

Institiúid Teicneolaíochta Cheatharlach



INSTITUTE of  
TECHNOLOGY  
CARLOW

At the heart of South Leinster



College of Engineering, Design, Art and Technology

Makerere University

# The Access Tier ISP, considering upstream connectivity

DIARMUID Ó BRÍAIN  
GameCORE / netLabs!UG

30 January 2018

**MUM**  
**Nairobi 2018**



gamecore  
engaging people with technology

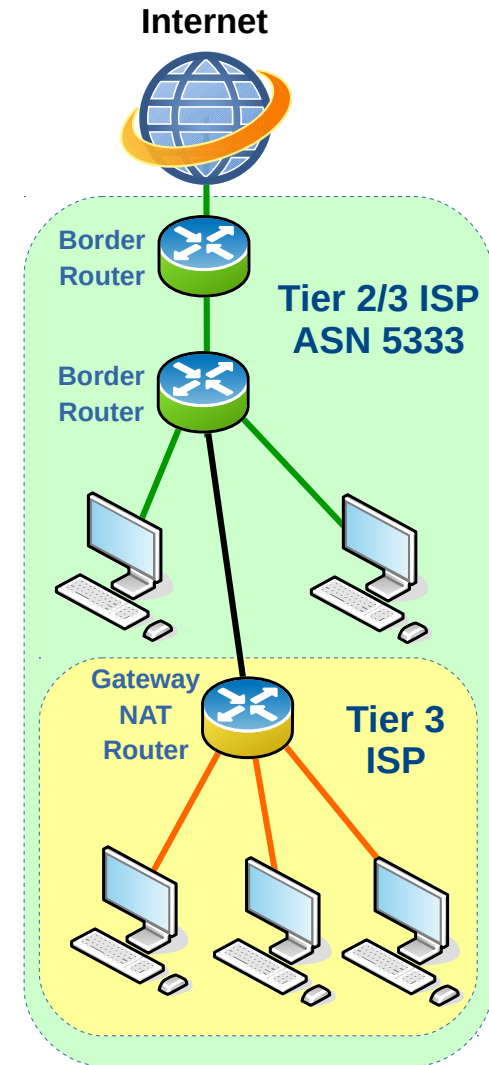
30 Jan 2018

MikroTik MUM, Nairobi 2018

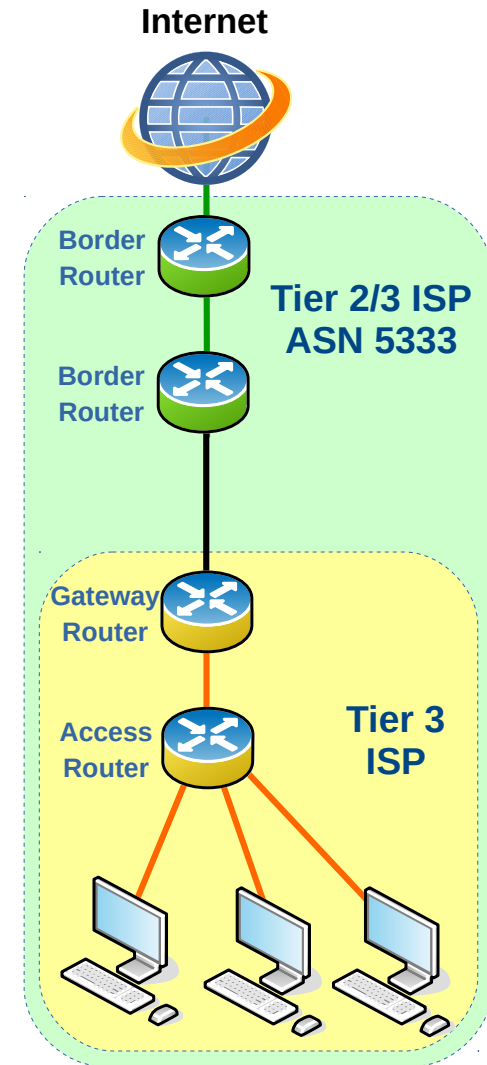
Copyright 2018 GameCORE

- Access ISPs connectivity to the Internet falls into one of these general categories:
  - ISP with minimal L3 address from upstream provider
  - ISP with L3 address from upstream provider
  - Full ISP, with ASN, IPv4 and IPv6 blocks
  - Full ISP with IXP Connection
  - Virtualised ISP

- Access customers connected to upstream provider using Network Address Translation (NAT)
- Easy to setup but limits size of network
- Gateway NAT router high workload
- Dependence from upstream provider/ISP
  - Renumbering headaches
    - When scaling up
    - When changing providers
- No ability to multi-home
- Dependency on NAT limits services to customers

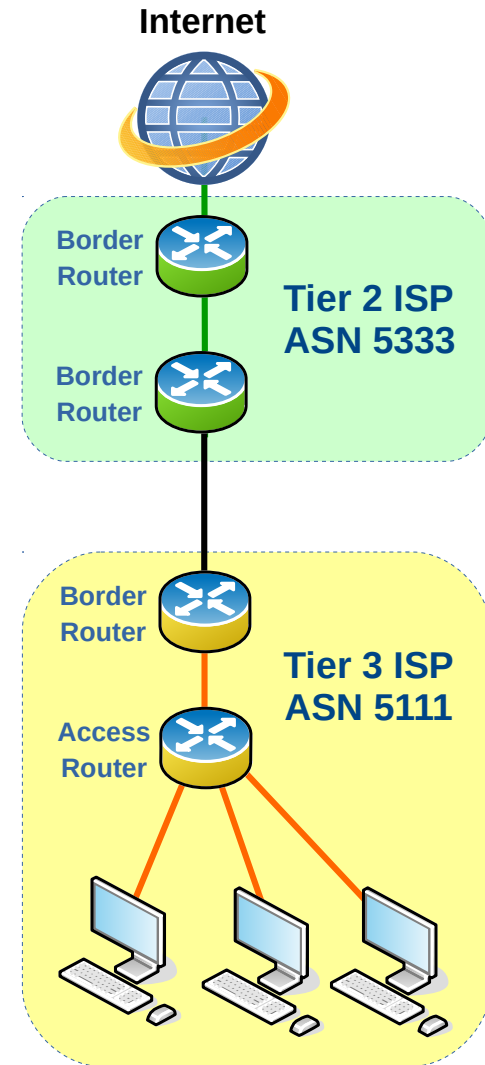


- Dependence from upstream provider/ISP
  - Renumbering headaches if changing providers
  - Difficult to change providers
- Cannot multi-home between different upstream ISPs
- Services to end customers limited by services of upstream provider
- Routing policies dictated by upstream provider

