

Reititys 2

Autonomisten järjestelmien välinen reititys
Monilähetysreititys

luvut 14 ja 16

Sanna Suoranta, email: T-110.4100@tkk.fi
<https://noppa.aalto.fi/noppa/kurssi/t-110.4100>

15.9.2011
1

Luennon sisältö

- Autonomiset järjestelmät ja kokonaiskuva
- Border Gateway Protocol (BGP)
- **Monilähetysreititys**
 - Internet Group Management Protocol (IGMP)
 - Monilähetyspuu
 - Monilähetysreititysprotokollat



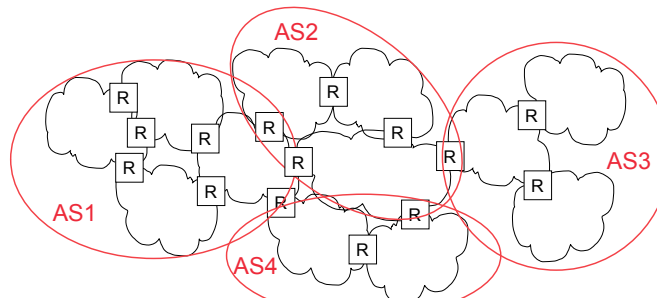
DVMRP	MOSPF
BGP-4	CBT
	PIM
TCP	UDP
	IGMP
	IP
	linkkikerros
	fyysinen kerros

A? Aalto-yliopisto
Tekninen korkeakoulu

Sanna Suoranta, email: T-110.4100@tkk.fi
<https://noppa.aalto.fi/noppa/kurssi/t-110.4100>

15.9.2011
2

Internetin rakenne



- Internet koostuu itsenäisistä verkoista, autonomisista järjestelmistä (autonomous systems (AS))

Autonominen järjestelmä (AS)

- AS on yhtenäistä reitityspolitiikkaa soveltava verkko
 - voi olla useamman organisaation muodostama verkko
 - verkkojen välillä käytetään BGP-protokollaa
- Jokaisella AS:llä on uniikki 32-bittinen tunniste
 - plain- (integer) ja dot-notaatio (1.10)
 - Internetissä nyt ~35000 ASää ja yli 300 000 verkkoprefiksiä
- Kolmenlaisia ASiä
 - stub-verkon kautta ei pääsyä muualle
 - transit-verkon kautta liikenne kulkee muihin verkkoihin
 - multihomed-verkolla on useampi yhteys muualle, mutta se ei välitä läpikulkuliikennettä

Miksi reitityksessä on kaksi tasoa?

- Skaalautuvuus, Internet on liian iso yhdelle reititysprotokollalle
 - Kaikki reitittimet eivät ole samassa verkossa, joten ne ei voi kommunikoida suoraan keskenään
- ASillä on omat hallinnolliset linjauksensa
 - aina ei käytetä lyhintä reittiä vaan esim. halvimman sopimuskumppanin tarjoamaa verkkoyhteyttä
- Asien välillä käytetään eri protokollia kuin sisäiseen reititykseen
 - (yleisnimi näille protokollille on exterior gateway protocol, EGP)
 - huomioi politiikan
 - nykyisin käytössä Border Gateway Protocol 4 (BGP-4)

Käytännössä vielä enemmän hierarkioita

- Kaksi tapaa maksaa toisten liikenteen välittämisestä
 - Transit: maksaa rahaa pääsystä tai välityksestä
 - Peer: "vaihtaminen" eli kumpikin naapuri huolehtii toisen kauttakulkuliikenteestä ilmaiseksi
 - Julkisia pareja (jakavat fyysisen verkon) tai yksityisiä pareja (fyysinen yhteys verkkojen välillä)
- Tier-1-operaattorit muodostavat nykyisin Internetin (hajautetun) runkoverkon
 - liikenteen välitys ilmaista toisen tier-1-operaattorin liikenteelle
- Tier-2-operaattorit myyvät verkkoyhteyttä ISP:lle (ja muille suurille asiakkaille)
 - maksavat Tier-1-operaattorin verkon käytöstä
- Tier-3-operaattorit myyvät verkkoyhteyttä yrityksille ja yksityisille
 - muistattehan viime luennolta OSPF:n alueet!

Seuraavaksi BGP-4



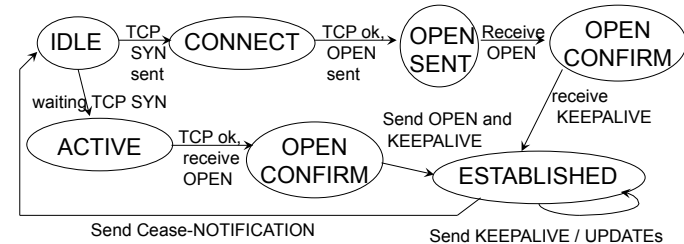
BGP-4 (RFC 4271)

- Välittää saavutettavuustietoa, ei reittien metriikkatietoa
 - Ei ole mahdollista vertailla reittien kustannuksia keskenään, sillä se edellyttäisi sisäverkon tuntemista
 - Mainostetaan siis reittejä, joita pitäisi käyttää, ei kaikkia reittejä
 - Kuormaa voi tasata vain verkkokohtaisesti
- Käytetään polkuvektoreita, jotka kertovat kohde-ASn lisäksi myös kauttakulku-ASn
- iBGP ASn eri reunareitittimien välillä
- BGP käyttää TCP:tä luotettavuuden vuoksi
 - Portti 179

BGP:n toiminnot ja viestit

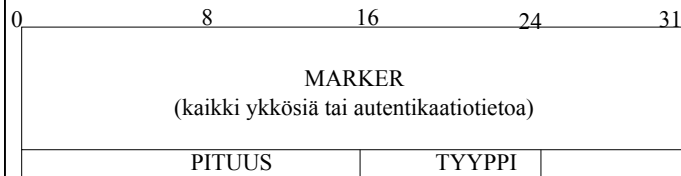
- Alustus
 - osapuolten hankinta ja tunnistus
 - OPEN-viesti yhteyden avaamiseen, KEEPALIVE kuittaa
- Tiedonvaihto
 - saavutettavuustiedon välitys ja päivitys
 - UPDATE-viesti uusien kohteiden ilmoittamiseen
- Tarkistus
 - nykyisen tiedon tarkistus ja testaus
 - KEEPALIVE saavutettavuuden tarkistamiseen
 - NOTIFICATION väärään viestiin vastaus
 - REFRESH pyyntö päivitykselle (RFC 2918)

