

Oskari Pellikka

# LANGATTOMAT NETWORK-ON-CHIPIT

Informaatioteknologian ja viestinnän tiedekunta

Kandidaattitutkielma

Joulukuu 2023

Tarkastaja: Joonas Multanen

# TIIVISTELMÄ

Oskari Pellikka: Langattomat Network-on-Chipit  
Kandidaattitutkielma  
Tampereen yliopisto  
Tieto- ja sähkötekniikan kandidaattiohjelma, tietotekniikka  
Joulukuu 2023

---

Network-on-Chip (NoC) on moniydinarkkitehtuurin yhteysteknologia. Johtuen kuitenkin sen huonosta skaalautuvuudesta, suuresta latenssista ja tehonkulutuksesta, tarvitaan ydinten määrän kasvun myötä tilalle ratkaisu, joka pystyisi vastaamaan teknologiakehityksen tarpeeseen. Tässä tutkimuksessa tarkastellaan NoCin uusia teknologioita sekä erilaisia arkkitehtuureja, joilla pyritään saavuttamaan teknologian vaatimat ominaisuudet. Tutkimuksessa selvitetään myös kunkin teknologian ja arkkitehtuurin mahdollisuudet ja haasteet. Tutkimuskysymyksenä on: Millaisia ovat tiedonsiirtoteknologiat ja arkkitehtuuri, joilla pyritään ratkaisemaan Network-on-Chipin rajoituksia? Tarkoituksena on, että tämän tutkimuksen luettuaan lukija tietää, mikä on NoCin teknologioiden ja arkkitehtuurien nykytilanne ja ongelmat, joita pyritään ratkaisemaan.

Työ on toteutettu kirjallisuuskatsauksena, jossa perehdytään NoCin ongelmien mahdollisiin ratkaisuihin. Kirjallisuuskatsaus toteutettiin perehtymällä olemassa oleviin tutkimuksiin ratkaisuihin. Työ esittelee ensiksi NoCin taustoja, jotta lukija ymmärtäisi perusteet. Tämän jälkeen työssä käsitellään NoCin uusia ratkaisuja, joita ovat radioaalloilla toteutettu NoC, fotoninen NoC, 3D-arkkitehtuuri sekä ohjelmallinen määrittäminen NoCissa. Lopuksi työssä käsitellään langattomien NoCien yhdistämistä 3D-rakenteeseen.

Tutkimus osoittaa, että kutakin teknologiaa voidaan käyttää NoCin skaalautuvuusongelmien ratkaisemiseksi ainakin joissakin tapauksissa. Kukin ratkaisu tarjoaa siis tapauskohtaisesti parannuksia. Yleensä parannukset ovat noin 60 prosentin luokkaa latenssissa ja tehonkulutuksessa suhteessa perinteiseen NoCiin vastaavassa ratkaisussa. Ratkaisut tarjosivat siis paremmin skaalautuvia vaihtoehtoja moniydinarkkitehtuurin yhteysteknologiaksi. Vaikka kullakin ratkaisulla on potentiaalia, ei yksikään niistä kuitenkaan ole selkeä valinta.

Avainsanat: Network-on-Chip, langaton, radioaalloilla toimiva, fotoninen, 3D-rakenne, ohjelmallisesti määritetty

Tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck –ohjelmalla.

# SISÄLLYSLUETTELO

<b>1</b>	<b>Johdanto .....</b>	<b>1</b>
<b>2</b>	<b>Network-on-Chip .....</b>	<b>3</b>
2.1	Rakenne	3
2.2	Toimintokerrokset	3
<b>3</b>	<b>Langaton toteutus Network-on-Chipille.....</b>	<b>6</b>
3.1	Tiedonsiirtoteknologiat	6
3.1.1	Langaton Network-on-Chip	6
3.1.2	Fotoninen Network-on-Chip	8
3.2	Arkkitehtuuri	10
3.2.1	3D-Network-on-Chip	10
3.2.2	Langaton 3D-rakenne	11
<b>4</b>	<b>Yhteenveto.....</b>	<b>13</b>
	<b>Lähdeluettelo.....</b>	<b>15</b>

## 1 Johdanto

Nykyajan mikroprosessorisuunnittelu rakentuu yhä enemmän moniydinarkkitehtuurin varaan. Nämä moniydinarkkitehtuurit koostuvat useiden itsenäisten prosessorien tai ydinten verkostosta. Nämä ytimet voidaan pakata samaan koteloon tai samalle piikiekolle, jolloin ne voivat tehokkaasti jakaa monitasoista välimuistia keskenään. Tämä parantaa omalta osaltaan kokonaissuorituskykyä. Ydinten ja muistirakenteiden väliset yhteydet, usein metallijohtimet, ovat moniydinarkkitehtuurin keskiössä, sillä moniydinprosessorin oikea toiminta tarvitsee moitteettomat yhteydet. Tämän lisäksi yhteyksien toimiminen on ratkaiseva tekijä moniydinprosessorin suorituskykyä määritettäessä. [1]

Network-on-Chip (NoC), jolla toteutetaan tällä hetkellä moniydinproessoreissa moniydinarkkitehtuurin yhteyksiä, koostuu ydinten ja muistirakenteiden välisistä johtimista, joita pitkin signaalit kulkevat. NoC ei kuitenkaan ole kovin skaalautuva ratkaisu. Ydinten määrän kasvaessa myös ydinten välisten johtimien pituus kasvaa, mikä taas johtaa kasvavaan latenssiin ja tehohäviöön. Latenssilla tarkoitetaan aikaa, joka signaalilla kestää kulkea ytimeltä toiselle ja takaisin. [1] Tämän takia on ideoitu konseptia, joka mahdollistaisi yhä useampien ydinten käytön aiheuttamatta kuitenkaan yhä suurempaa latenssia ja tehohäviötä. Mahdollisena ratkaisuna voisi olla langaton yhteys kaukana toisistaan sijaitsevien ydinten välille. Tämä langaton yhteys voitaisiin luoda joko radioaalloilla tai fotonien avulla, jolloin vältettäisiin turhat pitkät johtimet. Lisäksi yhteys voitaisiin toteuttaa niin sanottuna one-hop-ratkaisuna, eli yhteys olisi vain haluttujen ytimien välillä, jolloin välttyttäisiin datan lähettämiseltä ylimääräisille ytimille. [2]

Työn tarkoituksena on tuoda ilmi erilaisia teknologioita ja arkkitehtuureja, joilla NoCin rajoituksia voitaisiin ratkaista. Tutkimuskysymyksenä on: Millaisia ovat tiedonsiirtoteknologiat ja arkkitehtuuri, joilla pyritään ratkaisemaan Network-on-Chipin rajoituksia?

Tietoa on etsitty lähinnä Tampereen yliopiston Andor-palvelusta sekä Google Scholar-palvelusta. Parhaita tuloksia saatiin seuraavilla hakulauseilla: "Network-on-Chip" AND wireless, "wireless Network-on-Chip" AND structure, "wireless Network-on-Chip" AND latency, "photonic Network-on-Chip" ja " Network-on-Chip" AND 3D. Lähteille asetettiin vaatimukseksi, että niiden tulee olla uudempia kuin NoC-konsepti ja mieluummin alle 10 vuotta vanhoja. Lisäksi pyrittiin käyttämään vain vertaisarvioituja lähteitä, mutta joitakin pieniä asioita on myös vertaisarvioimattomista lähteistä.