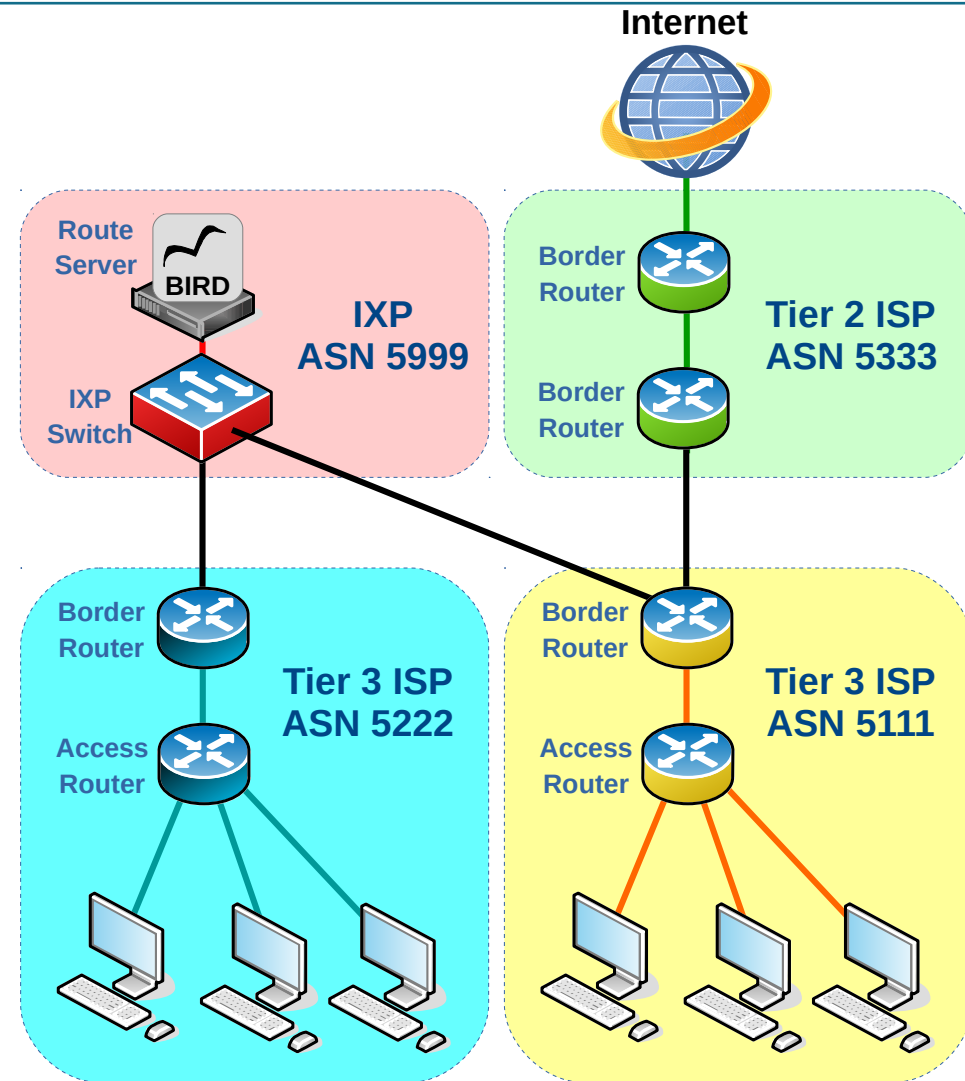


- Operational Advantage
- Independence from upstream provider/ISP
  - no renumbering headaches when changing providers
- Ability to multi-home
- No dependency on NAT, can offer enhanced services to end customers
- Simpler to implement routing policies that suit the company and services

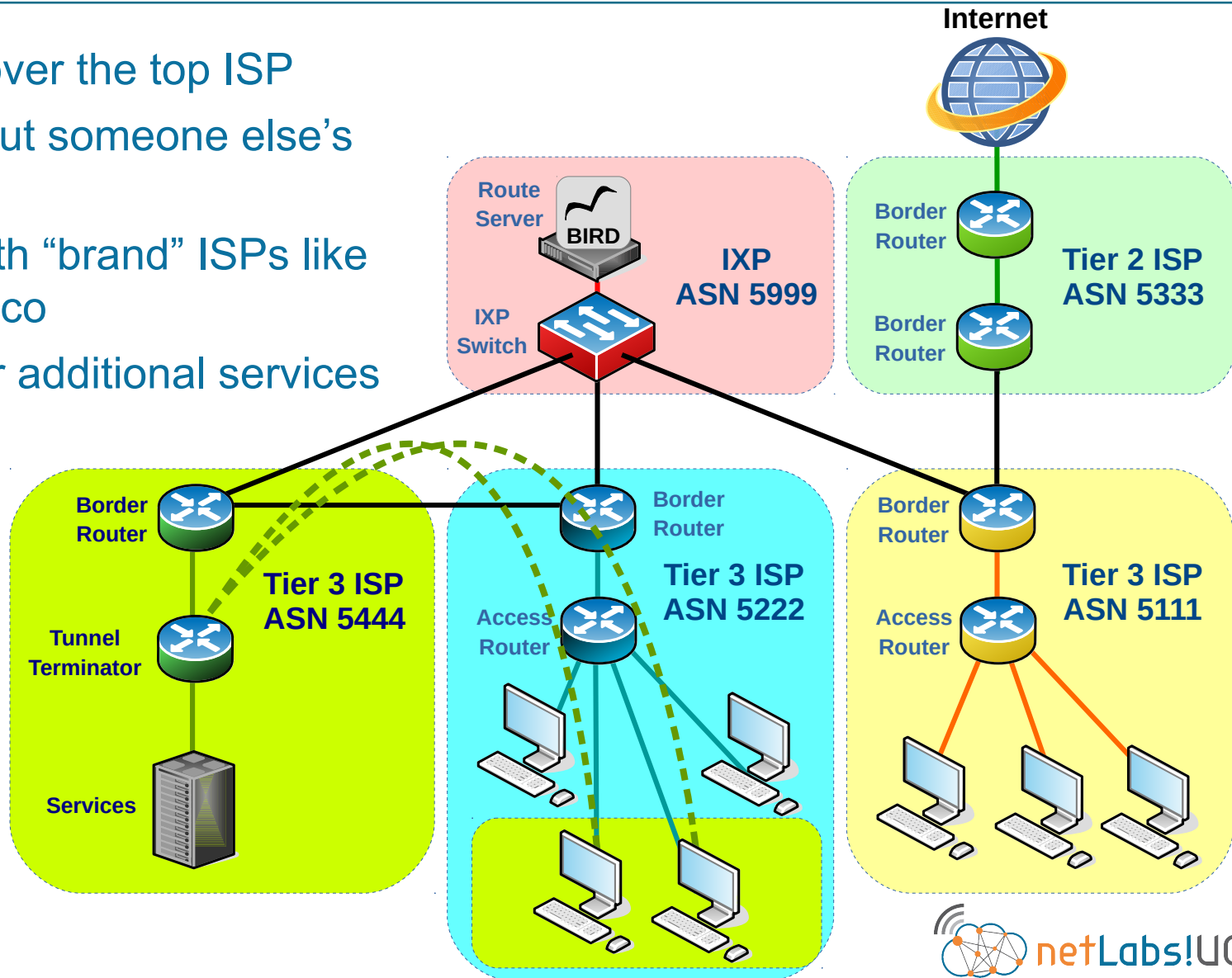
REF: <https://apps.afrinic.net/nmrp/>



- All of the previous advantages
- Plus;
  - Local traffic kept local
  - Reduced transit costs
  - Reduced latency to local content and content caches
  - Improved service to customers



- Virtualised over the top ISP
- Access rollout someone else's problem
- Common with "brand" ISPs like Sky and Tesco
- Usually offer additional services





# Inter Autonomous System routing

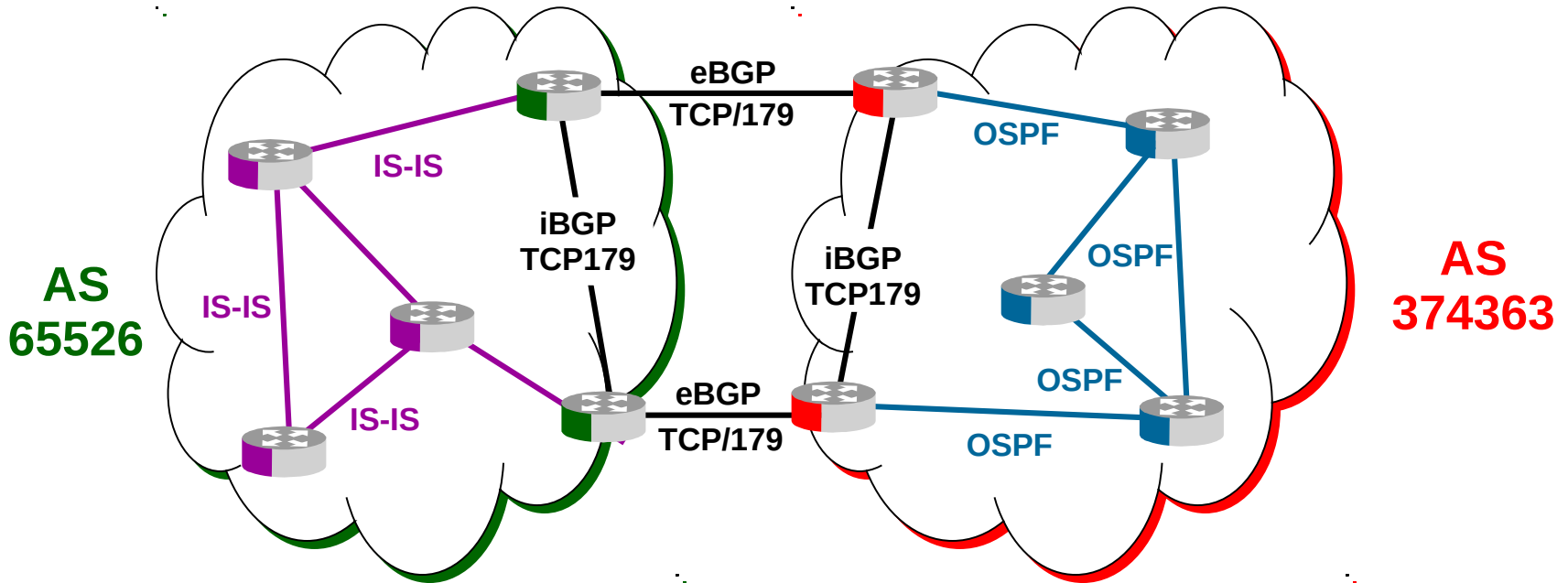
What exactly is an Autonomous System and how is routing carried out in one ?

