenta kohtaa uhkia, joita ei hyvin hallitussa pilviympäristössä ole (Stojmenovic ym., 2016). Uusista tietoturvariskeistä huolimatta sumulaskennan avulla pystytään joissain tapauksissa myös parantamaan tietoturvallisuutta (Alrawais ym., 2017). Seuraavaksi vertaillaan sumulaskennan tietoturvaongelmia pilvilaskennan tietoturvaongelmiin eri näkökulmista katsottuna. Tämän jälkeen tarkastellaan sumulaskennan tietoturvamahdollisuuksia IoT-sovelluksien alustana.

3.2.1 Sumusolmujen tietoturvahaasteet ja datan turvallisuus

Yi ym. (2015) nostavat artikkelissaan esille sumuresurssien luottamuksellisuutta koskevan ongelman. Pilvilaskennassa tietojenkäsittelyresurssit omistaa pääsääntöisesti palveluntarjoaja, kun taas sumulaskennassa resurssien omistaja voi olla palveluntarjoajan lisäksi loppukäyttäjä tai kolmas osapuoli. Sumulaskennassa datan ulkoistaminen palveluntarjoajan tai kolmannen osapuolen omistamille ja hallinnoimille sumusolmuille luo samankaltaisia turvallisuusriskejä datan eheydelle, kuin pilvilaskennassakin (Yi ym., 2015). Näiden riskien lisäksi datan ulkoistaminen eri toimijoiden hallinnoimille sumusolmuille tekee sumulaskennan palveluista potentiaalisia kohteita perinteisille hyökkäystyypeille, mutta ennen kaikkea man-in-the-middle -hyökkäyksille (Stojmenovic ym., 2016; Yi ym., 2015). Tämän tyyppisessä hyökkäyksessä hyökkääjä voi houkutella uhrejaan yhdistämään laitteensa epäluotettavaan sumusolmuun (engl. *rogue fognode*) (Yi ym., 2015). Kun uhrin laite on kerran yhdistetty epäluotettavaan sumusolmuun, niin hyökkääjän on mahdollista seurata uhrin verkkoliikennettä, manipuloida lähteviä ja tulevia pyyntöjä sekä laukaista uusia hyökkäyksiä (Stojmenovic ym., 2016; Yi ym., 2015). Yi ym. (2015) toteavat epäluotettavien sumusolmujen olemassaolon olevan suuri

uhka käyttäjän datan turvallisuudelle. Sumuympäristössä tällaisten sumusolmujen havaitseminen on hyvin hankalaa (Stojmenovic ym., 2016).

Toinen perinteinen hyökkäystyyppi, jolle sumusolmut ovat alttiita on palvelunesto- eli DOS-hyökkäys, joka tulee sanoista Denial of Service (Shropshire, 2014). Palvelunestohyökkäykset aiheuttavat suuren uhan sumulaskennan palveluiden saatavuudelle, sillä sumusolmujen laskentakapasiteetti on hyvin rajoitettu pilviresursseihin verrattuna, jolloin suhteellisen pieni määrä häirintää voi estää sumusolmun toiminnan (Hong, Lillethun, Ramachandran, Ottenwälder & Koldehofe, 2013). Shropshiren (2014) mukaan yksittäisen sumusolmun hetkellinen toimimattomuus ei vaaranna sumulaskennan palveluita, mutta jos suuri määrä sumusolmuja on poissa käytössä tietyltä alueelta, niin palveluiden saatavuus on vaarannettu. Tietyille alueille kohdistetut palvelunestohyökkäykset ovat mahdollisia sumusolmujen kontekstuaalisen sijaintitietoisuuden vuoksi (Shropshire, 2014).

