



S-38.3192 Verkkopalvelujen tuotanto

Luento: BGP

Osa materiaalista on kopioitu tai johdettu Olivier Bonaventure'n BGP luentosarjasta
www.info.ucl.ac.be/people/OBO/BGP/



Sisäiset reititysprotokollat

- **Tehtävä**
 - Muodostaa tarkan kuvan autonomisen alueen sisäisestä rakenteesta
 - Jokainen reititin ja linkki on kuvattu yhteiseen linkkitilietokantaan
- **Toiminta**
 - Väliittää topologista tietoa
 - Linkkitila algoritmi



Internet

- **Internetin arvo on globaalissa saavutettavuudessa**
 - Globaali saavutettavuus saadaan aikaiseksi eri operaattoreiden yhteistoiminnalla
 - Asiakassuhteilla
 - Paikallisella yhdysliikennetoiminnalla
- **Saavutettavuus rakentuu reitityksellä, johon ottavat osaa kaikki operaattorit (AS:t)**
 - 27593 AS-numeroa käytössä
 - 23304 tynkä ASia (yrityksia, ISP:tä jne)
 - 117 puhdasta transit operaattoria
 - 4172 yhdistelmää näistä kahdesta

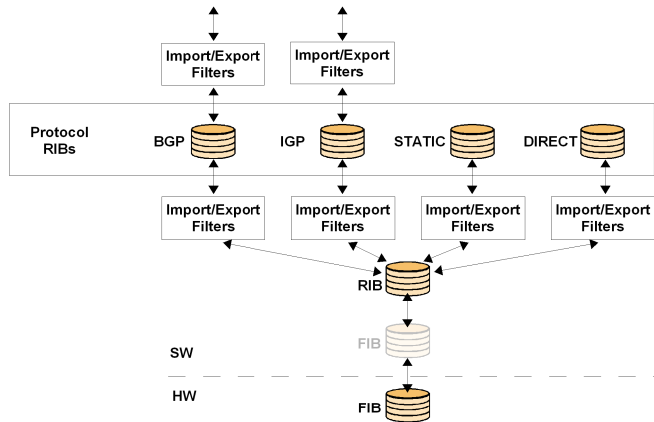


Ulkoiset reititysprotokollat

- **Tehtävä**
 - Keventää reititysprosessia eri autonomisten alueiden välillä poistamalla lokaalien muutosten aiheuttamat reittimuutokset muualla verkossa
 - Yksityiskohtaista tietoa autonomisten alueiden sisäisestä rakenteesta ei siirretä muille alueille
 - Mahdollistaa erilaisten politiikoiden käyttämisen lokaalin verkon reuna-alueilla
- **Toiminta**
 - Ei välitä topologista tietoa vaan saavutettavuus tietoa
 - ~ Etäisyysvektori algoritmi



Reititysprotokollien toimintamalli



BGP

- BGP vaatii toimiakseen tiedon
 - Ulkoisista naapureista**, joihin sen tulee olla yhteydessä eBGP:llä
 - IP osoite
 - AS numero
 - Sisäisistä naapureista**, joihin sen tulee olla yhteydessä iBGP:llä
 - IP osoite
 - AS numero



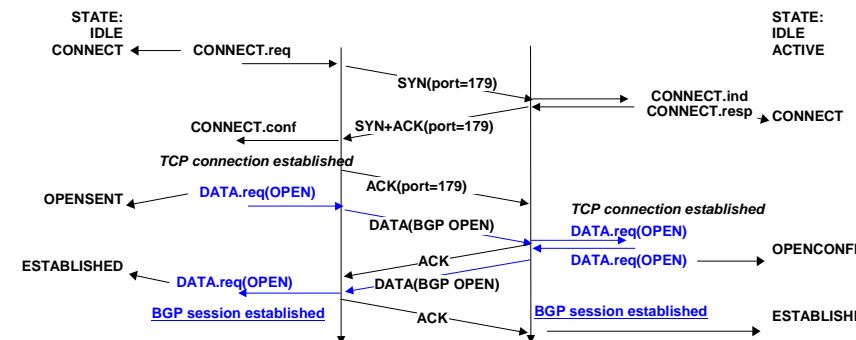
BGP

- Border Gateway Protocol**
 - Nykyinen versio BGP-4
 - Virallinen määrittely RFC 4271 (2006)
 - Lukuisia lisäyksiä
 - Reitityspeilit RFC 2796
 - Reitityssiirtoumat RFC 3065
 - MPLS leimojen välittäminen RFC 3107
 - MPLS VPN tuki RFC 2547
 - Reitinvalinnan vaimennus RFC 2439
 - etc