BGP:n (yksinkertaistettu) tilakone



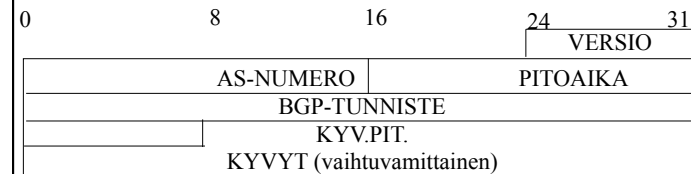
- Yhteyden avaamiseen ja ylläpitämiseen käytetään OPEN ja KEEPALIVE-viestejä
- ESTABLISHED-tilassa vaihdetaan saavutettavuustietoa käyttäen UPDATE-viestejä ja virheistä tiedotetaan NOTIFICATION-viestillä

BGP-otsikko



- MARKERia käytetään synkronointiin
 - kertoo, mistä seuraava BGP-viesti TCP-vuossa alkaa
- Viestin pituus okteetteina
- Viestin tyyppi: open, update, notification, keepalive, tai refresh

BGP Open (edellisen otsikon lisäksi)



- BGP-yhteys avataan lähettämällä Open-viestit, joka lopuksi kuittataan Keepalive-viestillä
- Jos kahden reitittimen välille avataan samaanaikaan kummastakin suunnasta yhteys, toinen niistä suljetaan
 - korkeamman BGP-tunnisteen omaavan aloittama yhteys jää

BGP Open (2)

- Versio = 4
- AS-numero, 16-bittinen tunniste
 - 32-bittiset AS-numerot välitetään optioissa
- HOLD TIME on ehdotus pitoajaksi (sekuntia)
 - kertoo kauanko yhteys on käytettävissä, jollei saa uutta keepalivea
 - käytetään pienempää omasta ja vastapuolen ehdottamasta
- BGP-tunniste: Reitittimen IPv4-osoite
- Kyvyt: lisäominaisuudet, joita BGP-reititin tukee
 - option tyyppi, option pituus, ja option arvo
 - RFC 5492: yleiskuvaus ja unsupported capability -virhekoodi
 - RFC 4893: 32-bittiset AS-numerot
 - RFC 4760: IPv6-yhteensopivuus (verkkokerrosriippumattomuus)

BGP Keepalive

- Viesti sisältää pelkän BGP-otsikon
- Viesti lähetetään viimeistään kun 1/3 pitoajasta on kulunut
 - korkeintaan sekunnin välein kuitenkin
- Tarkoitus on pitää TCP-yhteys auki
 - testaa samalla vastaanottajan olevan edelleen kunnossa
- Ei välitetä reititystietoa samalla
 - säästää verkon kapasiteettia

BGP Notification - virheistä viestiminen

0	8	16	24	31
ERR CODE	ERR SUBCODE	DATA (vaihtuvanmittainen)		

- Otsikossa vikaa
 - synkronointivirhe, joku kenttä vääränmittainen tai tuntematon
- OPEN-viestissä vikaa
 - tukematon versio, väärä AS-numero tai BGP-tunniste, hyväksymätön pitoaika, tuntematon optio
- UPDATE-viestissä vikaa
 - jokin kentistä väärin muodostettu, väärä vastaanottaja (polku rikki)

BGP Notification (2)

- Pitoaika täyttynyt -viesti lähetetään, jos vastaava ajastin laukeaa
- Tilakoneessa vikaa -viesti lähetetään, jos saadaan odottamaton viesti
- Cease-lopetusviestillä suljetaan BGP-yhteys
 - normaalitilanne, ei siis virheviesti oikeasti
- Notification-viestiä ei lähetetä ilmoitusviesteistä
 - virhe raportoidaan jotenkin muuten omalle ylläpitäjälle (ei mitenkään vastapuolelle)

Seuraavaksi BGP-4 reititystaulun rakentaminen ja tietojen välitys



BGP:n reititystaulu (routing information base, RIB)

- Tietokantaan kerätään kolmenlaista tietoa ja se koostuu NLRI-tietueista (Network layer reachability information)
- Local Routing Information Base (Loc-RIB)
 - BGP-reitittimen tietämää paikallista reititystietoa, kerätty itse, ja
 - saatu muilta BGP-reitittimiltä, mutta valittu perustuen omaan politiikkaan
- Adj-RIBs-In
 - muiden BGP-reitittimien lähettämää tietoa
 - saatu UPDATE-viestillä
- Adj-RIBs-Out
 - muille BGP-reitittimille lähetettyä tietoa
 - lähetetään muille UPDATE-viestillä

BGP - reitin valinta

- BGP:n päätösprosessi ei tarkkaan määrää, milloin reitti lisätään ulos mainostettaviin reitteihin ja itse käytettäviin reitteihin
 - pitää olla reitti next-hop-kentässä määrättyyn kohteeseen
 - vain yksi reitti / kohde valitaan
- Yleiset periaatteet reitin valintaan
 - lyhin AS_PATH eli pienin määrä välissä olevia Asiä
 - pienin ORIGIN attribuutti eli paikallinen tieto ensisijaisesti (paikallisista se jolla korkein paikallinen preferenssi)
 - pienin MULTI_EXIT_DISC (MED) arvo (erottaa naapurin tarjoamat entrypointit toisistaan) tai sama yhteisö (community)

BGP Update - Reiteistä ja niiden muutoksista tiedottaminen

0	8	16	24	31
LOPETETUT-PITUUS		LOPETETUT		
REITIT (vaihtuvanmittainen)				
POLUN ATTR.-PITUUS		POLKU-		
ATTRIBUUTIT (vaihtuvanmittainen)				
KOHDEVERKOT (vaihtuvanmittainen)				

- Lähetetään kun BGP-yhteys on avattu tai kun verkon tila muuttuu
- Viesti sisältää BGP-otsikon ja joko sekä lopetettuja reittejä että uusia reittejä tai vain jompia kumpia

BGP Update (2)

- Lopetettujen reittien kentän pituus oktetteina
 - jos nolla, ei lopetettuja reittejä lainkaan
- Lopetetut reitit