Yi ym. (2015) tuovat artikkelissaan esille sen, että sumusolmujen sijaintitietoisuus asettaa datan lisäksi myös käyttäjän yksityisyyden uhatukasi. Mukana kannettavat älylaitteet voivat mahdollistaa käyttäjän sijainnin paljastumisen. Jos tällainen laite on yhdistetty sumusolmuun, niin sumusolmun maantieteellisen sijainnin perusteella on mahdollista päätellä myös käyttäjän sijainti melko tarkasti, mikäli käyttäjä voidaan idenfidioida laitteen perusteella. Myös esimerkiksi kodin älykkäät IoT-laitteet voivat kerätä dataa, jonka perusteella on mahdollista päätellä onko laitteiden omistaja kotona vai ei. Tämän kaltaiset uhat voivat yksityisyyden lisäksi pahimmillaan vaarantaa käyttäjän fyysisen turvallisuuden (Yi ym., 2015).

3.2.2 Sumulaskennan tietoturva IoT-sovelluksissa

Alrawais ym. (2017) toteavat artikkelissaan sumulaskennan olevan ominaispiirteidensä vuoksi oivallinen alusta IoT-sovelluksille. Sumulaskennan yksi keskeisimmistä käyttökohteista on IoT-sovellukset, ja sen avulla on mahdollista tukea näiden sovellusten turvallisuustavoitteita (Alrawais ym., 2017; Chiang & Zhang, 2016).

Alrawais ym. (2017) pitävät monia IoT-laitteita hyvin haavoittuvaisina. Tämä johtuu siitä, että kaikilla laitteilla, kuten esimerkiksi antureilla ei ole tarvittavia resurssija tietoturvan ylläpitämiseen. Sumusolmut voivat kuitenkin tarjota IoT-sovelluksille sekä tietojenkäsittelyresursseja että tietoturvallisuutta, sillä sumusolmujen resurssien avulla voidaan tarjota kryptografisia palveluita (Alrawais ym., 2017; Chiang & Zhang, 2016). Alrawais ym. (2017) huomauttavat, ettei tämä kuitenkaan ole aivan ongelmatonta, sillä tietoturvapalveluiden tarjoaminen voi viedä suuren osan joidenkin sumusolmujen laskentakapasiteetista, mikä taas näkyy kasvavina viiveinä. Kryptografisten palveluiden lisäksi sumulaskennan avulla on mahdollista valvoa IoT-laitteiden tietoturvapäivitysten ajantasaisuutta (Alrawais ym., 2017).

Sumulaskenta voi helpottaa erilaisten hyökkäysten havaitsemista, sillä sen avulla on mahdollista seurata, esiintyykö verkossa epätavallisia ja normaalista

poikkeavia toimia (Stojmenovic ym., 2016). Koska sumulaskenta laajentaa pilvilaskennan palvelut verkon reunalle, niin pilvilaskentaa varten kehitetyt havaitsemisjärjestelmät on mahdollista ottaa käyttöön myös sumuympäristössä (Alrawais ym., 2017). Lisäksi sumusolmut voivat pitää huolen IoT-sovellusten kulunvalvonasta (Chiang & Zhang, 2016). Kulunvalvonta on tehokas työkalu suojaamaan järjestelmää luvattomilta toimilta (Yi ym., 2015).

Alrawais ym. (2017) toteavat artikkelissaan sumulasennan parantavan osaltaan tietoturvaa minimoimalla henkilökohtaisen datan siirtämisen pilviresursseihin. Tällöin datan ei tarvitse kulkea useiden sumusolmujen ja verkon lävitse, ja se pysyy koko elinkaarensa ajan lähellä loppukäyttäjää ja laitteita, jotka käyttävät tätä dataa (Alrawais ym., 2017). Yi ym. (2015) näkevät tämän kuitenkin eri tavalla. Heidän mukaansa on hyvin riskialtista säilyttää henkilökohtaista dataa haavoittuvassa ympäristössä (Yi ym., 2015). Chiang ja Zhang (2016) pitävät parhaana ratkaisuna henkilökohtaisen datan siirtämistä pilveen siten, että sumusolmut suojaavat datan jollakin salausmenetelmällä ennen siirtoa.