BGP

- Hyödyntää TCP:n palveluja reititysinformaation siirtoon
 - Sovellusportti 179



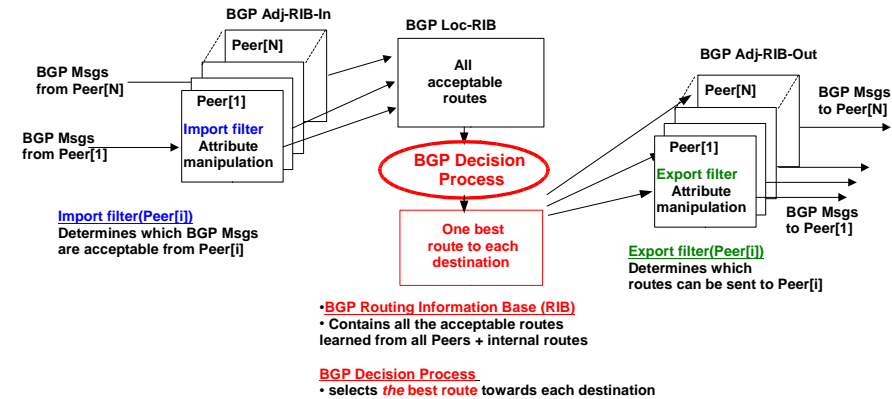


Tärkeimmät sanomat

- **UPDATE**
 - Välittää tiedot uusista verkoista, jotka on saavutettavissa kyseisen naapurin kautta
- **WITHDRAW**
 - Poistaa verkot, jotka eivät enää ole saavutettavissa kyseisen naapurin kautta
- **BGP sanomat etenevät globaalissa mittakaavassa muutospiireistä jokaisen naapurin suuntaan niin pitkälle, kunnes se saavuttaa reitittimen, jonka reitinlaskenta ei käytä kyseistä tietoa**
 - Jos muutos tapahtuu osoitteen lähteellä, etenee muutos globaalisti
 - **Stabiilisuuden saavuttaminen vaikeata**

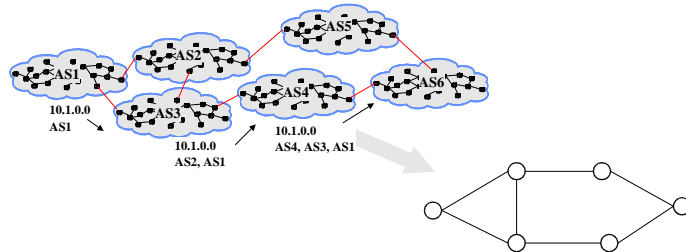


BGP:n toimintamalli



BGP

- Välitettävä Informaatio on vektoreita, jotka kuvaavat AS-tasolle abstraktoitua reittiä tietyn verkon (IP-osoitelohkon) ja kyseisen pisteen välillä



BGP

- Uutta verkkoa mainostava BGP sanoma, sisältää tiedot
 - **Polkuparametreista**
 - Reitti-ilmoituksen lähteestä
 - **Polkuvektorista**, joka sisältää kaikki ne AS:t, joiden kautta on kuljettu
 - Sarjana arvoja lähteestä käsin
 - Sarjana arvoja aggregoijasta käsin sekä joukkona sarjoja, jotka johtavat yksittäisiin aggregaatin osaverkkoihin
 - Reitittimestä, jonka kautta ko verkko on saavutettavissa
 - Preferenssistä, joka kyseiselle ilmoitukselle tulee antaa
 - Mahdollisesta aggregoinnista, joka ilmoitetulle verkolle suoritettiin
 - **Verkko-osoitteet**, jotka on saavutettavissa polkuvektorin toisesta päästä



Polkuvektori

- Polkuvektori kuvaa reittiä (kauttakulkualueita), joka verkossa kuljetaan kahden pisteen välillä
 - Välitetään karkea kuva etäisyydestä kahden pisteen välillä
 - Muut kustannusarvot olisivat keinotekoisia, koska yhteenliitettyjen verkkojen sisäiset kustannustekijät ovat vaihtelevia
- Polkuvektori mahdollistaa nopean [silrukoiden havainnointi](#) ja purkamismenetelmän etäisyysvektorityyppiseen toimintaan
 - **Autonomisen alueen reunareititin ei hyväksy reittiä, mikäli sen oma alue on jo polkuvektorissa**