PITUUS	PREFIX (vaihtuva)
--------	-------------------

 - lopetettujen verkkojen IP-osoitteet
 - oikeasti verkkomaskin pituus, ja verkkomaskin verran osoiteesta
- Uudet reitit: Kaikkien polun attribuuttien pituus
 - sekä polun attribuuttien että kohteiden pituus
- Uudet reitit: Kohdeverkot
 - Lista IP-osoitteiden verkko-osia, samaan tapaan kuin lopetetut
 - Kentän pituus: koko update-viestin pituus - 23 - polun attribuuttien pituus - lopetettujen reittien pituus

BGP Update - Polun attribuutit

0				8				16				24				31			
W	T	P	E	0	0	0	0	ATT. TYPE				ATT. PIT (1 tai 2)				ATT. DATA (v)			

- Käytetään UPDATE-viestissä
- Attribuutit, neljä erilaista
 - pakolliset ja harkinnanvaraiset tunnetut
 - valinnaiset transitiiviset ja ei-transitiiviset
- Bitit (optioiden käsittelyyn jne)
 - W: tunnettu (=0, well-known) tai vapaaehtoinen (=1)
 - T: transitiivinen (=1, hyväksytään ja välitetään eteenpäin) tai ei-transitiivinen (=0, ignoroidaan eikä välitetä eteenpäin)
 - P: jos valinnainen transitiivinen tieto on osittaista (=1), muuten 0
 - E: attribuutin pituus on yksi oktetti (=0) tai kaksi oktettia (=1)

BGP Polun attribuutit (2) - pakolliset tunnetut tyypit ja niiden arvot

- 1 = ORIGIN eli tiedon lähde
 - 0 = IGP, 1 = EGP (RFC 904), 2 = jotenkin muuten
- 2 = AS_PATH <tyyppi, pituus, arvo> tripletteinä,
 - tyyppi 1=järjestämätön ja 2= järjestetty lista Asistä
 - pituus = ASien lukumäärä
 - arvo = ASn tunniste (voi olla useita)
- 3 = NEXT_HOP = seuraavan reitittimen IP-osoite

BGP Polun attribuutit (3) - muut tyypit ja arvot

- 4 = MULTI_EXIT_DISC
 - Tunniste, jonka avulla voi erottaa vastaanottajaverkon rajapinnat toisistaan (useita rajapintoja tarjolla)
 - Valinnainen, ei-transitiivinen
- 5 = LOCAL_PREF
 - Samassa verkossa oleville tiedoille ilmoitetun reitin preferenssistä
 - tunnettu

BGP Polun attribuutit (4) - Reitin koostaminen

- Useiden reittien tiedoista voidaan yhdistää (aggregate) yksi reitti
 - Reitin pitää kulkea verkoista ulos saman reitittimen kautta
- 6 = ATOMIC AGGREGATE
 - Reitin ilmoittaminen ilman kaikkia sen varrella olevia AS-numeroita
 - Tunnettu, harkinnanvarainen
- 7 = AGGREGATOR = reitin koostajan tunniste
 - AS-numero + reitittimen IP-osoite

BGP - Communities Attribute (RFC 1997 ja RFC 4360)

- Reititystiedon ryhmittelyyn, yksinkertaistaa reititystauluja
 - joukko kohteita, joilla on sama ominaisuus, muodostaa yhteisön
 - oletusryhmä: "yleinen Internet"
- RFC 1997: Yhteistö tunnustetaan 32-bittisillä luvuilla
 - ekat 16 on AS:n tunniste, sitten AS-spesifistä
 - attribuutin tyyppi = 8
- RFC 4360: kolme erilaista attribuuttia
 - IANAn määrittämä, globaali ryhmä (sis. listan jäsenistä)
 - IPv4-osoitealue
 - geneerinen määritelmä ryhmälle
 - attribuutin tyyppi = 16

BGP - Refresh - Poliitiikan muutokseen reagoiminen (RFC 2918)

0	8	16	24	31
OSOITEPERHEEN ID		VARATTU (0)		SAFI

- Kun reitityspoliittika muuttuu, täytyy kaikkien reittien tiedot päivittää uuden reitityspoliittikan mukaiseksi
 - Ennen reitin käynnistettiin uudestaan ja pidettiin kopiot vanhasta reititystaulusta, mikä vei paljon muistia
- "Soft reset" toteutetaan Refresh-viestillä
 - molempien osapuolten pitää tukea (sovitaan yhteyden aluksi mitä tietoja tämä koskee)
 - Osoiteperheen tunniste (AFI address family identifier): IPv4=1, v6=2
 - Alitunniste (SAFI subsequent address family identifier), 1=yksilähetys, 2=monilähetys

Internal BGP (IBGP)

- Laajan AS:n reunareitittimet keskustelevat keskenään BGP:llä (IBGP)
 - Verkon pitää antaa itsestään ulos konsistentti kuva
 - Alunperin kaikkien reunareitittimien pitää olla kokoajan BGP-yhteydessä toisiinsa (full mesh)
- RFC 4456 määrittelee tavan klusteroida IBGP-reitittimiä
 - Esim joku reitittimistä voi vastata koko verkosta
 - reititystieto lähetetään "tähtimäisesti" muille saman AS:n reitittimille

Seuraavaksi BGP ja turvallisuus



BGP ja turvallisuus

- Usein hyökkäyksen lopputulema on sama kuin väärinkonfiguroinnin
 - Man in the middle ja salakuuntelu: uudelleenohjaus jonkin hyökkääjän verkon kautta salakuuntelun vuoksi tai verkon ruuhkauttamiseksi
 - Palvelunesto: joko verkon saattaminen epästabiliin tilaan tai jonkin kohteen liikenteen ohjaaminen väärään paikkaan mainostamalla väärää reittiä (musta aukko)
- Suojamekanismeja
 - Reititinvalmistajien ohjeet hyvistä käytännöistä
 - TTL=1, eli viestejä otetaan vastaan vain naapurireitittimeltä, sillä oikeasti vain naapuriin voi luottaa
 - TCP-session suojamekanismit
- Pari kehitteillä olevaa tapaa

PKI-pohjaiset: Secure BGP (S-BGP) ja Secure Origin BGP (soBGP)