**mum**  
**Nairobi 2018**

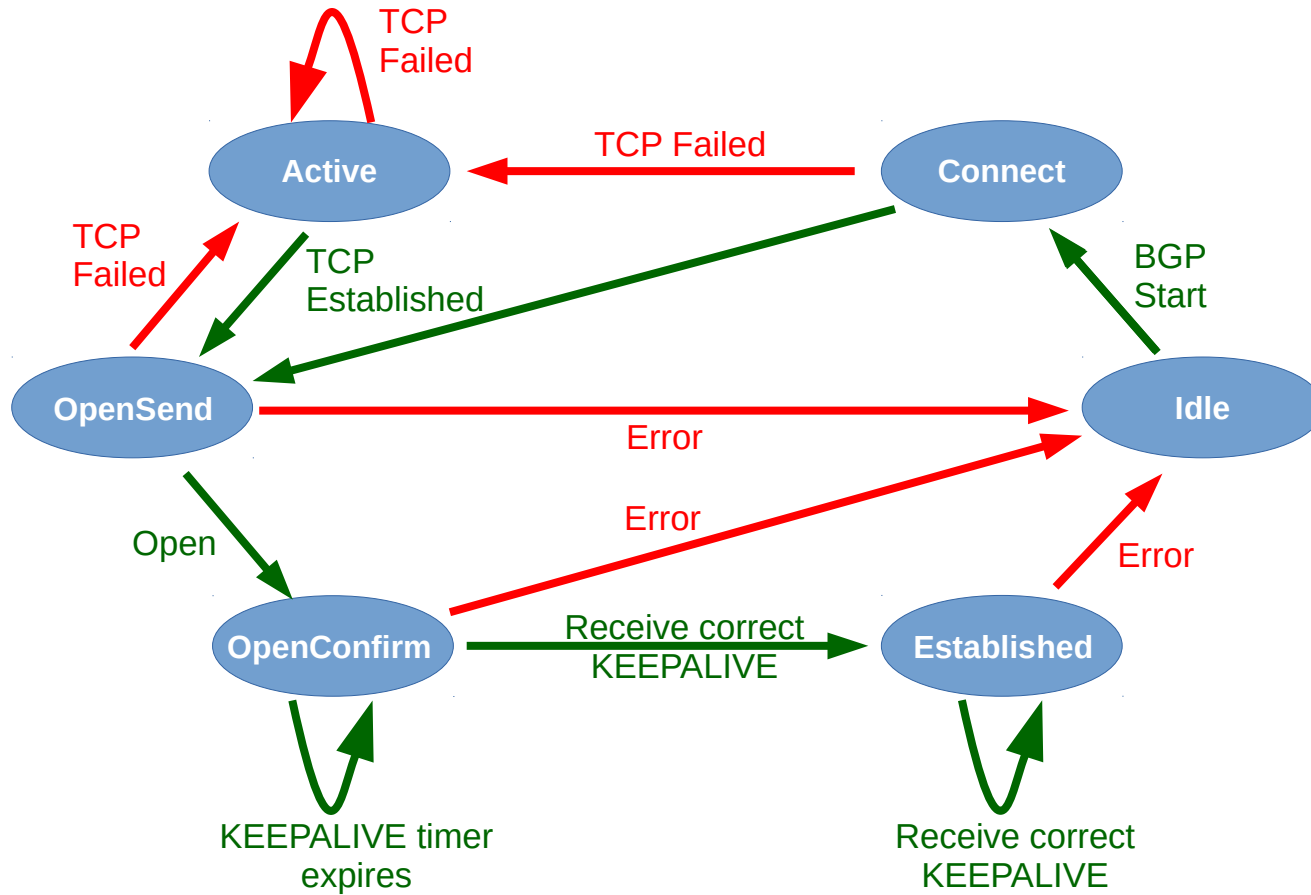




|         |      |       |              |        |         |          |
|---------|------|-------|--------------|--------|---------|----------|
| RFC1105 | 1989 | BGPv1 | RFC1654      | 16-bit | 2-Octet | AS65526  |
| RFC1654 | 1994 | BGPv4 | RFC4893/6793 | 32-bit | 4-Octet | AS374363 |



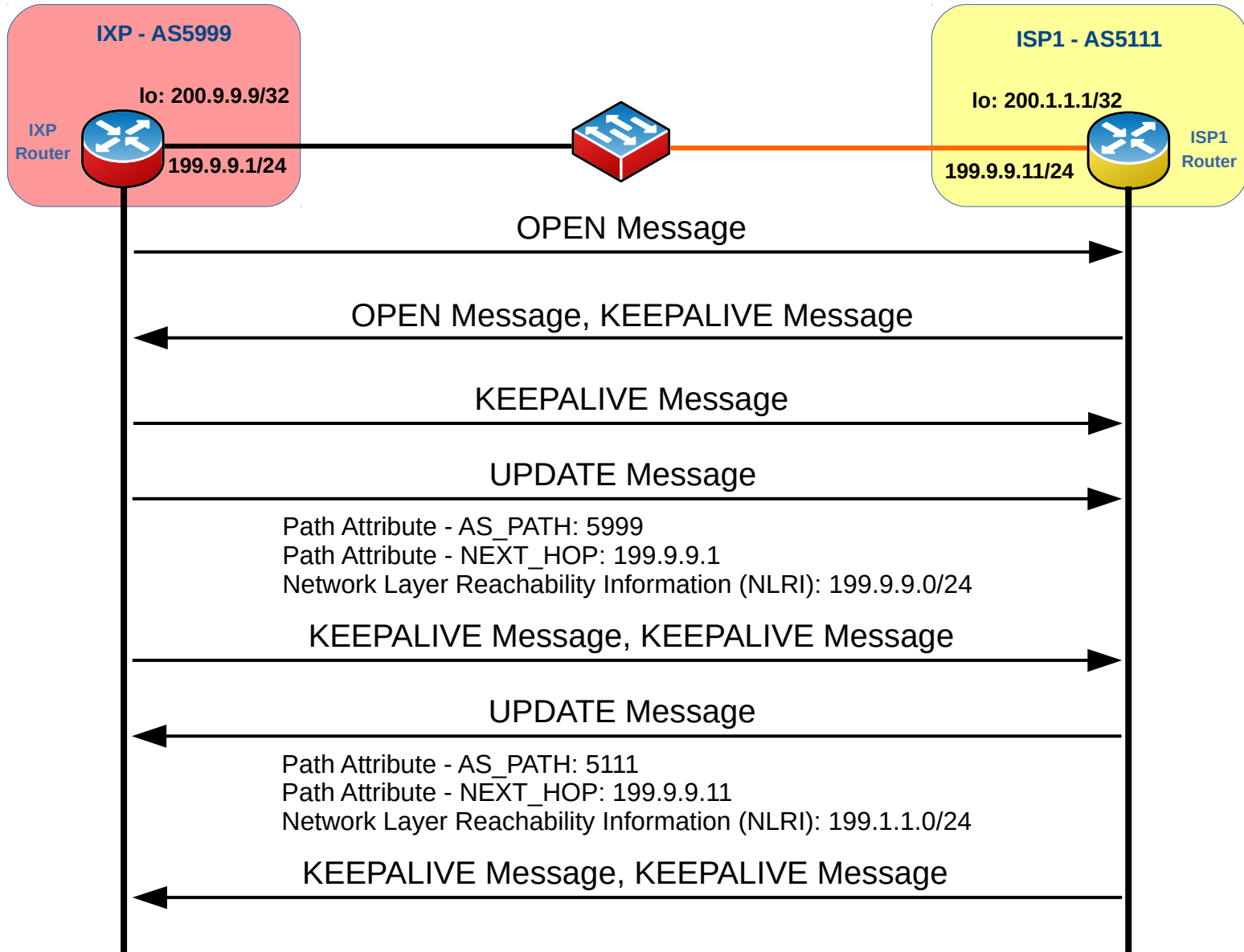
- **ORIGIN:** defines the origin of the path information
- **AS PATH:** sequence of AS path segments
- **NEXT HOP:** defines the IP address of the next hop router

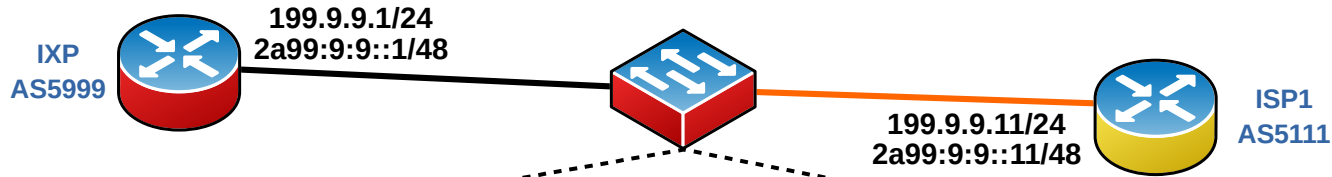


```

routing bgp peer print
Flags: X - disabled, E - established
#  INSTANCE          REMOTE-ADDRESS  REMOTE-AS
0  E  ASN5111         199.5.5.33     5333
1  E  ASN5111         2a99:5:5::33  5333
  
