Tampereen yliopisto

Kristian Skogberg

LOHKOKETJUTEKNOLOGIAN SKAALAUTUVUUS

Haasteet ja ratkaisut

Informaatioteknologian ja viestinnän tiedekunta
Kandidaatintyö
Toukokuu 2022

TIIVISTELMÄ

Kristian Skogberg: Lohkoketjuteknologian skaalautuvuus: Haasteet ja ratkaisut
Tampereen yliopisto
Kandidaatintyö
Tieto- ja sähkötekniikan tutkinto-ohjelma, Tietotekniikka
Toukokuu 2022

Lohkoketjuteknologia on ajankohtainen aihe varsinkin tekniikan alalla. Monet alan asiantuntijat pitävät sitä mullistavana teknologiana ja uskovat lohkoketjuteknologian tulevan osaksi jokapäiväistä elämäämme tulevaisuudessa. Tässä työssä käsitellään lohkoketjuteknologian toimintaperiaatteita ja skaalautuvuutta. Työn tavoite on selvittää lohkoketjun skaalautuvuusongelmia ja minkälaisia ratkaisuja niihin on kehitetty.

Työ on jaettu kahteen osaan. Aluksi käsitellään lohkoketjun toimintaperiaatteita ja yleisimpiä konsensusmekanismeja. Sen jälkeen esitetään lohkoketjun skaalautuvuuteen liittyviä ongelmia ja niihin kehitettyjä ratkaisuja. Työ on toteutettu kirjallisuuskatsauksena ja lähteiksi on valittu tieteellisiä julkaisuja aikaväliltä 2017–2022.

Lohkoketju on hajautettu järjestelmä, joka mahdollistaa esimerkiksi vaihtokauppojen tekemisen turvallisesti ilman kolmatta osapuolta. Nimensä mukaisesti lohkoketju koostuu peräkkäisistä lohkoista ja jokainen lohko voi sisältää dataa. Lohkoketjut ovat lähtökohtaisesti täysin avoimia, eli kuka tahansa voi tarkastella ketjun sisältämää dataa ja tapahtumia. Lukuisat eri palvelinkoneet vahvistavat lohkoketjun tapahtumat erilaisilla konsensusmekanismeilla, joista yleisimpiä ovat Proof of Work ja Proof of Stake. Kun uusi tapahtuma on vahvistettu, se kootaan uuteen lohkoon ja tämä lohko lisätään lohkoketjun loppuun.

Työssä havaittiin, että lohkoketjuteknologian skaalautuvuuden suurimmat haasteet liittyvät tapahtumien nopeuteen ja lohkojen tilaan. Jokainen lohkoketju pystyy käsittelemään tietyn verran tapahtumia sekunnissa. Tapahtumien käsitteleminen voi kestää pitkään, mikäli lohkoketjussa on ruuhkaa. Lohkojen tila on myös hyvin rajallinen ja lohkoketjun koko kasvaa aina, kun sinne lisätään uusia lohkoja.

Lohkoketjun skaalautuvuuden haasteita voidaan ratkaista joko lohkoketjun sisällä tai sen ulkopuolella. Lohkoketjun sisällä voidaan suurentaa lohkoja, jolloin lohkoihin mahtuisi enemmän dataa. Tapahtumien datasta on myös mahdollista siirtää tietty osuus lohkon ulkopuolelle, jotta lohkoon mahtuisi enemmän dataa. Useita tapahtumia on myös mahdollista suorittaa samanaikaisesti hyödyntämällä sirpalointia. Tapahtumien käsitteleminen voidaan myös ulkoistaa lohkoketjun ulkopuolelle maksukanaville ja sivuketjuille.

Avainsanat: lohkoketjuteknologia, skaalautuvuus, konsensusmekanismit, Proof of Work, Proof of Stake

SISÄLLYSLUETTELO

LYHENTEET JA MERKINNÄT

PoW	Proof of Work
PoS	Proof of Stake
SPV	Simple Payment Verification
FTS	Follow-the-Satoshi
SegWit	Segregated Witness

1. JOHDANTO

Lohkoketju (engl. blockchain) on saavuttanut suuren suosion viime vuosina ja sitä pidetäänkin yhtenä maailmaa mullistavana teknologiana ja innovaationa. Yhä useammat yritykset ovat alkaneet miettiä tapoja integroida lohkoketjuteknologiaa heidän liiketoimintaansa. Parhaimmillaan lohkoketjuteknologia tehostaa yrityksen tapahtumien kirjanpitoa, sekä tarjoaa turvallisen ja muuttumattoman alustan datan säilyttämiselle. Lohkoketjuteknologiaan liittyy kuitenkin tiettyjä skaalautuvuuteen liittyviä ongelmia, joita käsitellään tässä tutkielmassa.

Lohkoketju on hajautettu järjestelmä, joka mahdollistaa esimerkiksi vaihtokauppojen tekemisen turvallisesti ja läpinäkyvästi ilman kolmatta osapuolta. Perinteisessä keskitetyssä järjestelmässä on aina jokin kolmas osapuoli, kuten pankki, joka vahvistaa erilaiset maksutapahtumat. Lohkoketjun voidaan ajatella olevan suuri verkosto, jota ylläpitävät lukuisat palvelinkoneet eri puolilla maailmaa. Näitä palvelimia kutsutaan solmuiksi (engl. node) ja sama data on tallennettu jokaiselle solmulle ja tätä dataa päivitetään jatkuvasti [5]. Solmut siis ylläpitävät kopiota lohkoketjusta ja päivittävät sitä aina, kun lohkoketjuun lisätään uusia lohkoja.

Lohkoketju koostuu nimensä mukaisesti peräkkäisistä lohkoista, jotka voivat sisältää dataa [1, 5]. Suurin osa lohkoketjuista on julkisia, eli niiden sisältämä data ja tapahtumat ovat kaikkien saatavilla ja kuka tahansa voi halutessaan tarkastella niitä. Lohkoissa olevaa data on lähtökohtaisesti muuttumatonta [2], eli sitä ei voi enää muuttaa jälkeenpäin ja tämä lisää lohkoketjuteknologiaan ylimääräisen kerroksen turvallisuutta ja luotettavuutta.

