

Vertaisverkkojen tuomat ongelmat palveluntarjoajille

Vertaisverkkoja käytetään yleisesti suurten tiedostojen jakamiseen, tosin niillä on myös muita käyttötarkoituksia, kuten esimerkiksi Internet-puhelimet (*Skype* yms.). Esimerkkejä vertaisverkkoihin perustuvista sisällönjakomenetelmistä ovat esimerkiksi BitTorrent-protokolla ja Gnutella-verkko. Sisällönjakajat (*content distributor*) hyötyvät vertaisverkoista, sillä ne eivät vaadi kalliita infrastruktuurihankintoja kuten sisällönjakeluverkkoja (*CDN, Content Distribution Network*). Kysynnän kasvaessa ne myös pystyvät skaalautumaan itsestään, sillä vertaisverkoissa käyttäjät jakavat yhdessä kaistaa, tallennustilaa ja suoritustehoa. Lisäksi sisällönjakajan saatua jaettua kaikki sisällön koostavat tavut hajautetusti käyttäjien kesken, pystyvät nämä periaatteessa hoitamaan itse kaiken lopun sisällön jakamisesta. Jakamalla täten sisällön jakamisen usean eri käyttäjän kesken, vertaisverkot kestävät paremmin palvelunestohyökkäyksiä ja äkillisiä ruuhkatilanteita perinteisiin asiakas-palvelin malliin (*client-server model*) perustuviin ratkaisuihin verrattuna, parantaen täten sisällön saatavuutta. Käyttäjät hyötyvät tästä saatavuudesta, sekä vertaisverkkojen mahdollistamista nopeasta suurien tiedostojen latausnopeudesta.

Sisällön saatavuuteen liittyvät edut eivät ole kuitenkaan täysin yksipuolisia. Vertaisverkoissa uutta sisältöä jaetaan alussa ahkerasti, jolloin saatavuus ja latausnopeus ovat korkealla. Ajan saatossa kuitenkin vertaiset hiipuvat, ja jos heille ei ole mitään kannustimia jatkamaan sisällön jakamista, saattavat vertaiset kadota kokonaan. Tällöin sisältöä ei ole enää mahdollista ladata, mutta ongelma voidaan kuitenkin välttää mikäli sisällönjakaja tarjoaa ainakin yhden solmun vertaisverkkoon jakamaan sisältöä. Käyttäjän näkökulmasta tilanne ei juurikaan eroa asiakas-palvelin-mallista, mutta sisällönjakajalle se tulee paljon halvemmaksi kuin kaiken latausliikenteen keskittäminen omille palvelimille.

Vaikka vertaisverkot motivoivat käyttäjiä ostamaan nopeampia verkkoyhteyksiä, tuottavat ne paljon hankaluuksia palveluntarjoajille (*Internet Service Provider, ISP*). Nämä ongelmat johtuvat vertaisverkkojen tekemästä reitityksestä. Vertaisverkot luovat oman päällysverkkonsa (*overlay network*) alusverkkona (*underlay network*) toimivan Internetin päälle, ja tekevät reitityksen sovelluskerroksella kokonaan itse. Tämä hankaloittaa palveluntarjoajien verkon sisäistä reititystä, sillä mielivaltainen reititys saattaa esimerkiksi johtaa tilanteeseen, missä vertaisverkko ruuhkauttaa reitin. Tällöin palveluntarjoaja joutuu uudelleenreitittämään muiden sovelluste paketteja kulkemaan muualta, aiheuttaen näille viivettä. Lisäksi on mahdollista jos verkko kärsii erityisen pahasti vertaisverkkojen liikenteestä, palveluntarjoaja joutuu muuttamaan verkon rakennetta verkkokerroksella. Vertaisverkot voivat huomata nämä muutokset, ja muuttaa taas omaa reititystään. Tämä voi taas uudelleen vaatia verkon uudelleenjärjestämistä, johtaen ongelmalliseen sykliseen tilanteeseen.

Vertaisverkot eivät pääse suoraan käsiksi Internetin eri reittien tietoihin kuten viiveeseen yms., joten vertaisverkot joutuvat jättämään nämä ominaisuudet huomioimatta tai sitten estimoida niitä. Estimointiin tarvitaan luotainpakettien lähettämistä, mikä myös luo hieman ylimääräistä raskautta Internetille ja palveluntarjoajien verkoille. Vertaisverkot luovat myös paljon palveluntarjoajan verkolle ulkopuolista liikennettä, mistä aiheutuu ylimääräisiä kuluja palveluntarjoajille. Monesti kuitenkin samaa sisältöä lataavat ja lähettävät vertaiset saattavat löytyä lähempää, johtuen mm. sisällön kielestä ja eri maantieteellisissä alueissa jaetuista kiinnostusten kohteista. Esimerkiksi on mahdollista, että Helsingistä lataavan käyttäjän haluama sisältö löytyisi samasta palveluntarjoajan verkosta Espoosta, mutta vertaisverkko asettaa mielivaltaisesti käyttäjän lataamaan sisältöä Yhdysvalloissa sijaitsevalta vertaiselta. Tämä on selvästikin tehotonta

palveluntarjoajan näkökulmasta, ja myös käyttäjäkin pystyisi todennäköisesti lataamaan sisällön nopeammin Espoosta.