Luvussa Network-on-Chip esitellään Network-on-Chip-teknologia. Luvun tarkoituksena on tuoda lukijalle ilmi tarve uusille ratkaisuille ja pohja, joille nämä tulisi rakentaa.

Langaton toteutus Network-on-Chipille-luvun on tarkoitus esitellä lukijalle uusia langattomia teknologioita, joilla voidaan kehittää Network-on-Chipiä sekä arkkitehtuurinen ratkaisu, jolla voidaan entisestään edistää näitä teknologioita. Lisäksi luvussa käsitellään näiden kahden ratkaisun yhdistämistä.

## 2 Network-on-Chip

Network-on-Chip on moniydinarkkitehtuurin teknologia, jota käytetään ydinten välisten yhteyksien luomiseen ja käyttämiseen. Vuonna 2002 Luca Benini ja Giovanni De Micheli esittelivät NoCin, joka on sittemmin ratkaissut monia sirujen sisäisiä kommunikaatio- ja datanmuokkausongelmia. [3] Termiä Network-on-Chip käyttivät kuitenkin ensimmäisen kerran Kumar et al. tutkimuksessaan [4]. Tässä luvussa käydään läpi NoCin perusrakenne.

### 2.1 Rakenne

NoC on mahdollistanut useiden ydinten liittäminen samalle sirulle korjaamalla kommunikaatio-ongelmat. NoCin arkkitehtuureja ja niiden kehittämistä on tutkittu hyvin paljon. [5] NoCin rakenne koostuu tyypillisesti reitittimistä, verkkosovittimista ja yhteyksistä. Perustoimintaperiaatteena on, että reititin ohjaa datan kulkua valitun protokollan mukaisesti. Verkkosovittimet taas luovat yhteyden reitittimen ja siihen yhdistetyn elementin välille päätehtävänäan erottaa kommunikaatiolaskenta. Tämä tapahtuu protokollamuuntimella ja pakettien rakentamisella. [6] Tätä perusrakennetta on kuvattu kuvissa 1 ja 2 erilaisilla topologioilla.

NoCille on kehitetty erilaisia ratkaisuja. Näitä ovat esimerkiksi ÆTHREAL, SPIN sekä QnoC. ÆTHREALin on kehittänyt Philips Research Laboratories, ja se tarjoaa häviöttömän ja korruptiottoman tiedonsiirron sekä organisoidun siirtojärjestyksen. Näillä saavutetaan ennustettava latenssi ja taatut siirtonopeudet. Spin taas on kehitetty Pierre ja Marie Curien yliopistossa. Sen pääetuina ovat laaja puutopologia ja reitityspaketti, jolla varmistetaan tiedon siirtymisen sujuvuus NoCissa. QnoC on Israelin teknologiainstituutin kehittämä järjestelmä, joka perustuu epäsäännölliselle meshtopologialle, eli eri topologioiden yhdistelmille ja madonreiälle kytkentäteknikkana. Madonreiässä paketit jaetaan alipaketteihin. Hallintadataa pidetään otsikkoalipaketissa, jolloin pystytään lähettämään yksittäisiä paketteja eri solmukohdilta. Tästä seuraa pienempi latenssi, mutta tämä voi myös aiheuttaa verkostoon pullonkauloja. [6] Nämä ovat vain esimerkkejä erilaisista ratkaisuista ja hyödyistä, joita erilaisilla ratkaisuilla voi olla.

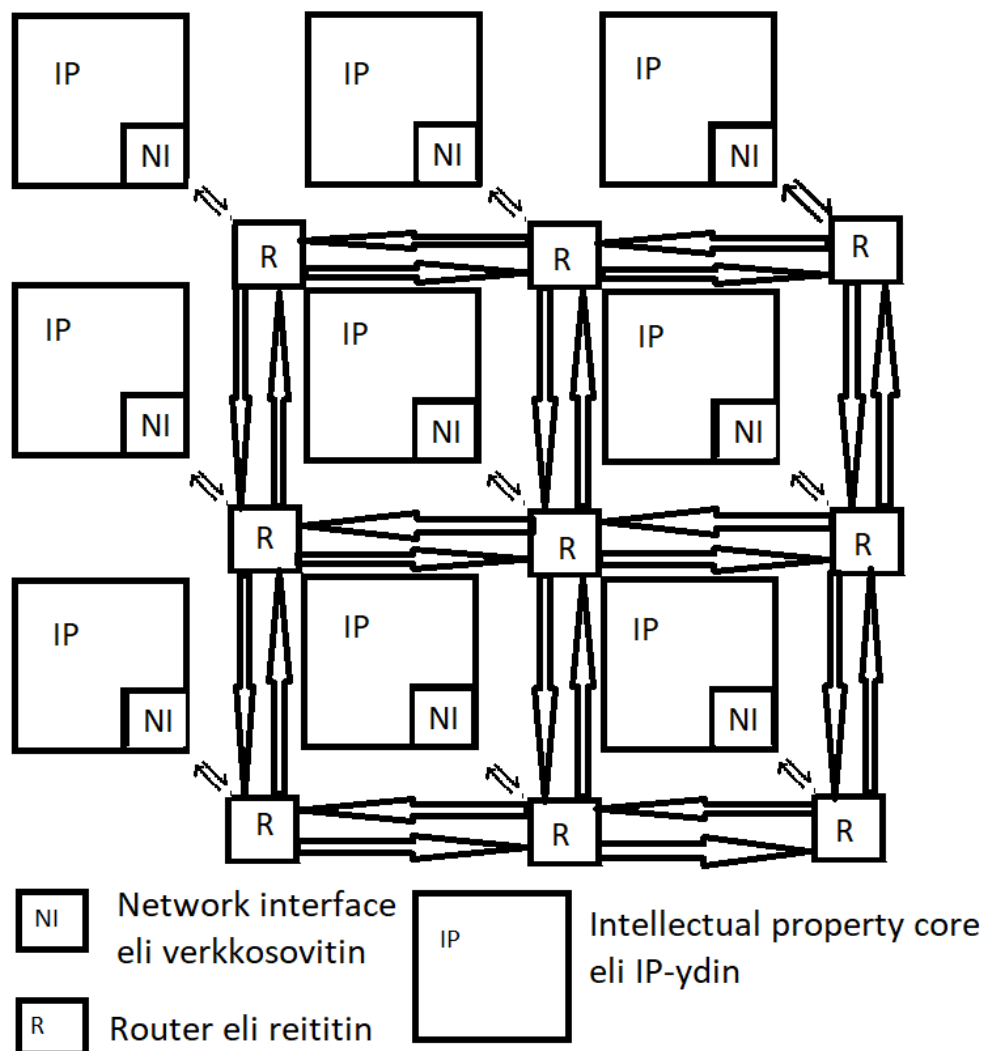
NoCin toiminnasta vastaavat toimintokerrokset. Kullakin toimintokerroksella on oma tehtävänsä, joka jollakin tavalla edesauttaa verkoston toimintaa. NoCin reitittimien tulee sisältää sekä ohjelmiston että laitteiston toteutukset, jotta ne kykenevät tukemaan kerrosten toimintoja. [6]