- S-BGP:ssä kolme turvamekanismia
 - X.509-sertifikaattipohjainen PKI IP-osoitteiden ja AS-numeroiden omistajien tunnistamiseksi
 - BGP transitive path attribute UPDATE-viestien allekirjoittamiseen
 - IPsec osallistuvien reitittimien tunnistamiseen ja liikenteen suojaamiseen
- soBGP:ssä AS-reitittimet muodostava luottamusverkon (web-of-trust)
 - Tarkistetaan AS:n oikeus mainostaa verkkoa sekä polku
 - Naapuri voi "vahvistaa" omien naapuriensa luomaa tietoa
- IETF:ssä drafteina 2003 ja 2006, ei enää voimassa

Secure Inter-Domain Routing (SIDR)

- Tavoite: kehittää tapa tiedon lähteen autentikoimiseen
 - Eli että tietty verkko kuuluu sille AS:lle, joka verkkoa mainostaa
 - Reunareitittimen "policy station" valvoo, että samaa politiikkaa käytetään AS:n sisällä ja kerätä lokia
 - Politiikassa määritellään mitkä verkko-prefiksit kuuluvat AS:ään
 - Draft määrittelee, miten verkko-prefiksejä mainostetaan ulos verkosta
 - Yhteensopiva SBGP:n ja soBGP:n kanssa
- Työryhmä IETF:ssä : <http://datatracker.ietf.org/wg/sidr/charter/>

Seuraavaksi BGP ja IPv6



BGP ja IPv6 (RFC 4760) eli Multiprotocol Extension for BGP-4

- IPv4-osoitespesifisiä ovat BGP:ssä vain
 - Seuraava hyppy
 - Aggregator eli ID polun koostavalle reitittimelle
 - NLRI eli IPv4 osoitteen verkko-osa
- Jotta BGP toimisi jonkin toisen verkkokerroksen protokollan kanssa, kaksi asiaa pitää lisätä
 - Seuraavan hypyn osoitteeksi ko verkkokerroksen osoite
 - Mahdollisuus ilmiottaa jonkin toisen verkkokerroksen teknologian protokollasta NLRI-kentässä
- Edelleen oletetaan, että BGP-reitittimellä on IPv4-osoite

BGP ja IPv6 - uudet attribuutit

- Valinnaisia ei-transitiivisia attribuutteja
 - MP-BGP:tä osaamaton reititin ignoroit attribuutit
- Tavoitettavuus ja tavoittamattomuus -viestit eriytetty omiksi saavutettavuustietoattribuuteikseen
- Multiprotocol reachable NLRI (MP_NLRI)
 - Käytetään kertomaan uudesta reitistä
 - Myös uudesta seuraavasta hypystä, jota tulisi käyttää
- Multiprotocol unreachable NLRI
 - Poistuneiden reittien tunnisteet (samoin kuin edellä uusien reittien tunnisteet)

BGP ja IPv6 - MP reachable NLRI

0	8	16	24	31
OSOITEPERHEEN TUNNISTE		ALIOSOITEP.ID	OSOIT. PIT.	
SEURAAVAN HYPYN OSOITE (vaihtuvanmittainen)				
VARATTU		VERKON SAAVUTETTAVUUS- TIE TOA (vaihtuvanmittainen)		

- Osoiteperheen (AFI) ja aliosoitteperheen tunniste (SAFI) kuten Refresh-viestissä
 - osoiteperheen tunnistimien perusteella tiedetään osoitteen tyyppi, silti sille myös oma pituus kenttä
- Seuraavan hypyn osoite

BGP ja IPv6 - Saavutettavuustieto

- Saavutettavuustieto koodataan yhteen tai useampaan <pituu, prefix> pariin, jossa
 - Pituu kertoo prefixin pituuden
 - Prefix kentässä on osoitteen verkko-osa (+täytetty)
- Jos reitin tukee useampaa erilaista osoitetta, saavutettavuustiedossa on jokaiselle oma osio

AS:ien välinen reititys - yhteenveto

- Käytetään BGP-4-protokollaa
- Ei välitetä etäisyyksiä, vain pelkkää saavutettavuustietoa
- Reitintä valintaan ei vaikuta vain sen pituus, vaan myös todelliset kustannukset ja muu politiikka

Lähteet - ASien välinen reititys

- RFC 4271 A Border Gateway Protocol 4 (BGP-4), 2006
- RFC 1773 Experience with the BGP-4 protocol, 1995
- RFC 1774 BGP-4 Protocol Analysis, 1995
- RFC 4272 BGP Security Vulnerabilities Analysis, 2006
- RFC 4893 BGP Support for Four-octet AS Number Space, 2007
- RFC 5492 Capabilities Advertisement with BGP-4, 2009
- RFC 2918 Route Refresh Capability for BGP-4, 2000
- RFC 5386 Textual Representation of Autonomous System (AS) Numbers, 2008
- RFC 1997 BGP Communities Attribute, 1996
- RFC 4360 BGP Extended Communities Attribute, 2006
- RFC 4760 Multiprotocol Extension for BGP-4, 2007
- RFC 4456 BGP Route Reflection: An Alternative to Full Mesh Internal BGP (iBGP), 2006
- RFC 5065 Autonomous System Confederations for BGP, 2007
- RFC 1997 BGP Communities Attribute, 1996
- IANA Address Family Numbers: <http://www.iana.org/assignments/address-family-numbers/address-family-numbers.xhtml>
- S-BGP: <http://www.ir.bbn.com/sbgp/>
- soBGP: <http://tools.ietf.org/html/draft-white-sobgp-architecture-02>
- Secure Inter-Domain Routing working group <http://datatracker.ietf.org/wg/sidr/charter/>

Seuraavaksi
monilähetys



Monilähetys

- Yhdeltä monelle -liikenne
 - verkko huolehtii viestin monistamisesta ja perillekuljettamisesta kaikille vastaanottajille
- Monilähetyksryhmä kertoo kohteet
 - mukaanliittymisilmoitusten perusteella, jolloin ryhmään voi liittyä tai siitä voi poistua dynaamisesti
 - tai etukäteen määriteltäjä ryhmiä
- Edut
 - vähentää verkon liikennettä

Monilähetys Internetin laajuisesti

- Mbone toimi monilähetyksen runkoverkkona 1990-luvulla
 - kokeellinen, mm. videokonferenssien lähettämiseen
 - Runkoverkon yli yksilähetyksenä muodostettu verkko
- Nykyisin BGP-reititin voi ilmoittaa monilähetysosoitteita
 - RFC 4760 MP-BGP (eli sama kuin IPv6:lle)
- Todellisuudessa monilähetystä ei välitetä Internetin laajuisesti
 - ei ole tehokasta tapaa monilähetyksien reittien pääsynvalvontaan
 - verkko-operaattorien on vaikea määrittää miten laskuttaa monilähetyksestä

