

Internetin ja WWW:n teknologioiden lyhyt historia

TURUN YLIOPISTO
Tietotekniikan laitos
TkK-tutkielma
Tietotekniikka
Kesäkuu 2025
Juho Nykänen

TURUN YLIOPISTO
Tietotekniikan laitos

JUHO NYKÄNEN: Internetin ja WWW:n teknologioiden lyhyt historia

TkK-tutkielma, 26 s.
Tietotekniikka
Kesäkuu 2025

Internet ja WWW ovat kaikille tunnettuja teknologioita. Niitä käytetään ympäri maailmaa, ja nykyään maailma pyörii niiden ympärillä. Yleinen tietämys niistä kuitenkin perustuu nykymaailman versioon, jolloin historian osuus jää useasti ihmisille tietämättä. Ilman historiaa niistä jää keskeneräinen ymmärrys, jonka takia niiden nykyistä versioita on vaikea ymmärtää.

Tämä tutkielma on kirjallisuuskatsaus Internetin ja WWW:n historiaan. Katsauksen tarkoituksena on selvittää niiden kehityksen alkua ja etenemistä. Internetistä käsitellään sen varhaisia konsepteja, patettikytkennän taustaa, ja ARPANET-tietokoneverkkoa. Käsittely rajataan TCP/IP-protokollapinoon, jonka standardisointi merkkasi modernin Internetin alun. WWW:ssä keskitytään sen nopeaan alkukehitykseen. Molemmista otetaan niiden keskeisimmät teknologiat, ja selvitetään niiden toiminnallisuudet. Lopputuloksena on tiivis historiakatsaus Internettiin ja WWW:hen, joita käsitellään tutkielmassa kronologisessa järjestyksessä.

Asiasanat: Internet, WWW, historia, kehitys, teknologian toiminnallisuus

Sisällys

1	JOHDANTO	1
2	INTERNETIN HISTORIA	3
2.1	Pakettikytkennän alkuideat (1960-1964)	3
2.2	Internetin juuret ja ensimmäinen laajaverkko (1962-1965)	5
2.3	ARPANETin alkuvaiheet (1966-1967)	6
2.4	ARPANETin perustaminen (1968-1969)	7
2.5	TCP/IP kehitys (1972-1983)	8
3	WWW:N HISTORIA	11
3.1	Enquire sovellus	11
3.2	WWW:n alkuvaiheet (1988-1989)	11
3.3	WWW:n kehitys (1990)	12
3.4	WWW:n kasvu (1991->)	14
4	Keskeiset teknologiat	16
4.1	Internetin teknologiat	16
4.2	WWW:n teknologiat	22
5	Yhteenveto	25
	Lähdeluettelo	27

1 JOHDANTO

Internet ja WWW (World Wide Web) ovat informaatiojärjestelmiä, joita ihmiset käyttävät päivittäin. Internet mahdollistaa kommunikaation ympäri maailmaa, ja WWW antaa käytännöllisen tavan jakaa informaatiota ihmisten välillä. Molemmat ovat hyvin olennaisia asioita ihmisten arkielämässä. Näistä järjestelmistä opetetaan yleensä niiden nyky maailman kontekstin kautta, jolloin useasti jää taustalle niiden historia. Ymmärtääkseen Internetiä ja WWW:tä paremmin, on hyvä tietää mistä ne ovat lähteneet, ja miten ne ovat toimineet niiden varhaisissa versioissa.

Tutkielman tarkoituksena on luoda tiivistetty historiakatsaus Internetistä, WWW:stä, ja niiden keskeisistä teknologioista. Olen tutkielmaa varten lähtenyt selvittämään miten Internetin ja WWW:n kehitys tapahtui, ja miten niiden keskeiset teknologiat toimivat. Olen rajannut teknologioiden toiminnallisuuden selvitystä, riippuen kuinka keskeisiä ne ovat olleet Internetin ja WWW:n kehityksessä.

Tutkielman aihetta ja rajaamista varten olen valinnut seuraavat tutkimuskysymykset:

- **TK1:** Miten Internetin ja WWW:n kehitys sai alkunsa ja eteni niiden alkuvaiheissa?
- **TK2:** Mitkä olivat keskeisiä teknologioita Internetin ja WWW:n kehityksessä?
- **TK3:** Miten nämä keskeiset teknologiat toimivat?

Ensimmäisen tutkimuskysymyksen tarkoituksena on selvittää Internetin ja WWW:n

kehitys niiden alussa, ja rajoittaa se alkuun. Toisella tutkimuskysymyksellä otetaan kohteeksi tärkeät teknologiat kehityksestä. Kolmannella tutkimuskysymyksellä avataan löydettyjen teknologioiden toiminnallisuus.

Tutkielma oli kirjallisuuskatsaus, jota varten lähdin etsimään tutkimuskysymyksiin vastauksia. Tutkielman tarkoituksena ei ole vertailla kaikkia aiheeseen löydettyä aineistoa, vaan hyvälaatuiset informaatiolähteet historiasta riittivät. Tiedonhakuprosessi jakautui kolmeen eri aiheeseen: Internetin historia, WWW:n historia, ja niiden keskeisten teknologioiden toiminnallisuuteen. Tiedonhakua tein tietokannoista Volter ja Web of Science. Internetin ja WWW:n historiaa varten käytin yksinkertaisia hakulauseita, kuten "(www OR world wide web OR web) AND (history OR development)". Hakulauseilla en löytynyt suoraan aiheesta paljon aineistoa, mutta löytyi muutama, jotka olivat tarpeeksi. Internetin historiaa varten käytin tietoverkkotekniikat kurssin täydentävää lukumateriaalia [1], ja siitä löydettyä lähdettä [2]. Historian osuuden jälkeen pystyin teknologioiden toiminnallisuutta varten etsimään tarkasti niiden kehityshetken lähivuosilta. Yleinen teema kaikkien aiheiden aineistoissa oli, että useasti niiden tekijänä oli päähenkilöitä Internetin ja WWW:n kehityksen ajalta.

Luvussa 2 käsittelen Internetin kehityksen alkua, ja sen etenemistä vuosien aikana. Luvussa 3 käsittelen WWW:n lyhyttä kehitystä. Luvussa 4 perustelen keskeisten teknologioiden valintoja, ja käsittelen niiden toiminnallisuutta. Luku 5 on yhteenveto, jossa olen vastannut tutkimuskysymyksiin lyhyesti, reflektoinut tutkielmaa, ja ehdottanut jatkotutkimusaiheita.

2 INTERNETIN HISTORIA

Internetin kehitys alkoi varhaisista tietokoneverkoista, ja niiden taustalla olevista konsepteista ja teknologioista. Tutkielmaa varten olen valinnut alkupisteeksi ARPANETin, ja siihen johtavat tekijät. ARPANETin luominen, ja sen jälkeinen kehitys, johti nykyään tunnettuun Internettiin. Loppupisteeksi olen rajannut TCP/IP-protokollapinon, jonka standardisointi merkkasi alun modernille Internetille. Tämän kehityksen aikaväliä käsittelen tässä luvussa kronologisessa järjestyksessä.

Luvussa mainittu Yhdysvaltojen DARPA:n (Defense Advanced Research Projects Agency) alkuperäinen nimi oli ARPA (Advanced Research Projects Agency), mutta vuosien aikana organisaatio on vaihtanut nimeänsä edestakaisin [3]. Käytän luvussa heidän nykyistä nimeä DARPAa.