### 2.2 Toimintokerrokset

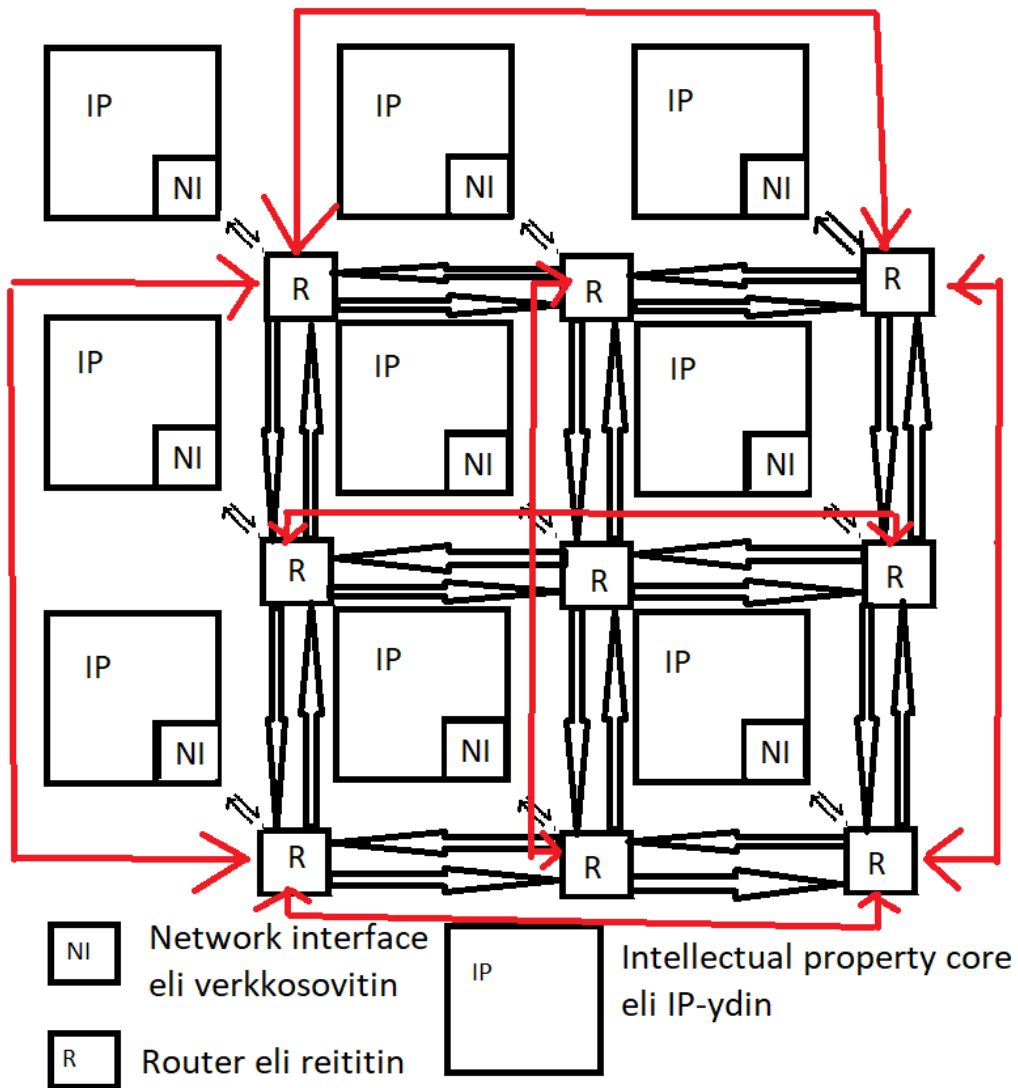
NoC jaetaan useisiin kerroksiin, joita ovat sovelluskerros, kuljetuskerros, verkkokerros, siirtoyhteyskerros ja fyysinen kerros. Sovelluskerros purkaa kohdesovellukset laskenta- ja kommunikaatiotehtäviksi, jotta esimerkiksi käytetty energia ja nopeus voidaan optimoida. NoCissa ydinten sijainti pyritään optimoimaan siten, että vähennetään kokonaiskommunikaatioita ja käytettyä energiaa samalla kuitenkin huomioiden yksittäisten yhteyksien rajoitukset. [7]

Kuljetuskerros vastaa ruuhka- ja virtauksenvalvontaongelmista. Koska NoCin tärkeimpiä ominaisuuksia on pieni latenssi ja korkea läpisyöttöaste, on huomattavan tärkeää, että ruuhkat ja pakettien ohjauksen ongelmat voidaan korjata ja ohjata perille nopeasti. [7]

Verkkokerros huolehtii verkkotopologiasta, eli siitä, miten verkoston resurssit yhdistetään toisiinsa. Käytännössä se siis vastaa siitä, että staattiset kanava- ja solmukohtajärjestelyt toimivat oikein yhteysverkossa. Viimeisimmät toteutukset käyttävät lähes poikkeuksetta torus- tai meshtopologioita, joilla saavutetaan parhaat skaalautuvuudet ja suorituskykyhyödyt. [7] Nämä topologiat esitetään kuvissa 1 ja 2.



Kuva 1: Tyylitelmä NoCin mesh-topologiasta. Tyylitelty lähteen [7] mukaan.



Kuva 2: Tyylitelmä NoCin torus-topologiasta. Tyylitelty lähteen [5] mukaan.

Siirtoyhteyserkerros lisää yhteyksien luotettavuutta, kun oletetaan, että fyysinen kerros ei itsessään ole riittävän luotettava. Siirtoyhteyserkerros siis lähinnä tukee fyysistä kerrosta. Fyysinen kerros taas painottaa signaaliajureita ja vastaanottimia sekä erilaisia suunnitteluteknologioita, joilla järjestään uudelleen ja ohjataan signaaleja johtimia pitkin. Teknologian edetessä jännitetasot ovat laskeneet, mikä on heikentänyt tehokohinasuhteita. Sähköinen kohina aiheuttaa ajoitus- ja datavirheitä vaikeuttaen luotettavien siruyhteyksien saavuttamista. [7]



### 3 Langaton toteutus Network-on-Chipille

Tässä luvussa esitellään Network-on-Chipin erilaisia langattomia tiedonsiirtoteknologioita sekä arkkitehtuuria. Luvun tarkoituksena on tutustuttaa lukija eri teknologioihin sekä selventää lukijalle teknologioiden eroavaisuuksia. Lisäksi luvun tarkoituksena on esitellä lukijalle, kuinka arkkitehtuurisilla ratkaisuilla voidaan vaikuttaa NoCin teknologioiden ominaisuuksiin.