```

# Basic BGP Message flow





```
IPv4, Src: 199.9.9.1, Dst: 199.9.9.11
TCP, Src Port: 40932, Dst Port: 179, Seq: 65, Ack: 65, Len: 52
```

BGP - **UPDATE Message**

Marker: ff

Length: 52

Type: UPDATE Message (2)

Withdrawn Routes Length: 0

Total Path Attribute Length: 25

**Path attributes**

Path Attribute - ORIGIN: IGP

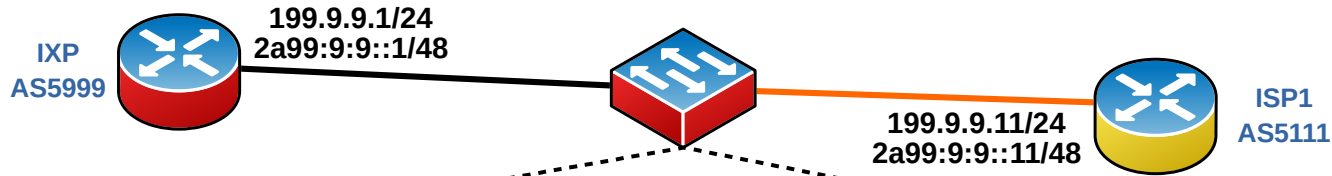
Path Attribute - AS\_PATH: 5999

Path Attribute - NEXT\_HOP: 199.9.9.1

Path Attribute - MULTI\_EXIT\_DISCRIMINATOR: 0

Network Layer Reachability Information

199.9.9.0/24



IPv6, Src: 2a99:9:9::1 , Dst: 2a99:9:9::11  
TCP, Src Port: 40932, Dst Port: 179, Seq: 65, Ack: 65, Len: 52

## BGP - UPDATE Message

Marker: ff

Length: 90

Type: UPDATE Message (2)

Withdrawn Routes Length: 0

Total Path Attribute Length: 67

### Path attributes

Path Attribute - ORIGIN: IGP

Path Attribute - AS\_PATH: 5999

Path Attribute - MULTI\_EXIT\_DISCRIMINATOR: 0

Path Attribute - MP\_REACH\_NLRI

Next hop network address (32 bytes)