Ratkaisuksi kyseisille ongelmille on ehdotettu, että vertaisverkot ottaisivat paikallisuuden huomioon, mahdollisesti tehden yhteistyötä palveluntarjoajien kanssa. Näin vertaisverkkojen liikenne voisi pysyä mahdollisimman paljon verkon sisällä, jolloin palveluntarjoajat säästävät rahaa kun pystytään välttämään verkon ulkopuolista liikennettä. Verkkojen pääsylinkit (*access link*) ovat myös yleensä Internetin suurimpia pullonkauloja, joten myös käyttäjät voisivat hyötyä paikallisuudesta nopeampien latausajojen muodossa. Mikäli vertaisverkot ja palveluntarjoajat tekisivät yhteistyötä, voisivat vertaisverkot hyötyä palveluntarjoajien tietämästä tarkasta verkon ominaisuuksien tiedoista, poistaen vertaisverkoilta tarpeen niiden estimoinnille.

Paperissaan *Should internet service providers fear peer-assisted content distribution? (Proceedings of the 5th ACM SIGCOMM conference on Internet Measurement, 2005)* Thomas Karagiannis et al. tutkivat vertaisverkkojen aiheuttamaa verkkoliikennettä ja miten paikallisuuden huomioon ottaminen vaikuttaisi siihen. Ensimmäiseksi tutkijat tarkastelivat pienen noin 20000 käyttäjän verkon pääsylinkkiä kolmena eri ajankohtana yli 24 tunnin sessioina, ja erottelivat verkkoliikenteen seasta BitTorrent-paketit. Tilastoitujen pakettien pohjalta tutkijat sitten laskivat kuinka verkon ulkopuolinen liikenne muuttuisi mikäli BitTorrent hyödyntäisi paikallisuutta. Toisessa tutkimuksessa tutkijat kasvattivat tarkasteltavaa skalaa lokaalista globaaliin. Yhden pääsylinkin sijasta tarkasteltiin yksittäisen BitTorrentilla jaettavan tiedoston seurantapalvelinta (*tracker*) viiden kuukauden ajan. Tutkijat saivat seurantapalvelimesta selville liikkuneen datan määrän sekä vertaisten IP-osoitteet. Näiden tietojen perusteella tutkijat pystyivät selvittämään vertaisten sijainnit eri palveluntarjoajien verkoissa, laskemaan jokaisen verkon läpi kulkeneen datan määrän sekä paikallistamisen tuomat mahdolliset säästöt.

Tutkimuksissa paikallisuutta tarkastellaan kahden eri skenaarion kautta. Ensimmäisessä skenaariossa vertaisverkossa sisältö ladataan aina palveluntarjoajan verkon sisältä, mikäli vain mahdollista. Jos sisältöä ei löydy laisinkaan verkon sisältä, tai jos jokin kyseistä sisältöä omaava vertainen ei ole lataushetkellä aktiivinen (eli ei lataa/lähetä sisältöä), joudutaan sisältö lataamaan verkon ulkopuolelta. Toisessa skenaariossa palveluntarjoajan verkossa on yksi välimuisti, joka sisältää kaiken vertaisverkkojen palveluntarjoajan verkon ulkopuolelta ladatun sisällön datan. Tällöin kun data on kerran saatu ulkopuolelta ladattua, voidaan vastaisuudessa sisältö ladata suoraan välimuistista. Jos ensimmäisen skenaarion paikallistamisratkaisussa olisi koko ajan vähintään yksi sisällön omaava vertainen aktiivinen, tilanne ei juurikaan eroaisi toisen skenaarion välimuistiratkaisusta.