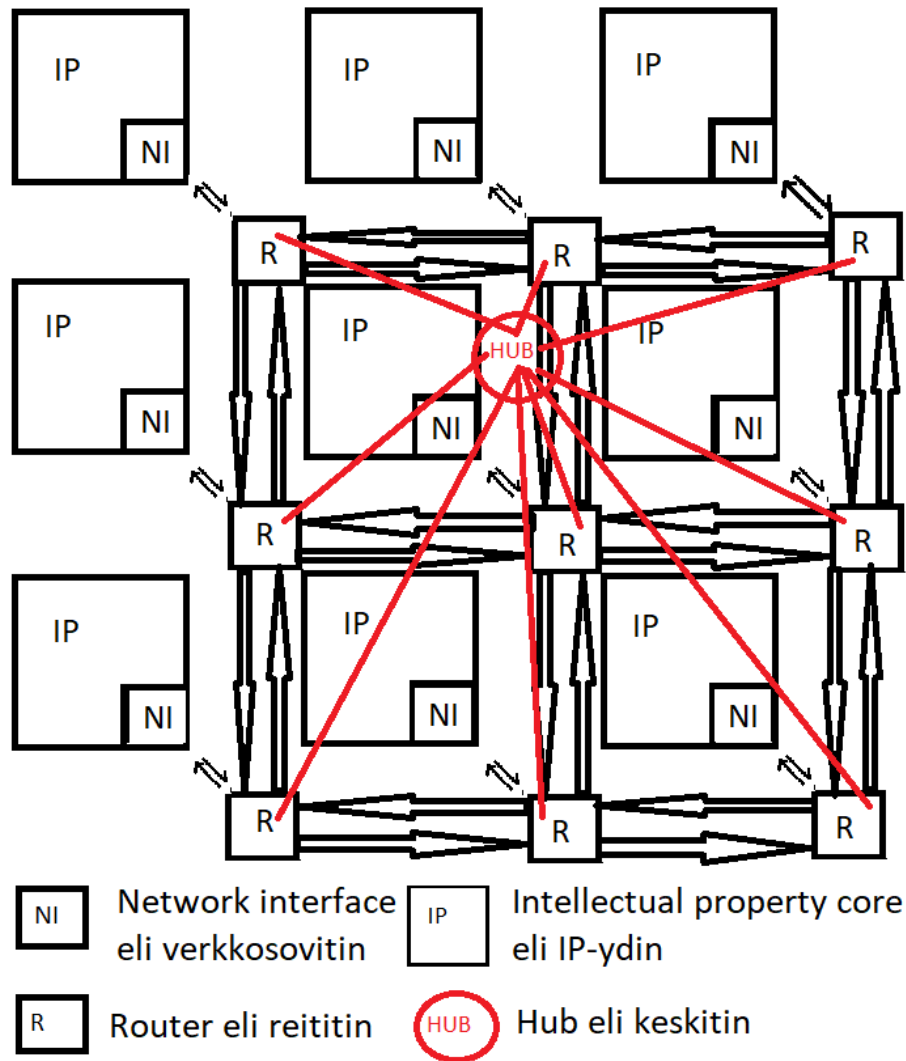
#### 3.1 Tiedonsiirtoteknologiat

Tässä aliluvussa käsitellään Network-on-Chipille suunniteltuja langattomia tiedonsiirtoteknologioita. Nämä teknologiat ovat langaton Network-on-Chip, eli radioalloilla toimiva Network-on-Chip, sekä fotoninen Network-on-Chip, joka puolestaan toimii optisesti fotoneilla.

##### 3.1.1 Langaton Network-on-Chip

Perinteisessä NoC-sovelluksessa yhteydet on toteutettu siten, että jokainen ytimen lähettämä signaali matkaa monen ytimen kautta johdinverkostoa pitkin ja monen kytkimen/reitittimen läpi päästäkseen lopulta määränpäähänsä. Tämä aiheuttaa suurta latenssia, läpisyöttöhäviötä sekä tehohäviötä. Ratkaisuna on kehitetty langaton rakenne, langaton Network-on-Chip eli Wireless Network-on-Chip (WiNoC), jossa yhteys kaukana toisistaan sijaitsevien ydinten välillä halutaan muodostaa käyttämällä mikroantenneja, jotka toimisivat kymmenistä gigahertseistä satoihin terahertzeihin. Näillä mikroantenneilla pystytään luomaan yhteyksiä suoraan ytimeltä ytimelle, jolloin ylimääräisten yhteyksien luoma latenssi ja tehohäviö poistuvat. [2] Vaikka langattomalla yhteydellä on suuri tehonkulutus, on silti energiatehokkaampaa luoda langaton yhteys, kuin käyttää pitkiä metallijohtimia [8].

Langattoman yhteyden luominen kaikkien ydinten välille ei ole kannattavaa johtuen langattoman yhteyden suuresta tehon kulutuksesta. Tästä syystä käytetään langatonta yhteyttä usein yhdistettynä perinteisten metallijohdinten kanssa. Tällaisessa ratkaisussa pystytään optimoimaan tehohäviö. Ganguly et al. ehdottavat ratkaisua, jossa keskittimet ovat yhteydessä toisiinsa sekä langattomalla että langallisella yhteydellä ja niiden aliverkoilla taas on käytössä pelkkä langallinen yhteys. Nämä langattomat keskittimet jakaisivat datapaketteja langattomasti toisilleen ja lopulta langallisesti omiin aliverkkoihinsa. Ytimet siis muodostaisivat aliverkkoja, joilla on yhteinen keskitin. Keskittimellä taas on langaton tukiasema. Jotta pystyisimme saavuttamaan suurinopeuksiset sekä matalaenergiset yhteydet, tulee langattomien keskittimien sijainti suunnitella tarkasti ja optimoida. [2] Tämä ehdotettu ratkaisu on kuvattu kuvassa 3. Kuvassa on siis mesh-topologiamuodossa aliverkko, jossa kukin ydin on langallisesti yhteydessä sekä toisiin ytimiin että keskittimeen. Tämä keskitin on yhteydessä muiden aliverkkojen keskittimiin langattomasti ja langallisesti.



Kuva 3: Kuvassa mesh-topologialla aliverkon langaton toteutus tyyliteltynä lähteen [2] mukaan.

Kun tutkimme WiNoCin suorituskykyä, tärkeäksi tutkittavaksi muodostuu, kuinka paljon dataa pystymme lähettämään ja millaisia kompromisseja se vaatii. Suurempi datanlähetyksenopeus vaatii yleensä suuremman tehon ja pinta-alan.

Fyysisen tason valinnat vaikuttavat lähetyksenopeuteen. Lähetyksenopeuden kaava  $R$  bitteinä sekunnissa (bps) on:

$$R=B \cdot S_E, \quad (1)$$

missä  $B$  on yhteyden kaistanleveys hertseinä (Hz) ja  $S_E$  on käytetyn modulaation spektritehokkuus bps/Hz. Tämän takia lähetyksenopeutta voi skaalata joko kasvattamalla kaistanleveyttä, mikä lisää tarvittua pinta-alaa ja tehonkulutusta lähes lineaarisesti, tai käyttämällä jotakin modulaatiota, jolla on suurempi spektritehokkuus, jolloin tuloksena voi olla epälineaarinen pinta-alan ja tehonkulutuksen kasvu johtuen monimutkaisemmasta modulaattorista. [1] Modulaattorilla tarkoitetaan laitetta tai piiriä,

joka suorittaa modulaation, eli signaalin muokkaamisen toisella signaalilla, jolloin alkuperäinen signaali saadaan haluttuun muotoon.