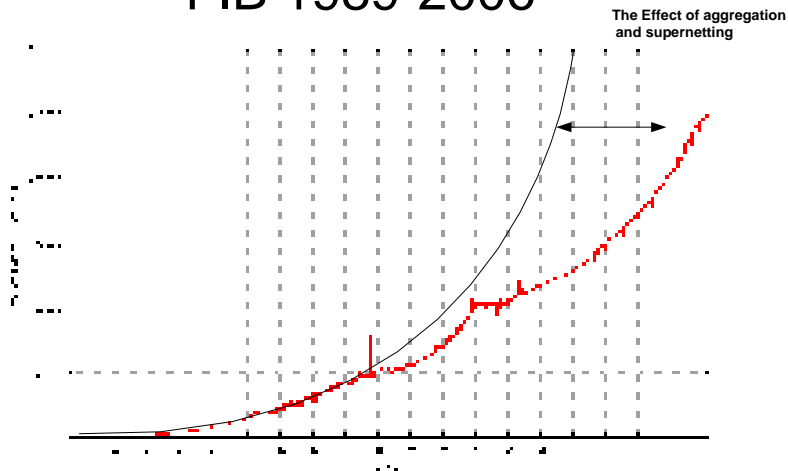


Polkuvektori

- **Ongelma:**
 - Verkkojen määrä suuri
 - Polkuvektoreiden määrä suuri
 - Reitittimet vaativat paljon muistia
 - Vektoreiden siirto hidasta
- **Ratkaisu**
 - Yliverkotus (supernetting) CIDR
 - Agregoidaan mahdollisimman monia verkkoja yhden prefixin taakse
 - **Yksittäisen polkuvektorin koko kasvaa**
 - **Kaikkiaan polkuvektoreiden määrä pienenee**



FIB 1989-2006



RIB vs FIB

- Current BGP statistics
 - FIB: 250280 <-- 278560 <-- 239211 entries
 - RIB: 735999 <-- 655837 <-- 557218 entries
 - With 114107 <-- 31961 different AS PATHs
 - Average length 5 <-- 3.5
 - Maximum length 13 <-- 9 (prepend 28 <-- 33)
 - Average prefix size 22 <-- 24



Polkuvektori

- Polkuvektori kuvaa reittiä kohdeverkon ja reitituspisteen välillä
 - Kaksi vaihtoehtoa
 - Jokainen verkko on oma vektori

Polku 1: 197.8.0.0/24 saavutetaan (T):n kautta
Polku 2: 197.8.2.0/24 saavutetaan (T,X):n kautta
Polku 3: 197.8.3.0/24 saavutetaan (T,Y):n kautta

- Yliverkotuksen kutistaa vektoreiden määrää mutta johtaa monimutkaisempaan rakenteeseen

Polku 1: 197.8.0.0/22 saavutetaan (T):n ja osajoukon (X,Y) kautta



Polkuvektori

- Yliverkotus voi tapahtua mielivaltaisen usein
 - Yksinkertaiset polkuvektorit
 - Jokainen yliverkko on oma vektori

Polku 1: 197.8.0.0/22 saavutetaan (Z,T):n ja osajoukon (X,Y) kautta
Polku 2: 197.8.5.0/23 saavutetaan (Z,R):n ja osajoukon (M,N) kautta
Polku 3: 197.8.7.0/24 saavutetaan (Z):n kautta

- Useampikertaisen yliverkotuksen jälkeen polkuvektori aivan samanlainen
 - Yksilölliset osat polkua on siirretty jälkijoukkoon

Polku 1: 197.8.0.0/21 saavutetaan (Z):n ja osajoukon (M,N,R,T,X,Y) kautta



NLRI

- Actual address representation is generic
 - Network Layer Reachability Information (NLRI)
 - Structured to
 - Identifying objects
 - Addressing objects
- Identifying objects describe the nature of addressing object
 - Address Family Identifier (AFI)
 - Subsequence Address Family Identifier (SAFI)



NLRI

- AFI
 - Describes the main family of the addressing type
 - 1: IPv4
 - 2: IPv6
 - 6: 802 Ethernet format
 - 16: DNS
- SAFI
 - Describes the usage within the family
 - 1-3: Unicast and multicast forwarding
 - 4: MPLS Label distribution
 - 65: VPLS
 - 128: MPLS VPN



BGP-verkottaminen

- BGP-verkotuksessa on keskeistä tietää alueiden rakenteet ja suhteet toisiin alueisiin
 - Tynkä AS
 - Monikotinen AS
 - Transit AS
 - Suora transitverkko
 - Looginen transitverkko
- Lisäksi tulee miettiä sisäisen reitityksen ja ulkoisen reitityksen suhdetta
 - eBGP <-> iBGP, IGP <-> BGP
 - eBGP <-> IGP