2.1 Pakettikytkennän alkuideat (1960-1964)

Yhdysvaltojen Dod (Department of Defense) halusivat 1950-luvun loppuvaiheissa kestävämmän kommunikaatioverkon, joka kestäisi ydinsotaa. Siinä aikana DoD käytti julkisia puhelinverkkoja. Soittaakseen lankapuhelimien välillä, yhteyden piti ensimmäiseksi mennä puhelinkeskukseen, joka sitten ohjasi yhteyden eteenpäin korkeamman tason puhelinkeskukseen, tai haluttuun lankapuhelimeen. Puhelinkeskusten välinen hierarkia osoitti haavoittuvuuksia, josta yhdenkin menettäminen saattaisi irrottaa puhelinverkkoja toisistaan. [1]

Suunnilleen vuonna 1960, ratkaistaakseen tämän ongelman, DoD palkitsi urak-

kasopimuksen RAND yhtiölle. Yksi heidän työntekijöistään, Paul Baran, teki suunnitelman uudentlaisesta järjestelmästä. Suunnitelmassa käytettiin hajautettua puhelinverkkoa, missä puhelinkeskukset olisivat yhteydessä moneen keskukseen samanaikaisesti, ja puhelinyhteys voisi mennä monta erilaista reittiä. Suunnitelmaa varten hän kehitti konseptin digitaalisesta pakettikytkentäisestä teknologiasta. Konseptin testausta varten pyydettiin AT&T:tä rakentamaan prototyyppi, joka oli siinä aikana Yhdysvaltojen puhelinverkon monopoli. AT&T hylkäsi sen ehdotuksen, jonka jälkeen Baranin idea uudesta puhelinverkosta lopetettiin. [1]

Artikkeli internetin historiasta kertoo, että pakettikytkennän teoriasta on kirjoittanut Leonard Kleinrock MIT:stä (Massachusetts Institute of Technology) [3]. Hän julkaisi ensimmäisen paperin aiheesta vuonna 1961, ja myöhemmin kirjan vuonna 1964 [3]. Artikkelissa mainitaan, että Kleinrock oli vakuuttanut Lawrence G. Robertsin teoreettisesta mahdollisuudesta käyttää paketteja, eikä piirejä kommuunkaatiassa [3]. Roberts toimi päähenkilönä ARPANET:in luomisessa, jossa käytettiin pakettikytkentäistä verkkoa [2]. John Naughton kertoo kirjassaan - joka kertoo internetin historiasta - että Roberts väittää Kleinrockin kirjoittaneen ensimmäisen paperin paketteihin liityen, ja jatkoi sen jälkeen kirjoittamalla kirjan aiheesta [2]. Naughtonin tarkistaessa nämä kaksi julkaisua, niissä ei ole kuitenkaan mainintaa pakettikytkennästä [2]. Kleinrock itse väittää, että hänen kirjassaan idea paketeista oli implisiittinen [2]. Lähteistä saan ymmärryksen, että Kleinrock on vaikuttanut Robertsin ymmärrykseen paketeista, ja myöhemmin mainittu NPL järjestelmä - joka käytti Baranin työtä - näytti tämän konkreettisemmin hänelle. Selvittääkseen kenelle antaa kunnia pakettikytkennän keksimisestä menee tutkielman ulkopuolelle, mutta molemmilla on ollut vaikutusta ARPANET:in luontiin.

2.2 Internetin juuret ja ensimmäinen laajaverkko (1962-1965)

Ensimmäinen kuvauksen Internetistä oli luonut J.C.R. Licklider MIT:stä vuonna 1962. Hän oli kirjoittanut muistilapuille konseptin "Galactic Network":istä, jossa hän visioi tietokoneita yhdistettyinä toisiinsa ympäri maailmaa, joilla pystyisi nopeasti etsiä tietoa ja ohjelmia mistä vaan sivustosta. Hänen aikainen konsepti tietokoneverkosta muistuttaa hyvin paljon tämänpäiväistä Internetiä. [3]

Licklider toimi ensimmäisenä johtajana DARPAN tietokonetutkimusohjelmassa IPTO (Information Processing Techniques Office) osastolla [3]. Hänen aikana siellä, hän oli vakuuttanut hänen seuraajiaan DARPassa tästä tietokoneverkon konseptin tärkeydestä [3]. Kyseiset seuraajat olivat Ivan Sutherland, Bob Taylor, ja Lawrence G. Roberts [3]. Sutherland, joka toimi Lickliderin seuraajana IPTO:ssa, kutsui Taylorin NASA:lta työskentelemään varajohtajana Sutherlandille vuonna 1965. Parin kuukauden jälkeen kuitenkin Taylor päätyi IPTO:n johtajaksi, kun Sutherland lähti osastolta. [2]

Selvittääkseen puhelinlinjojen kapasiteetin kantaa tietoa pidemmällä matkalla, Taylor aloitti tätä varten kokeellisen testin vuonna 1965 [2]. Testiä suorittamaan hän pyysi Robertsia, joka siinä hetkellä vielä työskenteli MIT:ssä DARPAN rahoituksella [2]. Yhteistyössä Thomas Merrillin kanssa, Roberts yhdisti TX-2 tietokoneen Massachusettsissa Q-32 tietokoneeseen Californiassa, käyttäen puhelinverkkoyhteyttä. Lopputuloksena huomattiin konseptin toimivan hyvin, mutta piirikytkennällä toimivat puhelinlinjaverkot eivät ole luotettavia viestintää varten. Kyseinen tietokoneverkko oli ensimmäinen rakennettu laajaverkko. [3]

2.3 ARPANETin alkuvaiheet (1966-1967)

Ennen ARPANET:in luontia, tietokoneilla oli mahdollisuus kommunikoida toistensa kanssa. Tätä kommunikaatiota varten piti olla suora yhteys tietokoneesta toiseen. Siinä aikana tietokoneita käytettiin erilaisten prosessien laskentaan, ja tätä varten oli käytössä tietokoneiden osituskäyttö. Osituskäytön tarkoituksena oli tietokoneen resurssien jakaminen monelle käyttäjälle. Osituskäyttöjärjestelmää varten yhdellä keskustietokoneella oli monta terminaalia, josta sitten pystyi päästä käsiksi keskustietokoneeseen. Näissä oli kuitenkin ongelmana, että osituskäyttöjärjestelmät olivat eristetty toisistaan. [2]

Kun Taylorista tuli IPTO:n johtaja, hän huomasi hänen toimistossaan kolme eri terminaalia, jotka kaikki yhdistivät eri osituskäyttöjärjestelmään. Hänelle ilmiselvä asia oli linkittää nämä yhteensopimattomat laitteet toisiinsa jollain tavalla. Tätä varten hän aloitti projektin ARPANET-tietokoneverkosta vuonna 1966. Hän sai rahoitusta projektille DARPA:n johtajalta, jonka jälkeen hän kutsui Robertsin toimimaan projektin ohjelmapäällikkönä. [2]

Roberts aloitti suunnittelemaan tätä tietokoneverkkoa. Alkuperäisenä ideana oli käyttää osituskäyttöön tarkoitettuja tietokoneita solmupisteinä, jossa ne alkuperäisen toiminnallisuuden lisäksi myös ohjaisivat tietokoneverkkoa. Suunnitelma sai kritiikkiä tästä idesta, mutta Clark Wenley ehdotti ratkaisua. Hänen ehdotuksessaan tietokoneiden väliin rakennettaisiin aliverkko, joka hoitaisi kommunikaation solmujen välillä, ja jätettäisiin osituskäyttöön tarkoitettut tietokoneet ulkopuolelle tästä aliverkosta. Uudessa suunnitelmassa aliverkkoa varten käytettiin minitietokoneita nimeltä IMP (Interface message processors). Robertsin alkuperäinen idea niiden toiminnallisuudelle vastasi aikasempaa laajaverkon testiä, jossa käytettiin piirikytkeistä puhelinlinjaverkkoa. [2]

Roberts esitti hänen suunnitelmansa konferenssissa vuonna 1967 [4]. Hänen yllätykseen samassa konferenssissa oli toinen paperi, jossa selitettiin pakettikytken-

täisestä verkosta. Paperin tekijänä oli Donald Davies ja Roger Scantlebury NPL:stä (National Physical Laboratory), joka sijaitsee Englannissa [3]. Paperissa oleva pakettikytkentäinen tietokoneverkko oli luotu jo NPL:ssä Daviesin johtamana, nimellä NPL järjestelmä. Kyseinen järjestelmä käytti Baranin aikaisempaa työtä pakettikytkennästä. NPL järjestelmä sai Robertsinkin uskomaan, että pakettikytkentäinen verkko olisi mahdollista rakentaa. Näillä tiedoilla Roberts lähti rakentamaan ARPANET-tietokoneverkkoa. [1]

2.4 ARPANETin perustaminen (1968-1969)

Vuonna 1968 ARPANETin aliverkon rakentamista varten laitettiin tarjous, ja konsulttiyhtiö BBN (Bolt Beranek and Newman) valittiin hommaan [3]. Heidän tehtävä oli kehittää pakettikytkentäiset IMP-laitteet. Tässä työssä Frank Heart ja Bob Kahn toimivat päähenkilöinä [2].