Lohkoketjut voidaan jakaa julkisiin ja yksityisiin lohkoketjuihin. Jako perustuu yksinkertaisesti siihen, ketkä tahot ovat oikeutettuja suorittamaan tapahtumia lohkoketjussa. Julkisessa lohkoketjussa kuka tahansa voi osallistua tapahtumien vahvistamiseen, kun taas yksityisessä lohkoketjussa vain ennalta määritellyt solmut osallistuvat tapahtumien vahvistamiseen [5, 12]. Tässä tutkielmassa keskitytään julkisiin lohkoketjuihin.

Tutkielma toteutettiin kirjallisuuskatsauksena. Tutkielmaan valittiin lähinnä vertaisarvioituja lähteitä kuten tieteellisiä julkaisuja ja konferenssijulkaisuja sekä yksi lohkoketjuteknologiaa käsittelevä kirja. Koska lohkoketju on teknologiana suhteellisen uusi, näkyy se myös lähteiden ajankohtaisuudessa. Tutkielmaan valittiin lähteitä vuosien 2017 ja 2022 väliltä. Lähteiden ajankohtaisuus on varsinkin tässä aiheessa tärkeää huomioida, sillä

lohkoketjuteknologia kehittyy jatkuvasti ja uusia menetelmiä ja toimintatapoja kehitetään nopeasti eteenpäin.

Lähteitä etsittiin Andorista, IEEE:stä ja ProQuestista. Suurin osa valituista lähteistä on Andorista sekä muutama lähde IEEE:stä. Haut rajattiin vertaisarvioituihin tieteellisiin artikkeleihin, konferenssijulkaisuihin ja kirjoihin. Seuraavia hakusanoja sekä niiden yhdistelmiä käytettiin aineiston keräämiseen:

- Blockchain (suom. lohkoketju)
- Proof of Work (suom. todistus työstä)
- Proof of Stake (suom. todistus osuudesta)
- Scalability (suom. skaalautuvuus)
- Consensus mechanism (suom. konsensusmekanismi)
- Cryptocurrency (suom. kryptovaluutta).

Tutkimuskysymyksenä on etsiä ratkaisuja lohkoketjujen skaalautuvuuteen. Skaalautuvuusongelmat liittyvät lähinnä ketjun nopeuteen ja lohkojen kokoon. Mitä enemmän tapahtumia lohkoketjussa tapahtuu samanaikaisesti, sitä kauemmin niiden käsitteleminen kestää ja sitä kalliimmaksi se tulee tapahtuman aloittajalle. Tämän seurauksena lohkoketjun käyttäjät voivat joutua odottamaan pitkään, kunnes heidän tapahtumansa on käsitelty. Lisäksi yksittäisiin lohkoihin mahtuu hyvin vähän dataa.

Lohkoketjun nopeuden parantamiseksi on kehitetty esimerkiksi erilaisia maksukanavia ja sivuketjuja, jotka käsittelevät tapahtumia lohkon pääketjun sijaan. Lisäksi voidaan hyödyntää sirpalointia, jolloin lohkoketjun solmut jaettaisiin ryhmiin ja nämä ryhmät käsittelisivät samanaikaisesti eri tapahtumia. Lohkojen kokoa on myös mahdollista suurentaa, jolloin niihin mahtuisi enemmän tapahtumia.

2. LOHKOKETJUTEKNOLOGIA

Lohkoketjut ovat yksinkertaisimmillaan hajautettuja järjestelmiä, jotka sisältävät valtavasti dataa. Lohkoketjuihin tallennettua dataa voi tarkastella milloin vain, mutta dataa ei voi poistaa lohkoketjusta. Tämä data koostuu lukuisista maksutapahtumista lohkoketjun käyttäjien välillä. Jokaisella kryptovaluutalla, kuten Bitcoinilla ja Ethereumilla, on lähtökohtaisesti oma lohkoketjunsa, jonka maksuvälineenä toimii kyseinen kryptovaluutta. Tässä kappaleessa käsitellään lohkoketjun toimintaa ja yleisimpiä konsensusmekanismeja, joiden avulla lohkoketjun tapahtumia käsitellään.

2.1 Lohkoketjun toiminta

Yksi lohkoketjun olennaisista ominaisuuksista on toimia järjestelmänä, jonne voi tallentaa dataa. Datan lisääminen onnistuu vain järjestelmän loppuun ja aikaisempaa dataa ei voi enää myöhemmin muokata. Dataa ei kuitenkaan välittömästi lisätä järjestelmän loppuun, vaan vireillä oleva data kootaan yhdeksi lohkoksi [5].

Nämä lohkot voivat sisältää yhden tai useamman tapahtuman tiedot, kuten lähettäjän ja vastaanottajan, sekä tiivisteen (engl. hash) edeltävään lohkoon [1, 2]. Jos lohkon tietoja yritettäisiin muokata, se tuottaisi täysin erilaisen tiivisteen, joka ei enää täsmäisi muiden lohkojen kanssa [5]. Tämä mahdollistaa datan muuttumattomuuden, sillä pienikin muutos aikaisempaan lohkoon muuttaisi kyseisen lohkon tiivisteen. Tällöin muut palvelimet eli solmut huomaisivat, että se ei enää täsmää heidän kopioidensa kanssa.

Lohkoketjun toiminta perustuu vertaisverkkoon (engl. peer to peer) [1, 5, 12]. Verkossa ei ole olemassa yksittäistä ylläpitäjää, vaan samanarvoiset käyttäjät pitävät yhdessä verkkoa pystyssä. Kryptografia mahdollistaa sen, että lohkoketjun käyttäjät voivat luottaa toisiinsa ja suorittaa vaihtokauppoja keskenään turvallisesti lohkoketjun välityksellä.