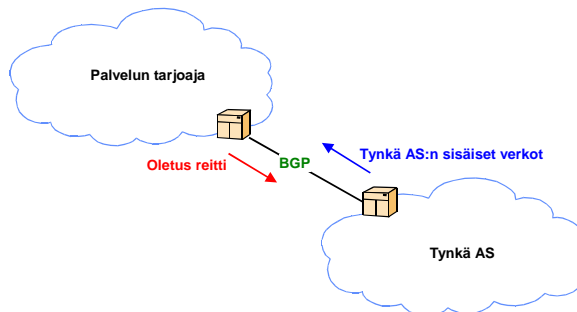


BGP-verkottaminen

- **Tynkä AS**
 - Autonominen alue, joka on yhteydessä muuhun Internetiin ainoastaan yhdestä pisteestä
 - Ei välttämättä vaadi BGP:n käyttöä
 - BGP:llä voidaan hallita joukkoa staattisia reittejä, jotka ovat seurausta hajanaisesta osoitteistuksesta
 - Käytetään privaatti AS-arvoja
 - AS 65412 - AS 65535
 - Tynkä AS näkyy muulle verkolle osana palvelun tarjoajaa
 - BGP:llä välitettävät reitit siirretään palvelun tarjoajan verkossa IGP:hen



BGP-verkottaminen

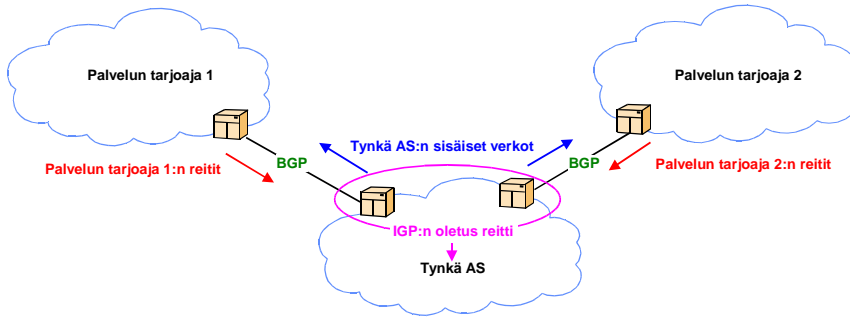


BGP-verkottaminen

- **Monikotinen AS**
 - Autonominen alue, joka on yhteydessä muuhun Internetiin useamman AS:n kautta
 - **Ei kuitenkaan toimi transit liikenteen välittäjänä**
 - BGP:n käyttöä suositellaan hallinnoimaan politiikkaa, jolla reitit valitaan
 - Sisäisesti tarvitaan erilliset reititystiedot verkoille, jotka saavutetaan eri AS:n kautta



BGP-verkottaminen



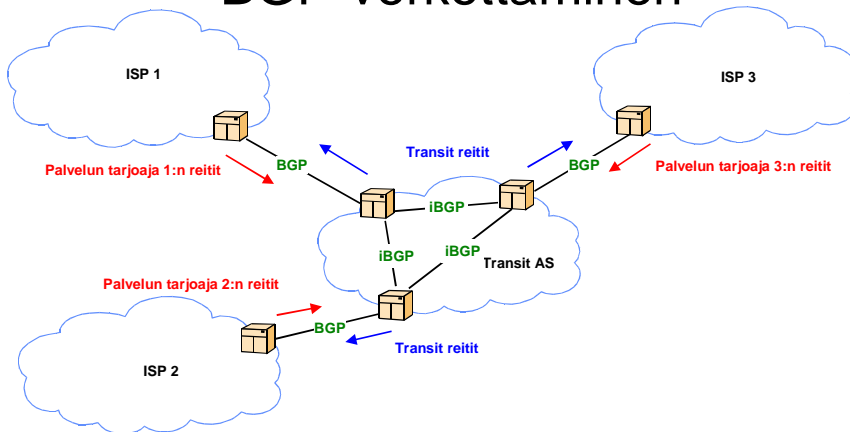
BGP-verkottaminen

- **Monikotinen transit AS**

- Operaattori, joka tarjoaa omaan verkkoonsa muiden operaattoreiden keskinäisen liikenteen välittämiseksi
 - Edellyttää operaattoreiden mainostamien reittien välittämistä ympäröiville operaattoreille
 - Edellyttää sisäisessä verkossa iBGP:n tai polkuvektoreiden välittämisen mahdollistavan IGP:n käyttöä
 - Reittien määrä voi olla niin suuri, että sisäisen protokollan käyttäminen reittien välittämiseen ei ole perusteltua
 - » OSPF:n linkkilatietokannan kasvu
 - » Jokainen ulkoinen reitti on oma rivi tietokannassa