Abadalin ja kumppaneiden mukaan tällä hetkellä sirukommunikaatiot käyttävät noin 60 GHz taajuuksia yksinkertaisilla modulaatioilla, jotta pinta-ala ja teho voitaisiin minimoida. He antavat esimerkkiarvoja erilaisista yhteyksistä. Eräällä 65 nm CMOS-implementaatiolla saatiin On-Off-Keying-modulaatiota (OOK) käyttämällä 16 Gbps, jolloin tehonkulutus oli 31,2 mW ja käytetty pinta-ala 0,25 mm<sup>2</sup>. Kun kaistanleveys noin kolminkertaistettiin, saatiin yhteensä 48 Gbps lähetyksenopeus, mutta tehonkulutus oli 97,5 mW ja käytetty pinta-ala 0,73 mm<sup>2</sup>. Kun modulaatio vaihdettiin Quadrature Phase-Shift Keying-modulaatioksi (QPSK), jolla on noin kaksinkertainen spektritehokkuus, saatiin lähetyksenopeudeksi 32 Gbps, tehokulutukseksi 96 mW ja pinta-alarapeeksi 0,4 mm<sup>2</sup>. [1]

Abadal et al. tutkivat hybridi-NoCin, eli sekä langatonta että langallista yhteyttä käyttävän NoCin, ja perinteisen NoCin latensseja, kun ytimiä on 256 ja muuttivat lähetyksen kapasiteetin käyttöastetta. He huomasivat, että kun lähetyksen kapasiteetista käytettiin noin 10 prosenttia, oli hybridimallin latenssi 20 prosenttia pienempi, kun kapasiteetista käytettiin 50 prosenttia, latenssi oli puolet vähemmän, ja kun kapasiteetista käytettiin 70 prosenttia, latenssi oli hybridimallissa enää kolmannes perinteisen mallin latenssista. [1]

### *3.1.2 Fotoninen Network-on-Chip*

Kuten WiNoC, myös fotoninen NoC, eli Photonic NoC (PNoC) tähtää parantamaan NoCin korkeaa latenssia ja suurta tehonkulutusta. PNoC:n arkkitehtuuri rakentuu fotonisten reititinten ympärille. Peruseriaatteeltaan PNoC toimii siten, että radiosignaalin sijaan lähetetään valoa eri aallonpituuksilla, jotka kuljettavat digitaalista dataa. Näitä aallonpituuksia tulkitaan vastaanottajan päässä. Tämä tapahtuu siten, että jokainen lähetetty aallonpituus suodatetaan, minkä jälkeen aalto saapuu valokennoon, joka absorboi valon luoden siitä sähkövirtaa. Tämä sähkövirta muutetaan jännitteeksi ja vahvistetaan. Tästä jännitetasosta tulkitaan data. Tämän jälkeen data lähetetään edelleen kohti määränpäättään. [5]

Jotta saadaan moniytimistä korkea suorituskyky, pitää moniprosessointijärjestelmän, eli järjestelmän, jossa kaksi tai useampi ydin jakaa muistin, osata hyödyntää parhaalla tavalla sekä laskenta- että kommunikaatioresursseja. Pitää siis osata käyttää langallista ja langatonta yhteyttä parhaalla mahdollisella tavalla. Lyhyissä yhden hypyn yhteyksissä langallinen on yleensä kannattavampi ratkaisu kuin langaton. Pitkissä yhteyksissä, joissa tarvittaisiin useampia hyppyjä, parempi valinta taas on langaton yhteys. Hamdi et al. ehdottavat tutkimuksessaan järjestelmää, jossa yhdistyvät konseptit software-defined networks (SDN) eli ohjelmallisesti määritetty verkko sekä PNoC. Tämä esitetty järjestelmä koostuu neljästä alikerroksesta, jotka toimivat yhdessä tai erikseen. [9]

Käsittelytaso koostuu kahdesta pohjakerroksesta. Ensimmäinen on Optinen (kytkentä) kerros, joka sisältää Fan-in ja -out prosessit. Toinen kerros on Ohjauskerros, joka ohjaa fotonisten kytkinten toimintaa. Verkostotaso koostuu kahdesta päälimmäisestä kerroksesta. Fotoninen (reititys) kerros vastaa fotonisten kytkinten reitityksestä. Jokainen fotoninen kytkin on yhteydessä reitittimeen, joka suorittaa reitityslaskennat ja kytkinjaot neljännen kerroksen, eli Hallintakerroksen määrittämien ohjelmien mukaisesti. Hallintakerros siis ohjaa Fotonisen kerroksen toimintaa. [9]

Tasolle on myös määritelty yhteiset kommunikaatioprotokollat, jotta ne pystyisivät toimimaan yhdessä. Kolme pääprotokollaa ovat sovellustapahtuma -ajastin, verkkoliitäntätoimintoprosessorin eli Network Interface Function (NIF)-prosessorin kommunikaatio sekä NIF-verkoston kommunikaatio. [9]

Järjestelmälle tarvitaan kuitenkin myös ohjelmistohallintataso. Tämä taso on algoritmi ohjelmiston hallintaan. Algoritmin tehtävänä on varmistaa, että viestit reititetään oikein. Tämä toteutetaan kolmella päätoiminnolla. Ensiksi toteutetaan reititys, sitten kilpailu, joka toimii kanavanvaraustekniikkana, ja lopuksi laitteen asennus määrätyn osoittamisen mukaan. [9]

Tulokseksi tälle järjestelmälle, kun käytössä ei ollut ohjelmallisesti määritettyä hallintaa, Hamdi et al. [9] saivat, että kun viestin koko on alle 500 tavua, on tavallisen NoCin latenssi pienempi, johtuen fotonisen verkoston vaatimasta modulaatioprosessista. Kuitenkin viestin koon kasvaessa, paranee fotonisen NoCin suorituskyky tavalliseen verrattuna. Latenssi pienenee joissakin tapauksissa jopa 46 prosenttia ja energiankulutus taas joissakin tapauksissa jopa 78,5 prosenttia. Kun ohjelmallisesti määritetty hallinta otettiin käyttöön, saatiin latenssista ja energiankulutuksesta vielä 3 prosenttia pois. [9]

Nämä tulokset ovat vain yhden tapauksen osalta, mutta ne toimivat silti suuntaa antavina myös yleisesti. Kyseisen tutkimuksen osalta voidaan sanoa, että sekä PNoC että ohjelmallinen määrittäminen parantavat verkoston latenssia ja energiankulutusta. Vaikka näitä hyötyjä ei suoraan voi olettaa kaikille systeemeille, voidaan tuloksista kuitenkin nähdä käytetyn järjestelmän hyödyt.

Käytetty järjestelmä on kuitenkin vain yksi vaihtoehto. Yhtenä mahdollisena vaihtoehtona on ehdotettu järjestelmää, joka koostuu kolmekerroksisesta arkkitehtuurista. Siitä puuttuu ohjelmallinen määrittäminen. Ensimmäisenä kerroksena on fotoninen topologia,

jolle on rakennettu fotoniset reitittimet ja yhdyskäytävät. Toisena kerroksena on sähköinen ohjauskerros, jolle on sijoitettu reitittimet ja johtimet vastaamaan ensimmäistä kerrosta. Kolmantena kerroksena on prosessointikerros. [5] Fotonisella NoCilla on siis erilaisia ratkaisutapoja. Mikään niistä ei kuitenkaan ole vakiintunut tapa, joten tarvitaan lisätutkimuksia, jotta fotonisesta NoCista saadaan paras hyöty irti.