Jokainen solmu ARPANET-tietokoneverkossa koostui isäntäkoneista (engl. host) - eli osituskäyttöön tarkoitettusta tietokoneesta - ja IMP:stä. Aliverkossa jokainen IMP olisi vähintään yhteydessä kahteen toiseen IMP:hen. Isäntäkone lähettää viestin IMP:lle, joka jakaa sen paketteihin. Paketit sen jälkeen lähtevät itsenäisesti aliverkkoa pitkin kohti päätysolmua, jossa ne kootaan takaisin yhtenäiseksi viestiksi. Verkkoa varten tarvittiin ohjelmistoa isäntäkoneille ja aliverkolle. [1]

- Aliverkon ohjelmisto koostui:
 - lähde IMP:ltä - kohde IMP:lle protokollasta
 - IMP:ltä - IMP:lle protokollasta
 - IMP:ltä - isäntäkoneelle yhteydestä
- Isäntäkoneiden ohjelmisto koostui:
 - isäntäkoneelta - IMP:ille yhteydestä

– isäntäkoneelta - isäntäkoneelle protokollasta

ARPANETin tietokoneverkkoa varten valittiin 4 sijaintia; "University of California, Los Angeles"(UCLA); "Stanford Research Insitute"(SRI); "University of California, Santa Barbara"; ja "University of Utah". BBN huolehtivat IMP:den väliestä kommunikaatiosta, mutta miten isäntäkoneet kommunikoivat toistensa välillä oli vielä selvittämättä. Tätä varten perustettiin NWG (Network Working Group), joka sisältäisi edustajia jokaiselta sijainnilta. [2]

NWG loi ekana TELNET ja FTP (File Transfer) protokollat. TELNET mahdollisti sisäänkirjautumisen tietokoneeseen etäyhteydellä. FTP salli tiedostojen siirtämisen verkon yli turvallisesti. Molemmat protokollat toimivat asiakas-palvelin-arkkitehtuurilla, ja olivat epäsymmetrisiä protokollia. NWG:n esittäessään nämä protokollat Robertsille, hän väitti, että enemmän tarvittiin. Tämän jälkeen NWG loi symmetrisen protokollan isäntäkoneiden välille, nimeltä "Network Control Program", joka myöhemmin tunnetaan nimellä NCP-protokolla (Network Control Protocol). [2]

Vuonna 1969 ensimmäinen IMP asennettiin UCLA:han, ja toinen SRI:hin. Näiden kahden paikan välillä lähetettiin ensimmäinen viesti isäntäkoneiden välillä. Jälkeenpäin samana vuonna IMP:t asennettiin loppusijainteihin, jonka jälkeen ensimmäinen versio ARPANETistä oli pystyssä neljällä isäntäkoneella. Tulevien vuosien aikana tietokoneita lisättiin ARPANETiin, ja ARPANET lähti kasvamaan nopeaa tahtia. [3]

2.5 TCP/IP kehitys (1972-1983)

Vuosien aikana muita tietokoneverkkoja oli kehityksessä, tai oli jo kehitetty ympäri maailmaa. Englannissa oli NPL-tietokoneverkko. Ranskassa oltiin kehittämässä Cyclades-tietokoneverkkoa Louis Pouzinin johtamana. Yhdysvaltojen sisälläkin

oltiin rakentamassa muita tietokoneverkkoja. Näissä ja ARPANETin tietokoneverkoissa oli ongelmana, että kaikki olivat sisäverkkoja (Intranet). Tietokoneverkoilla ei ollut mahdollisuutta kommunikoida toistensa kanssa, ja lisätäkseen laitteita verkkoihin, niiden piti seurata sen verkon vaatimuuksia. [2]

Vuonna 1972 Bob Kahn oli siirtynyt työskentelemään DARPalle. Hänen ideana oli lähteä keksimään tapaa yhdistämään tietokoneverkot toisiinsa. Hän kutsui vuonna 1973 Vinton Cerfin työskentelemään hänen kanssaan tätä varten. Tämän jälkeen Kahn pystytti tapaamisen, johon hän kutsui tutkijoita Euroopasta ja Yhdysvalloista keskustelemaan asiasta. Paikalla tapaamisessa oli Davies ja Scantlebury NPL:stä, Pouzin Ranskan Cyclades projektilta, ja monia muita. Osallistujia oli ainakin yli 30 ihmistä, joilla kaikilla oli vahvoja kiinnostuksia tietokoneverkoista. Kaikki olivat osana myös joissain tietokoneverkkojen projekteissa. Konferenssissa perustettiin INWG (International Network Working Group), jonka puheenjohtajaksi valittiin Cerf. [2]

Selvittäessä tapaa yhdistää tietokoneverkkoja toisiinsa, Cerfille tuli idea. Sen sijaan, että tietokoneverkot yritettäisiin saada toimimaan toistensa kanssa, tietokoneverkkojen välissä voitaisiin käyttää tietokoneita yhdyskäytävänä (engl. gateway). Kyseiset laitteet nykyään tunnetaan nimellä reititin. NCP-protokolla ei oltu suunniteltu tämänlaiseen työhön, joten he lähtivät suunnittelemaan uutta protokolla. Vuonna 1974 julkaistussa paperissa he esittivät ideansa reittittimistä ja TCP-protokollasta (Transmission Control Protocol) [5]. TCP:n tapa lähettää paketteja oli erilainen kuin aikaisemmissa IMP laitteissa. Paketit kapseloitiin ensin sähköisiksi kirjekuoriksi, jolloin paketin sisältö olisi vain luettavissa vastaanottavassa isäntäkoneessa. TCP olisi myös vastuussa pakettien päätyemisestä kohdepisteeseen, ja sen reitityksestä tietokoneverkkojen läpi. [2]

TCP sen alkuperäisessä muodossa ei kuitenkaan riittänyt perustamaan luotettavaa kommunikaatiota tietokoneverkkojen välille. Vuonna 1977 Cerf ja Kahn kutsui-

vat PARC:in (Xerox Palo Alto Resarch Center) osallistumaan INWG keskusteluihin. PARC tiimi oli tehnyt samanlaista tutkimusta tietokoneverkkojen liittämistä. PARC ja monet ihmiset INWG:llä olivat tulleet samaan tulokseen, että yksi protokolla joka yrittää tehdä kaiken ei sopinut näihin tarkoituksiin. TCP kävi neljä iteratiivista muutosta, ennen kuin se päätettiin jakaa kahteen protokollaan. Uusi versio TCP:stä, ja IP-protokolla (Internet Protocol). [2]

TCP sen uudessa versiossa oli ei ollut enään vastuussa reitityksestä. IP:n tehtävänä oli löytää kohdepiste kaikista miljoonista tietokoneista, ja perusti standardin tietokoneiden välisestä viestinnästä. TCP ja IP ovat vuosien aikana kehittyneet, ja ne tunnetaan nykyään TCP/IP-protokollapinona. Vuonna 1982 tehtiin päätös, että kaikki solmut ARPANETissä vaihdettaisiin vanhasta NCP-protokollasta TCP/IP-protokollapinon. Vuoden 1983 alussa NCP laskettiin historiaan, ja tulevaisuus kuului TCP/IP:lle. Tämä merkkasi alun modernille Internetille, mikä on vielä nykyään käytössä. [2]

3 WWW:N HISTORIA

Tässä luvussa käsittelen WWW:n kehityksen lyhyttä aikaväliä. Sen alkuperäisen version kehitti Tim Berners-Lee vuonna 1990, mutta vuosien aikana WWW:tä on huomattavasti jatkokehitetty [6]. Keskityin luvussa WWW:n alkuvaiheisiin.

