KRAUSS INTERNATIONAL

CALL FOR SALES: 9717387778 / 9910416231 EMAIL:SALES@KC-INDIA.COM

BGP 101 FOR ISP/TSP

PRESENTED BY MANKOMAL SINGH (KRAUSS INTERNATIONAL)

AT NEW DELHI MUM, SEPTEMBER 2016

ABOUT THE SPEAKER

- Mankomal Singh, Krauss International, New Delhi, India
- In networking field for 17 years.
- Certified trainer of MikroTik (MTCNA, MTCWE)
- Certified trainer of wireless networking and networking technologies
- Designed & implemented a wide array of networks for corporates, ISPs, DCs, BPOs and most recently Indian Armed Forces
- Running our own ISP in Punjab, Maharashtra, Tamil Nadu and Gujarat

OBJECTIVES

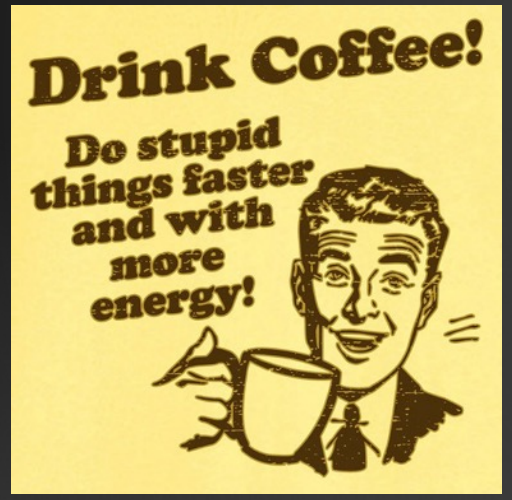
- What is BGP ? What is an ASN?
- Establishing simple BGP in MikroTik ROS
- Various scenarios of BGP implementation
- Route Filtering making BGP secure

DOWNLOADING THIS PRESENTATION

- This presentation will be available on http://mum.mikrotik.com under archives
- Or you can send me an email at mankomal@gmail.com I will forward you
- Or you can visit www.mankomal.com , the presentation was uploaded here this morning

BEFORE WE START

Please have a cup of coffee for some theoretical part of the session and don't doze off.



WHAT IS BGP?

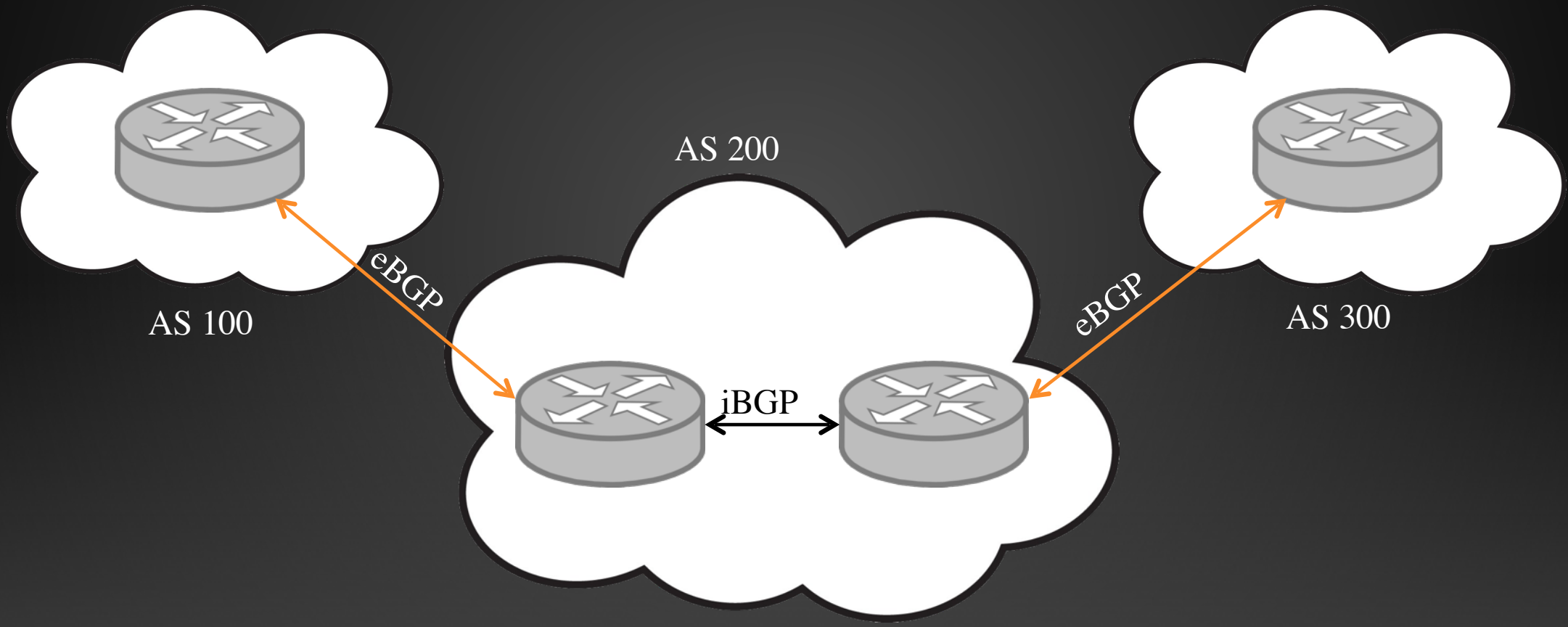
- BGP: Border Gateway Protocol
- Standardized exterior gateway protocol designed to exchange routing and reachability information
- Routing decision are made on path, network policies, or rule-sets defined
- BGP neighbours are referred as Peers
- When run between peers of same AS its known as iBGP and when run between different AS its known at eBGP
- By default found at TCP(179)

WHAT IS AN AS?

- AS: Autonomous System
- Within the internet an AS is a collection of connected IP routing prefixes under control of one or more network operators
- ASN, AS Number is a unique ID used for BGP routing.
- ASN as IPs are assigned by registrar, in India by IRINN

BGP PROTOCOL AND AS

AS talk to each other and share routing information using language of BGP



WHY BGP?

- Scalable protocol capable of handling huge amount of prefixes, which are constantly growing
- It is reliable and robust
- Provides tools to influence external traffic which may not be under direct control of the administrator
- No matter how big the network is , BGP doesn’t care about the internal topology, only how can the network be reached

HOW TO RUN BGP IN MY NETWORK?