Monilähetyksen käyttökohteita

- Tutkimuskäyttö
 - Paljon ideoita :)
- Sovelluskäyttö
 - Paikallisessa verkossa IPTV-sovelluksen liikenteen välitys
- Verkonhallinnallinen käyttö
 - IPv6-verkoissa mm. reitittimien ja naapurien löytämiseen, sillä IPv6 ei tue yleislähetystä
 - Myös IPv4-verkossa palvelun löytämiseksi

Monilähetys IP-verkoissa

- Monilähetysosoite voi olla vain vastaanottajalla
 - pysyvät ja väliaikaiset osoitteet
 - paikalliset ja globaalit osoitteet
 - sekä IPv4- että IPv6-verkoissa
- Tapa liittyä ryhmään
 - IPv4: Internet Group Management Protocol (IGMPv3, RFC 3376)
 - IPv6: Multicast Listener Discovery (MLDv2, RFC 3810)
- Monilähetyksreititys
 - Monilähetykspuun rakentaminen
 - Monilähetyksreititiedon välittäminen reititimille

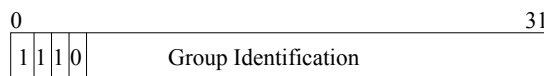
Seuraavaksi monilähetysosoitteet



Monilähetysosoitteet

- Monilähetysosoite toimii monilähetysryhmän tunnisteena
 - Ei kerro, missä ryhmän jäsenet sijaitsevat
- Kuka tahansa voi lähettää viestin, jonka vastaanottaja on monilähetysryhmä
 - Mutta myös tapa rajoittaa ryhmäkohtaisia lähettäjiä
- Sekä pysyviä että dynaamisesti jaettavia osoitteita
 - IANA allokoivat pysyvät
- Sekä vain paikalliseen verkkoon että Internetin laajuisesti välitettäviä osoitteita
 - IPv6:ssa osoitteen lisähierarkiana organisaation verkko
 - IPv4:ssä TTL tai RFC 2365:n avulla IPv6:n kaltainen

IPv4-monilähetysosoitteet



- “luokka D” tai CIRD-notaatiolla 224.0.0.0/4
 - 224.0.0.0 - 239.255.255.255
- Dynaamisesti varattavia osoitteita
 - 224.0.2.0 - 224.0.255.255, 224.3.0.0 - 224.4.255.255, ja
 - 233.252.0.0 - 233.255.255.255, jota tosin käytetään myös GLOB-blokin laajennukseen 32-bittisille AS:lle

IPv4: pysyvät monilähetysosoitteet

- Paikallisen verkon kontrollointiin: 224.0.0/24
 - ei välitetä ulos linkin verkosta
 - esim OSPF:n reitittimille 224.0.0.5
- Internetin laajuisen verkon kontrolliin: 224.0.1/24
 - voidaan välittää Internetiin
 - esim. 224.0.1.1 Network Time Protocol NTP (RFC 5905)
- Session Announcement Protocol (RFC 2974) avulla varattavat osoitteet: 224.2/16
 - Monilähetystä käyttävät monelta-monelle yhteydet (multicast multimedia conferences)
 - istuntojen perustamis- ja ylläpitoviestintää

IPv4: pysyvät monilähetysosoitteet (2)

- Source-specific multicast: 232/8, FF3x::/32 (RFC 4607)
 - kohta lisää
- GLOP-osoitteet: 233/8 (RFC 3180)
 - sisältää keskimmaisissa okteteissa AS:n numeron
 - pysyviä ryhmiä
 - AS määrittää viimeisen oktetin itse
- Hallinnollisesti rajatut osoitteet (Administratively scoped): 239/8 (RFC 2365) IPv6:n tyyliin IPv4:lle, ei globaaliin Internetiin
 - 239.255.0.0/16 IPv4 local: vain paikalliseen fyysiseen verkkoon
 - 239.192.9.9/14 IPv4 organisaation: vain saman organisaation verkkoon
 - aiemmin käytetty TTL:n avulla liikenteen leviäminen hankaloitti tehokasta monilähetysreitityspuun luontia

IPv6-monilähetysosoitteet (RFC 4291)

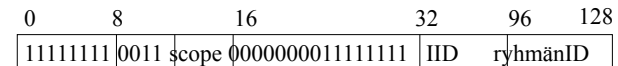


- Ryhmän ID: tunniste ryhmälle
- Scope rajoittaa viestin leviämistä verkkoon
 - 1 = paikallinen kone, 2=fyysinen verkko, 4=hallinnollisesti rajattu (useampi fyysinen), 5=paikallinen, 8 = organisaatio, E=global
- R (RFC 3956) - Rendezvous point PIM-reititykselle
- P (RFC 3306) - Unicast-osoitepohjainen monilähetysosoitteiden allokointi
- T = 0 =pysyvät ja 1= dynaamiset osoitteet

IPv6-monilähetysosoitteiden jako (RFC 3307)

- Pysyviä (IANA määrittelee, RFC 2375)
 - Node-local - esim. kaikki noodit ja kaikki reitittimet
 - Link-local - esim. DHCP ja (monilähetys)reititykseen liittyviä ryhmiä
 - Site-local - esim. DHCP ja kaikki reitittimet -ryhmät
 - vaihtuvan skoopin ryhmiä, mm. monilähetysreititykseen
- Dynaamisia
 - palvelimen allokioimia, käytetään Multicast Address Dynamic Client Allocation Protocol (MADCAP) -protokollaa (RFC 2730)
 - noodin itse allokioimia (RFC 4489) kuten seuraavassa esitetään

IPv6-monilähetysosoitteet (RFC 4489)



- IPv6 tarjoaa koneelle mahdollisuuden luoda dynaamisesti monilähetysosoitteita käyttämilleen sovelluksille
 - koska joka koneella on Interface Identifier (IID): 64-bittinen uniikki tunniste verkon rajapinnalle (IPv6-osoitteen koneosa)
- Ryhmän ID: 32-bittinen uniikki tunniste monilähetyssovellukselle