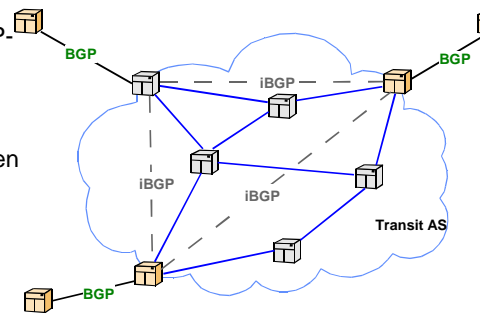


BGP-verkottaminen



Transit-verkot

- Transit-verkkojen sisäinen rakenne on yhtäläinen muiden autonomisten alueiden kanssa
 - Sisällä käytetään omaa IGP-protokollaa
 - Oletusreitti / reitit ulkomaailmaan
 - iBGP-sessio reunareitittimien välillä huolehtii BGP-polkuvektoreiden välittämisestä



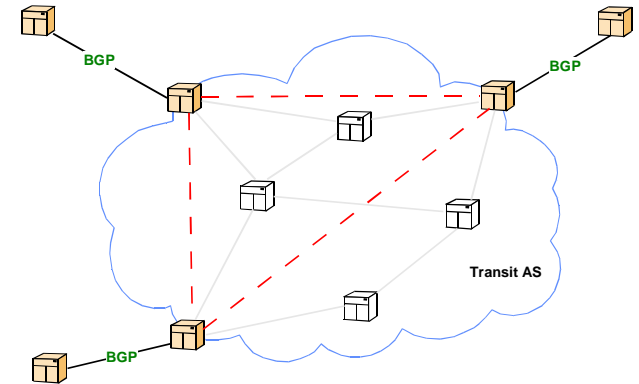


Transit-verkot

- **Ongelma:**
 - Sisäverkon reitittimet tarvitsevat kaikki reitit välittääkseen paketit reunareitittimien välillä
 - Tuhlaa FIBiä
 - Lisää IGP:n työtä
- **Ratkaisu:**
 - Rakentaa suora yhteys reunareitittimien välille
 - Suora siirtoyhteys
 - Tunnelointi
 - GRE
 - MPLS



Transit-verkot



BGP

- Stabiilisuuden kannalta on oleellista, että yhteys naapureiden välillä säilyy
 - BGP prosessi on TCP-yhteys naapureiden välillä
 - Yhteyden katkeaminen aiheuttaa kaikkien naapurin mainostamien reittien poistamisen
 - BGP prosessia ei kannata sitoa yksittäisen liitännän IP-osoitteeseen
 - Loopback -osoite
 - Reunareitittimeen pitää aina tulla useampia riippumattomia polkuja
 - Sisäverkossa
 - Ulkoverkossa



iBGP vs eBGP

- **eBGP** sessiossa reititin mainostaa **parasta reittiä** kaikkiin tuntemiinsa kohteisiin
 - Tähän vaikuttavat sessioon käytettävät suodattimet sekä levitys säännöt
- **iBGP** sessiossa reititin mainostaa ainoastaan **parhaita reittejä**, jotka se on **oppinut omista eBGP** seissioista
 - **iBGP** session kautta opittua reittiä ei koskaan mainosteta toiselle **iBGP** sessiolle
 - Suodattimia ei yleensä käytetä **iBGP** sessioissa



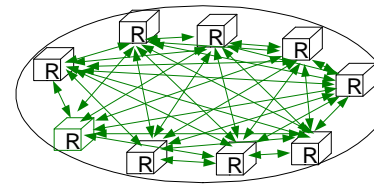
iBGP

- iBGP on autonomisen alueen reunareitittimien välillä käytettävä BGP-protokolla
 - Polkuvektoria ei voida käyttää ehkäisemään sisäisiä silmukoita
 - Kaikilla iBGP-reitittimillä on sama AS numero
 - Silmukoituminen voidaan estää poistamalla reititystietojen välitys iBGP:stä
 - Kaikki reunareitittimet on suorassa naapuruussuhteessa keskenään
 - **Täysin kytketty verkko**