3.1 Enquire sovellus

Enquire sovellus oli Berners-leen harrastusprojekti. Enquiren avulla Berners-lee piti muistiinpanoja erilaisista informaatioista. Yksi sivu sisälsi informaation ihmisestä, laitteesta, tai ohjelmasta. Jokainen sivu oli kuin solmu verkossa. Ainoa tapa tehdä uusi sivu, oli luoda linkki aikaisemmasta sivusta. Linkit sivuun ja sivusta olivat sivun alareunassa. Ainoa tapa löytää informaatiota oli selailla aloitussivulta. Enquinessa oli kahden tyyppisiä linkkejä, sisäisiä ja ulkoisia linkkejä. Sisäiset linkit menivät sivujen välillä yhden tiedoston sisällä. Ulkoiset linkit menivät toiseen tiedostoon. Näissä oli erona, että ulkoisen linkin kohdesivu ei pitänyt palautumislinkkiä sivun alareunassa. [6]

3.2 WWW:n alkuvaiheet (1988-1989)

Berners-Leen aikana CERNissä, hän näki Enquiren tyyppiselle järjestelmälle hyötyä siellä, josta tutkijat voisivat löytää informaatiota ja niiden yhteyksiä paremmin. Hänen visiona oli yhdistää Enquiren ulkoiset linkit, hyperteksti, ja myös hänen ke-

hittämä RMC-sovellus (remote procedure call). Hänelle tärkeä ominaisuus tälle järjestelmälle oli pitää se hajautettuna. Järjestelmällä ei saisi olla keskipistettä, koska sen kapasiteetti estäisi järjestelmän kasvua. [6]

Vuoden 1988 loppuvaiheissa Berners-Lee ehdotti hänen ideaansa esimiehelleen Mike Sendallille. Hänen mielestä se oli hyvä idea, mutta hän ohjasi Berners-Leen kirjoittamaan virallisen ehdotelman asiasta. Järjestelmästä vielä puuttui tapa yhdistää tietokoneita toisiin tietokoneisiin, joilla oli erilaisia käyttöjärjestelmiä. Berners-Lee sai tietää Internetistä hänen RMC projektimentorin Ben Segalin kautta. Siinä aikana CERNissä ja euroopassa ei käytetty vielä yleisesti Internetiä. Euroopassa tavoiteltiin erilaista kokoelmaa tietokoneverkon protokollista. Internet sai kuitenkin suosiota tulevien vuosien aikana. Berners-Lee näki Internetin, ja sen taustalla olevan TCP/IP:n olevan oikea vaihtoehto järjestelmälle. [6]

Hän alkoi lähettämään ehdotelmia ideasta vuonna 1989, mutta ne eivät edenneet CERNissä. Saadakseen projektia käyntiin hänen ideasta, hän pyysi Sendalilta NeXT-tietokoneen ostoa. Laitteessa nähtiin olevan hyödyllisiä ominaisuuksia, jotka saattaisivat auttaa projektia. Samalla hän sai myös perusteen käyttää aikaa projektiin NeXT laitteen testauksena. Samoihin aikoihin Berners-Lee etsi nimiä hänen projektilleen. Hän päätti käyttää lyhennettä "HT"hypertekstistä jokaiselle ohjelmalle projektissa. Hän myös löysi toisen nimen matematiikasta, jota hän päätti käyttää. Nimeä "World Wide Web". [6]

3.3 WWW:n kehitys (1990)

WWW:tä varten tarvittavien ohjelmien kehitys alkoi vuonna 1990. Saadaessaan pyydetyn NeXT-tietokoneen, Berners-Lee pystyi aloittamaan tekemään projektia. CERNissä oli mottona "Osta, älä rakenna", jonka takia hän lähti etsimään valmiita ratkaisua hypertekstin alalta, joita hän ei kuitenkaan löytänyt. Hän päätti mottosta huolimatta alkaa luomaan WWW:tä itse. [6]

Berners-Lee ensimmäiseksi alkoi kirjoittamaan verkkoa varten asiakasohjelmaa. Ohjelman tarkoituksena oli mahdollistaa hypertekstisivujen luomisen, selailun, ja muokkaamisen. Ohjelma näyttäisi käytännössä tekstienkäsittelyohjelmana. Hän käytti NeXTin käyttöjärjestelmää NeXTSteppiä, ja sen mukana tulleita työkaluja. Ohjelma tuli valmiiksi nopeasti, jonka jälkeen hän siirtyi tekemään tarvittavaa koodia ohjelmaa varten. Hän loi seuraavan parin kuukauden aikana:

- HTTP-protokollan (Hypertext Transfer Protocol), jolla tietokoneet pystyisivät jakaa resursseja Internetin yli.
- URI-merkkijonon (Universal Resource Identifier), joka toimi skeemana dokumenttien osoitteille (myöhemmin käytetään versiota "Uniform Resource Locator").
- HTML-merkkikielen (Hypertext Markup Language), joka pystyisi muotoilla sivuja, joissa oli hypertekstilinkkejä.

Parin kuukauden sisällä hänellä oli valmiina ensimmäinen verkkoselain, jota hän kutsui nimellä "WorldWideWeb". Selaimen jälkeen hän kirjoitti ensimmäisen verkkopalvelimen. Sille annettiin alias "info.cern.ch", ja sillä pääsi käsiksi ensimmäiseen sivustoon. Ensimmäinen versio WWW:stä oli pystyssä vuoden 1990 joulukuussa. Myöhemmin hän vielä ohjelmoi selaimen tavan käyttää FTP-protokolla linkkien kanssa, jonka avulla päästiin käsiksi uutisartikkeleihin ja uutisryhmiin. Sen avulla ne olivat saatavina hypertekstisivuina, ja päästiin käsiksi valtavaan määrään informaatiota, jota oli jo olemassa. [6]

WWW toimi vain alussa NeXT-tietokoneilla, joka rajoitti mahdollisia käyttäjiä. Samaan aikaan myöskin HTTP-protokollaa käyttävä palvelin vei aikaa Berners-Leeltä, joten tarvittiin ylimääräistä apua. Hommaan pyydettiin Nicola Pellowia, joka toimi harjoittelijana CERNissä. Nicolalle annettiin tehtäväksi luoda selain, joka

toimisi mahdollisimman monella tietokoneella. Selainta kutsuttiin line-mode selaimeksi. [6]

3.4 WWW:n kasvu (1991->)

WWW lähti alussa kasvamaan hitaasti. Berners-Lee yritti herättää kiinnostusta CERNin sisällä, mutta WWW ei saanut suosiota. Hän julkaisi vuonna 1991 maaliskuussa CERNin sisällä WorldWideWeb-ohjelman ihmisille, joilla oli NeXT-tietokoneet. Tämän avulla kiinnostuneet ihmiset sentään pystyisivät kirjoittamaan omia hypertextejä, ja pääsisivät käsiksi informaation, mitä laiteettiin osoitteeseen "info.cern.ch". Myöhemmin samana vuonna hän julkaisi Internettiin WorldWideWeb-ohjelman, line-mode selaimen, ja yksinkertaisen palvelimen joka toimisi jokaisella tietokoneella. Hän laittoi tästä ilmoituksia useaan Internetin uutisryhmään, joista isoin oli hypertextistä kiinnostuneille alt.hypertext. [6]