4 YHTEENVETO

Tässä tutkielmassa tutkittiin pilvi- ja sumulaskentaparadigmoja sekä niiden tietoturvaongelmia. Tutkielma suoritettiin kirjallisuuskatsauksena, ja sen tarkoituksena oli selvittää kuinka sumulaskennan tietoturvaongelmat poikkeavat pilvilaskennan tietoturvaongelmista. Tutkielman näkökulmaksi valikoitui tietoturva, sillä se on tärkein tekijä uusien teknologioiden menestymisen kannalta.

Tutkielman ensimmäisessä sisältöluvussa tutustuttiin sekä pilvi- että sumulaskentaan. Ensimmäisenä tutustuttiin pilvilaskentaan, joka on laskentaparadigma, jossa käyttäjälle tarjotaan tietojenkäsittelyresursseja verkon välityksellä ajasta ja paikasta riippumatta. Pilvilaskennalle ominaista on se, että dynaamisesti skaalautuvat ja nopeasti käyttöönotettavat tietojenkäsittelyresurssit sijaitsevat keskitetysti datakeskuksissa. Mell ja Grance (2011) ovat määritelleet pilvilaskennan kirjallisuudessa yleisesti hyväksytyllä tavalla. Tämän määritelmän mukaan pilvilaskenta muodostuu viidestä ominaispiirteestä, kolmesta palvelumallista ja neljästä käyttöönottomallista.

Seuraavaksi tutustuttiin sumulaskentaan, joka on pilvilaskentaa täydentävä laskentaparadigma. Sen tarkoituksena on tukea pilvilaskentaa ja laajentaa sen palveluita lähemmäs loppukäyttäjää ja verkon reunaa. Sumulaskenta on kehitetty tukemaan pilvilaskennan puutteita, joista keskeisin on ydinverkkojen kuormituksen myötä kasvavat viiveet. Pilviresurssien lisäksi sumulaskenta hyödyntää maantieteellisesti hajautettuja sumusolmuja, jotka sijaitsevat pilviresurssien ja verkon reunan välillä. Sumusolmut ovat heterogeenisiä ja ne voivat olla joko virtuaalisia tai fyysisiä laitteita. Lisäksi sumusolmut ovat tietoisia sekä maantieteellisestä että loogisesta sijainnistaan. Iorga ym. (2018) ovat tunnistaneet sumulaskennalle kuusi ominaispiirrettä, jotka ovat kontekstuaalinen sijaintitietoisuus ja matala viive, maantieteellinen jakautuminen, heterogeenisyys, yhteentoimivuus ja yhdistäminen, reaaliaikainen vuorovaikutus sekä yhdistettujen sumusolmu-klustereiden skaalautuvuus ja ketteryys.

Toisen sisältöluvun aluksi tutustuttiin pilvilaskennan tietoturvaongelmiin. Keskeistä pilvilaskennan tietoturvassa on datan ja palveluiden ulkoistaminen, jolloin käyttäjä menettää datan fyysisen hallinnan ja on riippuvainen palveluntarjoajasta myös tietoturvan osalta. Ulkoistaminen altistaa datan luottamuksellisuuden, eheyden ja saatavuuden. Erityisen suuria haasteita aiheuttavat datan maantieteellinen sijainti, moniasiakkuusmalli ja palvelunestohyökkäykset. Tämän lisäksi pilvilaskennan eri palvelumallit kohtaavat erilaisia tietoturvaongelmia. IaaS-palvelumallilla tarjotaan käyttäjällä vain infrastruktuuria, joten palvelumallin tietoturvaongelmat koskevat virtualisointia. Muilta osin käyttäjä kontrolloi itse käyttämiensä järjestelmien tietoturvallisuutta. PaaS-palvelumallin turvallisuus koostuu sekä PaaS-alustan turvallisuudesta että käyttäjän ohjelmistojen turvallisuudesta. Palveluntarjoajan vastuulla on alustan tietoturvasta huolehtiminen ja käyttäjä huolehtii omien