Seuraavaksi monilähetyksen lähettäminen



Monilähetyksen lähettäjän näkökulmasta

- IP-kerroksen luoma monilähetysosoitteeseen pakattu viesti kapseloidaan linkkikerroksen kehykseen ja lähetetään linkkikerroksen monilähetysosoitteeseen, jos se tukee monilähetystä, muuten yleislähetyksenä
 - Sekä paikallisen verkon että Internetin monilähetysosoitteeseen
- Paikallisverkossa monilähetysosoite määritetään linkkikerroksen monilähetysosoitteeksi, jota kaikki ryhmään kuuluvat kuuntelevat
- Reititin kuuntelee monilähetysosoitteita, ja välittää viestin eteenpäin perustuen reititystaulussaan olevaan tietoon
 - jos siis kyse osoitteesta, joka välitetään ulos

Monilähetyksen - virhetilanteiden käsittely

- IPv4-monilähetyksen virhetilanteissa ei lähetetä ICMP-virheviestejä
 - monilähetysosoitetta ei voi pingata
 - tapahtuneista virheistä ei kerrota lähettäjälle
- IPv6-monilähetyksen virhetilanteissa lähetetään ICMP-viesti
 - alkuperäiselle lähettäjälle
 - koska käytetään mm verkkoa koskevan tiedon välitykseen

Luotettava monilähetyksen

- Luotettava: kaikki paketit menevät perille järjestyksessä ilman virheitä
 - toteutetaan esim. TCP:ssä kuittauksilla
 - Kuittauksia tulisi ihan liikaa, jos ryhmässä paljon kuulijoita
- Negatiiviset kuittaukset
 - Kerrotaan, jos jokin odotettu paketti ei tullut perille
- Kuittauksien silti mahdollinen
 - Kuittauksia voisi yhdistää hierarkisen reitityksen määrittämissä pisteissä
- Edelleen kehitteillä IETF:ssä

Seuraavaksi monilähetyksen vastaanotto (paikallinen verkko)



vierevinU otaA
asneisä to loorisä
vgolonisat bna



Sanna Suoranta, email: T-110.4100@tkk.fi
<https://noppa.aalto.fi/noppa/kurssi/t-110.4100>

15.9.2011
57

Tunnetut monilähetyksryhmät

- Jos laite kuuluu tunnettuun monilähetyksryhmään, se kuuntelee ko. ryhmän lähetyksiä ilman ryhmään liittymismenettelyäkin
 - jos linkkiverkko tukee monilähetystä, kuunnellaan ko monilähetyksryhmän osoitetta vastaavaa linkkikerroksen osoitetta
 - jos linkki ei tue monilähetystä, kuunnellaan kaikkea linkkikerroksen yleislähetystä (broadcast), ja IP-kerroksella kehyksen purun jälkeen havaitut kuunneltavan monilähetyksryhmän paketit käsitellään (muut hylätään)

A? Aalto-universitetet
Tekniska högskolan

Sanna Suoranta, email: T-110.4100@tkk.fi
<https://noppa.aalto.fi/noppa/kurssi/t-110.4100>

15.9.2011
58

Internet Group Management Protocol IGMPv3 (RFC 3376)

- IPv4-verkossa käytetään dynaamisiin ryhmiin liittymisissä IGMP-protokollaa
 - ICMP:n kaltainen, toimii suoraan IP:n päällä
- Yksittäinen kone käyttää ryhmään liittymiseen
 - IGMP-viesti lähetetään halutun ryhmän osoitteeseen
 - Paikallinen reititin huomaa viestin ja kertoo muille reitittimille
- Reititin pitää tietoa ryhmistä, joissa on jäseniä
 - tarkistus säännöllisesti, onko reitittimen hallitsemassa verkossa yhä kuulijoita (yksikin vastaus riittää, kaikkien ei tarvi vastata)
 - jos ei ole, kerrotaan muille reitittimille ja päivitetään reititystaulua

A? Aalto-universitetet
Tekniska högskolan

Sanna Suoranta, email: T-110.4100@tkk.fi
<https://noppa.aalto.fi/noppa/kurssi/t-110.4100>

15.9.2011
59

IGMPv3-protokolla (2)

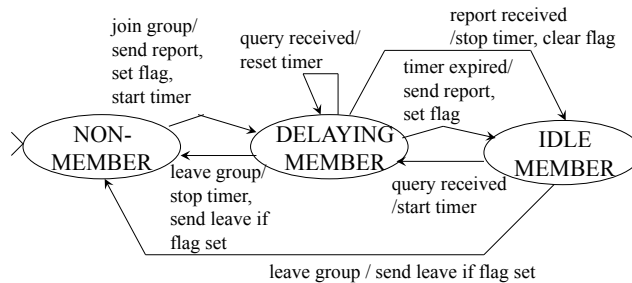
- IGMPv3 tarjoaa lähettäjän suodatuksen
 - taaksepäin yhteensopiva sekä v1 (RFC 1112) että v2 (RFC 2236) kanssa
- Kaksi viestiä
 - QUERY: reititin kysyy onko sen verkossa ryhmän kuulijoita
 - REPORT: ryhmän jäseneltä reitittimelle raportti kuunneltavista ryhmistä
 - (IGMPv2 lisäksi myös LEAVE, v3 käyttää REPORT-viestiä)

A? Aalto-universitetet
Tekniska högskolan

Sanna Suoranta, email: T-110.4100@tkk.fi
<https://noppa.aalto.fi/noppa/kurssi/t-110.4100>

15.9.2011
60

IGMP: Isäntäkoneen yksinkertaistettu tilakone

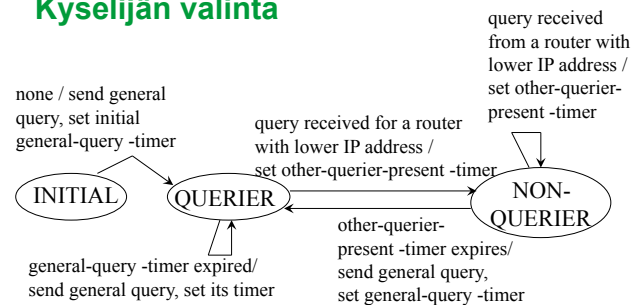


- IGMPv2 (RFC 2236), eli ei lähettäjän suodatus (RFC 3376)