Lohkoketjussa tapahtumat vahvistetaan kryptografialla, jotta voidaan varmistaa, että tietty käyttäjä on aloittanut tapahtuman ja kukaan ei voi väärentää tätä tapahtumaa [1]. Toisin sanoen henkilö A ei voi väärentää henkilön B suorittamaa tapahtumaa, eikä hän voi suorittaa tapahtumaa kenenkään muun puolesta. Kryptografiset funktiot tuottavat aina saman lopputuloksen samalla sisääntulolla ja pienikin muutos alkuperäiseen sisääntuloon tuottaisi täysin erilaisen tuloksen [1, 5].

Kuka tahansa voi halutessaan liittyä julkiseen lohkoketjuun tai poistua julkisesta lohkoketjuun. Liittyessään palvelin eli solmu tallentaa itselleen kopion koko lohkoketjusta eli tiedot jokaisesta lohkosta [5]. Jokainen uusi tapahtuma jaetaan verkon jokaisen solmun

kesken [1, 5, 6], joten kaikki solmut ovat jatkuvasti ajan tasalla jokaisesta tapahtumasta. Eli jokaisella solmulla on täysin identtinen kopio lohkoketjusta [1, 5].

Uusien lohkojen luomiseen tarvitaan kuitenkin menetelmiä tai periaatteita, joita kaikki solmut noudattavat. Jokaisella verkon solmulla tulee olla yhteisymmärrys (engl. consensus) jokaisesta lohkosta. Tämä on tärkeää, jotta esimerkiksi haitalliset solmut eivät pystyisi manipuloimaan lohkoketjun tapahtumien paikkansapitävyyttä. Solmujen yhteisymmärryksen avulla verkosta saadaan myös vakaa ja kestävä, sekä sen avulla voidaan torjua monia erilaisia hyökkäyksiä. Lohkoketjun tapahtumat käsitellään erilaisten konsensusmekanismien (engl. consensus mechanism) avulla, joista yleisimpiä ovat Proof of Work ja Proof of Stake. [1]

2.2 Proof of Work

Proof of Work (PoW) on yksi yleisimmistä konsensusmekanismeista [4]. Se perustuu tietokoneiden tehokkaaseen laskentakykyyn ja siihen, että uusien lohkojen luominen on mahdollisimman työlästä. Lohkoketjun turvallisuuden kannalta on tärkeää, että uusien lohkojen luominen on työlästä ja siten epäkannattavaa toimintaa, sillä se estää monia mahdollisia hyökkäyksiä. [1]

PoW:issä tehokkaat tietokoneet ratkaisevat monimutkaisia kryptografisia ongelmia [2, 13]. Näitä tietokoneita kutsutaan louhijoiksi (engl. miner). Se louhija, joka lopulta ratkaisee tämän ongelman, valitaan satunnaisesti [2, 5]. Valinta kuitenkin suosii louhijoita, joilla on suurin laskentateho suhteessa verkon kokonaislaskentatehoon [2, 5]. Jokainen solmu vahvistaa ja hyväksyy uuden lohkon, ennen kuin siirrytään uudelle kierrokselle protokollassa [5].

Laskentateho määritellään käyttämällä useita mittareita, joita tarvitaan kryptografisten ongelmien ratkaisemiseen. Toisin sanoen laskentateho riippuu tietokoneen tehokkuudesta, minkä takia PoW-menetelmä kuluttaa paljon energiaa [7]. Louhijat voivat parantaa laskentatehoaan lisäämällä tai päivittämällä tietokoneeseensa uusia komponentteja, jolloin tietokone pystyy yrittämään samassa ajassa enemmän erilaisia yhdistelmiä kryptografisten ongelmien ratkaisemiseksi [2]. Näytönohjaimet ovat louhinnan kannalta tärkeimpiä komponentteja. Louhinnan yleistyminen on johtanut muun muassa siihen, että tietokonekomponenttien kysyntä on kasvanut huomattavasti. Parhaan louhimistehon saavuttamiseksi tulisi käyttää näytönohjaimia, jotka ovat mahdollisimman tehokkaita, mutta kuluttavat mahdollisimman vähän energiaa.

Louhijat voivat myös työskennellä yhteistyössä, jolloin muodostuu louhintaryhmiä (engl. mining pool). Kuka tahansa voi yhdistää oman tietokoneensa osaksi valitsemaansa louhintaryhmää ja osallistua louhimiseen. Näissä ryhmissä louhinnasta aiheutuneet tuotot jaetaan jokaiselle louhijalle sen perusteella, kuinka paljon laskentatehoa he tuovat ryhmään. Tällä tavalla louhijan ei tarvitse luoda itselleen täydellistä solmua, jotta hän pystyisi osallistua louhimiseen. [2]

Tapahtuman aloittajalle aiheutuu aina pieni kustannus tapahtumasta (engl. transaction fee). Se louhija, joka lisää lohkoketjuun uuden tapahtuman sisältävän lohkon, saa työstään palkkioksi tämän tapahtumamaksun. Tapahtumamaksut ovat pieni osa kryptovaluuttaa, kuten Bitcoinia tai Ethereumia. Tyypillisesti julkisissa lohkoketjuissa tapahtumamaksujen suuruus ei ole vakio, vaan se riippuu verkossa olevien tapahtumien määrästä sillä hetkellä. Useimmat lohkoketjut tarjoavat tämän tapahtumamaksun lisäksi louhijoille palkinnoksi pienen määrän kryptovaluuttaa, joka on juuri luotu (engl. minted) uuden lohkon yhteydessä. Tämä juuri luotu kryptovaluutta kasvattaa kryptovaluutan tarjontaa. [5]

2.3 Proof of Stake

Toinen suosituimmista konsensusmekanismeistä on nimeltään Proof of Stake (PoS). PoS perustuu siihen, että satunnaisesti valitut validaattorit (engl. validators) vahvistavat lohkojen tapahtumat. Validaattorit ovat verkon jäseniä, jotka ovat asettaneet osan kryptovaluutastaan verkon tukemistilaan (engl. stake), jolloin niitä ei voi käyttää. Tämän seurauksena nämä verkon jäsenet voivat osallistua tapahtumien vahvistukseen. Kryptovaluutat pitää asettaa tukemistilaan hyvissä ajoin ennen kuin verkon jäsenestä voi tulla mahdollinen validaattori, ja ne on pidettävä tukemistilassa tietyn aikaa. Tämä estää tiettyjä hyökkäyksiä verkossa. [1, 5]

Validaattorit vastaavat PoW-menetelmän louhijoita ja palkinnot jaetaan samalla periaatteella. Se validaattori, joka luo uuden lohkon ketjuun, saa palkinnoksi tapahtuman siirrosta aiheutuneet kulut [5]. Kuten PoW:issä, palkinto on pieni määrä lohkoketjun natiivia kryptovaluuttaa.