sovellustensa turvallisuudesta. PaaS- malli perii myös IaaS-mallin tietoturvaongelmat, sillä PaaS rakentuu IaaS-mallin päälle. Sovelluspalveluita tarjoavassa SaaS-mallissa käyttäjä on täysin riippuvainen palveluntarjoajan tietoturvaratkaisuista. SaaS-malli kohtaa samankaltaisia turvallisuusuhkia, kuin perinteiset viestintäteknologiatkin. Suurimmat SaaS-mallin turvallisuusongelmat liittyvät datan ulkoistamiseen sekä sovelluskerrokseen kohdistuviin hyökkäyksiin. Tämän lisäksi SaaS perii alempien palvelumallien turvallisuusongelmat.

Tämän jälkeen tutustuttiin sumulaskennan tietoturvaongelmiin. Kirjallisuus osoitti, että sumulaskenta sisältää kaikki pilvilaskennan olemassa olevat tietoturvaongelmat, sillä se on pilvilaskentaa täydentävä paradigma. Tämän lisäksi sumulaskenta tuo erityispiirteidensä mukana useita kokonaan uusia tietoturvaongelmia. Hajautetut järjestelmät ovat yleisesti haavoittuvaisempia, kuin keskitetyt järjestelmät, ja maantieteellisesti hajautetut sumusolmut toimivatkin pääsääntöisesti heikosti suojatuissa olosuhteissa, kun taas pilvilaskennan resurssit ovat keskitetty tiukasti suojattuihin datakeskuksiin. Sumulaskenta on pilvilaskentaa alttiimpi erilaisille hyökkäystyypeille, kuten DOS- ja man-in-the-middle -hyökkäyksille. Syy tähän on yksittäisten sumusolmujen resurssien vähyys tietoturvariskeiltä suojautumiseen sekä sumusolmujen useat eri omistajat. Sumulaskenta voi altistaa datan lisäksi myös käyttäjän yksityisyyden. Sumusolmujen sijaintitietoisuus voi paljastaa käyttäjän sijainnin idenfidioidun älylaitteen kautta. Lisäksi kodin IoT-laitteet voivat kerätä dataa, jonka perusteella on esimerkiksi mahdollista päätellä onko käyttäjä kotona. Pahimmillaan tämä voi vaarantaa myös käyttäjän fyysisen turvallisuuden.

Toisen sisältöluvun tarkoituksena oli vastata tutkimuskysymykseen kuinka sumulaskennan tietoturvaongelmat poikkeavat pilvilaskennan tietoturvaongelmista. Näiden laskentaparadigmojen tietoturvaongelmien vertailu osoitti, että sumulaskenta on pilvilaskentaan verrattuna tietoturvaltaan melko haavoittuvainen laskentaparadigma, vaikka sen avulla onkin mahdollista tarjota IoT-sovelluksille tietoturvapalveluita. Tutkielma osoitti myös, että sumulaskennan tietoturvaongelmien ratkaisu on haastavaa, sillä monet pilvilaskennan tietoturvaratkaisut eivät ole sellaisinaan sovellettavissa sumulaskennan palveluihin.

Kirjallisuudessa on tutkittu hyvin kattavasti pilvilaskennan tietoturvaa sekä keinoja tietoturvallisuuden parantamiseksi. Sumulaskennan osalta näin ei kuitenkaan ole. Sumulaskenta on teknologiana aivan elinkaarensa alkupäässä ja suuri osa sitä käsittelevästä kirjallisuudesta pohjautuu muutamaan keskeisimpään artikkeliin. Näin ollen kirjallisuuden tarjoama kuva sumulaskennasta on melko suppea ja vain osa artikkeleista tarjoaa uutta informaatiota sumulaskentaan liittyen. Sumulaskennan tietoturvaongelmista on kuitenkin jonkin verran tutkimusmateriaalia, mutta näiden ongelmien ratkaisumahdollisuuksia ei juurikaan ole tutkittu. Suurin syy tähän on todennäköisesti se, että sumulaskenta on vielä hyvin tuore laskentaparadigma.