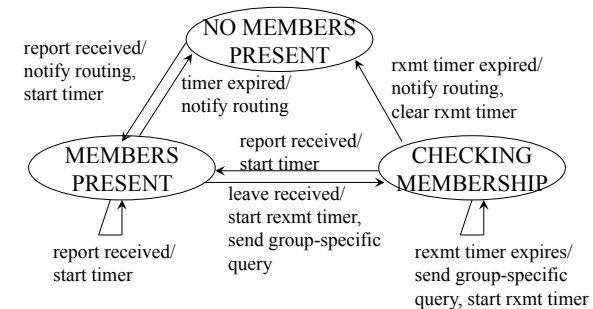
Reitittimen tehtävä ryhmien hallinnassa

- Verkossa voi olla useampi monilähetysreititin
 - Ensin valitaan pienimmän IP-osoitteen omaava reititin kyselijäksi
 - Kyselijä selvittää aina välillä, onko muita monilähetysreitittejä verkossa
- Kyselijä selvittää
 - onko ylipäänsä missään ryhmässä kuulijoita
 - onko tietyssä ryhmässä kuulijoita
 - onko tietyssä lähettäjäspesifisessä ryhmässä kuulijoita
- Pidetään kirjaa ryhmistä, joissa kuulijoita
 - "nopein" vastaa, yksi vastaus riittää (ei pidetä kirjaa tarkemmin)
 - myös ei-kyselijätilassa olevat reitittimet tallettavat tiedon

IGMP: Reitittimen tilakone (RFC 2236) Kyselijän valinta



IGMP Kyselijä-reitittimen tilakone (RFC 2236)



IGMPv3-kysely (query, RFC 3376)

0	8	16	31
TYPE = 0x11	MAX RESP TIME	CHECKSUM	
ZERO OR GROUP ADDRESS			
0000	S Q RV	Q QIC	NUMBER OF SOURCES
SOURCE ADDRESS [1]			
...			
SOURCE ADDRESS[n]			

- Suoraan IPv4-otsikon sisään pakattuna
 - protokollan numero IGMP:lle on 2
 - TTL = 1, eli ei välitetä ulos verkosta
 - Yleiskysely lähetetään osoitteeseen 224.0.0.1 "kaikki",
 - muut kyselyt ryhmän omaan monilähetysosoitteeseen

IGMPv3-kysely (2)

- Tyyppi: query
- Maksimivastausaika 0.1 s tarkkuudella (ing tai float)
- Joko tietylle ryhmälle tai kaikille, jolloin osoite nolli
- Kyselynhallintabittejä
 - S (Suppress Router-Side Processing) käytetään tavallisia arvoja ajastimille
 - QRB (Querier's Robustness Variable) reitittimelle kyselyn lähettämiseksi
 - Q|QIC (Querier's Query Interval Code) käytetään kyselyille ajastusta
- Lähdeosoitteiden lukumäärä
 - 0, jos yleis- tai ryhmäkohtainen kysely
 - jotain muuta, jos ryhmä+lähdekohtainen kysely
- Sallittujen lähteiden IP-(yksilähetys)osoitteet
 - Niin monta, kuin mahtuu MTU:hun

IGMPv3-raportti (Membership report)

0	16	31
TYPE=0x22	ZERO	CHECKSUM
RESERVED=ZERO	# of Group Records	
GROUP RECORD[1]...		

- Kohde voi kertoa kaikki kuuntelemaansa monilähetysryhmät yhdellä viestillä
- Raportin otsikko kertoo lähinnä kuinka monta ryhmäraporttia viesti sisältää
 - vanhemmilla IGMP-viesteillä on omat raportointiviestinsä, joita tuetaan edelleen

IGMPv3 ryhmäraportti (Group Report)

TYPE	Aux Data Len	Number of Sources
MULTICAST ADDRESS		
Source Address [1]		
... [n]		
AUXILIARY DATA		

- Sen ryhmän monilähetysosoite, jota kuunnellaan
 - myös lähteet, jos ne on määritelty
- Tyyppi: lisätään tai ei lisätä uusia lähteitä, lähdelista on muuttunut, tai noodi filteröi /ei filteröi lähteitä jatkossa
- Apudataa ei ole (vielä) määritelty

Multicast Listener Discovery v2 (RFC 3810)

- IGMP:tä vastaava ryhmien hallintaprotokolla IPv6:lle
 - Samaan tapaan selvitetään onko ryhmissä kuuntelijoita
 - Yksi kyselee, muut kuuntelevat ja joku niistä vastaa
 - Viestit on aika samannäköisiä, lähinnä osoitteet on pidempiä
- Samoin MLDv2 on ICMPv6:n aliprotokolla, kuten IGMP noudattaa ICMP-viestien rakennetta
 - Next-header arvo IPv6-otsikossa 58



Seuraavaksi monilähetysreititys

Monilähetysreititys

- Yksilähetystauluja ei voi käyttää suoraan monilähetystykseen, eikä yksilähetystyksen reititysprotokollia voi helposti muokata monilähetysreititysprotokolliksi
 - Ryhmään liittyminen ja siitä poistuminen tapahtuu useammin kuin mitä yksilähetystyksen verkossa tapahtuu muutoksia
 - Lähettäjän sijainnin lisäksi vastaanottajien sijainti vaikuttaa reititykseen
 - Viestien monistuminen niiden kulkiessa eri kautta pitää ratkaista
 - Viesti saattaa kulkea sellaisen verkon läpi, jossa ei ole kuulijoita
 - Viesti saatetaan tunneloida monilähetystä osaamattoman läpi
- Yksilähetystyksen luomaa kuvaa verkon lyhimistä reiteistä voi käyttää apuna, mutta se ei yksin riitä monilähetyspuun luomiseen

Kaksi periaatetta lähettämiseen

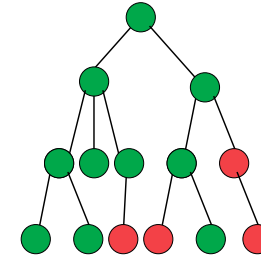
- Kuka vaan voi lähettää viestin monilähetysosoitteeseen
 - Any source multicast (ASM)
 - Alkuperäinen malli
- Vain tietty lähde voi lähettää viestin
 - Single-source multicast (SSM)
 - yksinkertaisempi, nykyisin IGMPv3:ssa tuettu malli

Yksinkertainen monilähetys: Käänteispolkureititys (Reverse Path)