Verrattuna Internetin kehitykseen, missä uudet teknologiat kasvattivat sen suosiota, WWW:n kasvu alkoi hitaasti vuodesta 1991. Seuraavien vuosien aikana sen suosio kuitenkin alkoi kasvamaan eksponentiaalisesti. Vuonna 1991 kesällä osoitteella "info.cern.ch" oli 100 kävijää. Vuonna 1992 kesällä 1000 kävijää. Vuonna 1993 kesällä 10000 kävijää. Ihmiset ja laitokset alkoivat hitaasti kiinnostua WWW:stä ympäri maailmaa. [6]

Vuosien aikana uusia selaimia alkoi ilmestymään. Vuonna 1991 ilmestyi selaimia, kuten Suomesta Erwise, ja Yhdysvalloista ViolaWWW ja Midas selaimet. Vuoden 1992 alussa erilaisia selaimia alkoi olemaan noin 50:ntä. Yksi tunnetuista selaimista oli Mosaic, joka julkaistiin vuonna 1993. [6]

Internetille tärkeä taphtuma, joka merkkasi nykyään tunnetun vapaan WWW:n, tapahtui vuonna 1993. Minnesotan yliopisto oli kehittänyt vuosia aikasempaan järjestelmän nimeltä Gopher, joka toimi kampuksen informaatiojärjestelmänä. Gopher oli saanut siinä aikana suosiota nopeammin kuin WWW. Vuoden 1993 keväällä Min-

nesotan yliopisto kuitenkin päätti muuttaa ohjelman lisenssiä. Laitos päätti jättää järjestelmän palvelin-ohjelman käytön ilmaiseksi voittoa tavoittelemattomalle yhteisölle ja opetuskäyttöön. He alkaisivat kuitenkin veloittamaan muilta käyttäjiltä vuosittaista maksua. Internetin yhteisö ei ottanut tätä vastaan hyvin, ja Gopherin suosio alkoi tipahtamaan. Tämä kuitenkin herätti kysymyksen WWW:n tilanteesta, että suunnitteleeke CERN tehdä samalla tavalla WWW:lle. Saman vuoden huhtikuun 30 päivä, CERN julkaisi, että WWW-protokolla ja sen koodi, olisi kaikille ilmaista käyttää ilman minkäänlaista sitomusta. [6]

WWW:n kasvavan suosion takia ja Gopherin tyylisestä ongelmasta lisenssien takia, Berner-Lee näki tarpeen perustaa jonkinlainen organisaatio valvoakseen WWW:n kehitystä. Organisaation tarkoituksena oli auttaa palvelimien ja selaimien kehittäjiä yhteenymmärrykseen siitä, että miten WWW:n pitäisi toimia. Tätä varten perustettiin konsortiota WWW:lle vuonna 1994 lokakuussa, jonka nimeksi tuli "World Wide Web Consortium"(W3C). Tätä ennen tapahtui myös ensimmäinen kansainvälinen WWW konferenssi, joka järjestettiin toukokuussa vuonna 1994. [6]

4 Keskeiset teknologiat

Tässä luvussa käsittelen Internetin ja WWW:n keskeisten teknologioiden toiminnallisuutta. Teknologiat ovat jaettu luvussa Internetin ja WWW:n osalukujen alle.

4.1 Internetin teknologiat

Internetin kehityksen keskeisimpiä asioita olivat ARPANET-tietokoneverkko ja sen jälkeen kehitetty TCP/IP-protokollapino. ARPANETin tärkeimpiä teknologioita olivat sen IMP aliverkko ja NCP-protokolla. Kyseiset teknologiat olivat Internetin kehityksen ja toiminnallisuuden kannalta elintärkeitä. Tässä osaluvussa käsittelen näitä teknologioita.

Interface message processor (IMP)

IMP:t olivat "Honeywell DDP-516"tietokoneita. Niitä kutsuttiin minitietokoneiksi, koska olivat pienempiä kuin ositusjärjestelmien suurtietokoneet. Ne olivat kuitenkin suunnilleen jääkaapin kokoisia. Jokaisella IMP:llä on vakio-osa, joka on osana jokaista IMP:tä. Erilaisia *isäntäkoneita* (engl. host) varten niihin sitten lisätään räätälöity erityisliitettä. Se toimii asynkronisesti IMP:n muun laitteiston kanssa käyttäen kättelymenettelyä. [7]

IMP:t muodostivat ARPANETissä IMP aliverkon, joka oli suunniteltu toimimaan itsenäisesti. Tarkoittaen, että aliverkko oli vastuussa sen sisäisestä kommunikatiosta, ja niihin kiinnitetyt isäntäkoneet eivät vaikuttaneet aliverkon toimintaan.

Aliverkon kommunikaatiota varten ne käyttivät sen aikaista puhelinlinjaverkkoa, jolle vuokrattiin omat linjat. Ne yhdistettiin puhelinlinjaverkkoon modeemien avulla. [7]

Viestin lähettäminen alkaa isäntäkoneesta. Se kiinnittää jokaisen viestin alkuun tunnistavan otsakkeen, ja lähettää viestin IMP:lle, joka vastaanottaa sen osissa. Osat paketoidaan, ja lähetetään erikseen tietokoneverkkoon. IMP:t muodostavat ylätunnisteen paketeille otsakkeella ja lisäämällä siihen tietoa verkon käytöstä. Paketin liikkua aliverkossa, IMP:t säilyttävät paketit, kunnes saavat kuittauksen vastaanottavalta IMP:ltä. IMP:t voivat kieltäytyä paketeista, mutta eivät kuittauksen lähettämisen jälkeen. Jos IMP ei saa kuittausta, se lähettää paketin eri reittiä. Päätyipisteen IMP kokoaa paketit takaisin viestiksi, ja lähettää viestin vastaanottavalle isäntäkoneelle osissa, joka vastaanottaa ne yhtenä osana. Viestin osittaminen on siis käytännössä näkymätön isäntäkoneelle. IMP:t ilmoittavat myös isäntäkoneelle tiedon menetyksestä tai aliverkon kommunikaation katkaisusta. [7]

Aliverkko hyväksyy vain yhden viestin kahden isäntäkoneen välillä. Viestin saapuessa kohdeisäntäkoneen IMP:lle, se lähettää lähdeisäntäkoneelle takaisin viestin, joka ilmoittaa valmiudesta seuraavalle viestille. Sitä ennen lähde IMP estää uusia viestejä sen isäntäkoneelta. [7]

IMP:t pitävät reititystaulukkoa reititys-algoritmiä varten. Ne tekevät puolen sekunnin välein viivearvion kaikille pisteille aliverkossa, kommunikoi tästä niiden naapureiden kanssa, ja yhdistää nämä informaatiot yhteen. IMP:t tekevät myös kätteleymenetelmällä tarkistusta naapurilinjoille, ja kuollut linja merkataan reititystaulukkaan. Se päivitetään takaisin eläväksi, jos on tarpeeksi monta kättelyä tehty. [7]

Network control protocol (NCP)

NCP-protokollaa käytettiin isäntäkoneiden välisessä kommunikaatiossa ARPANETissä. IMP aliverkko ohjaa viestit isäntäkoneiden välillä, mutta NCP-protokolla

muodosti yhteydet isäntäkoneiden välillä. Yhteyksien avulla pystyttiin ohjaamaan viestit oikeisiin prosesseihin, eli vastaanottava isäntäkone tiesi mitä saaduilla viesteillä tehdään. [8]

Isäntäkoneiden erilaisia prosessinimiä varten luotiin välille abstraktiota nimeltä soketit (engl. socket). Soketit olivat 40-bittisiä numerosarjoja, jotka koostuivat kolmesta osasta: 24-bittisestä käyttäjänumerosta, 8-bittisestä isäntäkoneen numerosta, ja 8-bittisestä sokettinumerosta. Käyttäjänumeron avulla pystyttiin uniikisti tunnistamaan tietokoneverkon käyttäjiä. Se yleensä koostui 8-bittisestä koti-isäntäkoneen numerosta, ja omasta 16-bittisestä numerosta, joka tunnistaa hänen uniikisti koti-isäntäkoneella. Isäntäkoneen numeron avulla erotettiin isäntäkoneet toisistaan. Sokettinumeron avulla annettiin jokaiselle käyttäjälle 256 sokettiä käytettäväksi. Sokettinumeron viimeinen bitti merkkasi soketin joko, 0 = vastaanottavaksi soketiksi, tai 1 = lähettäväksi soketiksi. [8]