Suurin osa sumulaskennan tutkimuksesta painottuu vuosille 2014-2016, minkä jälkeen aiheen tutkiminen on selkeästi vähentynyt. Syitä tälle ilmiölle voi

olla useita, mutta yksi mahdollinen syy voi olla juurikin sumulaskennan runsas tietoturvaongelmien määrä, mikä on mahdollisesti laskenut mielenkiintoa aiheen tutkimista kohtaan. Tämän vuoksi jatkossa olisikin tärkeää tutkia sumulaskennan tarjoamien mahdollisuuksien sijasta keinoja sumulaskennan tietoturvaongelmien ratkaisemiseksi. Tutkielman suurimmaksi ongelmaksi osoittautui liian laaja tutkimuskysymys. Aihe olisi voitu rajata koskemaan vain pilvi- tai sumulaskennan tietoturvaongelmia. Toisaalta sumulaskennan tietoturvaongelmia ei olisi ollut mielekästä käsitellä tutustumatta ensin pilvilaskentaan.

LÄHTEET

Aazam, M. & Huh, E. (2014). Fog computing and smart gateway based communication for cloud of things. Teoksessa *Future Internet of Things and Cloud (FiCloud), 2014 International Conference* (464-470) Suwon: Kyung Hae University.

Ali, M., Khan, S. U. & Vasilakos, A. V. (2015). Security in cloud computing: Opportunities and challenges. *Information Sciences, 305*, 357-383.

Alrawais, A., Alhothaily, A., Hu, C. & Cheng, X. (2017). Fog computing for the internet of things: Security and privacy issues. *IEEE Internet Computing, 21*(2), 34-42.

Armbrust, M., Fox, A., Griffith, R., Joseph, A. D., Katz, R., Konwinski, A., . . . Stoica, I. (2010). A view of cloud computing. *Communications of the ACM, 53*(4), 50-58.

Bonomi, F., Milito, R., Natarajan, P. & Zhu, J. (2014). Fog computing: A platform for internet of things and analytics. Teoksessa *Big data and internet of things: A roadmap for smart environments* (169-186) Springer: Cham.

Bonomi, F., Milito, R., Zhu, J. & Addepalli, S. (2012). Fog computing and its role in the internet of things. Teoksessa *Proceedings of the first edition of the MCC workshop on Mobile cloud computing* (13-16). ACM

Chen, D. & Zhao, H. (2012). Data security and privacy protection issues in cloud computing. Teoksessa *Computer Science and Electronics Engineering (ICCSEE), 2012 International Conference* (647-651) Sheyang: Northeastern University

Chiang, M. & Zhang, T. (2016). Fog and IoT: An overview of research opportunities. *IEEE Internet of Things Journal, 3*(6), 854-864.

Cisco. (2015). Fog computing and the internet of things: Extend the cloud to where the things are. Haettu osoitteesta https://www.cisco.com/c/dam/en_us/solutions/trends/iot/docs/computing-overview.pdf

Dastjerdi, A. V. & Buyya, R. (2016). Fog computing: Helping the internet of things realize its potential. *Computer, 49*(8), 112-116.

Dastjerdi, A. V., Gupta, H., Calheiros, R. N., Ghosh, S. K. & Buyya, R. (2016). Fog computing: Principles, architectures, and applications. *Internet of things* (61-75) Elsevier.

Datta, S. K., Bonnet, C. & Haerri, J. (2015). Fog computing architecture to enable consumer centric internet of things services. Teoksessa *Consumer Electronics (ISCE), 2015 IEEE International Symposium* (1-2) IEEE.