- Both administrators run BGP peering
- A session using port TCP 179 is established
- Both sides exchange routing information until total convergence
- After this only information about new and withdrawn routes are exchanged

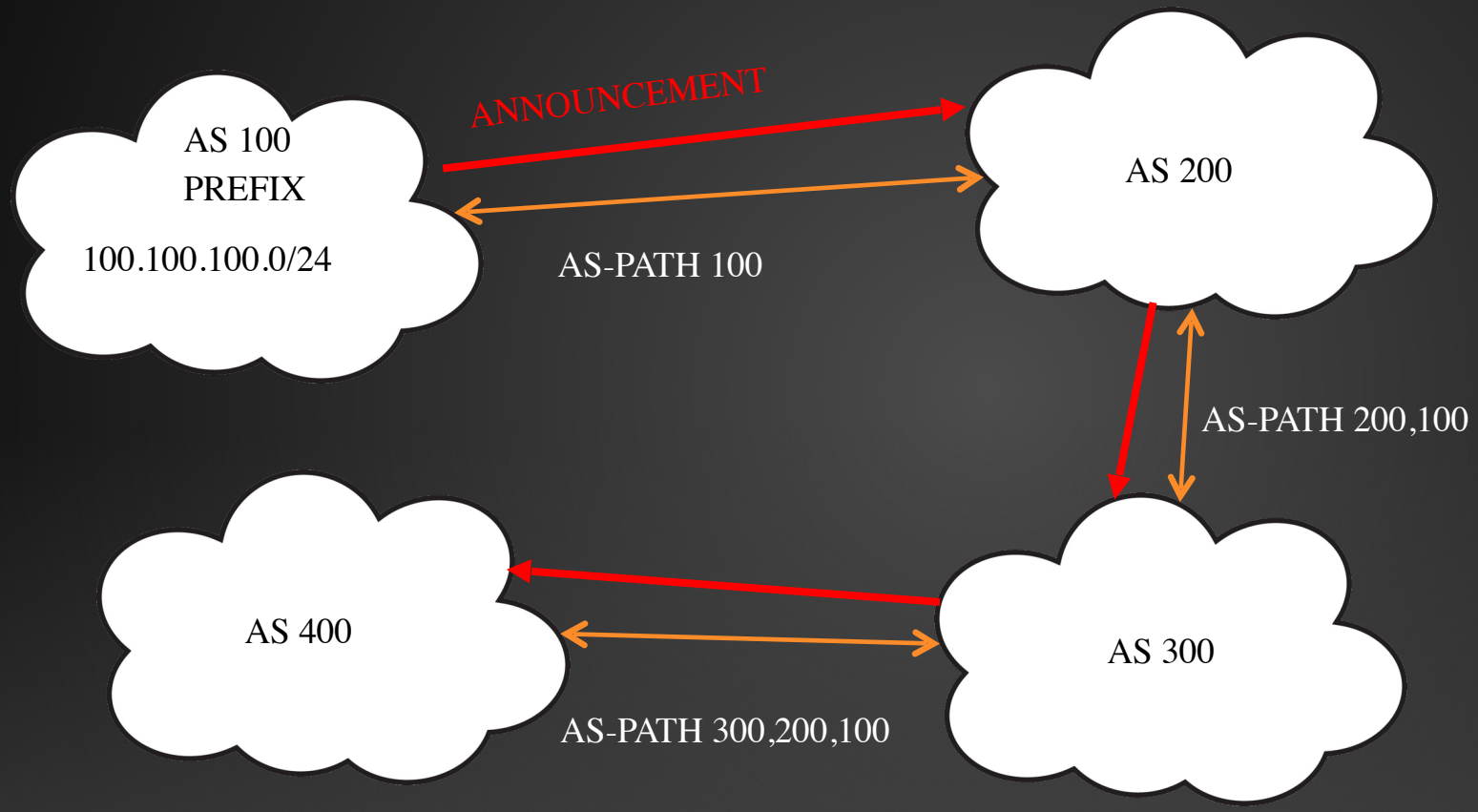
BGP MESSAGES

- OPEN: Establish a peering session, confirmed with a keep alive
- KEEP ALIVE: Handshake at regular intervals to check peer state
- NOTIFICATION: Shuts down a peering session, when an error occurs
- UPDATE: Announcing new routes or withdrawing previously announced routes

BGP ATTRIBUTES

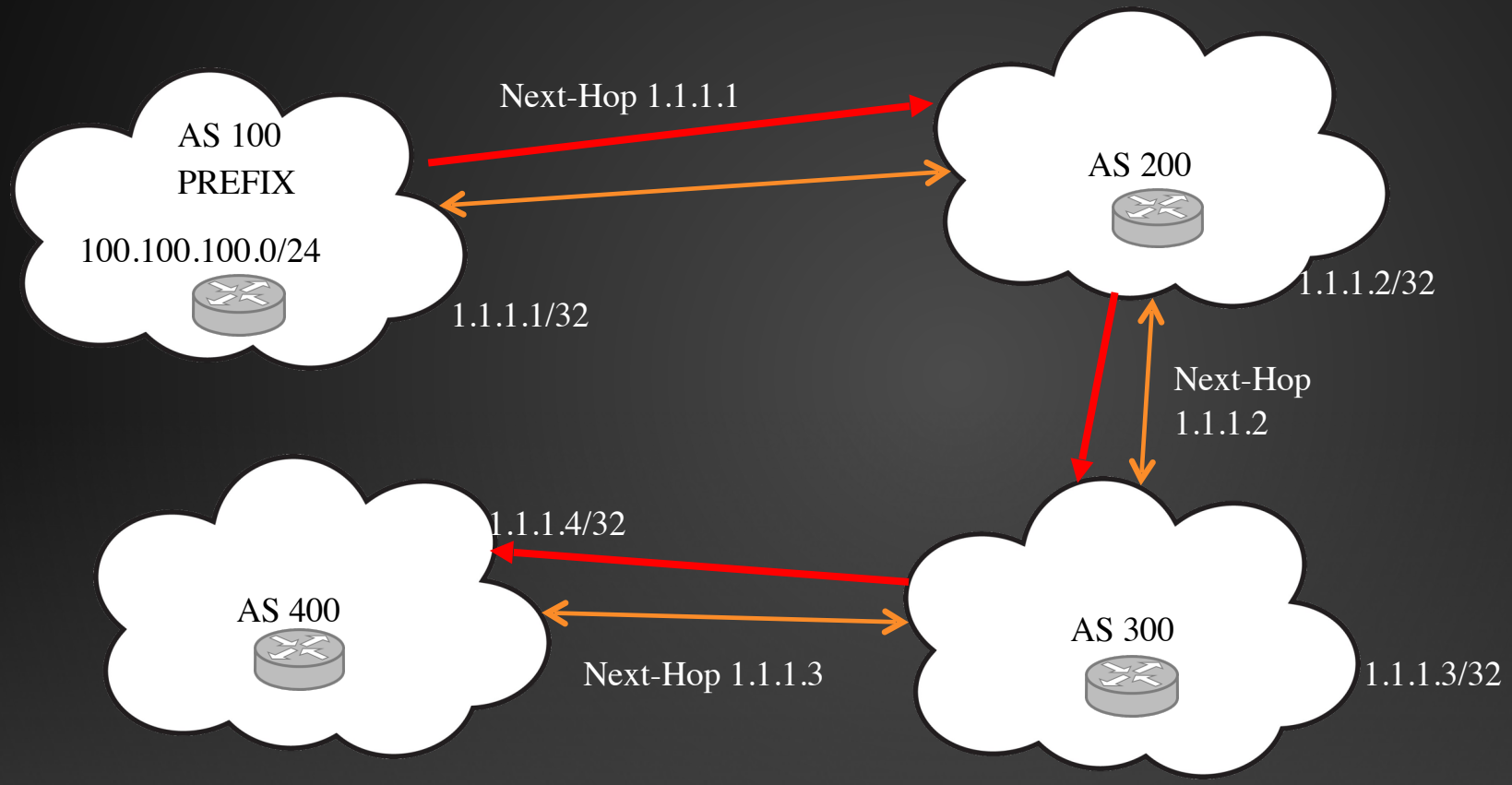
- Attribute describes the characteristics of a particular prefix
  - AS_PATH: AS Sequence thru which the network is reachable
  - NEXT_HOP: IP Address of the next hop router
  - LOCAL_PREF: Used to chose outbound path inside an AS
  - Community: Numeric value that can be attached to a prefix with some specific purpose
  - Multi-exit-discriminator(MED): Attribute to influence inbound route when multi-homed with same AS