PoS:lla pyritään ratkaisemaan PoW:in ongelmia, kuten sen energiatehokkuutta [5, 13] ja nopeutta [2]. Koska PoS perustuu laskentatehon sijaan tukemistilaan asetettujen kryptovaluuttojen määrään, toimii se huomattavasti nopeammin ja energiatehokkaammin PoW:iin verrattuna [1]. PoS pyrkii myös siihen, että tapahtumia voitaisiin vahvistaa samanaikaisesti huomattavasti enemmän, mikä lisää lohkoketjun tehokkuutta [2].

Todennäköisyys sille, että validaattori pääsee luomaan uuden lohkon, on suoraan riippuvainen hänen tukemistilaan asettamaansa kryptovaluutan määrään. Esimerkiksi jos

verkon jäsen omistaa 3 % verkossa olevasta Ethereum-kryptovaluutasta, pääsee hän myös vahvistamaan 3 % Ethereum-lohkoketjun tapahtumista [1]. Siis todennäköisyys sille, että yksittäinen verkon jäsen valitaan validaattoriksi, on äärimmäisen pieni. Lisäksi solmun pitää olla jatkuvasti yhdistettynä verkkoon, jotta se voidaan tarvittaessa valita validaattoriksi [7]. Tämän seurauksena tukijat voivat muodostaa tukiryhmiä (engl. staking pool) eli periaatteessa yhdistää tukemansa kryptovaluutat suuremmaksi kokonaisuudeksi. Jos lohkoketju valitsee kyseisen tukiryhmän validaattoriksi, jaetaan palkinto tukiryhmän jäsenten kesken [8].

PoS-menetelmässä valitaan satunnaisesti tukijoiden joukosta tukija, jolle annetaan oikeus päivittää lohkoketjua [4]. Yksinkertaisin PoS toteutus on niin sanottu FTS-algoritmi (engl. Follow-the-Satoshi), jossa tukijoiden joukosta valitaan satunnaisesti ja yhdenmukaisesti yksi poletti (engl. token). Se osakas, joka sattuu omistamaan tämän poletin, saa oikeuden päivittää lohkoketjua lohkopalkintoa (engl. block reward) vastaan. Toisin sanoen mitä enemmän kryptovaluuttaa osakas on asettanut tukemistilaan, sitä suurempi todennäköisyys sille, että lohkoketju valitsee juuri hänen tukemansa kryptovaluutan. [8]

Koska PoS perustuu tukemistilaan asetettujen kryptovaluuttojen määrään, toimii se huomattavasti nopeammin ja energiatehokkaammin kuin PoW [1]. PoS on periaatteessa myös turvallisempi, koska se vaatii, että vähintään kaksi kolmasosaa kaikista tuetuista kryptovaluutoista tulisi kuulua ”rehellisille” solmuille [5]. Toisin sanoen uusi lohko luodaan vain, jos vähintään kaksi kolmasosaa kaikista validaattoreista on keskenään samaa mieltä lohkon tapahtumista. PoS myös torjuu paremmin hyökkäyksiä, jossa yksittäiset validaattorit tai solmut yrittäisivät väärentää lohkoketjun tapahtumia.

3. SKAALAUTUVUUSONGELMAT

Lohkoketjun skaalautuvuuden ratkaiseminen on hankalaa vaarantamatta sen turvallisuutta, hajautuneisuutta tai luotettavuutta [12]. Koska lohkoketju on hajautettu järjestelmä ja jokainen solmu osallistuu uusien tapahtumien vahvistukseen [5], aiheuttaa se myös tiettyjä skaalautuvuuteen liittyviä ongelmia. Lisäksi lohkojen tila on hyvin rajallinen ja tapahtumien käsitteleminen vie aina kaistaa. Tässä kappaleessa käsitellään lohkoketjuteknologian skaalautuvuuteen liittyviä ongelmia.

3.1 Nopeus

Tällä hetkellä Bitcoin-lohkoketju pystyy prosessoimaan noin kolmesta seitsemään tapahtumaa sekunnissa [11, 12]. Ethereum-lohkoketju pystyy puolestaan suorittamaan noin 15 tapahtumaa sekunnissa [12]. Vastaavasti luottoyhtiö VISA pystyy käsittelemään keskimäärin 1600-2000 tapahtumaa samassa ajassa [11, 12]. Tämä antaa osviittaa sille, kuinka hitaasti tapahtumat käsitellään lohkoketjussa verrattuna keskitettyihin järjestelmiin.

Lohkoketjun tapahtumien nopeus on suhteellinen lohkojen luomisaikaan sekä siihen, kuinka monta tapahtumaa yhteen lohkoon mahtuu [5]. Näitä kahta tekijää on kuitenkin tarkoituksellisesti rajoitettu, jotta lohkoketju pysyisi mahdollisimman turvallisena ja hajautettuna [5]. Hyvin suuret lohkot kuormittavat lohkoketjun solmuja ja ne voivat altistaa verkon erilaisille palveluhyökkäyksille [12]. Jos lohkoja puolestaan luotaisiin nopeammin, aiheuttaisi se solmujen välille synkronointiongelmia, sillä yksittäisten solmujen suorituskyky voi vaihdella huomattavasti [12].