Next Hop: 2a99:9:9::1

Next Hop: fe80::216:c7ff:fe9c:4860

Network Layer Reachability Information

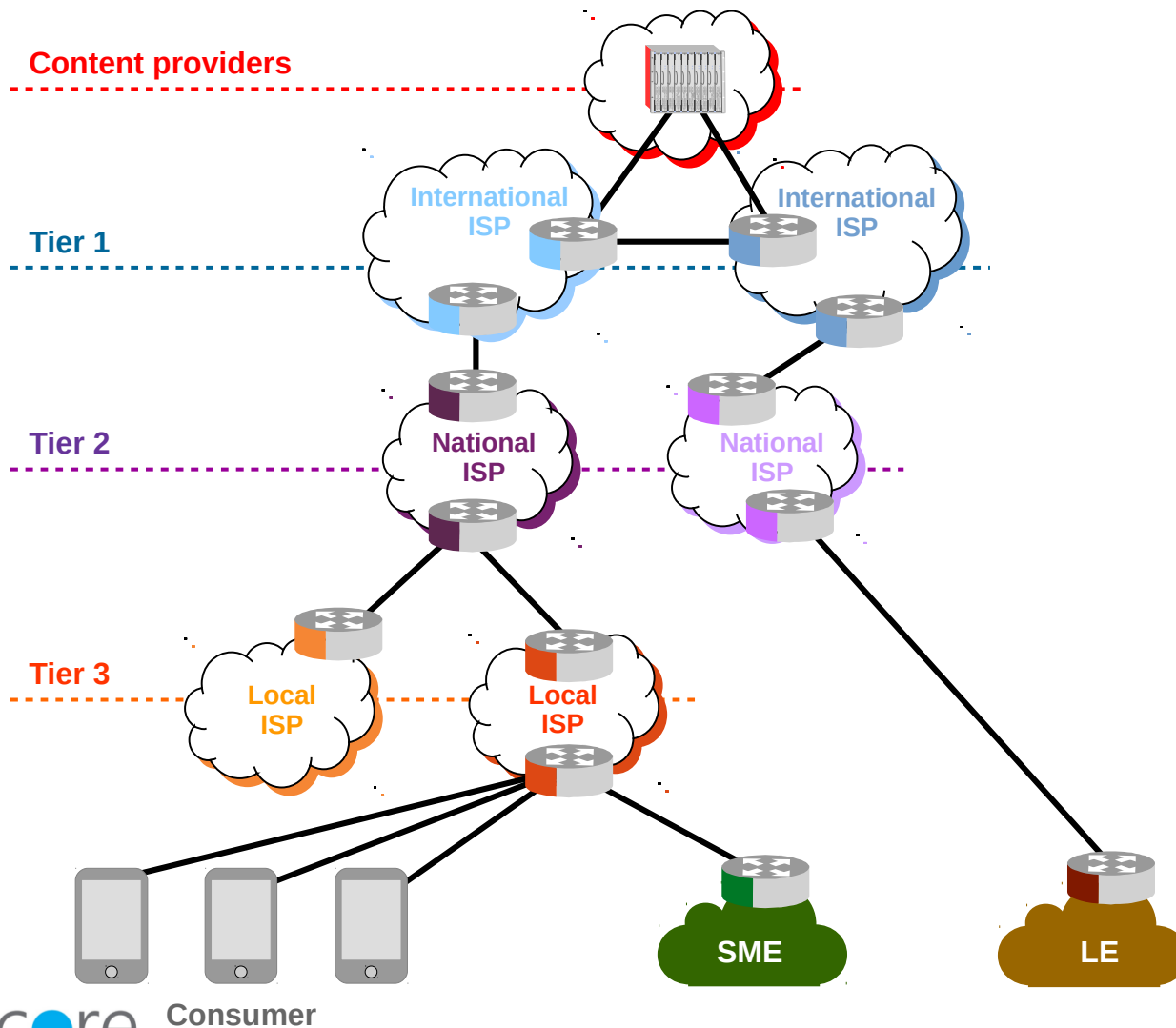
2a99:9:9::/48

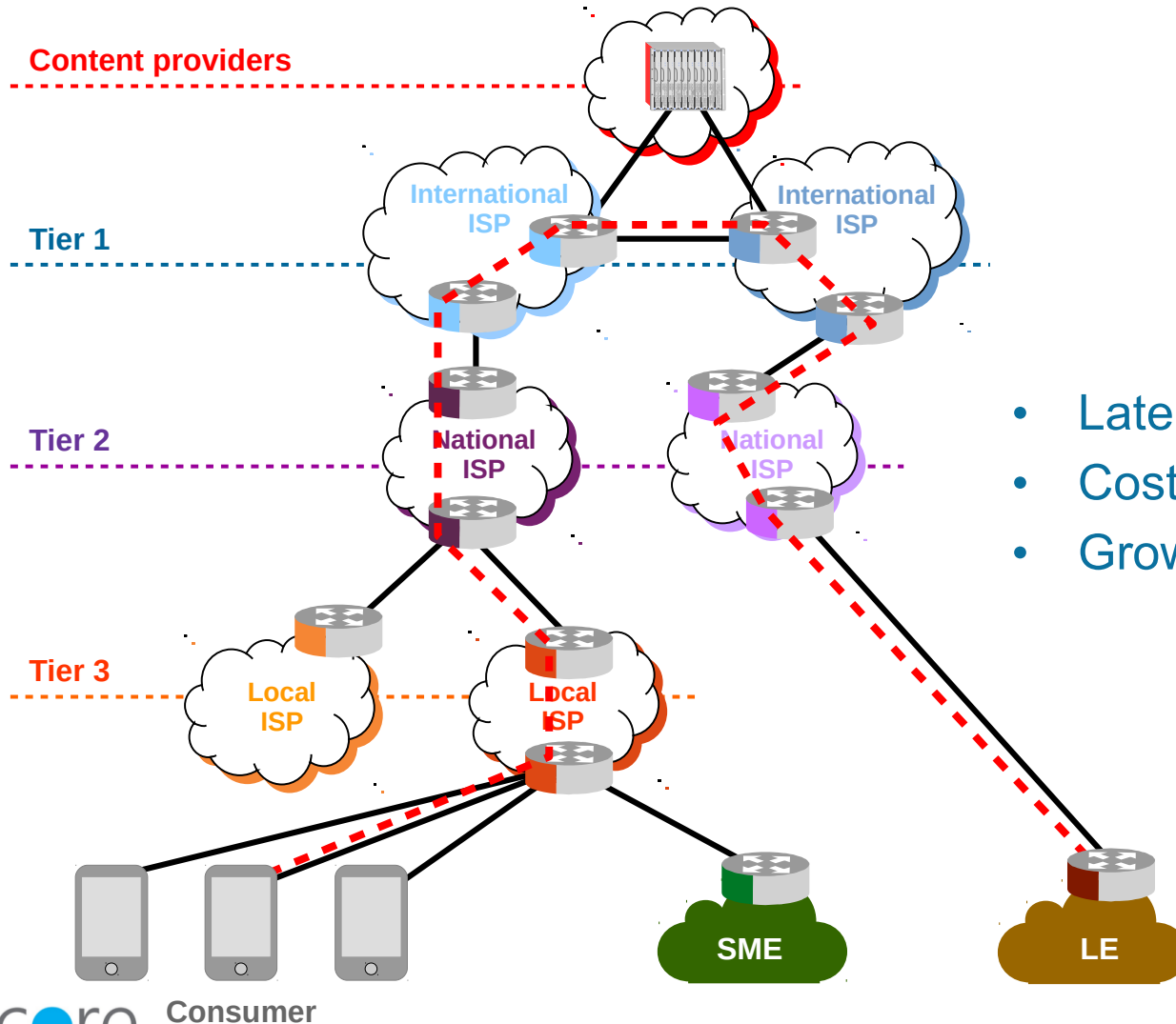


## Internet eXchange Points

How can an IXP help me ? Are they  
not just for the big boys ?