Dawoud, W., Takouna, I. & Meinel, C. (2010). Infrastructure as a service security: Challenges and solutions. Teoksessa *Informatics and Systems (INFOS), 2010 the 7th International Conference* (1-8) IEEE.

Dillon, T., Wu, C. & Chang, E. (2010). Cloud computing: Issues and challenges. Teoksessa *Advanced Information Networking and Applications (AINA), 2010 24th IEEE International Conference* (27-33) IEEE.

Farooq, M. U., Waseem, M., Khairi, A. & Mazhar, S. (2015). A critical analysis on the security concerns of internet of things (IoT). *International Journal of Computer Applications, 111*(7)

Foster, I., Zhao, Y., Raicu, I. & Lu, S. (2008). Cloud computing and grid computing 360-degree compared. *Teoksessa Grid Computing Environments Workshop, 2008. GCE'08* (1-10) IEEE.

Fox, A., Griffith, R., Joseph, A., Katz, R., Konwinski, A., Lee, G., . . . Stoica, I. (2009). Above the clouds: A berkeley view of cloud computing. *Dept.Electrical Eng.and Comput.Sciences, University of California, Berkeley, Rep.UCB/EECS, 28*(13), 2009.

Garfinkel, T. & Rosenblum, M. (2005). When virtual is harder than real: Security challenges in virtual machine based computing environments.

Hashizume, K., Rosado, D. G., Fernández-Medina, E. & Fernandez, E. B. (2013). An analysis of security issues for cloud computing. *Journal of Internet Services and Applications, 4*(1), 5.

Hong, K., Lillethun, D., Ramachandran, U., Ottenwälder, B. & Koldehofe, B. (2013). Mobile fog: A programming model for large-scale applications on the internet of things. Teoksessa *Proceedings of the second ACM SIGCOMM workshop on Mobile cloud computing* (15-20) ACM.

Iorga, M., Feldman, L., Barton, R., Martin, M. J., Goren, N. S. & Mahmoudi, C. (2018). Fog Computing Conceptual Model. *NIST Special Publication,* (NIST SP)-500-325).

Jaeger, T. & Schiffman, J. (2010). Outlook: Cloudy with a chance of security challenges and improvements. *IEEE Security & Privacy, 8*(1)

Ju, J., Wang, Y., Fu, J., Wu, J. & Lin, Z. (2010). Research on key technology in SaaS. Teoksessa *Intelligent Computing and Cognitive Informatics (ICICCI), 2010 International Conference* (384-387) IEEE.

Kattepur, A., Dohare, H., Mushunuri, V., Rath, H. K. & Simha, A. (2016). Resource constrained offloading in fog computing. Teoksessa *Proceedings of the 1st Workshop on Middleware for Edge Clouds & Cloudlets* (1) ACM.

Liu, F., Tong, J., Mao, J., Bohn, R., Messina, J., Badger, L. & Leaf, D. (2011). NIST cloud computing reference architecture. *NIST Special Publication, 500*(2011), 1-28.

Lombardi, F. & Di Pietro, R. (2011). Secure virtualization for cloud computing. *Journal of Network and Computer Applications, 34*(4), 1113-1122.

Luan, T. H., Gao, L., Li, Z., Xiang, Y., Wei, G. & Sun, L. (2015). Fog computing: Focusing on mobile users at the edge. *arXiv Preprint arXiv:1502.01815,*

Marston, S., Li, Z., Bandyopadhyay, S., Zhang, J. & Ghalsasi, A. (2011). Cloud computing—The business perspective. *Decision Support Systems, 51*(1), 176-189.

Mather, T., Kumaraswamy, S. & Latif, S. (2009). *Cloud security and privacy: An enterprise perspective on risks and compliance* " O'Reilly Media, Inc.".

Matt, C. (2018). Fog computing. *Business & Information Systems Engineering*, 1-5.