Lohkoketjun käyttäjät voivat joutua odottamaan pitkään, ennen kuin heidän tapahtumansa on suoritettu onnistuneesti. Tämän lisäksi lohkoketjun käyttäjien on suositeltavaa odottaa muutamien uusien lohkojen luomista heidän tapahtumansa sisältävän lohkon päälle, jotta voidaan varmistua datan muuttumattomuudesta. Esimerkiksi Bitcoin yhteisö suosittelee, että Bitcoin-lohkoketjun käyttäjät odottaisivat noin 60 minuuttia heidän tapahtuman käsittelemisen jälkeen, jotta uusia lohkoja ehditään luoda tarpeeksi [5].

Viive on myös iso ongelma lohkoketjun skaalautuneisuudessa. Hajautetuissa järjestelmissä uusien palvelimien eli solmujen lisääminen lisää viivettä, kun taas perinteisissä keskitetyissä järjestelmissä vaikutus on päinvastainen. Toisaalta tämä pätee lähinnä julkisiin lohkoketjuihin, sillä yksityiset lohkoketjut hallitsevat solmujen tehoa ja kaistaa. [1]

Bitcoinin yleistymisen seurauksena myös sen volyymi on kasvanut huomattavasti. Lohkojen pieni koko ei usein vastaa samanaikaisesti tapahtuvien tapahtumien määrää, jolloin louhijat suosivat tapahtumia, joilla on korkea tapahtumamaksu. Tämän seurauksena tapahtumat, joilla on pieni tapahtumamaksu joutuvat odottamaan kauemmin ja se lisää viivettä. Ethereumin kohdalla tilanne on vielä pahempi, sillä monet sen päälle kehitetyt sovellukset (engl. DApps) vievät kaistaa koko verkolta. [10]

Ethereum-lohkoketjun tämänhetkinen toteutus vaatii, että jokainen verkon validaattori vahvistaa jokaisen verkon tapahtuman [9]. Eli voi kulua pitkä aika, ennen kuin validaattorien välinen yhteysymmärrys tapahtumasta saavutetaan. Lisäksi mitä enemmän tapahtumia suoritetaan samanaikaisesti, sitä kauemmin niiden käsitteleminen kestää.

3.2 Tila

Vuoden 2020 elokuussa Bitcoin-lohkoketjun koko oli yli 280 GB ja Ethereumin vastaavasti 562 GB [12]. Lohkoketjun koko kasvaa jatkuvasti ja koska solmut joutuvat ylläpitämään kopiota lohkoketjusta, vaatii se solmuilta lisää tallennustilaa [5, 12, 13]. Lisäksi esilatausaika (engl. bootstrap time) kasvaa lineaarisesti lohkoketjun kasvaessa, mikä hidastaa uusien solmujen liittymistä verkkoon [10]. On kuitenkin mahdollista luoda osittaisia solmuja (engl. partial node), jotka ylläpitäisivät vain tiettyä osaa lohkoketjun datasta [13].

Lohkoketjun suorituskyky kasvaa, kun lohkojen tila kasvaa [12]. Esimerkiksi Bitcoinissa yhteen lohkoon mahtuu vain 1 MB dataa [12, 13]. Jos tämä kapasiteetti ylitetään, verkko hylkää koko lohkon [9, 13]. Lisäksi solmuilla on tietty kapasiteetti, joka tulisi myös huomioida. Lohkoketjun kasvava koko rajoittaa myös esimerkiksi lohkoketjujen soveltamista sulautettuihin järjestelmiin, sillä sulautetuilla järjestelmillä on muutenkin hyvin rajattu muisti [12].

Jotta lohkoketju pysyisi avoinna mahdollisimman monelle, on myös tärkeää, että lohkojen kokoa rajoitetaan jollain tavalla. Muuten olisi riski sille, että vain ne tahot, jotka omistavat tehollisesti ja säilytystilallisesti parhaimmat solmut, ylläpitäisivät verkkoa. Tämä olisi lohkoketjun hajautuneisuutta ja avoimuutta vastaan. [5]

4. RATKAISUJA SKAALAUTUVUUSONGELMIIN

Lohkoketjuteknologian skaalautuvuutta voidaan lähteä ratkaisemaan joko lohkoketjun sisällä (engl. on-chain) tai lohkoketjun ulkopuolella (engl. off-chain). Tässä kappaleessa esitellään mahdollisia ratkaisuja edellisessä kappaleessa mainittuihin lohkoketjuteknologian skaalautuvuusongelmiin.

On-chain ratkaisut suoritetaan kokonaan lohkoketjun sisällä. Off-chain ratkaisut perustuvat siihen, että lohkoketjuun tallennetaan lähinnä lopputulokset, ja raskas työ tehdään sen ulkopuolella. Sen voidaan ajatella olevan lohkoketjun ylempi kerros, joka tekee raskaimmat työt. [1]

4.1 Lohkojen koon suurentaminen

Lohkojen kokoa on mahdollista suurentaa [12, 13, 14]. Tällöin lohkoihin mahtuu enemmän dataa ja tapahtumia. Bitcoin-cash on yksi esimerkki tästä, ja sen avulla lohkojen kooksi saatiin 8 MB ja myöhemmin 32 MB [12]. Suuremmat lohkot parantavat ketjun suoritustehoa.

Suuret lohkot voivat kuitenkin vaarantaa lohkoketjun turvallisuutta, sekä haarautumiskohtien (engl. fork) ja palvelunestohyökkäysten lisääntymistä [11, 12]. Suurin ongelma on kuitenkin se, että suurten lohkojen jakaminen kaikkien verkon solmujen kesken vie paljon aikaa ja voi aiheuttaa viivettä [14]. Lisäksi jos lohkojen kokoa suurentaa, kuormittaa se myös lohkoketjun solmuja.