### 3.2 Arkkitehtuuri

Moniydinarkkitehtuuria on mahdollista rakentaa myös Z-suuntaan eli rakentaa useita tasoja päällekkäin, näin mahdollistaen lyhyemmät yhteydet. Kun Network-on-Chipiin sovelletaan 3D-rakennetta, kutsutaan sitä 3D-Network-on-Chipiksi (3D-NoC).

#### 3.2.1 3D-Network-on-Chip

3D-NoC:n tavoitteena on vähentää pitkien johdinten aiheuttamaa latenssia ja tehohäviötä. Tähän tavoitteeseen pyritään pääsemään käyttämällä 3D-rakennetta. Tällöin pystytään pienentämään 2D-rakennetta ja siten lyhentämään ydinten väliin tarvittavaa johdinta. [10] Lyhyemmät johtimet saavutetaan sillä, että johdin voi olla vain yhden ytimen verran Z-akselilla sen sijaan että johtimen pituus olisi mahdollisesti monen ytimen verran XY-akselilla. Tämä saavutetaan käyttämällä eri tasolle rakennettujen kerrosten yhdistämiseen Through Silicon Via (TSV) -läpivientejä [11]. Kun yhdistetään 3D-integraatio ja NoCin skaalautuvuus, voidaan saavuttaa pienempi yhteysviive, kanoninen yhteysrakenne, lisääntynyt mukautuvuus sekä mahdollisuus integroida erilaisia systeemejä ja teknologioita [10]. Lisäksi 3D-NoCin rakenne mahdollistaa ydinten välisten hyppyjen vähentämisen, mikä taas entisestään laskee verkoston latenssia [11].

3D-NoC:lla on kuitenkin omat haasteensa. Erilaisia ongelmia, jotka vielä odottavat ratkaisuaan on esimerkiksi ylikuuluminen, joka aiheuttaa systeemeihin erilaisia virheitä. Lisäksi tämän hetken reititysmallit omaavat yleensä haittoja, jotka johtavat kasvaneeseen pinta-alarapeeseen sekä tehonkulutukseen, joista kumpikaan ei ole haluttu ominaisuus on-chip-verkostoissa. [12] 3D-NoC käyttää 3D-reitittimiä, joilla on suuri tehonkulutus. Tämä tehonkulutus myös aiheuttaa suurta lämmöntuottoa, joka heikentää reitittimien suorituskykyä. 3D-NoCin lämpöhaasteisiin on mahdollisia ratkaisuja. Ensimmäinen on niin kutsuttu Approximate communication -tekniikka, jossa paketeista poistetaan osia, joita algoritmi ei arvioi tarvittavan. Tämä johtaa vähentyneeseen verkoston rasitukseen ja pienempään latenssiin. Toinen keino on jänniteskaalaus. Siinä komponenttien jännitetasoja säädelään tarpeen mukaan, mikä johtaa alentuneeseen kokonaistehonkulutukseen ja sitä kautta laskevaan lämmöntuottoon. Nämä tekniikat kuitenkin aiheuttavat kasvua bittivirhesuhteessa. Näin ollen suorituskykyä ja tehonkulutusta parannetaan, mutta virhetarkkuus kärsii hieman. [11]

3D-rakenteisen NoCin suorituskyvyn vertaileminen perinteiseen 2D-rakenteeseen on suhteellisen yksiselitteistä. On luonnostaan selvä asia, että kun rakennetaan korkeampia rakenteita, saadaan noin samalla pinta-alalla enemmän toiminnallisia yksiköitä. Feero et al. selvittivät tutkimuksissaan 3D-rakenteen vaikutusta latenssiin ja pinta-alaan erilaisilla topologioilla. Kun he vertailivat Butterfly Fat Tree (BFT) -topologiaa, eli eräänlaista puuarkkitehtuurillista topologiaa käyttävää 2D- ja 3D-rakennetta, he saivat samalle pinta-alalle ja samalle energiankulutukselle yli puolitettua pisimmän johtimen viiveen. [13]

Pavlidis ja Friedman tutkivat 2D- ja 3D-rakenteisten arkkitehtuurien latenssia ja tehonkulutusta. He huomasivat, että kun topologioita vertasi 2D- ja 3D-rakenteina käyttäen 128 ja 256 solmukohtaa, saatiin latenssi laskettua 40 prosenttia ja 36 prosenttia eri solmukohtamäärillä ja tehonkulutusta 62 prosenttia ja 58 prosenttia. [14]

Tiivistetyksi 3D-arkkitehtuuria käyttämällä saavutetaan tutkitusti hyötyjä niin latenssin, käytetyn pinta-alan kuin tehonkulutuksenkin suhteen. Kuitenkin 3D-arkkitehtuurilla on myös omat ongelmansa, sillä tiivis rakenne aiheuttaa ylikuulumista ja lämmön kerääntymistä. Näitä ongelmia voi ratkaista erilaisilla tekniikoilla, mutta yleensä se tapahtuu jonkinlaisella vaihdolla. Esimerkiksi lämmöntuottoa voidaan vähentää, mutta silloin bittivirhesuhde kasvaa hieman.

### *3.2.2 Langaton 3D-rakenne*

3D-rakennetta on ehdotettu käytettäväksi myös yhdessä langattomien ratkaisuiden kanssa. Feng et al. tutkivat yhdistetyn WiNoCin ja 3D-NoCin suorituskykyä. He käyttivät toteutuksessaan ratkaisua, jossa aliverkoilla on langaton tukiasema, johon ytimet ovat langallisesti yhteydessä. Nämä tukiasemat ovat toisiinsa langattomasti yhteydessä. Eri kerroksien väliset yhteydet on toteutettu pääosin näiden langattomien tukiasemien kautta. TSV-yhteyksiä, eli piikiekon läpi vietyjä yhteyksiä käytetään lähinnä virtalähteenä ja kelloverkoston ylläpitämiseen. Tämä järjestelmä on voitu toteuttaa erityisten 3D-reitittimien avulla, joilla on seitsemän porttia. Viisi niistä on langallisia yhteyksiä varten ja kahta käytetään langattomiin yhteyksiin kerrosten välillä. Yksi langallisista porteista on reitittimen ja ytimen välille, neljä muuta langallista porttia ovat yhteydessä saman aliverkon reitittimiin. [15] Tämä mahdollistaa reitittimelle yhteyden pidon joka suuntaan samalla ylläpitäen langatonta yhteyttä.