UNDERSTANDING AS-PATH



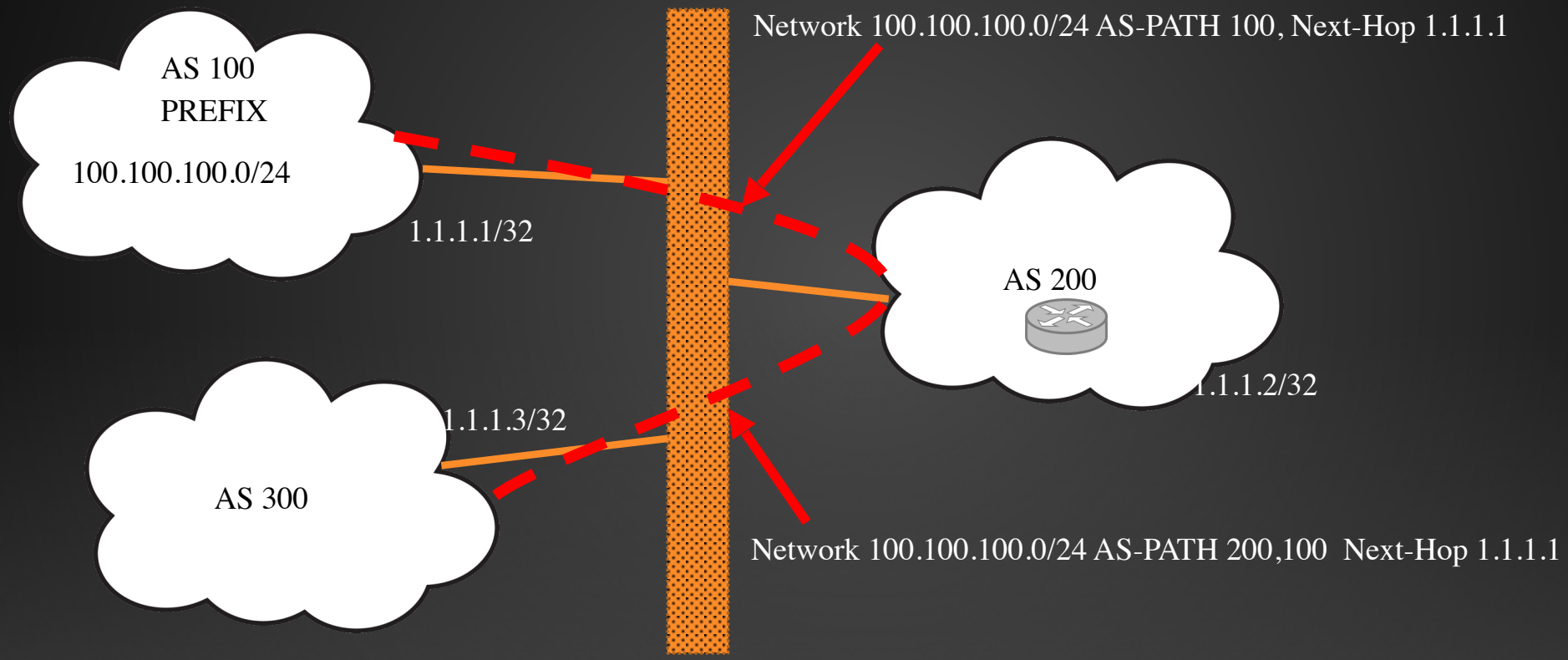
AS-PATH FROM AS400 TO REACH 100.100.100.0/24 WILL BE 300,200,100

UNDERSTANDING NEXT-HOP



To reach network 100.100.100.0/24 from AS400 the next-hop address will be 1.1.1.3

NEXT-HOP IN AN IXP SCENARIO



BGP CRITERIA FOR DECISION

BGP will compare identical prefixes in the following order:

- Prefers the path with highest weight(default = 0)
- Prefers the path with highest Local-Preference (default = 100)
- Prefers the path locally originated via aggregate or BGP network announce
- Prefers the path with shortest AS-Path
- Prefers the path with lowest origin (Preference=igp<egp)
- Prefers the path with lowest MED (default=0)
- Prefers the path learned by eBGP over the ones by iBGP
- Prefers the path received from the router with lowest router ID
- Prefers the path with shortest route reflection cluster list
- Prefers the path that comes with the lowest neighbour address

*details can be found at http://www.noction.com/blog/bgp_bestpath_selection_algorithm

INFLUENCING BGP DECISION(ROUTE FILTERS)

The way to influence BGP decisions is to configure route filters

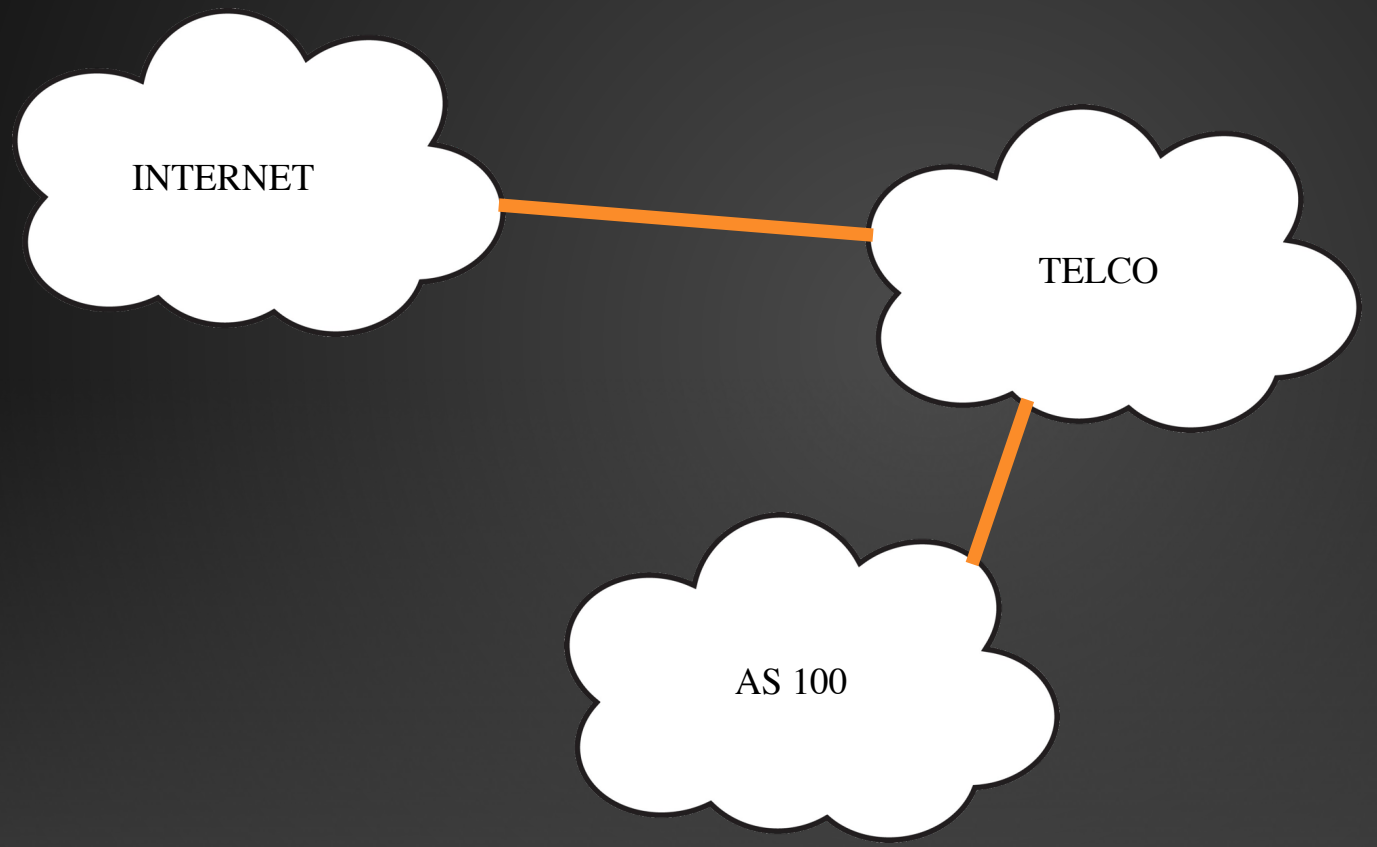
- Filtering **incoming** routes will change how we see the external prefixes advertised by our peers, thus influencing how we **send** traffic