SegWit (engl. Segregated Witness) on protokolla, joka pyrkii parantamaan Bitcoinin lohkojen kapasiteettia sekä tapahtumien muokattavuutta. Jokainen lohkoketjun tapahtuma vaatii aina digitaalisen allekirjoituksen, joka vie noin 65-70% lohkojen tilasta. SegWit-protokollan avulla nämä allekirjoitukset voidaan poistaa tapahtumien datasta ja siirtää lohkon ulkopuolelle. Tällöin lohkoihin mahtuu jopa neljä kertaa enemmän tapahtumia, ja SegWitin avulla Bitcoinin lohkojen kooksi saadaan 4 MB. [12, 13, 14]

Näiden lisäksi on myös ehdotettu ratkaisuja, joissa lohkoketjusta karsittaisiin tarpeetonta dataa [10]. Näin lohkoketju veisi vähemmän tilaa ja solmujen kapasiteetti kasvaisi. Voi olla kuitenkin vaikeaa määritellä, mikä data on tarpeetonta ja missä tätä dataa säilytettäisiin.

4.2 Sirpalointi

Sirpalointia (engl. sharding) on hyödynnetty jo pitkään tietojenkäsittelytieteessä. Siinä data jaetaan pienempiin osiin esimerkiksi eri palvelimille, jolloin järjestelmän suoritusnopeus paranee [10, 12].

Sirpalointia on mahdollista hyödyntää niin, että solmut jaettaisiin pienimpiin ryhmiin. Nämä ryhmät käsittelisivät vain tietyt tapahtumat, jolloin useita tapahtumia voitaisiin käsitellä samanaikaisesti huomattavasti enemmän. Tapahtumat jaetaan ryhmien sisällä yksittäisille solmuille hyödyntämällä Bysantin konsensusalgoritmia (engl. Byzantine consensus algorithm). Toisin sanoen lohkoketjun nopeus paranisi huomattavasti, sillä tapahtumia käsiteltäisiin samassa ajassa paljon enemmän [10, 12, 13]

Kehityksessä oleva Ethereum-lohkoketjun versio 2.0 hyödyntää sirpalointia ja Proof of Stake -menetelmää. Ethereum 2.0:ssa solmut on jaettu 64 ryhmään ja jokainen ryhmä käsittelee tapahtumia ja tallentaa dataa samanaikaisesti. Validaattorit valitaan kahdeksan sekunnin välein jokaisessa ryhmässä. On ennustettu, että Ethereum 2.0:n avulla tapahtumien määrä sekunnissa kasvaisi jopa 100 000 asti. [12]

Sirpalointiin liittyy kuitenkin myös ongelmia. Yksi näistä ongelmista on se, miten tai millä perusteella tapahtumia jaetaan eri solmuille [10]. Toinen ongelma liittyy siihen, miten tapahtumien prosessoimisesta saataisiin mahdollisimman tehokasta [10]. On myös vaikeaa muodostaa sopivan kokoisia ryhmiä. Mitä enemmän ryhmiä, sitä enemmän jouduttaisiin ajamaan Bysantin konsensusalgoritmia [13]. Jos ryhmiä olisi vähän, silloin jokainen ryhmä koostuisi liian monesta solmusta, mikä voi aiheuttaa turvallisuusriskejä [13].

4.3 Maksukanavat

Lohkoketjun tapahtumien käsittelemisen nopeuttamiseksi on myös kehitetty niin sanottuja maksukanavia (engl. payment channel). Ne ovat väliaikaisia vaihtokauppakanavia, jotka prosessoivat osan verkon tapahtumista [10, 14]. Mikrotapahtumia suoritetaan lohkoketjun “pääketjun” ulkopuolella ja ne kootaan lopuksi yhdeksi tapahtumaksi pääketjuun [12]. Tämän seurauksena tapahtumamaksut pienenevät ja dataa ladataan pääketjuun vähemmän [12].

Salamaverkko (engl. lighting network) ja Raiden verkko (engl. Raiden network) ovat suosituimpia maksukanavatoteutuksia. Salamaverkko kehitettiin parantamaan Bitcoin-lohkoketjun suoritustehoa [10, 12]. Verkko mahdollistaa tapahtumien välittömän käsittelemisen, suuren suoritustehon sekä alhaiset tapahtumamaksut [14]. Lisäksi salamaverkko

hyödyntää älysopimuksia (engl. smart contract) välittömien tapahtumien mahdollistamiseksi [14]. Älysopimukset sisältävät suoritettavaa koodia ja erilaisia tiloja (engl. state) [1].

Salamaverkossa on kuitenkin myös huonoja puolia. Salamaverkon suurin heikkous on se, että sekä lähettäjän että vastaanottajan pitää olla verkossa samanaikaisesti [10]. Vaikka salamaverkon avulla pystytään käsittelemään huomattavasti enemmän tapahtumia, haasteena on käsitellä usean eri käyttäjän tapahtumia samanaikaisesti. Lisäksi tapahtumien käsitteleminen ei ole yhtä turvallista, kuten alkuperäisessä Bitcoin-lohkoketjussa. Salamaverkko toimii ainoastaan Bitcoin-lohkoketjun kanssa. [14]

Raiden-verkko pyrkii parantamaan Ethereum-lohkoketjun suorituskykyä. Raiden-verkko on nopea, edullinen ja helposti skaalattava verkko, jossa käsitellään tapahtumia. Tapahtumien käsitteleminen on edullisempaa Raiden-verkossa Ethereumin pääketjuun verrattuna. Lisäksi Raiden-verkko tukee kaikkia ERC20 poletteja [10, 12, 14].

Raiden-verkolla on myös heikkoutensa. Raiden-verkko vaatii, että osa poleteista tulisi olla lukittuna lopullisesti älysopimukseen. Lisäksi suurten polettimäärien siirtäminen voi olla haastavaa. [14]

4.4 Sivuketjut

Sivuketjut (engl. sidechains) ovat lohkoketjuja, jotka ovat yhteydessä lohkoketjun pääketjuun. Sivuketjut mahdollistavat tapahtumien suorittamisen lohkoketjun pääketjun ulkopuolella [10 12, 14]. Ne parantavat pääketjun tehokkuutta, yksityisyyttä ja turvallisuutta. Mikäli sivuketjuissa esiintyy turvallisuusongelmia, eivät ne kuitenkaan vaaranna pääketjua. Lisäksi sivuketjut voivat hyödyntää pääketjusta poikkeavia konsensusmekanismeja. [12]