Feng et al. käyttivät tutkimuksessaan järjestelmää, jossa käytettiin kolmea kerrosta, kullakin 16 ydintä jaettuna neljään aliverkkoon. He analysoivat saamiaan tuloksia

vertaamalla niitä vastaavan TSV-yhteyksillä rakennettuun 3D-NoCiin. Tutkimuksessa selvisi, että kun datan määrä kasvaa, saadaan langatonta yhteyttä hyödyntämällä selkeitä hyötyjä niin datansiirtonopeuden kuin latenssin osalta. Parhaimmillaan siirtonopeus parani jopa 125 prosenttia. Latenssi taas parani parhaimmillaan yli 500 prosenttia. [15] Nämä tulokset lupaavat huomattavia hyötyjä langattoman teknologian ja 3D-rakenteen yhdistämisestä.

Myös fotonisen NoCin ja 3D-rakenteen yhdistämistä on tutkittu. Zhixun et al. ehdottavat ratkaisua PNoCin ja 3D-NoCin yhdistämiseksi. Heidän ratkaisussaan on kolme kerrosta ja jokaisella kerroksella on oma valosähköinen kerroksensa yhteistä kartoitusta varten. Nämä kolme kerrosta ovat yhteydessä toisiinsa TSV-yhteydellä. Järjestelmä toimii siten, että reitittimet toimivat kahden reitittimen kokonaisuuksina. Ensimmäinen reititin yhdistää ytymet toisiinsa, kun taas toinen reititin huolehtii yhteyksistä kerrosten välillä. Tämä vähentää yksittäisten optisten reitittimien monimutkaisuutta sekä vähentää tietojen menetystä ja ylikuulumista. Lisäksi järjestelmän suorituskykyä on parannettu algoritmilla, jolla pyritään parantamaan järjestelmän luotettavuutta. [16] Järjestelmän suorituskykyä ja ominaisuuksia on vaikea analysoida, koska tutkimuksessa ei ole arvoja vastaavasta järjestelmästä ilman optista yhteyttä. Tutkimuksen olemassaolo kuitenkin kertoo, että PNoCin ja 3D-rakenteen yhdistelmiä tutkitaan, ja niiden potentiaalia moniydinarkkitehtuurin tiedonsiirtoratkaisuna pyritään selvittämään ja kehittämään.

## 4 Yhteenveto

Tässä työssä on esitelty kaksi Network-on-Chipin uutta teknologiaa sekä arkkitehtuurivaihtoehto: Langaton Network-on-Chip, fotoninen Network-on-Chip sekä 3D-arkkitehtuuri. Lisäksi tarkasteltiin hieman ohjelmallisen määrittämisen vaikutusta PNoCin ominaisuuksiin sekä teknologioiden ja 3D-arkkitehtuurin yhdistämistä. Kustakin tutkimisen kohteesta käytiin läpi perusrakenne sekä tutkittiin, millaisia vaikutuksia niillä on latenssiin ja tehonkulutukseen. Uusien teknologioiden ja arkkitehtuurin lisäksi tutkittiin, millainen on NoC, jonka pohjalta nämä uudet teknologiat ja arkkitehtuuri on suunniteltu. Tarkoituksena tällä oli selvittää lukijalle yhteistä pohjaa, jolle uudet ratkaisut on suunniteltu.

Network-on-Chip on moniydinarkkitehtuurin käytetyin yhteysteknologia. Sen pääperiaatteena on rakentaa verkosto, jossa informaation on mahdollisimman nopeaa siirtyä ytimeltä toiselle sekä muistiin. NoC kuitenkin kärsii skaalautuvuusongelmista. Ydinten määrän kasvaessa ja niiden etäisyyden kasvaessa myös johdinten pituus kasvaa aiheuttaen suurta latenssia ja tehonkulutusta. [1] Ydinten määrän jatkuvan lisääntymisen takia on kuitenkin entistä tärkeämpää, että NoCin skaalautuvuusongelmiin löydetään ratkaisu. Tämän takia onkin esitetty uusia langattomia teknologioita ja arkkitehtuuri, joilla pyritään ratkaisemaan nämä skaalautuvuusongelmat.

Ensimmäisenä uutena teknologiana työssä käsiteltiin radioaalloilla toimiva langaton teknologia, langaton Network-on-Chip, englanniksi Wireless Network-on-Chip. Teknologia toimii yhdistelmänä langattomia yhteyksiä ja langallisia yhteyksiä, ja tärkeänä osana järjestelmän suunnittelua on optimoida eri yhteyksien käyttö. Tämä johtuu siitä, että suuren tehonkulutuksen langattomia reitittimiä kannattaa hyödyntää vain pidemmissä yhteyksissä. [2] Työssä tarkasteltiin myös yhtä mahdollista ratkaisua WiNoCista. Tämä ratkaisu antoi varsin lupaavia tuloksia. Käyttämällä Gangulyn et al. esittämää ratkaisua saavutettiin merkittävää latenssin pienenemistä, ja muuttamalla modulaatiota ja kaistanleveyttä pystyttiin säätelemään datansiirtonopeutta ja tehonkulutusta [2]. Nämä ominaisuudet optimoimalla tarpeen mukaan voidaan luoda järjestelmiä, jotka on optimoitu tarpeita vastaaviksi.

Toisena teknologiana käsiteltiin fotoneilla toimiva langaton tekniikka, fotoninen Network-on-Chip. Tässä teknologiassa yhdistyvät perinteinen NoC sekä fotoniset reitittimet, jotka välittävät langattomasti signaalia reitittimeltä toiselle. Myös fotonisessa NoCissa on tärkeää, että perinteisten yhteyksien ja langattomien yhteyksien optimoiminen on tehty oikein, jotta latenssi ja tehonkulutus saadaan optimoitua. [9] Työssä perehdyttiin ratkaisuun, jossa yhdistettiin fotoninen NoC ja ohjelmallinen määrittäminen. Työssä havaittiin, että kun viestin koko ylittää 500 tavun koon, on kannattavampaa käyttää fotonista langatonta yhteyttä kuin perinteistä metallijohdinta. Lisäksi työssä havaittiin, että joissakin tapauksissa latenssia saatiin pienennettyä jopa 46 prosenttia ja tehonkulutusta 78,5 prosenttia. Ohjelmistollisella määrittämisellä saatiin sekä latenssista että tehonkulutuksesta vielä 3 prosenttia pois.



Kolmantena käsiteltiin 3D-rakenteista NoCia. 3D-rakenteella tarkoitetaan XY-koordinaattien lisäksi rakentamisesta Z-suuntaan, eli ylöspäin. Tämä mahdollistaa ydinten välisten johdinten pituuden lyhenemisen, mikä taas johtaa niiden tuottaman latenssin ja tehonkulutuksen pienenemiseen [10]. Lisäksi 3D-reitittimien käyttö vähentää ydinten välisten hyppyjen määrää, mikä johtaa pienempään latenssiin [11]. 3D-rakenne tuo mukanaan omat ongelmansa, joita ovat ylikuuluminen, huonot reititysmallit sekä suuri lämmöntuottaminen [12], [11]. Ylikuuluminen aiheuttaa mahdollisia signaalivirheitä, huonot reititysmallit taas lisäävät latenssia ja tehonkulutusta [12]. Suuri lämmöntuottaminen aiheuttaa järjestelmän kuumenemista, jolloin se ei pysty toimimaan täydellä suorituskyvyllä. Ongelmiin on kehitetty ratkaisuja. Esimerkiksi suureen lämmöntuottoon on kehitetty tekniikoita, joilla joko vähennetään järjestelmän tekemää työtä Approximate communication -tekniikan avulla, jossa algoritmi poistaa turhiksi arvioimiaan paketteja, jolloin reitittimien tulee hallita vähemmän paketteja. Tämä pienentää niiden tehonkulutusta ja siten lämmöntuottoa. Toinen tekniikka on jänniteskaalaus. Siinä ajatuksena on säätää komponenttien jännitetasoja tarpeen mukaan. Tämä laskee komponenttien kokonaistehonkulutusta ja siten myös lämmöntuottoa. Nämä tekniikat kuitenkin kasvattavat bittivirhesuhdetta. [11] Työssä tutkittiin tutkimusta, jossa vertailtiin 2D- ja 3D-rakenteisten topologioiden latenssia ja tehonkulutusta. Parhaimmillaan latenssia saatiin laskettua 40 prosenttia käyttämällä 3D-rakennetta. Tehonkulutusta taas saatiin laskettua parhaimmillaan 60 prosenttia.