- Filtering **outgoing** routes will change how the peers see our prefixes, thus influencing how we **receive** traffic

SCENARIOS AND DISCUSSION

- Single homed network
- Single homed + Peering from IXP
- Multihomed + Peering from IXP

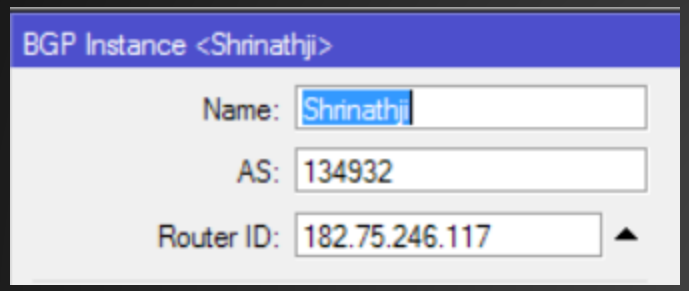
SINGLE HOMED NETWORK

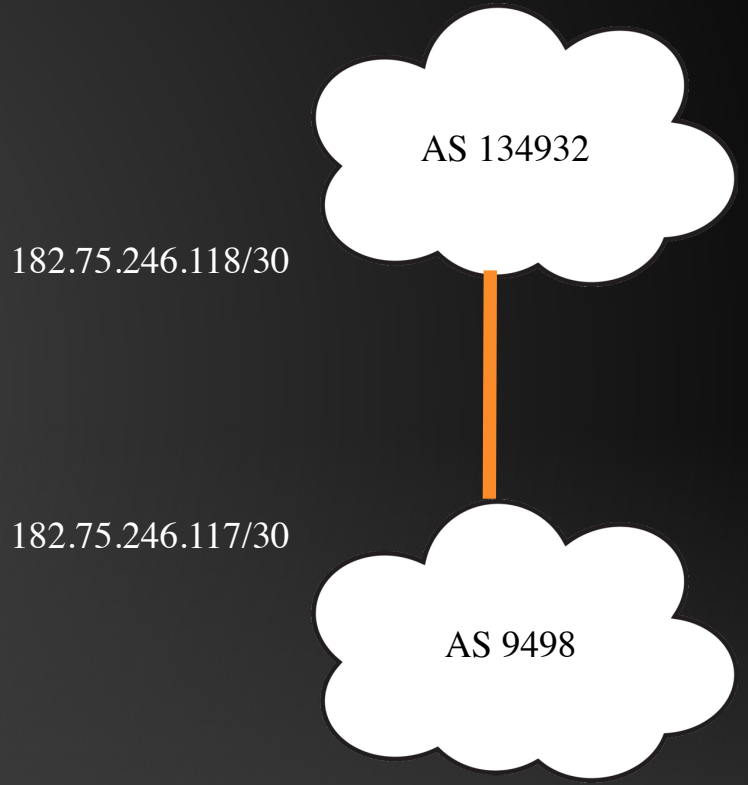


BEFORE BGP CAN BE ESTABLISHED

- Your whois should be updated on APNIC and RAdB
- Decide if you want full routes, partial routes or just the gateway routes
- Tell your Transit Operator which prefixes you will be announcing, its better to announce both the summarized routes and subnetted routes(Telco will only accept /24 and above summarizations), make sure your RAdB is updated
- Decide if you want a default route or not

PRACTICAL SCENARIO

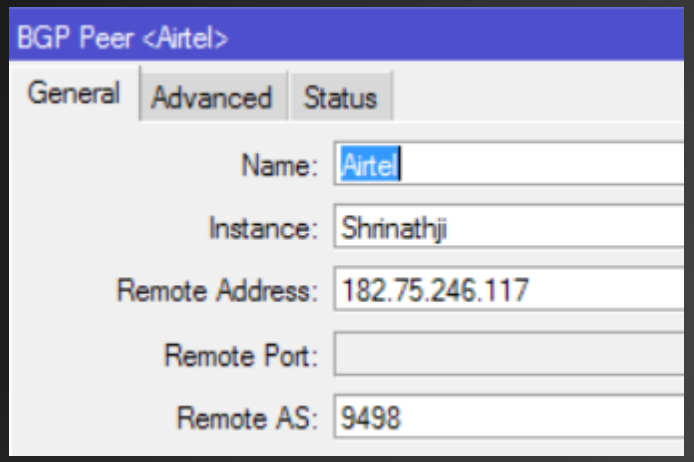


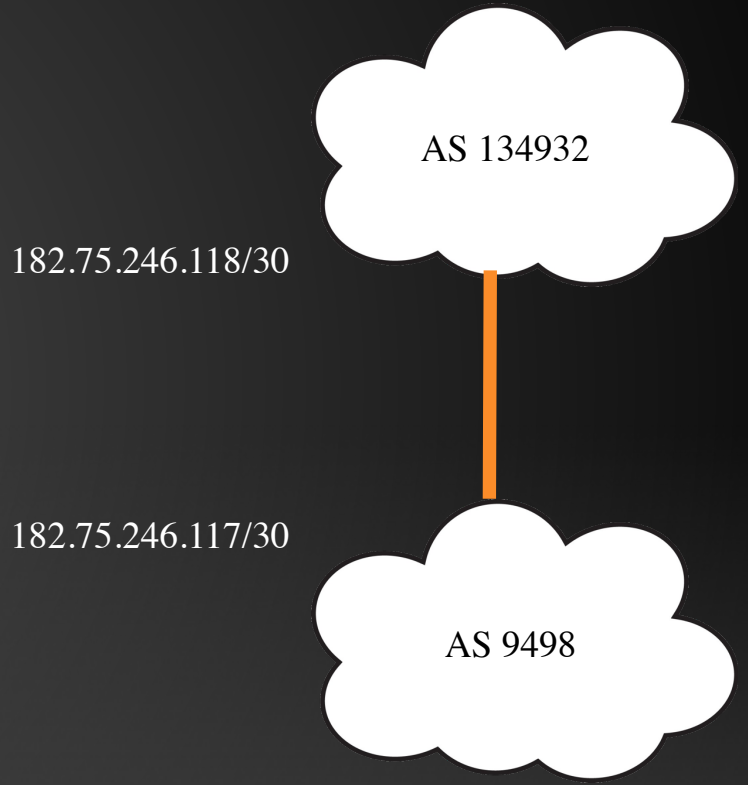


Configuration required at our end

- Create a BGP Instance
- Add your ASN
- Router ID is optional but recommended

PEERING WITH BGP

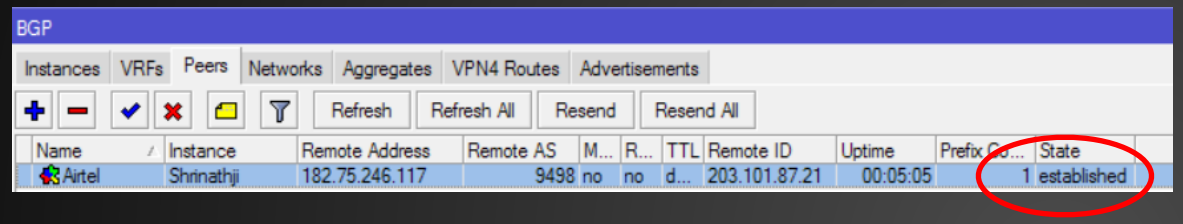




Configuration required at our end:

- Define which instance we will use for peering
- Remote Address that we will peer with
- Remote AS that we are peering with

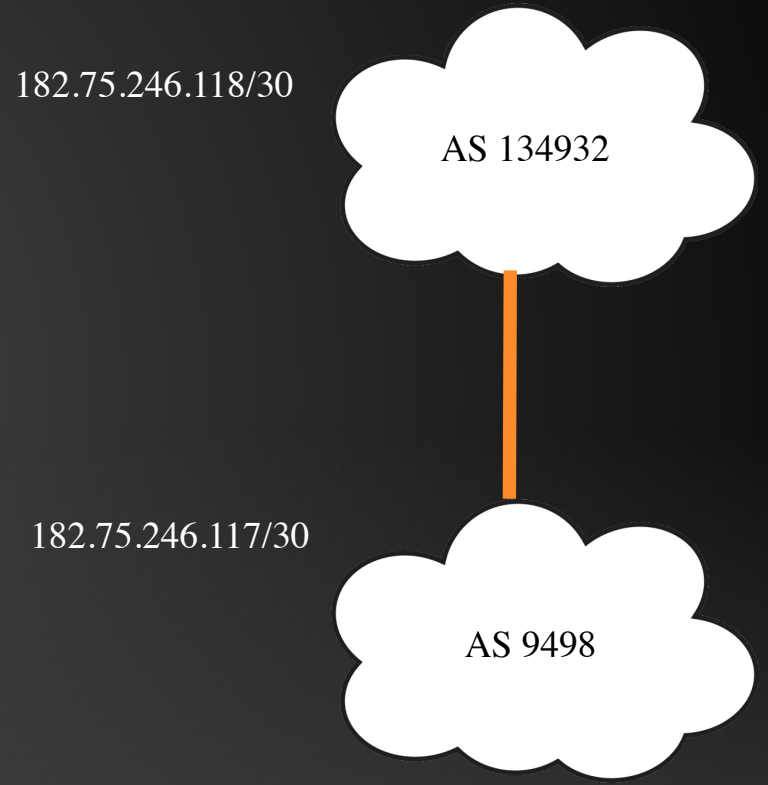
ESTABLISHED PEER



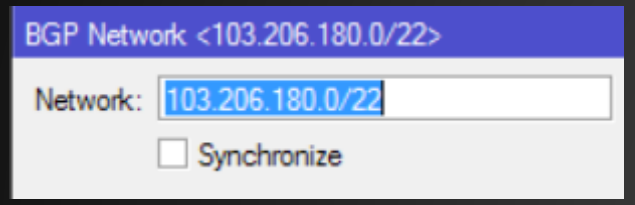
Our BGP is established and we are getting one Prefix from AS9498 (Airtel)

```
[mankomal@MikroTik] > ip route print where bgp
Flags: X - disabled, A - active, D - dynamic,
C - connect, S - static, r - rip, b - bgp, o - ospf, m - mme,
B - blackhole, U - unreachable, P - prohibit
 #      DST-ADDRESS        PREF-SRC        GATEWAY            DISTANCE
 0  Db  0.0.0.0/0                          182.75.246.117           20
[mankomal@MikroTik] >
```

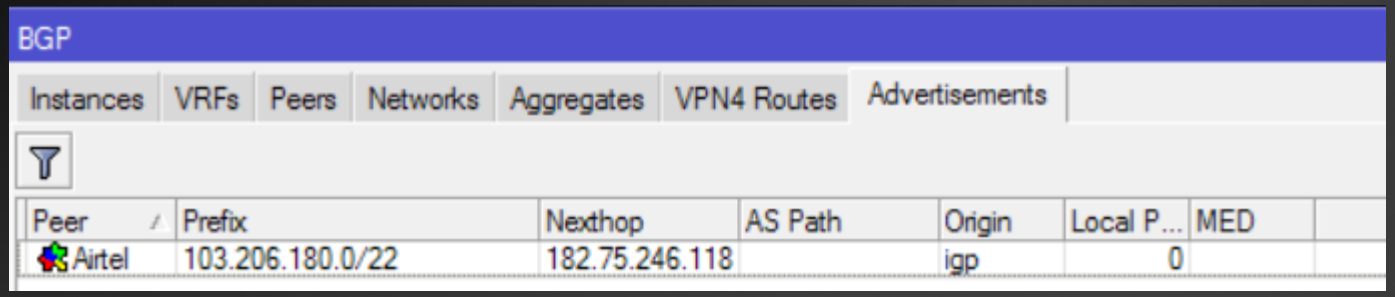
What we see from Airtel



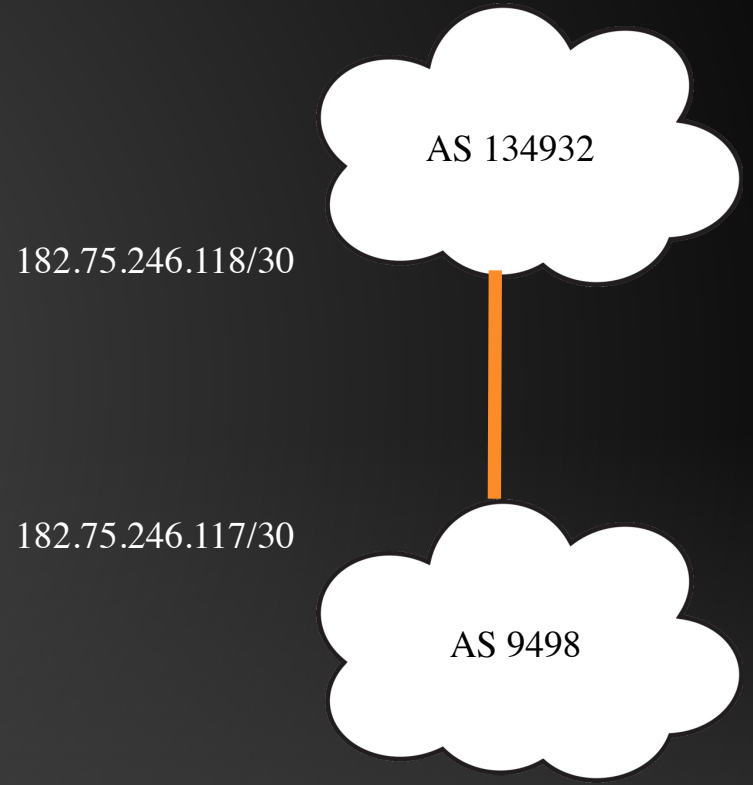
ANNOUNCE PREFIX



Advertising 103.206.180.0/22 network



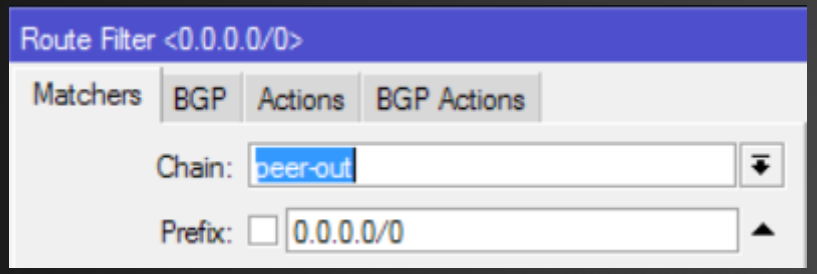
Our advertisement as visible to AS9498/Airtel