Pääketjun käyttäjät voivat siirtää kryptovaluuttojaan sivuketjuihin, jolloin ne lukitaan tietyksi aikaa. Jotta käyttäjät voivat suorittaa tapahtumia sivuketjuissa, täytyy heidän pystyä todistamaan, että he ovat lukinneet osan kryptovaluutoistaan. Yksinkertaisen maksuvahvistuksen (engl. Simple Payment Verification, SPV) avulla käyttäjät voivat tehdä tämän todistuksen. SPV antaa käyttäjälle sen verran sivuketjun kryptovaluuttoja, kuin mitä käyttäjä oli alun perin siirtänyt sivuketjuun lukituksi. Nyt käyttäjä pääsee suorittamaan tapahtumansa sivuketjussa. [6, 10, 12]

Plasma on Ethereum-lohkoketjun ensisijainen sivuketjutoteutus [10, 12]. Plasma muistuttaa tietojenkäsittelytieteen puutietorakennetta, eli on olemassa yksi vanhempi (engl. parent chain) ja sillä on lapsia (engl. child chain), ja kaikki lapset osoittavat vanhempaan. Lapsilla voi olla omia lapsia, eli muodostuu hierarkinen rakenne. Vanhempi ja jokainen

lapsi on oma lohkoketjunsa, ja jokaisella lapsella on omat ominaisuutensa. Lapsien ominaisuuksia ja toteutuksia on mahdollista muuttaa haluamallaan tavalla. [14]

Ne tapahtumat, jotka hyödyntävät älysopimuksia, suoritetaan Plasman avulla pääketjun sijaan. Tapahtumia voidaan suorittaa samanaikaisesti eri lapsien lohkoketjuissa [14]. Lisäksi Plasma mahdollistaa tapahtumien suorittamisen ilman, että jokaisen osallistujan tarvitsisi olla verkossa samaan aikaan. Plasma lähettää pääketjulle tietyin väliajoin Plasmassa luotujen lohkojen tunnisteita, jolloin pääketju pystyy varmistamaan Plasman tapahtumien paikkansapitävyyden [10]. Vaikka Plasma nopeuttaakin tapahtumien käsittelemistä, pääketjun täytyy kuitenkin tarkistaa kaikki sivuketjujen lohkot, mikä lisää pääketjun työkuormaa [10, 12]. Lisäksi käyttäjät joutuvat odottamaan jopa 7-14 päivää, ennen kuin he voivat nostaa kryptovaluuttansa Plasmasta pääketjuun.

Sivuketjuilla on kuitenkin myös huonoja puolia. Pääketjun louhijat suorittavat SPV:n, jolloin jouduttaisiin luottamaan siihen, että pääketjun louhijat eivät varastaisi sivuketjussa olevia kryptovaluuttoja [6]. Yksittäisen tapahtuman suorittaminen sivuketjussa on hidasta, sillä käyttäjät joutuvat odottamaan, kunnes heidän siirtämänsä kryptovaluutat lukitaan ja myöhemmin vapautetaan.

5. TULOKSET

Edellisissä kappaleissa huomattiin, että lohkoketjuteknologian skaalautuvuusongelmien ratkaiseminen on haastavaa. Skaalautuvuuden parantaminen vaatii tiettyjä kompromisseja joko lohkoketjun turvallisuuden, avoimuuden tai luotettavuuden osalta. Nämä kaikki ovat myös lohkoketjuteknologian tärkeimpiä ominaisuuksia. Jos jokin näistä ominaisuuksista menetettäisiin skaalautuvuuden parantamisen vuoksi, voisiko sellaista järjestelmää enää kutsua hyväksi lohkoketjuksi?

Lohkoketjun nopeus on yksi keskeisimmistä skaalautuvuusongelmista. Jos uusien lohkojen luomiseen käytettyä aikaa lyhennetään, myös lohkoketjun nopeus kasvaa. Tämä kuittenkin vaatii solmuilta enemmän suorituskykyä ja voi periaatteessa poissulkea solmut, joilla on heikoin suorituskyky [5]. Lisäksi se voi vaarantaa lohkoketjun turvallisuuden, sillä mitä nopeammin uusia lohkoja tehdään, sitä helpommin solmujen välille voi syntyä erilaisia synkronointivirheitä.

Lohkoketjun nopeutta voidaan parantaa esimerkiksi hyödyntämällä sirpalointia. Tällöin yksittäiset solmut jaettaisiin ryhmiin, jolloin olisi mahdollista käsitellä monia tapahtumia samanaikaisesti. Sirpaloinnin suurin haaste on kuitenkin se, millä perusteella ja minkälaisella algoritmilla data jaetaan eri solmuille. Näiden asioiden toteuttaminen vaatii kuitenkin jonkin verran ylimääräistä laskentatyötä. Lohkoketjun nopeutta voidaan parantaa myös hyödyntämällä maksukanavia ja sivuketjuja. Osa tapahtumista siirrettäisiin pääketjun ulkopuolelle, jolloin tapahtumia voitaisiin käsitellä samanaikaisesti enemmän.

Lohkojen pieni koko nousi toiseksi skaalautuvuusongelmaksi. Datan kerääminen on usein helppoa, mutta haasteita syntyy siinä vaiheessa, kun se pitää tallentaa jonnekin. Lohkoketjun yksittäisiin lohkoihin mahtuu tyypillisesti vain hyvin vähän dataa, eivätkä ne sellaisenaan sovi suurten datamäärien säilytyspaikaksi. Lohkojen kokoa on kuitenkin mahdollista suurentaa, jolloin sinne mahtuisi enemmän tapahtumia. Vaihtoehtoisesti lohkojen tapahtumista on mahdollista siirtää digitaaliset allekirjoitukset lohkon ulkopuolelle, jolloin samankokoiseen lohkoon mahtuisi huomattavasti enemmän tilaa.