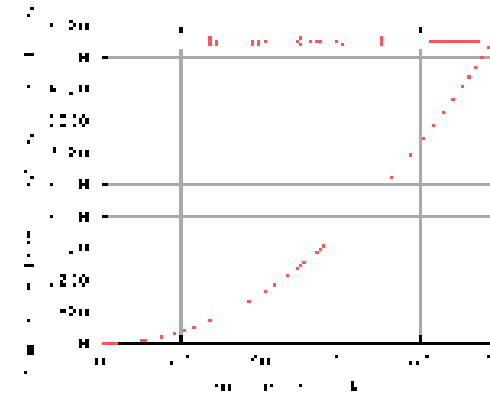


iBGP

- **Ongelma**
 - $N*(N-1)/2$ iBGP sessiota
N:llä reitittimellä



←→ iBGP sessio



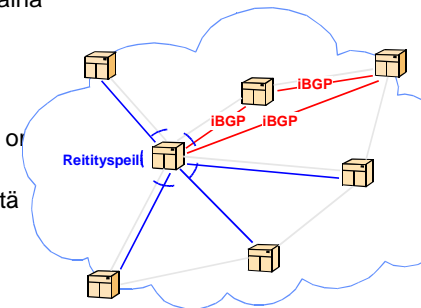
iBGP

- **Täysi kytkentä aiheuttaa ongelmia**, jotka ilmenevät lähinnä erillisten tilakoneiden hallinnan mahdottomuutena
 - Yhdellä operaattorilla saattaa olla kymmeniä liitäntäpisteitä, joiden välillä on tarvetta siirtää BGP-informaatiota
- **Kaksi ratkaisua** (ei toisensa poissulkevia)
 - **Reitityspeili** (route reflector)
 - iBGP-informaation välityspiste, joka on suorassa yhteydessä muihin iBGP-reitittimiin
 - Vertaa juurireititin OSPF:ssä
 - **Reitityслиiittouma**
 - Autonominen alue paloitellaan paikallisiin osiin, joiden välillä käytetään todellista BGP-protokollaa



Reitityspeili

- **Reitityspeili** toimii iBGP-reitityssanomien välittäjänä
 - Purkaa täyden kytkennän vaatimuksen niiden reitittimien osalta, jotka ovat kyseisen peilin asiakkaina
 - Client ↔ Server
- Verkossa voi olla samanaikaisesti
 - Alueita, joilla käytetään peilejä
 - Jokaisella peilialueella on o **ryhmätunnus**
 - Alueita, joilla käytetään perinteistä iBGP:tä





Reitityspeili

- Välytys perustuu yksinkertaiseen logiikkaan
 - Reitityspeilin asiakkaalta tuleva BGP-sanoma välitetään kaikille
 - iBGP:llä tuleva BGP-sanoma välitetään peilin kaikille asiakkaille
- Hierarkiset reitityspeilit ja mahdolliset väärät konfiguraatiot voivat aiheuttaa reitityssilmukoiden syntymistä
 - Kaksi uutta polkuvektoriparametria, jotka ovat AS:n sisäisiä
 - **Lähtöaseman tunnus (Originator ID)**
 - Reitityspeilin tunnus, joka sanoman ensimmäisen kerran peilasi sanoman AS:n sisällä
 - **Ryhmälista (Cluster List)**
 - Jokainen reititinpeili lisää oman peilausalueensa tunnuksen listaan
 - » Mahdollistaa silmukoiden havainnoinnin



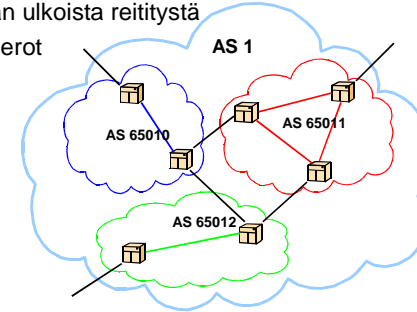
Reititysluotto

- Vaatii polkuvektorin muuttamista luottoalueella
 - Normaali polkuvекtori sisältää
 - Järjestetyn listan yhteisistä AS:sta aggregaatiopisteeseen
 - Osajoukon AS:sta, jotka ovat matkalla aggregaatin osiin
 - Reititysluotto sisäisessä polkuvektorissa on myös
 - Järjestetty lista sisäisistä AS:sta luotto sisäiseen aggregaatiopisteeseen
 - Osajoukko sisäisistä AS:sta, jotka ovat matkalla aggregaatin osiin



Reititysluotto

- **Reititysluotto** on joukko autonomisia alueita, jotka mainostetaan ulkopuoliselle yhtenä autonomisena alueena
 - Mahdollistaa suuren transit-verkon paloittelun pienempiin osaverkkoihin, joiden välillä käytetään ulkoista reititystä
 - Luotto muodostavat AS-numerot ovat privaattista avaruudesta
 - Luotto sisällä käytetään
 - iBGP:tä täysin kytketysti
 - Reitityspeilejä



BGP:n reitinvalinnan parametrusointi

- Reitinvalinta BGP:ssä perustuu pääasiallisesti kolmeen tekijään:
 - **Mainostettavan verkon prefixin pituuteen**
 - **Pisin prefix valitaan automaattisesti käyttöön**
 - **Paikalliseen politiikkaan**
 - Autonomisella alueella voi olla jokin paikallinen politiikka, jonka perusteella se valitsee tietyt reitit käyttöön
 - **Ulkoiseen politiikkaan**
 - Naapurialueella voi olla jokin ulkoinen politiikka, jolla se pyrkii vaikuttamaan autonomisen alueen reitinvalintaprosessiin