Mell, P. & Grance, T. (2011). The NIST definition of cloud computing. *Gaithersburg, MD, United States: National Institute of Standards & Technology*

Okoli, C. & Schabram, K. (2010). A guide to conducting a systematic literature review of information systems research.

Popović, K. & Hocenski, Ž. (2010). Cloud computing security issues and challenges. Teoksessa *MIPRO, 2010 proceedings of the 33rd international convention* (344-349) IEEE.

Ramgovind, S., Eloff, M. M. & Smith, E. (2010). The management of security in cloud computing. Teoksessa *Information Security for South Africa (ISSA)* (1-7) IEEE.

Reuben, J. S. (2007). A survey on virtual machine security. *Helsinki University of Technology, 2*(36)

Rong, C., Nguyen, S. T. & Jaatun, M. G. (2013). Beyond lightning: A survey on security challenges in cloud computing. *Computers & Electrical Engineering, 39*(1), 47-54.

Ryan, M. D. (2013). Cloud computing security: The scientific challenge, and a survey of solutions. *Journal of Systems and Software, 86*(9), 2263-2268.

Saharan, K. P. & Kumar, A. (2015). Fog in comparison to cloud: A survey. *International Journal of Computer Applications, 122*(3)

Shropshire, J. (2014). Extending the cloud with fog: Security challenges & opportunities.

Singh, G. (2013). A study of encryption algorithms (RSA, DES, 3DES and AES) for information security. *International Journal of Computer Applications, 67*(19)

Stojmenovic, I., Wen, S., Huang, X. & Luan, H. (2016). An overview of fog computing and its security issues. *Concurrency and Computation: Practice and Experience, 28*(10), 2991-3005.

Subashini, S. & Kavitha, V. (2011). A survey on security issues in service delivery models of cloud computing. *Journal of Network and Computer Applications, 34*(1), 1-11.

Takabi, H., Joshi, J. B. & Ahn, G. (2010). Security and privacy challenges in cloud computing environments. *IEEE Security & Privacy,* (6), 24-31.

Tordera, E. M., Masip-Bruin, X., García-Almiñana, J., Jukan, A., Ren, G., Zhu, J. & Farre, J. (2016). What is a fog node A tutorial on current concepts towards a common definition. *arXiv Preprint arXiv:1611.09193,*

Vaquero, L. M. & Rodero-Merino, L. (2014). Finding your way in the fog: Towards a comprehensive definition of fog computing. *ACM SIGCOMM Computer Communication Review, 44*(5), 27-32.

Wang, C., Wang, Q., Ren, K., Cao, N. & Lou, W. (2012). Toward secure and dependable storage services in cloud computing. *IEEE Transactions on Services Computing, 5*(2), 220-232.

Wu, L. & Buyya, R. (2012). Service level agreement (SLA) in utility computing systems. Teoksessa *Grid and cloud computing: Concepts, methodologies, tools and applications* (286-310) IGI Global.

Yi, S., Li, C. & Li, Q. (2015). A survey of fog computing: Concepts, applications and issues. Teoksessa *Proceedings of the 2015 workshop on mobile big data* (37-42) ACM.

Yi, S., Qin, Z. & Li, Q. (2015). Security and privacy issues of fog computing: A survey. Teoksessa *International conference on wireless algorithms, systems, and applications* (685-695) Springer.

Zhang, Q., Cheng, L. & Boutaba, R. (2010). Cloud computing: State-of-the-art and research challenges. *Journal of Internet Services and Applications, 1*(1), 7-18.

Zhou, M., Zhang, R., Xie, W., Qian, W. & Zhou, A. (2010). Security and privacy in cloud computing: A survey. Teoksessa *Semantics Knowledge and Grid (SKG), 2010 Sixth International Conference* (105-112) IEEE.

Zissis, D. & Lekkas, D. (2012). Addressing cloud computing security issues. *Future Generation Computer Systems, 28*(3), 583-592.