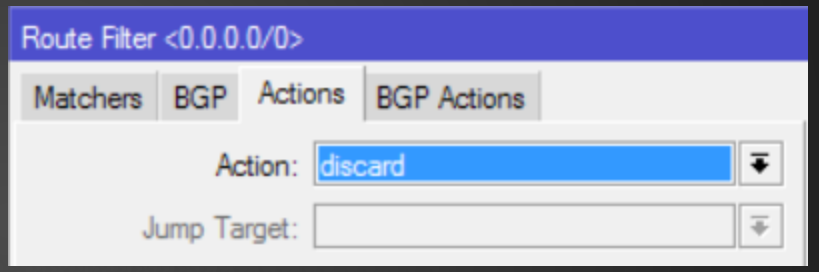


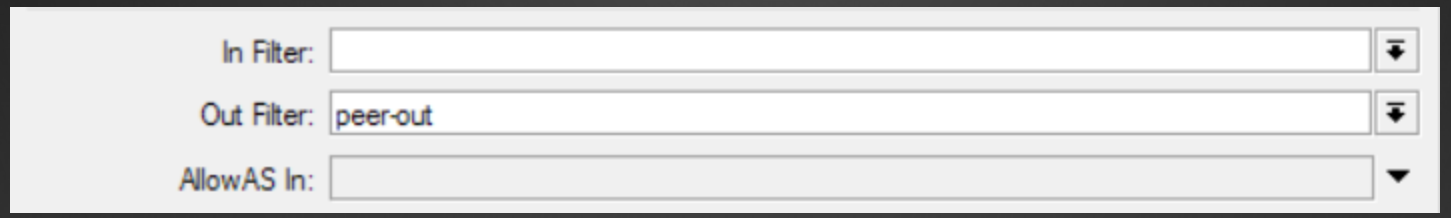
ROUTE FILTERS

- In the current scenario that we have done till now we are receiving all the broadcasted prefixes and we are broadcasting all prefixes in our network, i.e. nothing is being filtered
- Router filters allow ingress and egress announcements and can effect the way we can receive traffic
- Route filters can be used for varied reasons of security, economics and technical. But is a good practice that should be done so you hear what you want to hear and make your peer hear what they should hear from you

EXAMPLE: WE DON'T WANT TO SHARE OUR GATEWAY ROUTE TO THE PEER







BOGON

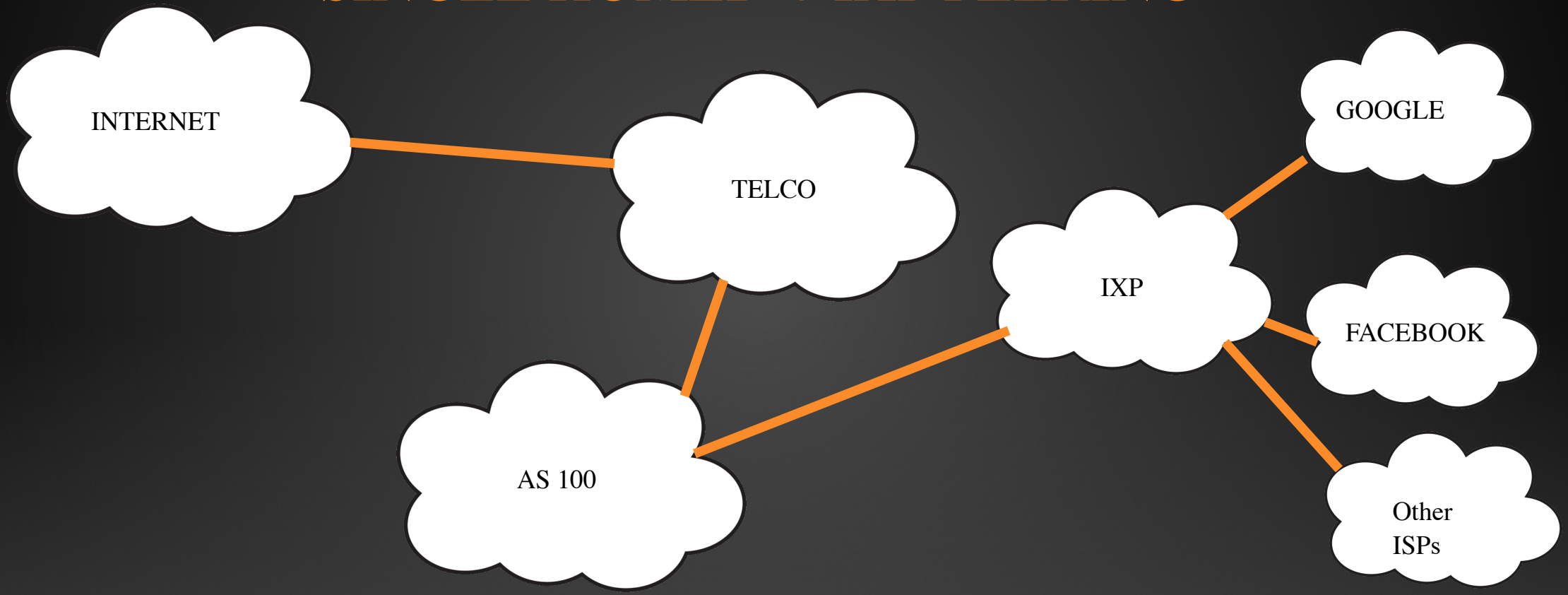
- Aka bogus IPs
- These are set of prefixes which are invalid and should not be advertised e.g. 10.0.0.0/8 which is a private IP space, also these are set of prefixes which have not been allotted and/or are reserved and yet not allocated or delegated by IANA(Internet Assigned Network Authority) or delegated RIR(Regional Internet Registrar) like IRINN
- You can use CYMRU BGP peering to get a partial or complete list of BOGONs, more details of configuring can be found at my blog www.mankomal.com

CYMRU PEERING EFFECT

Name	Instance	Remote Address	Remote AS	Multihop	Route Reflect	TTL	Remote ID	Uptime	Prefix Count	State
Airtel	Airgenie	182.75.141.233	9498	no	no	default	202.123.47.115	1d 06:23:04	1	established
CYMRU	Airgenie	38.229.66.20	65332	yes	no	default	38.229.66.20	00:10:47	3725	established
Google	Airgenie	72.14.208.85	15169	no	no	default	66.249.94.203	21:24:55	418	established

Inserted all received routes to blackhole routes so that customers cannot communicate to malicious IPs

SINGLE HOMED + IXP PEERING



WHAT IS AN IXP?

- IXP – Internet Exchange Point
- It is a physical infrastructure where ISPs and CDNs exchange internet traffic between their network
- IXP reduces some portion of an ISPs traffic which is delivered by upstream provider, thereby reducing the cost, increase quality.
- In India we have NIXI unfortunately for local traffic only not for google, facebook etc.