Tutkimusten perusteella tutkijat pystyivät vetämään lukuisia johtopäätöksiä ensimmäisen paikallisuuden huomioonottamisen puolesta. Ensimmäisen tutkimuksen mukaan päivän aikana yli 10% kaikista tiedostoista ladataan kokonaan verkon ulkopuolelta useammin kuin kerran, ja ainakin 20% kaikesta ladatusta sisällöstä löytyisi verkon sisäisten aktiivisten käyttäjien keskuudesta. Lisäksi 30% ajasta verkon sisältä löytyi samaa tiedostoa lataavia/lähetäviä aktiivisia vertaisia. Toisen tutkimusten tulosten mukaan paikallistamisen tuomat säästöt kasvavat logaritmisesti vertaisverkon aktiivisten käyttäjien kasvaessa. Ensimmäisen skenaarion mukainen paikallistaminen olisi siis hyvin perusteltavissa. Toisessa tutkimuksessa todettiin, että tämä paikallistamisratkaisu hyödyttäisi palveluntarjoajia puolittamalla kaiken ulkoa tulevan vertaisverkkoliikenteen. Vertaisverkkojen ei myöskään tarvitsisi lähettää kuin keskimäärin hieman yli yhden kokonaisen tiedoston palveluntarjoajan verkon ulkopuolelle. Tutkijoiden mukaan tämä ei häittäisi käyttäjiä sisällön saatavuuden suhteen. Latausajatkin lyhenisivät ensimmäisen tutkimuksen mukaan 70%.

Sisällönjakajatkin hyötyisivät paikallistamisesta, sillä palvelimien kapasiteettitarve puolittuisi verrattuna asiakas-palvelin-malliin.

Välimuistiratkaisu voisi olla tulosten mukaan myös varteenotettava, sillä ensimmäisen tutkimuksen mukaan vertaisverkossa ladataan yksittäisiä tiedostoja lyhyillä aikaväleillä. Tällöin välimuistin tarvitsema kokonaistallennustila voisi olla melko kohtuullinen, sillä tiedostoja ei tarvitse säilyttää kovinkaan kauan. Toisen tutkimuksen tuloksissa välimuisti pärjasi myöskin yleisesti paikallistamisratkaisua paremmin. Vertaisverkoille dedikoitujen välimuistien rakentaminen olisi kuitenkin todennäköisemmin hankalampaa kuin olemassa olevien vertaisverkkosovellusten algoritmien ja protokollien muokkaus, tai jokin muu kevyempi ratkaisu kuten Aggarwar et al.:in ehdottama oraakkeli. Lisäksi vertaisverkoissa usein levitetään tekijänoikeuksia rikkovaa sisältöä, joka voisi tehdä välimuistista hyvin kyseenalaisen ratkaisun.

Toisen tutkimuksen johtopäätöksissä todettiin, että ensimmäisen skenaarion paikallistaminen voisi olla hyvin kelpo ratkaisu vertaisverkkojen aiheuttamiin ongelmiin. Vaikka paikallistamisella verkon ulkopuolinen liikenne olisi viisinkertainen välimuistiratkaisuun verrattuna, olisi se nykyisten vertaisverkkojen ja asiakas-palvelin-ratkaisujen ulkoliikenteeseen verrattuna varsin pientä. Paikallistaminen tarvitsisi vain puolitoistakertaisen määrän maksimikapasiteettia välimuistiin verrattuna. Lisäksi verkkoliikenteen ollessaan ruuhkaisimmillaan, 95:n persentiilin perusteella paikallistamisessa verkkoliikenteen overheadin määrä saattaa olla jopa pienempi kuin välimuistia hyödyntämällä.

Välimuistiin verrattuna paikallistamisen implementoiminen olisi varmaankin järkevämpi ratkaisu, mutta ei sekään silti täysin triviaalia olisi. Vaikka paikallistamisen tuomat hyödyt näyttäisivät tutkimusten mukaan lupaavilta, siitä seuraava ulkoliikenteen väheneminen ja sisäliikenteen kasvu vaatisi varmastikin myös muutoksia Internetin rakenteeseen ja reititykseen. Tästä huolimatta paikallisuuteen perustuvat ratkaisut helpottaisivat huomattavasti palveluntarjoajien taakkaa. Tutkijat toteavat lopuksi että vaikka monet vertaisverkot eivät ota vertaisten läheisyyttä huomioon, ei tämä ole mikään fundamentaalinen ongelma mitä ei voisi helpottaa pienilläkin muutoksilla.

Lähteet:

Karagiannis, Thomas, Pablo Rodriguez, and Konstantina Papagiannaki. "Should internet service providers fear peer-assisted content distribution?." *Proceedings of the 5th ACM SIGCOMM conference on Internet Measurement*. 2005.

Aggarwal, Vinay, Anja Feldmann, and Christian Scheideler. "Can ISPs and P2P users cooperate for improved performance?." *ACM SIGCOMM Computer Communication Review* 37.3 (2007): 29-40.