BGP:n reitinvalinnan parametrisointi

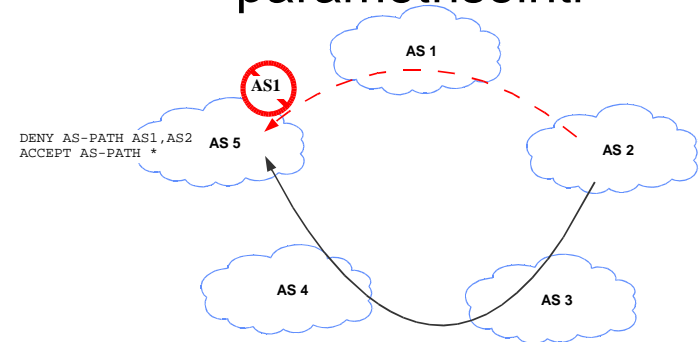
- Paikallinen politiikka voi perustua
 - **Polkuvektorin suodatukseen**
 - Hylätään ne polkuvektorit, jotka sisältävät tiety(n/t) AS-arvo(n/t)

```
DENY AS-PATH AS1,AS2  
ACCEPT AS-PATH *
```

- Kustannustekijä
 - » Tietyt transit-operaattorit ovat kalliimpia kuin toiset
- Luotettavuustekijä
 - » Tietyt transit-operaattorit eivät toimi avoimesti



BGP:n reitinvalinnan parametrisointi

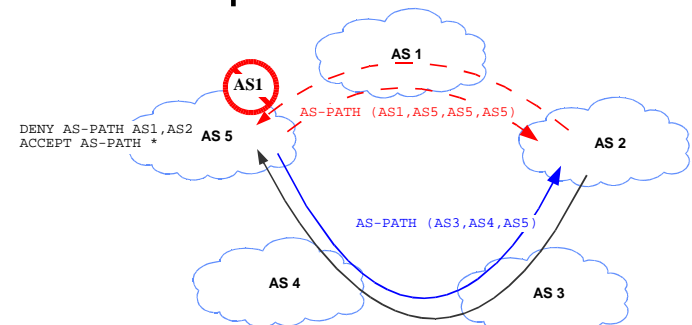


BGP:n reitinvalinnan parametrisointi

- Ulkoinen politiikka voi perustua
 - **Polkuvektorin kasvattamiseen**
 - Pyritään saattamaan tietyt polkuvektorit epäedulliseen asemaan muihin vektoreihin nähden
 - Sisällytetään oma AS-tunnus useita kertoja polkuvektoriin
 - Johtaa helpolla kilpavarusteluun, jos vastaava politiikka on toisaalla verkossa
 - Kumpikin automaattisesti kasvattaa vektorin pituutta yhdellä yli vaihtoehdoisen reitin



BGP:n reitinvalinnan parametrisointi



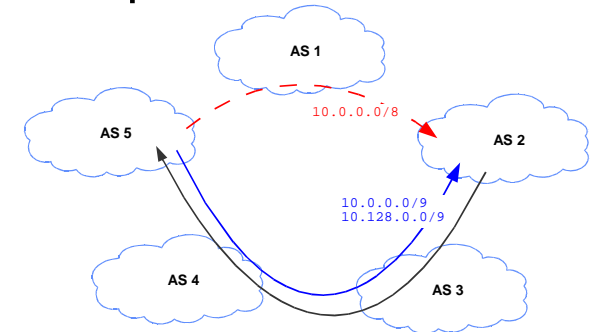


BGP:n reitinvalinnan parametrisointi

- Ulkoinen politiikka voi perustua
 - **Pidempien prefien mainostamiseen**
 - Mainostetaan omaa verkkoa pidemmällä prefixillä haluttuun suuntaan
 - **Etuna**
 - Mahdollistaa lyhyemmän prefixin jäämisen varareitiksi
 - **Haittana**
 - Kasvattaa verkon reititystaulujen kokoa
 - » Syö CIDR:llä saavutettua etua