NCP-protokollan funktio oli perustaa yhteyksiä, leikata yhteyksiä, vaihtaa yhteyksiä, ja hallita tiedonkulkua. Yhteyttä varten muodostettiin kaksivälinen yhteys kahden sokettiparin avulla. Yhteyden perustaminen tapahtui edestakaisin lähetettyjen viestien avulla, jossa molemmat isäntäkoneet lähettivät informaatiota heidän vastaanottavasta ja lähettävästä soketeista toisilleen. Yhteyden perustamista varten jokaisella isäntäkoneella käynnistettäisiin kuuntelijaprosessi, jonka tehtävänä oli vastaanottaa yhteydenottoviesti, ja vaihtaa vastaanottava kommunikaatio toiselle soketille. Kuuntelijaprosessia varten käytettiin käyttäjänumeroa 0, ja sokettinumeroa 0. [8]

Internet protocol (IP)

IP-protokollan selitystä varten valitsin määrittelyn vuodelta 1981, joka oli versio 4 IP-protokollasta [9]. Eli sama IPv4 mikä on nykyään vielä käytössä.

IP-protokollan tarkoituksena oli mahdollistaa kommunikaatio erilaisten tieto-



Kuva 4.1: Verkkojen välinen reititys IP-protokollalla [Internet protocol (IP)]

koneverkkojen yli, joissa käytettiin omia protokollia kommunikaatiota varten. IP-protokolla toteutti tämän käyttämällä paikallisia protokollia ja reitittäjiä – eli yhdyskäytäviä (engl. gateway) – verkkojen välissä [Kuva 4.1]. IP-protokolla tarvittiin toteuttaa vain lähtö ja kohde isäntäkoneissa, ja verkkojen välisissä reitittämissä. Viestejä kutsuttiin datagrammeiksi (engl. datagram), jotka olivat viestin siirtoyksiköitä. IP-protokollan lähettäessä datagrammi paikalliseen tietokoneverkkoon, paikallinen protokolla vastaanottaa, paketoit, ja kuljettaa sen saatuun osoitteeseen. Saatu osoite oli joko paikallinen isäntäkone, tai verkkojen välissä oleva reititin. Jälkimmäisessä tapauksessa reititin purkaa paikallisen paketin, selvittää uuden kohteen ja tietokoneverkon paketissa olevalla osoitteella, ja lähettää datagrammin valitun tietokoneverkon protokollaa käyttäen. Tarvittaessa IP-protokolla jakoi datagrammin osiin, jos valittu tietokoneverkko sallii vain pienempiä paketteja. [9]

IP-protokolla toi uuden osoitemuodon nimeltä IP osoitteet (nykyään tunnettu IPv4 osotteina). Osoitteet olivat 32-bittisiä numerosarjoja, jotka koostuivat tietokoneverkon numerosta ja paikallisesta osoitteesta. Tietokoneverkon numerolla eroteltiin tietokoneverkot toisistaan, kun taas paikallisella osoitteella tunnistettiin verkon sisällä oleva isäntäkone. IP-osoitteet tunnetaan muodossa "192.168.100.1". Osoitteet jaettiin kolmeen luokkaan:

- Luokka A osoitteessa ensimmäinen bitti oli 0, seuraavat 7 bittiä olivat tietokoneverkon numero, ja loput 24 bittiä olivat paikallinen osoite. Jaetut IP-osoitteet olivat "0.0.0.0 - 127.255.255.255".

- Luokka B osoitteessa ensimmäiset bitit olivat 10, seuraavat 14 bittiä olivat tietokoneverkon numero, ja viimeiset 16 bittiä olivat paikallinen osoite. Jaetut IP-osoitteet olivat "128.0.0.0 - 191.255.255.255".
- Luokka C osoitteessa ensimmäiset bitit olivat 110, seuraavat 21 bittiä olivat tietokoneverkon numero, ja loput 8 bittiä olivat paikallinen osoite. Jaetut IP-osoitteet olivat "192.0.0.0 - 223.255.255.255".

Eri luokkien avulla pystyttiin jakaa osoitteet tasaisemmin riippuen tietokoneverkon koosta. Luokka A oli laajoille tietokoneverkoille, joilla oli paljon isäntäkoneita. Luokka C oli pienille tietokoneverkoilla, joilla oli vähän isäntäkoneita. [9]

Transmission control protocol (TCP)

TCP-protokollan selitystä varten valitsin määrittelyn vuodelta 1981, jolloin sen toiminnallisuus oli osittain siirretty IP-protokollaan [10]. TCP-protokolla oli parannus aikasemmalle NCP-protokollalle, mutta käyttää samanlaisia konsepteja, kuten yhteyden muodostamista ja soketteja.

TCP-protokollan tarkoituksena oli tarjota luotettava yhteys kahden prosessin välillä, ja otti vastuuksi samanlaisia paketoititehtäviä mitä IMP laitteisto teki. TCP-protokolla jakoi viestin paketteihin, ja antoi paketin IP-protokollalle kuljetettavaksi. Paketin vastaanottava TCP-protokolla lähetti vastauksena kiittauksen (ACK) paketin vastaanotosta. Jos kiittäus ei saapunut alkuperäiselle lähettäjälle takaisin tietyssä määräajassa, se lähettää paketin uudestaan. Paketin mukana myös lähetettiin tarkistussumma, jolla vastaanottava TCP-protokolla tekee oman tarkistussumman vastaanotetusta paketista, ja jos se ei vastaa lähetettyä tarkistussummaa, niin paketti hylättiin. [10]

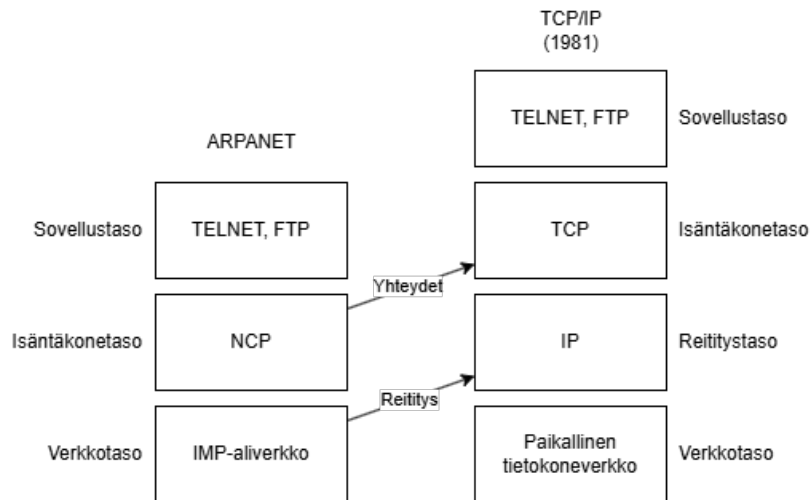
TCP-sokit toimivat samanlaisella periaatteella, kuin NCP-protokollassa. Niiden tarkoituksena on ohjata isäntäkoneiden välistä kommunikaatiota oikeihin prosesseihin. Sokit koostuivat kahdesta osasta: IP osoitteesta ja porttinumerosta. IP

osoitteen avulla pystyttiin tunnistamaan jokainen soketti uniikiksi verkossa. Porttinumerolla ohjataan kommunikaatio oikeaan prosessiin. Soketit pystyivät kommunikoimaan molempiin suuntiin, jolloin yhteydettä varten tarvittiin vain yksi soketti molemmista päistä. Yhteyden perustaminen tapahtuu lähettämällä yhteydenottoa tiettyyn sokettiin. Vastaanottavalla puolella pitää olla sama soketti valmiina kuuntelemassa. Yleensä vastaanottavalla puolella on pyörimässä palvelin-ohjelmia, jotka avaavat soketteja kuuntelua varten. Yhteydenotosta syntyy kolmesuuntainen kättely: Yhteydenottava TCP-protokolla lähettää yhteydenpyynnön, vastaanottava TCP-protokolla lähettää kiittauksen yhteydenotosta, ja yhteydenottava TCP-protokolla lähettää kiittauksen takaisin. [10]