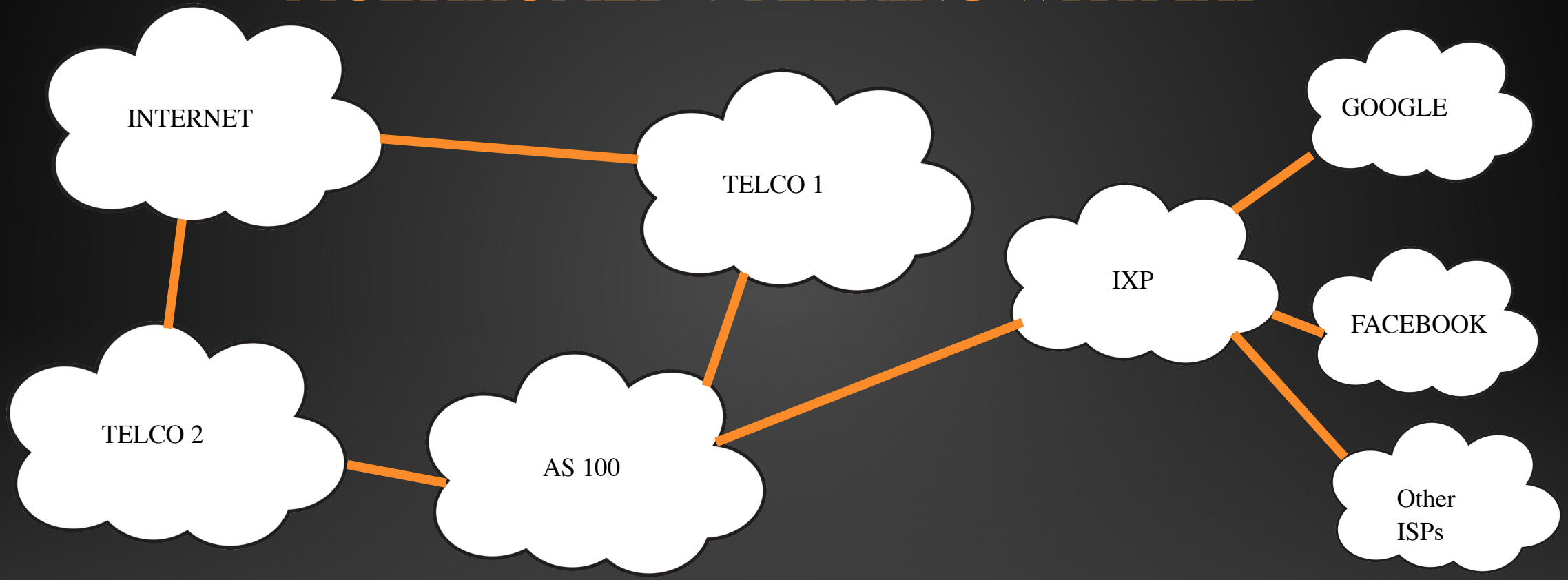
EXAMPLE OF DIRECT PEERING WITH GOOGLE

Name	Instance	Remote Address	Remote AS	Multihop	Route Reflect	TTL	Remote ID	Uptime	Prefix Count	State
Airtel	Airgenie	182.75.141.233	9498	no	no	default	202.123.47.115	1d 06:23:04	1	established
CYMRU	Airgenie	38.229.66.20	65332	yes	no	default	38.229.66.20	00:10:47	3725	established
Google	Airgenie	72.14.208.85	15169	no	no	default	66.249.94.203	21:24:55	418	established

Here we see Google advertising 418 prefixes to our AS, effect can be seen below, all google traffic going thru the interface where peering is present approx. 250Mbps of data

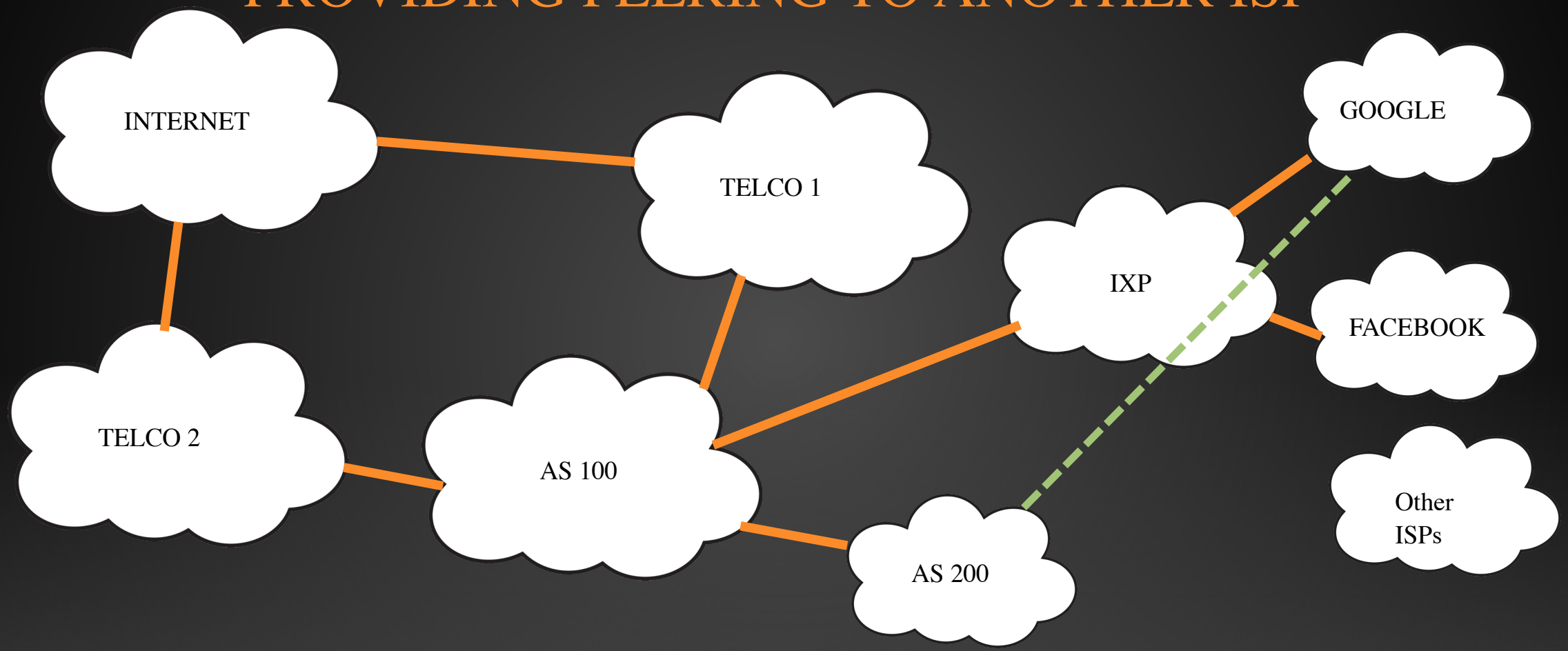
;;; GOOGLE PEERING										
R	ether7	Ethernet	1500	1580	35.5 Mbps	235.5 Mbps	15 940	29 085	0 bps	235.5 Mbps

MULTIHOMED + PEERING WITH IXP



Our objective is to only receive prefixes and advertise our prefixes not let our ASN become transit Point for Other ISPs to reach Internet

MULTIHOMED + PEERING WITH IXP + PROVIDING PEERING TO ANOTHER ISP



Our objective is to redistribute only Google routes to AS200 and not default route we receive from Telcos as it will result in their internet traffic passing thru us.