BGP:n reitinvalinnan parametrisointi

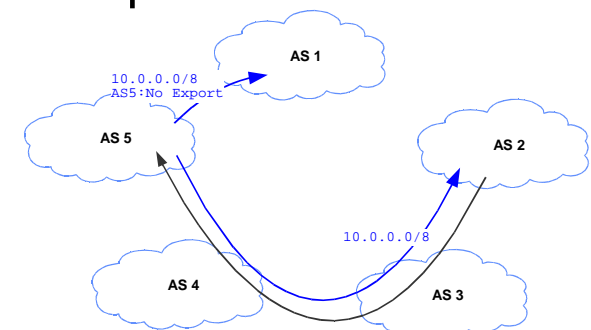


BGP:n reitinvalinnan parametrisointi

- Sisäinen / ulkoinen politiikka voi perustua
 - **BGP yhteisön käyttöön**
 - BGP-yhteisö on polkuparametri, joka mahdollistaa
 - Sisäisen politiikan levittämisen iBGP:ssä
 - Vaikutuksen ulkoisiin naapureihin (vain yhden hypyn parametri)
 - Yhteisöparametri
 - Kokoaa useita polkuvektoreita yhteisen politiikan piiriin
 - Mahdollistaa politiikan siirron naapureille
 - Koostuu
 - » AS-numerosta
 - » Poliittikanumerosta

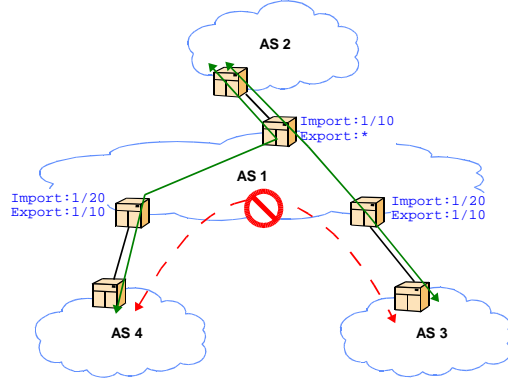


BGP:n reitinvalinnan parametrisointi





BGP:n reitinvalinnan parametrusointi

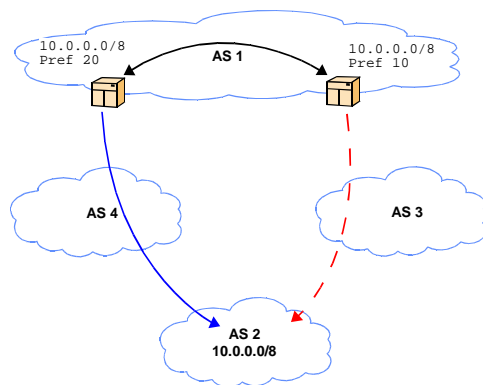


BGP:n reitinvalinnan parametrusointi

- Sisäinen politiikka voi perustua
 - **Paikallisen preferenssin**
 - Paikallinen preferenssi indikoi mitä liityntäpistettä tulee käyttää pyrittäessä kohti tiettyä verkkoa
 - Merkityksellinen ainoastaan iBGP:ssä



BGP:n reitinvalinnan parametrusointi

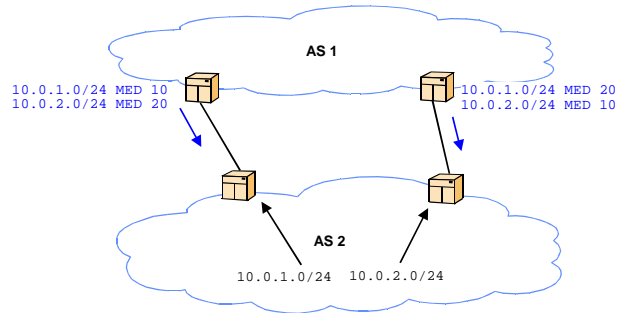


BGP:n reitinvalinnan parametrusointi

- Ulkoinen politiikka voi perustua
 - **Ulkaisen reitityspisteen valintaan**
 - Multi-Exit Discriminator (MED) mahdollistaa polkuvektorikohtaisen liitäntäpisteen preferenssin signaloinnin
 - Mahdollistaa kuorman jakamisen useiden liitäntäpisteiden välillä
 - Mikäli naapurialueen paikallinen politiikka ei sodi MEDin osoittamaa politiikkaa vastaan



BGP:n reitinvalinnan parametrusointi



BGP:n reitinvalinnan parametrusointi

- Sisäinen politiikka voi perustua
 - **Reitinvalinnan vaimennus**
 - BGP-toimii globaaleissa verkoissa, joiden reititysmuutokset ovat laaja-alaisia
 - Yksittäisten reittien aktivoitumiset ja passivoitumiset voivat olla jatkuvia prosesseja
 - Kuormittaa turhaan reitintä ja koko verkkoa
 - Voidaan vaimentaa liittämällä jokin vaimennusmenetelmä
 - » Rankaisu: oskilointi käynnistää ajastimen, jonka aikana reittiä ei huolita
 - » Hallinnollinen: oskiloiva reitti poistetaan pysyvästi