Oli mielenkiintoista tutkia lohkoketjun skaalautuvuushaasteita ja -ratkaisuja. Monet ratkaisut vaikuttavat aluksi hyviltä, mutta ne tuovat usein mukanaan myös huonoja puolia. Lohkoketjun turvallisuus ja avoimuus osoittautuivat tekijöiksi, joista joudutaan usein tekemään kompromisseja skaalautuvuuden parantamiseksi. Varsinkin turvallisuus on kuitenkin sellainen tekijä, josta ei kannata tehdä kompromisseja, vaikka se parantaisikin skaalautuvuutta.

6. YHTEENVETO

Työssä käsiteltiin aluksi lohkoketjuteknologiaa yleisesti ja sen toimintaperiaatteita. Lohkoketjussa tapahtumat vahvistetaan konsensusmekanismien avulla, joista yleisimpiä ovat Proof of Work ja Proof of Stake. Työssä havaittiin, että Proof of Work on vanhempi menetelmä, joka perustuu tietokoneiden tehokkaaseen laskentatehoon. Proof of Stake on puolestaan huomattavasti nopeampi ja energiatehokkaampi menetelmä. PoS perustuu laskentatehon sijaan verkon tukemistilaan asetettujen kryptovaluuttojen määrään.

Tutkimusten perusteella huomattiin, että lohkoketjuteknologian skaalautuvuusongelmat liittyvät lähinnä tapahtumien nopeuteen ja lohkojen rajattuun tilaan. Suosituimmat lohkoketjut, Bitcoin ja Ethereum, pystyvät käsittelemään todella vähän tapahtumia sekunnissa. Tämä ei vastaa samanaikaisesti tapahtuvien tapahtumien määrää, jolloin tapahtumien käsittelemisessä esiintyy viivettä.

Lohkoketjun nopeuden parantamiseksi on kehitetty useita keinoja. Sirpalointia on mahdollista hyödyntää niin, että palvelimet eli solmut jaettaisiin pienimpiin ryhmiin. Nämä ryhmät käsittelisivät vain tietyt tapahtumat ja ylläpitäisivät vain niille olennaista osaa lohkoketjusta. Lohkoketjun nopeutta on mahdollista parantaa myös niin, että ainakin osa tapahtumista käsiteltäisiin pääketjun ulkopuolella. Maksukanavat ja sivuketjut perustuvat tähän ajatukseen.

Lohkojen kokoa voidaan suurentaa. Tällöin lohkoihin mahtuisi enemmän tapahtumia, jolloin myös lohkoketjun nopeus kasvaisi. Tämän seurauksena lohkojen jakaminen verkon solmujen kesken kestäisi kuitenkin kauemmin, ja se vaatisi solmuilta enemmän suorituskykyä.

Vaikka tutkielmassa esitellyt ratkaisut vaikuttavat suhteellisen yksinkertaisilta, on niiden toteuttaminen usein hankalaa. On tärkeää tehdä paljon tutkimuksia skaalautuvuusratkaisuista, jotta voimme saada paremman käsityksen siitä, minkälaisia seurauksia niillä voi olla tulevaisuudessa. Lohkoketjuteknologia on kuitenkin hyvin uusi tieteenala ja siitä ei ole vielä kovin paljoa tietoa saatavilla. Toisaalta se tekeekin sen tutkimisesta ja soveltamisesta uutta ja mielenkiintoista. Onkin mielenkiintoista nähdä, miten lohkoketjut kehittyvät tulevaisuudessa ja minkälaisia sovelluksia niiden avulla kehitetään.

LÄHTEET

[1] Singhal B, Dhameja G, Panda PS. Beginning Blockchain A Beginner's Guide to Building Blockchain Solutions. 1st ed. 2018. Berkeley, CA: Apress; 2018.

[2] Islam N, Marinakis Y, Olson S, White R, Walsh S. Is BlockChain Mining Profitable in the Long Run? IEEE transactions on engineering management. 2021;1–14.

[3] Nofer M, Gomber P, Hinz O, Schiereck D. Blockchain. Business & information systems engineering. 2017;59(3):183–7.

[4] Gu W, Li J, Tang Z. A Survey on Consensus Mechanisms for Blockchain Technology. In: 2021 International Conference on Artificial Intelligence, Big Data and Algorithms (CAIBDA). Piscataway: IEEE; 2021. p. 46–9.

[5] Kolb J, AbdelBaky M, Katz R, Culler D. Core Concepts, Challenges, and Future Directions in Blockchain: A Centralized Tutorial. ACM computing surveys. 2020;53(1):1–39.

[6] Worley C, Skjellum A. Blockchain Tradeoffs and Challenges for Current and Emerging Applications: Generalization, Fragmentation, Sidechains, and Scalability. In: 2018 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData). IEEE; 2018. p. 1582–7.

[7] Nguyen CT, Hoang DT, Nguyen DN, Niyato D, Nguyen HT, Dutkiewicz E. Proof-of-Stake Consensus Mechanisms for Future Blockchain Networks: Fundamentals, Applications and Opportunities. IEEE access. 2019;7:85727–45.

[8] Saleh F. Blockchain without Waste: Proof-of-Stake. The Review of financial studies. 2021;34(3):1156–90.

[9] Mazlan AA, Mohd Daud S, Mohd Sam S, Abas H, Abdul Rasid SZ, Yusof MF. Scalability Challenges in Healthcare Blockchain System-A Systematic Review. IEEE access. 2020;8:23663–73.

[10] Zhou Q, Huang H, Zheng Z, Bian J. Solutions to Scalability of Blockchain: A Survey. IEEE access. 2020;8:16440–55.

[11] Xie J, Yu FR, Huang T, Xie R, Liu J, Liu Y. A Survey on the Scalability of Blockchain Systems. IEEE network. 2019;33(5):166–73.

[12] Sanka AI, Cheung RC. A systematic review of blockchain scalability: Issues, solutions, analysis and future research. Journal of network and computer applications. 2021;195:103232–.

[13] Khan D, Jung LT, Hashmani MA. Systematic literature review of challenges in blockchain scalability. Applied sciences. 2021;11(20):9372–.

[14] Hafid A, Hafid AS, Samih M. Scaling Blockchains: A Comprehensive Survey. IEEE access. 2020;8:125244–62.