WE NEED TO USE ROUTE FILTERS

To Telcos we just need to advertise our prefixes and discard everything else

3	airtel-out	103.54.220.0/22				accept
4	airtel-out	202.168.156.0/24				accept
5	airtel-out	103.220.36.0/22				accept
6	airtel-out	202.168.157.0/24				accept
7	airtel-out	202.168.159.0/24				accept
8	airtel-out	202.168.158.0/24				discard
9	airtel-out					discard

To ISPs seeking peering from us we will distribute everything other than 0.0.0.0/0

0	peer-out	103.54.220.0/22				discard
1	peer-out	202.168.156.0/22	22-24			discard
2	peer-out	0.0.0.0/0				discard

UNDERSTANDING COMMUNITIES

- BGP community attribute can be used in order to control the routing policy in an upstream provider
- Even though AS200 is now peered to us and can listen to google routes from us, we still need to tell Google that their server's should upload thru AS100 route, as they can see AS200 now from multiple locations and not necessarily the route from AS100 may be used.
- This can be achieved by 2 ways:
  - By using the most direct routes for advertising thru AS100
  - Or, by using communities, we will use this for the example

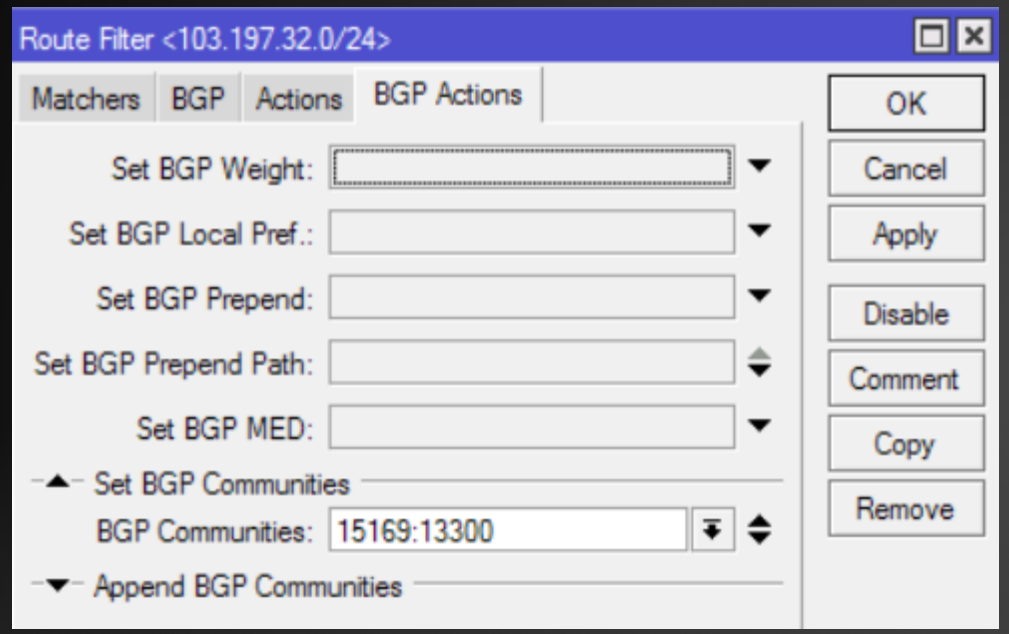
GOOGLE COMMUNITIES

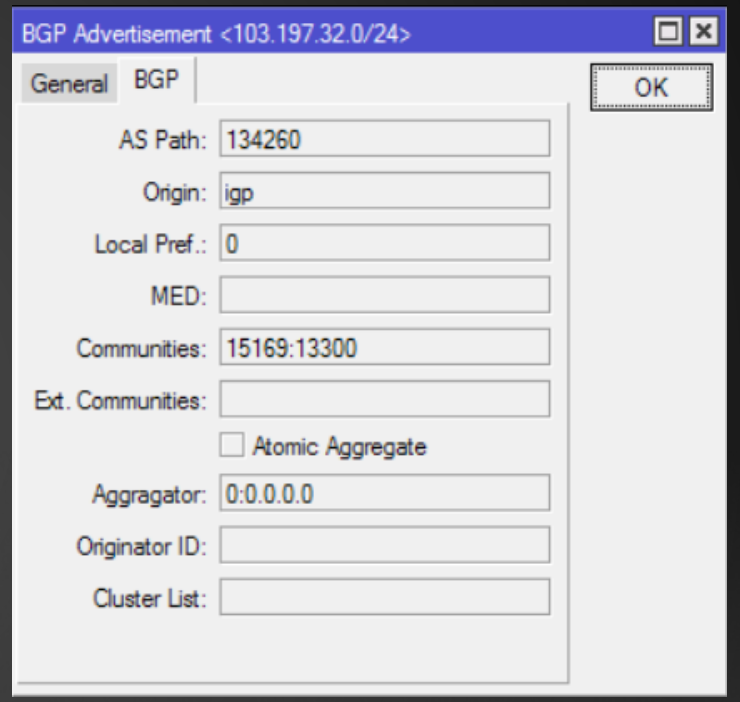
Table 2.2: Preferred Ingress Signalling Communities

Community	Preferred Ingress Signalling Range
15169:13000	Lowest preference to receive traffic for this prefix at this interconnection point (try to not serve traffic here). Attempt to serve traffic on an indirect path (through other upstreams or peers) before using this prefix.
...	15169:13001 - 15169:13099 indicate very low preference (the higher the tag the higher the preference). Any prefix tagged in this range is less preferred than an indirect path.
15169:13100	Default priority of traffic on an indirect path. Tagging with this community indicates that the preference is equal to receiving traffic over an indirect path.
...	15169:13101 - 15159:13199 indicate low preference. Any prefix tagged in this range is preferred over indirect paths but not preferred to an interconnection point where the prefix is untagged.
15169:13200	Default priority to receive traffic for this prefix at this interconnection point (the same as if the prefix is untagged).
...	15169:13201 - 15169:13299 indicate high preference (the higher the tag the higher the preference).
15169:13300	Highest preference to receive traffic for this prefix at this interconnection point (try to serve traffic here).

As evident the community 15169:13300 Will have the highest preference to Receive traffic for the prefix at the Specified interconnection

SETTING COMMUNITY IN FILTER





THE RESULT

INFLUENCING TRAFFIC OR MANIPULATING TRAFFIC

The way to influence BGP decisions is to configure route filters

- Filtering **incoming** routes will change how we see the external prefixes advertised by our peers, thus influencing how we **send** traffic

- Filtering **outgoing** routes will change how the peers see our prefixes, thus influencing how we **receive** traffic

CHECKING RESULTS

How to check what is the result of filters and different attributes

- Tools that don't tell all the truth but are still needed
  - Ping, Traceroute, Torch etc

- Use routing tables to see your upload policy
- Download policy can seen by other ASNs looking glass

Lot of free looking glass tools are available so as to view the results in case your Upstream is not able to/providing you the results

Q & A

THANKS TO …

Airgenie Communication

Zerone Networks

Spacenet Internet Services

Shrinathji Netsol

SOURCES

- https://en.wikipedia.org/wiki/Border_Gateway_Protocol
- https://en.wikipedia.org/wiki/Autonomous_system_(Internet)
- http://http://www.noction.com/blog/bgp_bestpath_selection_algorithm
- http://wiki.mikrotik.com/wiki/Manual:BGP_HowTo_%26_FAQ
- http://wiki.mikrotik.com/wiki/Manual:Routing/BGP
- http://www.evilrouters.net
- http://www.enterprisenetworkingplanet.com/netsp/article.php/3615896/Networking-101-Understanding-BGP-Routing.htm

COPYRIGHT NOTICE