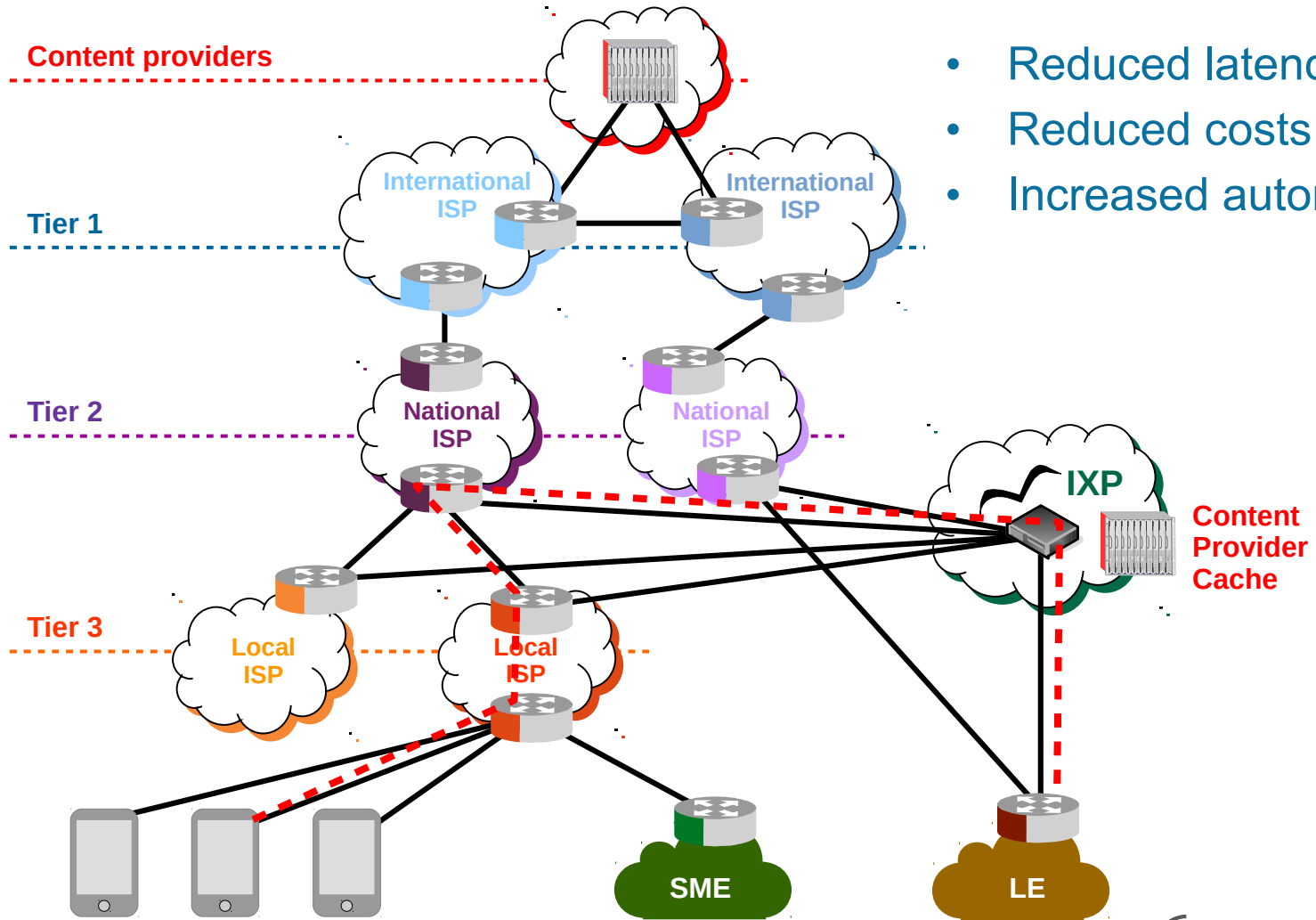
**mum**  
**Nairobi 2018**



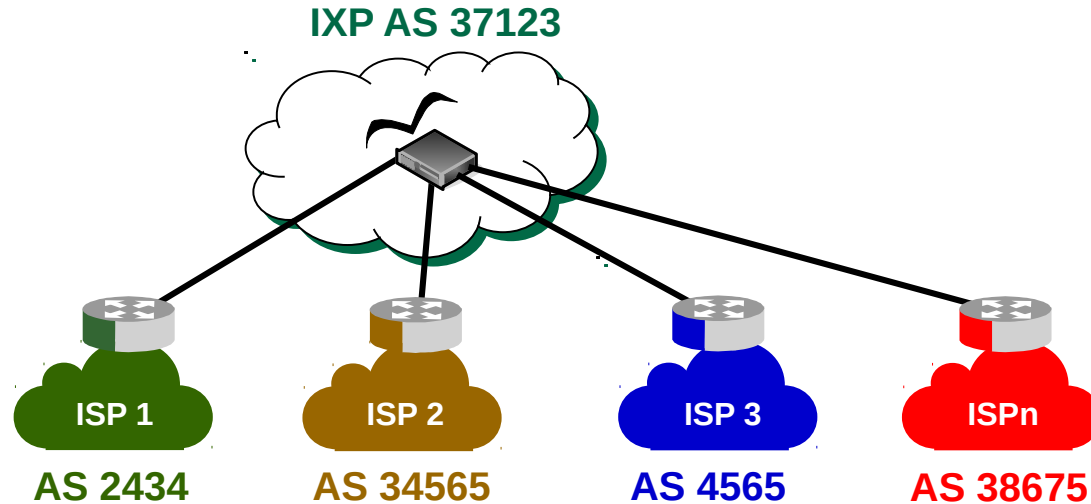


- Latency
- Cost
- Growth of ecosystem



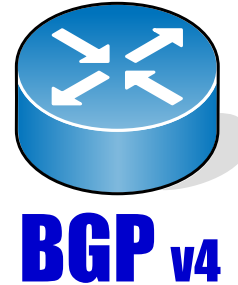


- Reduced latency
- Reduced costs
- Increased autonomy



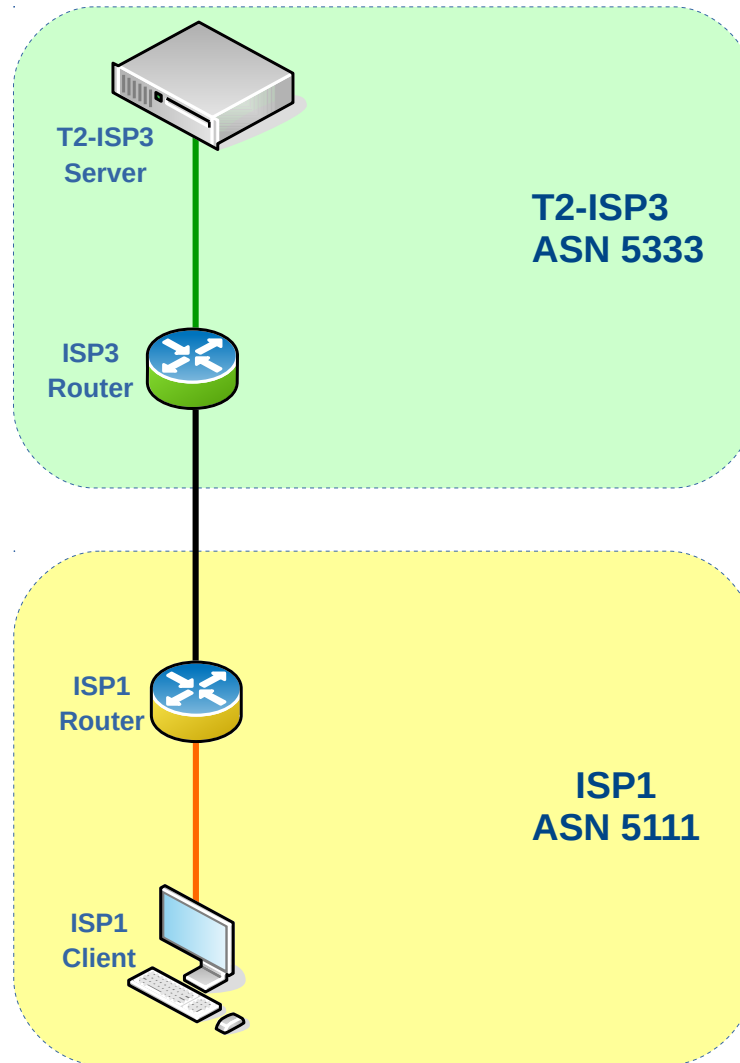
```
birdc> show route
165.253.0.0/23 via 197.243.54.38 on eth0 [ISP1 2017-05-15] * (100) [AS2434i]
105.179.200.0/22 via 197.243.54.51 on eth0 [ISP2 2017-05-15] * (100) [AS34565i]
198.51.100.0/24 via 197.243.54.46 on eth0 [ISP3 2017-05-15] * (100) [AS4565i]
41.221.89.0/24 via 197.243.54.33 on eth0 [ISP4 2017-05-15] * (100) [AS38675i]
```

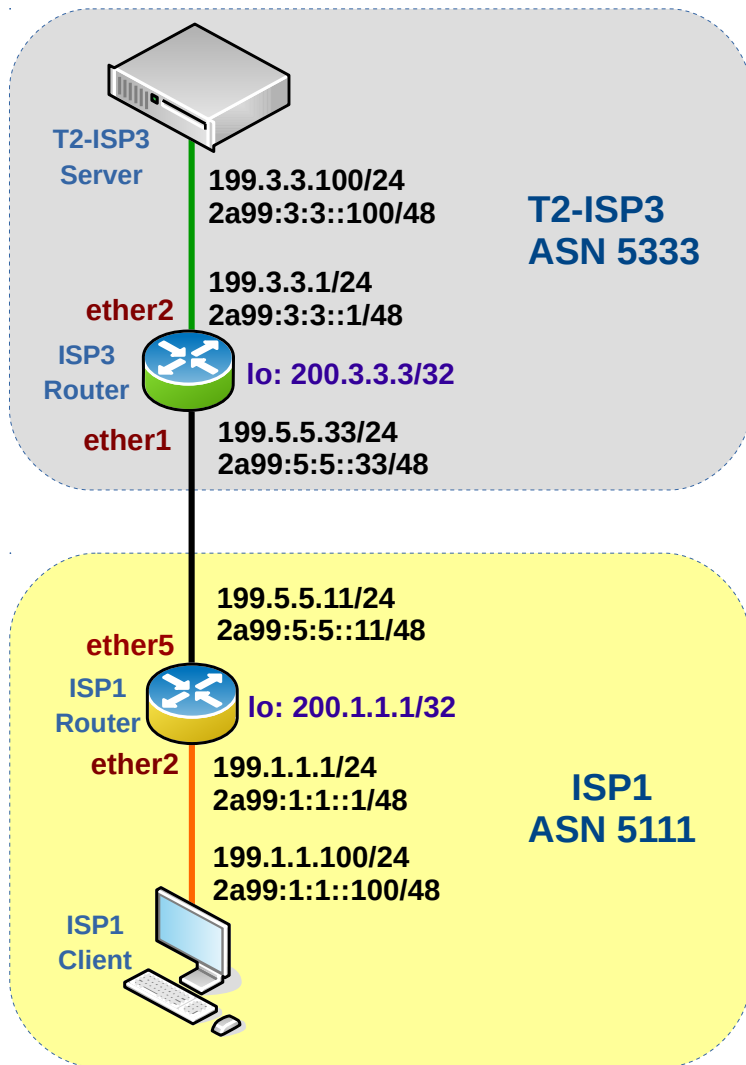
```
birdc> show route export ISP4
41.221.89.0/24 via 196.243.54.33 on eth0 [ISP4 2017-05-15] * (100) [AS38675i]
  Type: BGP unicast univ
  BGP.origin: IGP
  BGP.as_path: 38675
  BGP.next_hop: 196.243.54.33
  BGP.med: 0
  BGP.local_pref: 100
```



# BGP Peering Configuration

**mum**  
Nairobi 2018





- **Enable IPv6**

```
system package print
```

```
Flags: X - disabled
```

| #  | NAME            | VERSION | SCHEDULED |
|----|-----------------|---------|-----------|
| 0  | routeros-mipsbe | 6.38.7  |           |
| 1  | system          | 6.38.7  |           |
| 2  | X ipv6          | 6.38.7  |           |
| 3  | wireless        | 6.38.7  |           |
| 4  | hotspot         | 6.38.7  |           |
| 5  | dhcp            | 6.38.7  |           |
| 6  | mpls            | 6.38.7  |           |
| 7  | routing         | 6.38.7  |           |
| 8  | ppp             | 6.38.7  |           |
| 9  | security        | 6.38.7  |           |
| 10 | advanced-tools  | 6.38.7  |           |

```
system package enable 2
```

- **Reset configuration**

```
system reset-configuration no-defaults=yes
```

- **System identity**

```
system identity set name=ISP1
```

- Add IP Addresses to the interfaces

```
interface bridge add name=loopback0
ip address add address=200.1.1.1/32 interface=loopback0
```

```
ip address add address=199.1.1.1/24 interface=ether2
ipv6 address add address=2a99:1:1::1/48 interface=ether2
```

```
ip address add address=199.5.5.11/24 interface=ether5
ipv6 address add address=2a99:5:5::11/48 interface=ether5
```

### ip address print

```
Flags: X - disabled, I - invalid, D - dynamic
#   ADDRESS                NETWORK                INTERFACE
0   200.1.1.1/32           200.1.1.1             loopback0
1   199.9.9.11/24         199.9.9.0             ether1
2   199.5.5.11/24         199.5.5.0             ether5
```

### ipv6 address print

```
Flags: X - disabled, I - invalid, D - dynamic, G - global, L - link-local
#   ADDRESS                FROM-POOL INTERFACE ADVERTISE
0 DL fe80::20c:42ff:fec2:117c/64          ether2      no
1 DL fe80::20c:42ff:fec2:117e/64          ether4      no
2 DL fe80::20c:42ff:fec2:117f/64          ether5      no
3 IG 2a99:9:9::11/48                ether1      yes
4 G 2a99:5:5::11/48
```

- Create a BGP instance and add networks to be routed

```
routing bgp instance add name=ASN5111 as=5111 router-id=200.1.1.1
```

```
routing bgp network add network=199.1.1.0/24  
routing bgp network add network=2a99:1:1::/48
```

```
routing bgp instance print from=ASN5111
```

```
Flags: * - default, X - disabled
```

```
0 name="ASN5111" as=5111 router-id=200.1.1.1 redistribute-connected=no  
  redistribute-static=no redistribute-rip=no redistribute-ospf=no  
  redistribute-other-bgp=no out-filter="" client-to-client-reflection=yes  
  ignore-as-path-len=no routing-table=""
```

```
routing bgp network print
```

```
Flags: X - disabled
```

| # | NETWORK       | SYNCHRONIZE |
|---|---------------|-------------|
| 0 | 199.1.1.0/24  | yes         |
| 1 | 2a99:1:1::/48 | yes         |



- Ingress filters are used to filter advertisements into the network
- Best practice for ingress filters for all peers are:
  - Discard receiving ones own prefix
  - Discard receiving a default route as we are doing full routing
  - Discard special purpose address registry entries stated at RFC 6890
- RFC6890 - Special-Purpose IP Address Registries
  - IETF reserved blocks that should not be received in routing update messages