- Käytetään hyväksi lähettäjän osoitetta ja yksilähetysreititystä (reverse path forwarding RPF)
 - joss paketti saapui lähettäjän osoitteen osoittamasta suunnasta, lähetetään se eteenpäin kaikkialle muualle
- Typistävä versio (Truncated RPF)
 - edellinen, mutta lisäksi tarkistetaan, että
 - lähetetään paketti eteenpäin vain niihin verkkoihin, jossa on ryhmän vastaanottajia, muuten ei lähetetä
- Reverse path multicast (RPM)
 - reititysprotokolla edellisten periaatteiden tapaan
 - IGMP:llä selvitetään paikallisverkon osallistujat

Kaksi tapaa luoda monilähetyspuu

- Datalähtöinen (data-driven)
 - lähetetään kaikkialle, kunnes saadaan tieto, ettei lähetystä kukaan tarvitse tietyssä verkossa
- Tarvelähtöinen (demand-driven)
 - lähetetään vain, kun saadaan tieto, että lähetystä haluaisi joku kuunnella jossain tietyssä verkossa



Distance Vector Multicast Routing Protocol (DVMPR, RFC 1075)

- mrouterd unixeissa
- Pohjautuu RIP-protokollaan
 - Lisäksi tietoa monilähetyksestä <ryhmä, lähde>-pareina
 - Sisäverkon reititykseen
- Data-lähtöinen
 - Määritteli omia IGMP-viestejä reitittimien ryhmäänliittymistä varten
 - IGMP-viestit ryhmästä lähtemiseen
- Puutteita
 - pitää muistaa myös vanhat reitit, jos joku liittyy takaisin
 - ei skaalautu kovin hyvin

Core Based Trees (CBT, RFC 2189)

- Tarvelähtöinen sekä sisä- että ulko-verkon reititykseen
 - Paikallisverkossa ryhmään liitytään normaalisti IGMP:llä
 - Jakaa Internetin alueisiin, joissa jokaisessa vastuureititin (core)
 - Reititin lähettää pyynnön vastuureitittimelle, mutta "lähin" matkalla oleva jo ryhmää vastaanottava reititin keskeyttää viestin ja lähettää kiittauksen takaisin, muut päivittävät reititystaulunsa kiittausviestin perusteella
 - puuta pidetään hengissä (keepalive)
- Tieto talletetaan <ryhmä, incoming interface, outgoing interface>, eli tiedetään puun juuren ja lehtien suunta
- Ulkopuolinen lähettäjä: viestit vastuureitittimen kautta
- Miten ryhmään kuuluva lähettäjä?

Protocol Independent Multicast (PIM)

- Kaksiosainen protokolla
 - Koska kumpikaan monilähetyspuun luontitavoista ei ole oikein hyvä
 - mikä vaan yksilähetysreititys alle
 - erikseen sekä kaikille (ASM) että tietyille (SSM) lähettäjiille
- PIM Dense mode (RFC 3973)
 - tarkoitettu verkkoon, jossa paljon (tiheästi) kuulijoita
 - käytetään datalähtöistä tapaa
- PIM Sparse mode (RFC 4601)
 - tarkoitettu verkkoon, jossa reunoilla on paljon kuuntelijoita
 - käytetään tarvelähtöistä tapaa, CBT:n tyyliin, eli
 - valitaan yksi reitittimistä rendezvous pointiksi
- (Paljon lisää RFC:tä mm autentikointiin)

Multicast OSPF (MOSPF, RFC 1584)

- OSPF:n laajennus monilähetystykseen
- Käyttää tarvelähtöistä tapaa
- Verkko on jaettu alueisiin
 - reunareititin (multicast area border router) pitää huolen ryhmätiedon levittämisestä

Muuta monilähetystyksen reitityksestä

- BGP:hen multiprotocol BGP myös monilähetystiedon reititykselle ASien välillä
- IETF:ssä työryhmiä kehittämässä monilähetystä
 - multicast mobility
 - multicast security
 - reliable multicast transport
 - protocol independent multicast
- Wikipediassa yli 20 erilaista monilähetysreititysprotokollaa
http://en.wikipedia.org/wiki/Multicast_routing_protocol#Multicast_routing
 - Paljon on tutkittu ja ehdotettu tapoja reititykseen

Monilähetys - yhteenveto

- Tehokas tapa välittää tietoa ryhmälle vastaanottajia
- Vaatii uusia reititysprotokollia tai muutoksia vanhoihin
- Kasvattaa huomattavasti reititystaulujen kokoa
- Ei juurikaan käytössä
 - vaikea rahastaa, mukava t(h)utkia

Lähteet - monilähetys

- RFC 5777 IANA Guidelines for IPv4 Multicast Address Assignments, 2010
- RFC 2975 Session Announcement Protocol, 2000
- RFC 3180 GLOB Addressing in 233/8, 2001
- RFC 5757 Multicast Mobility in Mobile IP Version 6 (MIPv6): Problem Statement and Brief Survey, 2010
- RFC 3376 Internet Group Management Protocol v3, 2002
- RFC 4604 Using Internet Group Management Protocol Version 3 (IGMPv3) and Multicast Listener Discovery Protocol Version 2 (MLDv2) for Source-Specific Multicast, 2006
- RFC 3810 Multicast Listener Discovery Version 2 (MLDv2) for IPv6, 2004
- RFC 5519 Multicast Group Membership Discovery MIB, 2009
- RFC 4670 Source-specific multicast, 2000
- RFC 1075 Distance Vector Multicast Routing Protocol DVMRP, 1988
- RFC 4601 Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised), 2006
- RFC 5796 Authentication and Confidentiality in
- Protocol Independent Multicast Sparse Mode (PIM-SM) Link-Local Messages, 2010
- RFC 5059 Bootstrap Router (BSR) Mechanism for Protocol Independent Multicast (PIM), 2008
- RFC 2189 Core Based Trees (CBT version 2) Multicast Routing -- Protocol Specification, 1997
- RFC 2201 Core Based Trees (CBT) Multicast Routing Architecture, 1997
- RFC 1585 MOSPF: Analysis and Experience, 1994
- RFC 1584 Multicast Extensions to OSPF, 1994

Seuraavaksi

- Miten löydetään liikkuva kone?
- Mitä tehdä, jos koneella on useampi verkkoyhteys (multihomed)
- Miten toimitaan, kun verkon muodostaa joukko samanarvoisia noodeja (ad hoc-verkko), eikä reititintä ole?