ARPANET protokollapinot

ARPANETin toiminnallisuutta voidaan kuvailla modernilla protokollapinon konseptilla, jolloin se koostuisi 3-tasosta. Alimmalla tasolla oli IMP aliverkko, joka oli vastuussa viestin kuljetuksesta laitteiden välillä. Toisella tasolla oli NCP-protokolla, joka hoiti isäntäkoneiden välisten prosessien kommunikaatiosta. Kolmanella tasolla oli sovellukset, kuten tiedostosiirto (FTP) ja etäyhteys (TELNET). Myöhemmin ARPANET siirtyi käyttämään TCP/IP-protokollapinoa. [Kuva 4.2]

TCP ja IP protokollien kehityksessä niitä alettiin esittämään TCP/IP-protokollapinon muodossa, jossa oli 4 tasoa. Alimmalla tasolla oli paikallinen tietokoneverkon protokolla, joka kuljetti viestejä paikasta toiseen. Toisella tasolla oli IP-protokolla, joka oli vastuussa tietokoneverkkojen välisestä viestinnästä. Kolmannella tasolla oli TCP-protokolla, joka hoiti isäntäkoneiden välisten prosessien kommunikaatiosta. Se myös jakoi viestin paketteihin, ja oli vastuussa pakettien saapumisesta kohdepisteeseen. Neljännellä tasolla oli sovellukset. [Kuva 4.2] [10]



Kuva 4.2: ARPANETin protokollapinojen muutokset [ARPANET protokollapinot]

4.2 WWW:n teknologiat

Tässä osaluvussa käsittelen WWW:n toiminnallisuuden kannalta tärkeimpiä teknologioita, joita ovat URL-merkkijonot, HTTP-protokolla, ja HTML-merkintäkieli. WWW koostuu nykyään monesta muusta teknologiasta, mutta mainitut kolme teknologiaa ovat WWW:n toiminnallisuuden ytimessä.

Uniform resource locator (URL)

URL selitystä varten valitsin määrittelyn vuodelta 1994, joka oli ensimmäisiä määrittelyitä sille [11].

URL on kompakti merkkijono, jolla esitetään resurssin sijaintia Internetissä. URL-merkkijonot ovat yleisesti kirjoitettu tavalla "<skeema>:<skeema-kohtainen-osa>". Skeemoja on erilaisia protokollia tai ohjelmia varten, kuten FTP, Telnet, Gopher, HTTP, jne. [11]

HTTP-protokollan URL-skeema on muotoa: "http://<isäntäkone>:<portti>/<reitti>?<haku>". "<isäntäkone>" kertoo haettavan palvelimen sijainnin Internetissä. Sen voi kirjoittaa IP osoitteen tai verkkotunnuksen (engl. Domain Name) muodossa. "<portti>" kertoo mihin porttiin numeroon yhdistää. Sen voi jättää tyh-

jäksi, jolloin oletusporttina on 80. "<reitti>"on sivun sijainti palvelimen sisällä. "<haku>"kohtaan voi kirjoittaa merkkijonon etsimistä varten.

URL-merkkijonoa varten otin esimerkiksi Internetin historiasta kertova wikipedian sivun "https://en.wikipedia.org/wiki/History_of_the_Internet". Jaettuna skeeman osiin sen URL näyttäisi seuraavalta: "https" = uudempi ja turvallisempi versio http-protokollasta, "en.wikipedia.org" = isäntäkoneen sijainti käyttäen DNS-palvelua, portti on https oletusportti 443, "wiki/History_of_the_Internet" = reitti sivun sijaintiin. Hakukohta on tyhjä. [11]

Hypertext transfer protocol (HTTP)

HTTP-protokollan selitystä varten valitsin määrittelyn vuodelta 1991, joka oli versio 0.9 protokollasta [12].

HTTP-protokolla on sovellustason protokolla, joka käyttää TCP/IP-protokollapinon muodostavaa linkkiä hakeakseen HTML dokumentteja HTTP-palvelimilta. Asiakas tässä tapauksessa on dokumenttia hakeva isäntäkone. Asiakas muodostaa yhteyden TCP-protokollalla palvelimelle käyttäen HTTP URL:n annettua palvelimen osoitetta – IP tai verkkotunnusta. Lisäksi myös käyttää porttinumeroa (tai oletusporttia). Yhteyden perustamisen jälkeen asiakas lähettää pyynnön palvelimelle, joka sisältää sanan "GET" ja dokumentin sijainnin. Sijainti sisältää URL:stä osan <reitti>?<haku>, joka voi olla myös tyhjä, jolloin dokumentin sijainti on palvelimen juuressa. Palvelin palauttaa vastaukseksi viestin HTML-kielellä. Viestin siirtämisen jälkeen palvelin katkaisee TCP-yhteyden. [12]

Hypertext markup language (HTML)

HTML selitystä varten valitsin määrittelyn vuodelta 1992, joka oli ensimmäisiä versioita HTML:stä, jolloin sillä ei ollut vielä tarkkaa versionumeroa [13].

HTML on merkintäkieli, millä pystytään luomaan strukturoituja hyperteksti-

dokumentteja. HTML koostuu tägeistä, joiden avulla merkataan tekstin ominaisuuksia. Teksti yleensä laitetaan kahden tagin väliin, jolloin sisällä oleva teksti muokkaantuu tarkoitetulla tagillä. Tägejä kirjoitetaan muodossa "<tagi>", jossa haluttu tagi laitetaan merkkien "<" ja ">" väliin. Lopettavaan tagiin laitetaan myös "/" merkki muodossa "</tagi>". Yhdistettynä tämä näyttää esimerkiksi "<H1>Hei maailma</H1>". Joillekin tägeille pystyy myös antamaan parametrejä – joita kutsutaan attribuuteiksi – muodossa "<tagi attribuutti=xxx attribuutti2=xxx>", mutta vain aloittavassa tagissä. [13]

HTML:n tärkeimpiä ominaisuuksia oli ankkuri-tagin "<a>", joka toimi linkkinä sivujen ja tekstikohtien välillä. Ankkuri-tagille pystyi antamaan attribuutit NAME, HREF, ja TYPE. NAME-attribuutilla merkitään ankkurikohta linkin kohteeksi käyttäen arvoa "#xx". HREF-attribuutilla ankkuritekstistä muodostettiin linkki. Sen arvolla pystyttiin osoittamaan joko; kohtaa samasta sivulla, jolloin arvoksi annettiin toisen ankkurikohdan NAME-attribuutti; toista dokumenttia samalla palvelimella, jolloin arvoksi annettiin dokumentin suhteellinen sijainti nykyiseen dokumenttiin verrattuna; tai toista dokumenttia toisella palvelimella, jolloin arvoksi annettiin toisen palvelimen osoite URL-muodossa. TYPE-attribuutilla pystyttiin antamaan kuvausta linkin suhteesta sen kohteeseen. Myöhemmissä versioissa se korvattiin REL-attribuutilla [14]. Mainitut attribuutit olivat vaihtoehtoisia, mutta ilman niitä ankkuri-tagistä ei ollut hyötyä. [13]

5 Yhteenveto

Luvussa 2 ja 3 vastaan **tutkimuskysymykseen 1**, joissa käsittelin Internetin ja WWW:n historiaa. Internetin kehitys tapahtui monissa eri vaiheissa, joissa se eteni uusilla teknologioilla. Yleinen syy lähteä tekemään uutta kehitystä oli yhdistää vanhaa teknologiaa toisiinsa. Tietokoneet ensimmäiseksi, ja sen jälkeen tietokoneverkot. Internetin kehitys alkoi varhaisista tietokoneista, pakettikytkennän teoriasta, ja Internetin konseptista. Keskeisimpiä tietokoneverkkoja Internetin kehityksessä oli ARPANET. Se toteutti ensimmäisen laaja-alueen pakettikytkentäisen tietokoneverkon, jota varten luodut teknologiat – IMP aliverkko ja NCP-protokolla – ovat toimineet pohjana jatkokehitykselle. TCP/IP-protokollapino mahdollisti tietokoneverkkojen välisen kommunikaation, jonka standardisointi merkkasi alun modernille Internetille.