- Create a BGP ingress filter for IPv4
  - Discard own network
  - Discard Default route
  - Jump the filter to the RFC-6890 IPv4 filter

```
routing filter add chain=IN-ISP3-IPv4 prefix=199.1.1.0/24 action=discard
routing filter add chain=IN-ISP3-IPv4 prefix=0.0.0.0/0 action=discard
routing filter add chain=IN-ISP3-IPv4 action=jump jump-target=IN-RFC-6890-IPv4
```

```
routing filter add chain=IN-RFC-6890-IPv4 prefix=0.0.0.0/8 action=discard
routing filter add chain=IN-RFC-6890-IPv4 prefix=10.0.0.0/8 action=discard
routing filter add chain=IN-RFC-6890-IPv4 prefix=100.64.0.0/10 action=discard
routing filter add chain=IN-RFC-6890-IPv4 prefix=127.0.0.0/8 action=discard
routing filter add chain=IN-RFC-6890-IPv4 prefix=169.254.0.0/16 action=discard
routing filter add chain=IN-RFC-6890-IPv4 prefix=172.16.0.0/12 action=discard
routing filter add chain=IN-RFC-6890-IPv4 prefix=192.0.2.0/24 action=discard
routing filter add chain=IN-RFC-6890-IPv4 prefix=192.88.99.0/24 action=discard
routing filter add chain=IN-RFC-6890-IPv4 prefix=192.168.0.0/16 action=discard
routing filter add chain=IN-RFC-6890-IPv4 prefix=198.18.0.0/15 action=discard
routing filter add chain=IN-RFC-6890-IPv4 prefix=198.51.100.0/24 action=discard
routing filter add chain=IN-RFC-6890-IPv4 prefix=203.0.113.0/24 action=discard
routing filter add chain=IN-RFC-6890-IPv4 prefix=240.0.0.0/4 action=discard
routing filter add chain=IN-RFC-6890-IPv4 prefix=255.255.255.255/32 action=discard
```

- Create a BGP ingress filter for IPv6
  - Discard own network
  - Discard Default route
  - Jump the filter to the RFC-6890 IPv6 filter

```
routing filter add chain=IN-ISP3-IPv6 prefix=2a99:1:1::/48 action=discard
routing filter add chain=IN-ISP3-IPv6 prefix::/0 action=discard
routing filter add chain=IN-ISP3-IPv6 action=jump jump-target=IN-RFC-6890-IPv6
```

```
routing filter add chain=IN-RFC-6890-IPv6 prefix::1/128 action=discard
routing filter add chain=IN-RFC-6890-IPv6 prefix::/128 action=discard
routing filter add chain=IN-RFC-6890-IPv6 prefix=64:ff9b::/96 action=discard
routing filter add chain=IN-RFC-6890-IPv6 prefix::ffff:0:0/96 action=discard
routing filter add chain=IN-RFC-6890-IPv6 prefix=100::/64 action=discard
routing filter add chain=IN-RFC-6890-IPv6 prefix=2001::/23 action=discard
routing filter add chain=IN-RFC-6890-IPv6 prefix=2001::/32 action=discard
routing filter add chain=IN-RFC-6890-IPv6 prefix=2001:2::/48 action=discard
routing filter add chain=IN-RFC-6890-IPv6 prefix=2001:db8::/32 action=discard
routing filter add chain=IN-RFC-6890-IPv6 prefix=2001:10::/28 action=discard
routing filter add chain=IN-RFC-6890-IPv6 prefix=2002::/16 action=discard
routing filter add chain=IN-RFC-6890-IPv6 prefix=fc00::/7 action=discard
routing filter add chain=IN-RFC-6890-IPv6 prefix=fe80::/10 action=discard
```

- Egress filters are used to filter advertisements from the network, limiting it to only advertise the specific local networks
  - ***invert-match=yes*** statement, this instructs the filter to discard all but the prefix specified

```
routing filter add chain=OUT-IPv4 prefix=199.1.1.0/24 invert-match=yes  
action=discard
```

```
routing filter add chain=OUT-IPv6 prefix=2a99:1:1::/48 invert-match=yes  
action=discard
```

- Create the peer link to the other BGP peer at IXP3 for IPv4 and IPv6
  - Note that IPv4 family is the default and for IPv6 the family must be specified

```
routing bgp peer add name=isp3 instance=ASN5111 remote-as=5333  
remote-address=199.5.5.33 in-filter=IN-ISP3-IPv4 out-filter=OUT-IPv4
```

```
routing bgp peer add name=isp3 instance=ASN5111 remote-as=5333  
remote-address=2a99:5:5::33 address-families=ipv6  
in-filter=IN-ISP3-IPv6 out-filter=OUT-IPv6
```

- Confirm that the routes have been learnt from ISP3

## routing bgp peer print

Flags: X - disabled, E - established

| # | INSTANCE  | REMOTE-ADDRESS | REMOTE-AS |
|---|-----------|----------------|-----------|
| 0 | E ASN5111 | 199.5.5.33     | 5333      |
| 1 | E ASN5111 | 2a99:5:5::33   | 5333      |

## ip route print

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 | ADC 199.1.1.0/24 | 199.1.1.1  | ether2     | 0        |
| 1 | ADb 199.3.3.0/24 |            | 199.5.5.33 | 20       |
| 2 | ADC 199.5.5.0/24 | 199.5.5.11 | ether5     | 0        |
| 3 | ADC 200.1.1.1/32 | 200.1.1.1  | loopback0  | 0        |

- <http://www.team-cymru.org/>
- A bogon prefix should never appear in the Internet routing table.
  - bogon route-servers
    - 65333:888
  - fullbogon route-servers
    - 65332:888