3D-rakennetta tutkittiin myös yhdistettynä edellä mainittuihin langattomiin teknologioihin. Radioaalloilla toimivan langattoman teknologian eli WiNoCin kohdalla tulokseksi saatiin selkeitä parannuksia niin latenssin kuin datansiirtonopeuden kohdalla, mikä lupaa hyvää yhdistelmän tulevaisuuden kannalta [15]. Fotonisen NoCin kohdalla ei saatu vastaavia suorituskykytuloksia johtuen vaikeudesta verrata tuloksia, mutta tutkimus kuitenkin ilmaisee sen, että asiaa tutkitaan ja sen mahdollisuuksia kartoitetaan. Tutkimuksista voidaan päätellä, että asia on mahdollista toteuttaa, ja sillä on tutkijoiden kiinnostus.

Voidaan siis todeta, että kutakin työssä käsiteltyä teknologiaa ja arkkitehtuuria voidaan käyttää NoCin ominaisuuksien parantamiseksi. Pitää kuitenkin osata valita oikea arkkitehtuuri sekä teknologia tilanteen ja tarpeen mukaan. Lisäksi tarvitaan uusia tutkimuksia, jotta niiden haasteisiin löydetään ratkaisuja ja niiden käyttö saadaan optimoitua. Network-on-Chipin tulevaisuudessa on kuitenkin näkyvissä varsin lupaavia ratkaisuja.

## Lähdeluettelo

- [1] Abadal, S., Mestres, A., Nemirovskym M., Lee, H., Gonzalez, A., Alarcon, E., Cabellos-Aparicio, A. (2016). Scalability of Broadcast Performance in Wireless Network-on-Chip. *IEEE Transactions on Parallel and Distributed Systems*, 27(12), s. 3631-3645. <https://doi.org/10.1109/TPDS.2016.2537332>
- [2] Ganguly, A., Chang, K., Deb, S., Pande, P. P., Belzer, B., Teuscher, C. (2011). Scalable Hybrid Wireless Network-on-Chip Architectures for Multicore Systems. *IEEE Transactions on Computers*, 60(10), s. 1485-1502. <https://doi.org/10.1109/TC.2010.176>
- [3] Benini, L. & De Micheli, G. (2002). Networks on chips: a new SoC paradigm. *Computer*, 35(1), s. 70-78. <https://doi.org/10.1109/2.976921>
- [4] Kumar, S., Jantsch, A., Sojininen, J.-P., Forsell, M., Millberg, M., Oberg, J., Tiensyrja, K., Hemani, A. (2002). *A network on chip architecture and design methodology. Proceedings IEEE Computer Society Annual Symposium on VLSI. New Paradigms for VLSI Systems Design. ISVLSI*, s. 117-124. <https://doi.org/10.1109/ISVLSI.2002.1016885>
- [5] Asadi, B., Zia, S. M., Al-Khafaji, H. M. R., Mohamadian, A. (2023). Network-on-Chip and Photonic Network-on-Chip Basic Concepts: A Survey. *J Electron Test*, 39, s. 11-25. <https://doi-org.libproxy.tuni.fi/10.1007/s10836-023-06046-x>
- [6] Achballah, A. B. & Saoud, S. B. (2013). A Survey of Network-On-Chip Tools. *International Journal of Research and Reviews in Computer Science*, 4(9), s. 61-67. <https://doi.org/10.14569/IJACSA.2013.040910>
- [7] Tsai, W.-C., Lan, Y.-C., Hu, Y.-H., Chen, S.-J. (2011). Networks on Chips: Structure and Design Methodologies. *Journal of Electrical and Computer Engineering*. <https://doi.org/10.1155/2012/509465>
- [8] Devanathan, M., Ranganathan, V., Sivakumar, P. (2020). Congestion-aware wireless network-on-chip for high-speed communication. *Automatika*, 61(1), s. 92-98. <https://doi.org/10.1080/00051144.2019.1674511>
- [9] Hamdi, D. A., Ghoniemy, S., Dakroury, Y., Sobh, M. A. (2021). A Scalable Software Defined Network Orchestrator for Photonic Network on Chips. *IEEE Access*, 9, s. 35371-35381 <https://doi.org/10.1109/ACCESS.2021.3058238>
- [10] Tatas, K., Siozios, K., Soudris, D., Jantsch, A. (2014). Designing 2D and 3D Network-on-Chip Architectures. *Springer New York, NY*. <https://doi-org.libproxy.tuni.fi/10.1007/978-1-4614-4274-5>
- [11] Momeni, M. & Shahhoseini, H. S. (2022). Energy efficient 3D network-on-chip based on approximate communication. *Computer Networks*, 203. <https://doi.org/10.1016/j.comnet.2021.108652>
- [12] Tyagi, S., Maheshwari, P., Agarwal, A., Avasthi, V. (2017). Exploring 3D Network-

- on-Chip Architectures and Challenges. *International Conference on Computer and Applications (ICCA), Doha, Qatar*, s. 97-101. <https://doi.org/10.1109/COMAPP.2017.8079768>
- [13] Feero, B. S. & Pande, P. P. (2009). Networks-on-Chip in a Three-Dimensional Environment: A Performance Evaluation. *IEEE Transactions on Computers*, 58(1), s. 32-45. <https://doi.org/10.1109/TC.2008.142>
- [14] Pavlidis, V. F. & Friedman, E. G. (2007). 3-D Topologies for Networks-on-Chip. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 15(10), s. 1081-1090. <https://doi.org/10.1109/TVLSI.2007.893649>
- [15] Feng, Q., Ban, D., Li, H., Han, G., Dou, W. (2012). Performance analysis of wireless 3D network on chip. *International Symposium on Instrumentation & Measurement, Sensor Network and Automation (IMSNA), Sanya, China*, s. 232-235. <https://doi.org/10.1109/MSNA.2012.6324556>
- [16] Zhixun, L., Chuanpei, X., Lvqing, B., Yunying, S., Yunfei, Y., Hu, C. (2022). Modeling and performance analysis of a fault-tolerant 3D photonic network-on-chip based on hybrid Photonics–Plasmonics. *Computational Intelligence and Neuroscience*, <https://doi.org/10.1155/2022/9615610>