Verrattuna Internetin kehitykseen, WWW:n kehitys oli huomattavasti yksinkertaisempi. WWW:n kehitti Tim Berners-Lee vuonna 1990. Sen ensimmäisen version kehitys kesti vain pari kuukautta, joka oli käytännössä viikonloppuprojekti Internetin skaalalla. Siinä aikana Berners-Lee kehitti HTML-merkintäkielen, URL-merkkijonon, HTTP-protokollan, ja ensimmäisen selaimen ja palvelimen. WWW:tä ei olisi kuitenkaan ollut mahdollista kehittää ilman taustalla olevia teknologioita. Siinä kohdassa tietokoneet ja Internet olivat jo tarpeeksi kehittyneet, että WWW:tä varten tarvittiin vain siitä idea ja kiinnostunut henkilö.

Luvussa 4 vastaan **tutkimuskysymyksiin 2 ja 3**, ja kävin läpi keskeisten tek-

nologioiden toiminnallisuutta. Niiden keskeisenä teemana oli välttää vanhan teknologian muuttamista, ja rakentaa jotain uutta päälle. Monet niistä teknologioista ovat vieläkin käytössä, ja vaikka niitä ollaan jatkokehitetty, niiden tarkoitukset ovat pysyneet samana.

Tutkielman koon ja laajan aiheen takia tutkielmaa on pitänyt rajata, jonka takia on jäänyt paljon asioita Internetin ja WWW:n historiasta käsittelemättä. Internetistä on jäänyt esimerkiksi modeemit, reitittimet, ja "Domain Name System"toiminnallisuudet selvittämättä. Valittujen teknologioiden toiminnallisuuden selitys on myös pintapuoleista, jonka takia siitä kiinnostuneet henkilöt eivät välttämättä saa tarpeeksi kattavaa ymmärrystä. Näistä löytyy kuitenkin tarkempaa informaatiota lähteistä. Historian osuus on myös laajasti selitetty, joka rajasi toiminnallisuuden selvitystä.

Samanlaisen tutkielman aiheena ehdottaisin keskittymistä pelkästään Internetin tai WWW:n kehitykseen, jolloin pystyisi kattavammin selittämään yhdestä. Luvussa 2 mainitusta pakettikytkennän teorian alkuperäisestä tekijästä pystyisi tehdä kattavampaa tutkimusta. Loydetyistä lähteissä siitä oli ristiriitaisia väittämiä. Yksittäisen teknologian kehitystä voisi myös seurata, kuten HTML:n versioiden tuomia muutoksia.

Lähdeluettelo

- [1] A. S. Tanenbaum, *Computer networks*, Global edition., Sixth edition., yhteistyössä N. Feamster ja D. Wetherall. Harlow: Pearson, 2021, ISBN: 978-1-292-37406-2.
- [2] J. Naughton, *A brief history of the future: the origins of the internet*, 4. impr. London: Phoenix, 2001, 332 s., ISBN: 978-0-7538-1093-4.
- [3] B. M. Leiner et al., ”A brief history of the internet”, en,
- [4] L. G. Roberts, ”Multiple computer networks and intercomputer communication”, teoksessa *Proceedings of the ACM symposium on Operating System Principles - SOSPP '67*, Not Known: ACM Press, 1967, s. 3.1–3.6. DOI: 10.1145/800001.811680. viitattu 14. kesäkuuta 2025. url: <http://portal.acm.org/citation.cfm?doid=800001.811680>.
- [5] V. Cerf ja R. Kahn, ”A protocol for packet network intercommunication”, *IEEE Transactions on Communications*, vol. 22, nro 5, s. 637–648, toukokuu 1974, ISSN: 0090-6778, 1558-0857. DOI: 10.1109/TCOM.1974.1092259. viitattu 16. kesäkuuta 2025. url: <https://ieeexplore.ieee.org/document/1092259/>.
- [6] T. Berners-Lee ja M. Fischetti, *Weaving the web: the original design and ultimate destiny of the World Wide Web by its inventor*, 1. paperback ed., [repr.] New York, NY: HarperBusiness, 2011, 246 s., ISBN: 978-0-06-251587-2.

-
- [7] F. E. Heart, R. E. Kahn, S. M. Ornstein, W. R. Crowther ja D. C. Walden, ”The interface message processor for the ARPA computer network”, teoksessa *Proceedings of the May 5-7, 1970, spring joint computer conference*, sarja AFIPS ’70 (Spring), New York, NY, USA: Association for Computing Machinery, 5. toukokuuta 1970, s. 551–567, ISBN: 978-1-4503-7903-8. DOI: 10.1145/1476936.1477021. viitattu 17. kesäkuuta 2025. url: <https://dl.acm.org/doi/10.1145/1476936.1477021>.
- [8] C. S. Carr, S. D. Crocker ja V. G. Cerf, ”HOST-HOST communication protocol in the ARPA network”, teoksessa *Proceedings of the May 5-7, 1970, spring joint computer conference*, sarja AFIPS ’70 (Spring), New York, NY, USA: Association for Computing Machinery, 5. toukokuuta 1970, s. 589–597, ISBN: 978-1-4503-7903-8. DOI: 10.1145/1476936.1477024. viitattu 17. kesäkuuta 2025. url: <https://dl.acm.org/doi/10.1145/1476936.1477024>.
- [9] ”Internet Protocol”, Internet Engineering Task Force, Request for Comments RFC 791, syyskuu 1981, Num Pages: 51. DOI: 10.17487/RFC0791. viitattu 22. kesäkuuta 2025. url: <https://datatracker.ietf.org/doc/rfc791>.
- [10] ”Transmission Control Protocol”, Internet Engineering Task Force, Request for Comments RFC 793, syyskuu 1981, Num Pages: 91. DOI: 10.17487/RFC0793. viitattu 22. kesäkuuta 2025. url: <https://datatracker.ietf.org/doc/rfc793>.
- [11] T. Berners-Lee, L. M. Masinter ja M. P. McCahill, ”Uniform Resource Locators (URL)”, Internet Engineering Task Force, Request for Comments RFC 1738, joulukuu 1994, Num Pages: 25. DOI: 10.17487/RFC1738. viitattu 22. kesäkuuta 2025. url: <https://datatracker.ietf.org/doc/rfc1738>.

-
- [12] T. BERNERS-LEE. ”The Original HTTP as defined in 1991”, viitattu 22. kesäkuuta 2025. url: <https://www.w3.org/Protocols/HTTP/AsImplemented.html>.
- [13] ”HyperText Mark-up Language”, viitattu 22. kesäkuuta 2025. url: <https://www.w3.org/History/19921103-hypertext/hypertext/WWW/MarkUp/MarkUp.html>.
- [14] T. Berners-Lee ja D. W. Connolly, ”Hypertext Markup Language - 2.0”, Internet Engineering Task Force, Request for Comments RFC 1866, marraskuu 1995, Num Pages: 77. DOI: 10.17487/RFC1866. viitattu 22. kesäkuuta 2025. url: <https://datatracker.ietf.org/doc/rfc1866>.