<http://www.team-cymru.org/bgp-examples.html#mikrotik-full>



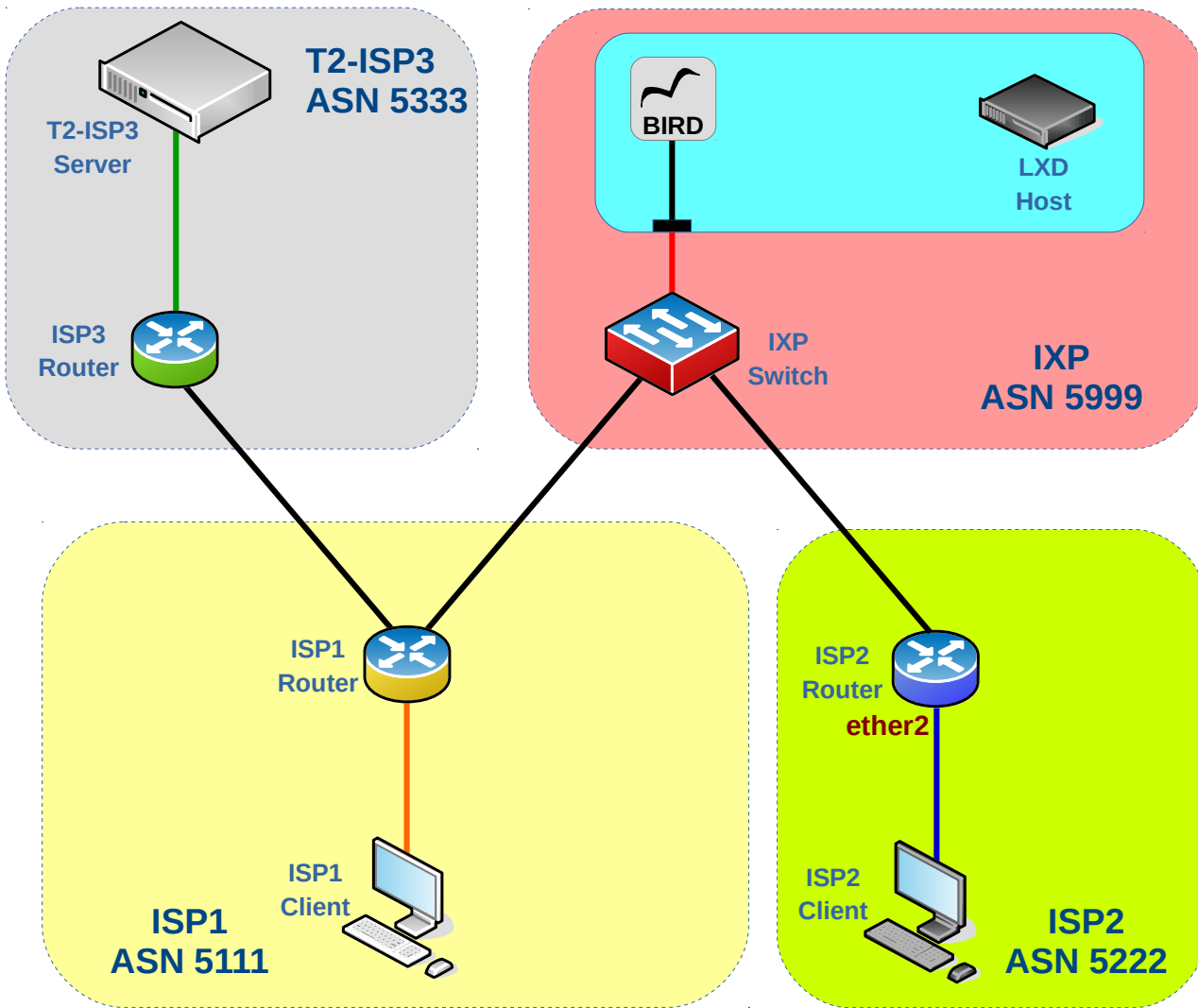


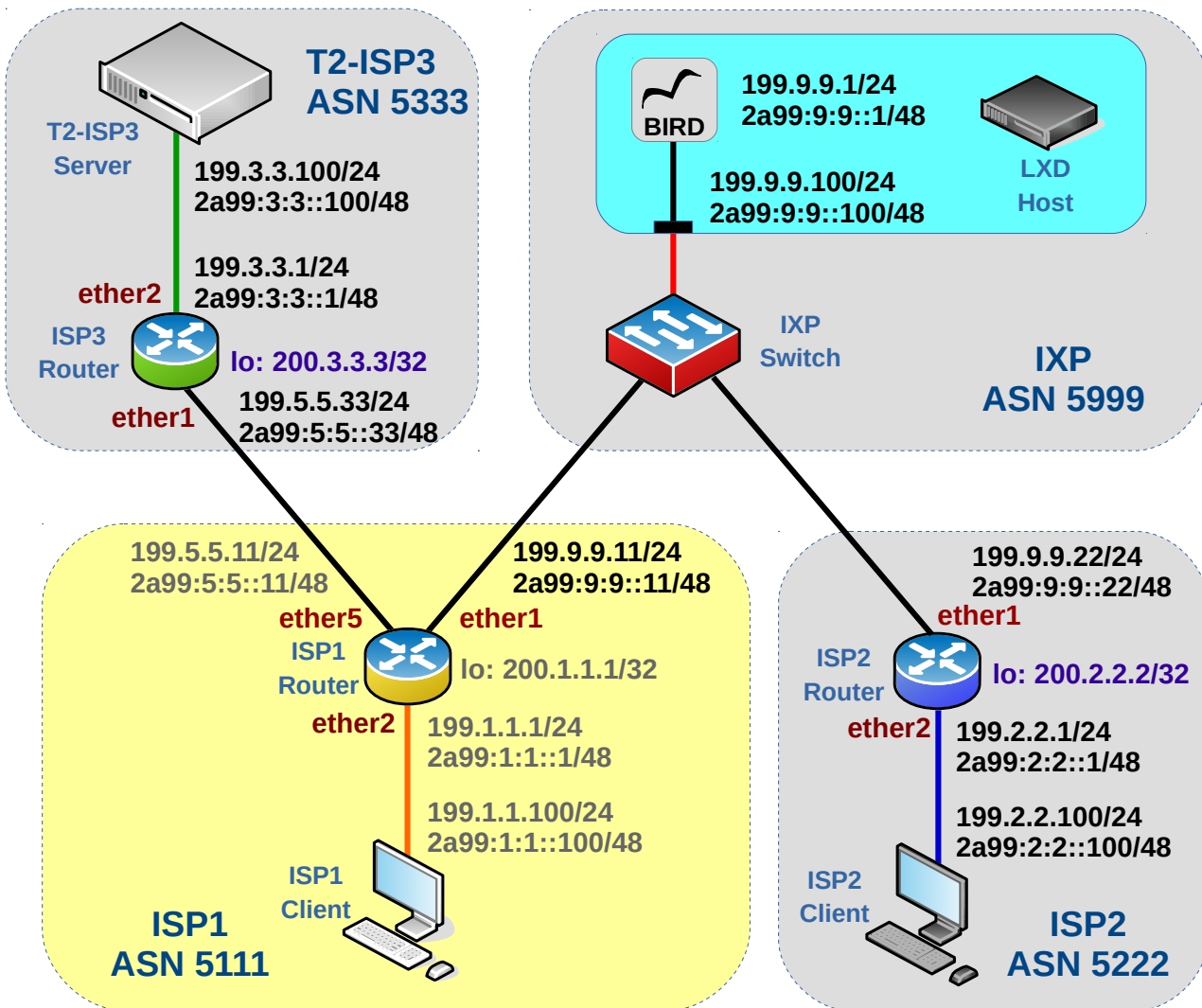
# Internet eXchange Point (IXP) Configuration

**mum**  
**Nairobi 2018**









- Add IP Addresses to the interface facing the IXP

```
ip address add address=199.9.9.11/24 interface=ether1
```

```
ipv6 address add address=2a99:9:9::11/48 interface=ether1
```

### ip address print

Flags: X - disabled, I - invalid, D - dynamic

| # | ADDRESS       | NETWORK   | INTERFACE |
|---|---------------|-----------|-----------|
| 0 | 200.1.1.1/32  | 200.1.1.1 | loopback0 |
| 1 | 199.5.5.11/24 | 199.5.5.0 | ether5    |
| 2 | 199.1.1.1/24  | 199.1.1.0 | ether2    |
| 3 | 199.9.9.11/24 | 199.9.9.0 | ether1    |

### ipv6 address print

Flags: X - disabled, I - invalid, D - dynamic, G - global, L - link-local

| # | ADDRESS                        | FROM-POOL | INTERFACE | ADVERTISE |
|---|--------------------------------|-----------|-----------|-----------|
| 0 | DL fe80::20c:42ff:fec2:117c/64 |           | ether2    | no        |
| 1 | DL fe80::20c:42ff:fec2:117e/64 |           | ether4    | no        |
| 2 | DL fe80::20c:42ff:fec2:117f/64 |           | ether5    | no        |
| 3 | G 2a99:5:5::11/48              |           | ether5    | yes       |
| 4 | DL fe80::20c:42ff:fec2:117d/64 |           | ether3    | no        |
| 5 | G 2a99:1:1::1/48               |           | ether2    | yes       |
| 6 | DL fe80::20c:42ff:fec2:117b/64 |           | ether1    | no        |
| 7 | G 2a99:9:9::11/48              |           | ether1    | yes       |

- Chains IN-IXP-IPv4 and IN-IXP-IPv6 are input filters that discards receiving ones own prefix or a default route as full routing is taking place
  - IPv4

```
routing filter add chain=IN-IXP-IPv4 prefix=199.1.1.0/24 action=discard
routing filter add chain=IN-IXP-IPv4 prefix=0.0.0.0/0 action=discard
routing filter add chain=IN-IXP-IPv4 action=jump jump-target=IN-RFC-6890-IPv4
```

- IPv6

```
routing filter add chain=IN-IXP-IPv6 prefix=2a99:1:1::/48 action=discard
routing filter add chain=IN-IXP-IPv6 prefix::/0 action=discard
routing filter add chain=IN-IXP-IPv6 action=jump jump-target=IN-RFC-6890-IPv6
```

- Create the peer link to the IXP for IPv4 and IPv6
- Note that IPv4 family is the default and for IPv6 the family must be specified

- IPv4

```
routing bgp peer add name=ixp instance=ASN5111 remote-as=5999  
remote-address=199.9.9.1 in-filter=IN-IXP-IPv4 out-filter=OUT-IPv4
```

- IPv6

```
routing bgp peer add name=ixp instance=ASN5111 remote-as=5999  
remote-address=2a99:9:9::1 address-families=ipv6 in-filter=IN-IXP-IPv6  
out-filter=OUT-IPv6
```

- Confirm that the routes have been learnt from IXP

## routing bgp peer print

Flags: X - disabled, E - established

| # | INSTANCE  | REMOTE-ADDRESS | REMOTE-AS |
|---|-----------|----------------|-----------|
| 0 | E ASN5111 | 199.5.5.33     | 5333      |
| 1 | E ASN5111 | 2a99:5:5::33   | 5333      |
| 2 | E ASN5111 | 199.9.9.1      | 5999      |
| 3 | E ASN5111 | 2a99:9:9::1    | 5999      |

## ip route print

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 | ADC | 199.1.1.0/24 | 199.1.1.1  | ether2     | 0        |
| 1 | ADb | 199.2.2.0/24 |            | 199.9.9.22 | 20       |
| 2 | ADb | 199.3.3.0/24 |            | 199.5.5.33 | 20       |
| 3 | ADC | 199.5.5.0/24 | 199.5.5.11 | ether5     | 0        |
| 4 | ADC | 199.9.9.0/24 | 199.9.9.11 | ether1     | 0        |
| 5 | ADC | 200.1.1.1/32 | 200.1.1.1  | loopback0  | 0        |



Thank you for your attention

Complete configurations can be found at:

<http://www.obriain.com/mikrotik>

diarmuid.obriain@itcarlow.ie  
diarmuid.obriain@netlabsug.org  
<http://www.netlabsug.org>

**MUM**  
**Nairobi 2018**

