INEX

BGP Best Practice at IXPs

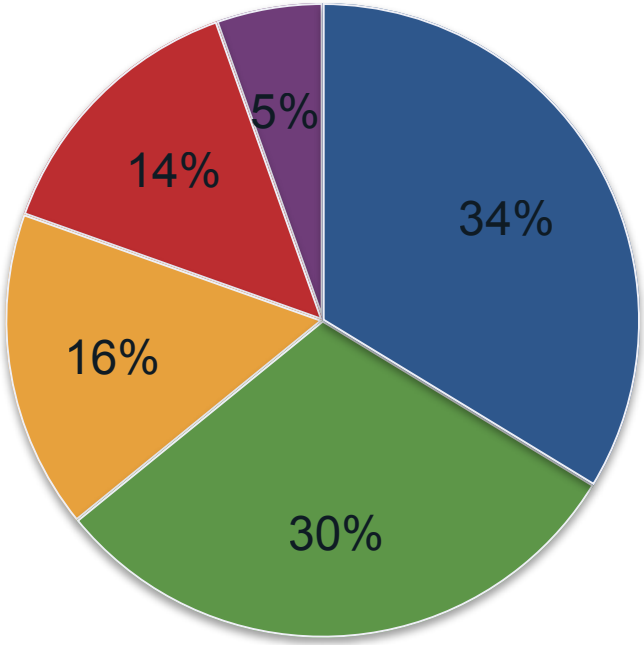
Beirut, March 2017

Nick Hilliard

Chief Technical Officer
Internet Neutral Exchange Association
Company Limited by Guarantee

OBLIGATORY INFOGRAPHIC

What Routers do People Use?



Cisco
Juniper
PC / VMware
Mikrotik
Others

INEX

BGP Hygiene at IXPs

BGP Speed Optimisation

- Don’t use as-path lists if possible
- Use BGP communities where possible
- On IOS, use prefix-lists instead of access-lists
- On IOS, use peer-groups or peer-templates for CPU efficiency

BGP Hygiene at IXPs

BGP Security - maximum prefixes

- Tears down BGP session once prefix limit is reached
- Helps to stop problems due to third party mistakes
- Large organisations are just as likely to make mistakes as small
- Some IXP operators maintain a list of maxprefixes on IXP Manager
- Otherwise sensible to limit to 200 generally, with exceptions

BGP Hygiene at IXPs

BGP Security - maximum prefixes

- Cisco Configuration

```
router bgp 64512
 address-family ipv4
  neighbor X.Y.Z.W maximum-prefix 200
```

- Juniper Configuration

```
set protocols bgp group <group> family inet unicast prefix-limit maximum 200
```

INEX

BGP Hygiene at IXPs

BGP Security - IRRDB Prefix Filters

- RIPE IRRDB route: objects generally not used for filtering prefix lists at IXPs
- Should not be used unless you're a programmer
- Well-managed Route Server system will use IRRDB data automatically
- No need to tell other members about new announcements

BGP Hygiene at IXPs

BGP Security - RPKI

- RPKI data currently not good enough to use in production networks
- Won’t fulfil aims until prefixes can be dropped on RPKI policy
- Unclear whether this can be achieved in practice
- Will never work properly until signed paths are supported
- Signed paths are incompatible with route servers
- Basic origin validation supported on IOS and JUNOS

BGP Hygiene at IXPs

BGP Security - MD5 passwords

- Some companies have security policies which require it
- Not as useful as some people claim
- Primary use at IXP is to stop session hijacking when addresses are re-used
- Formally obsoleted by TCP-AO since June 2010
- Still no production TCP-AO implementations
- MD5 will continue to be the only option in future

BGP Hygiene at IXPs

BGP Security - Deny Own Prefixes

- Accepting your prefixes from third party networks will cause problems
- Can cause internal BGP announcements to be dropped
- Default Cisco BGP settings will allow on-net hijacking of traffic

BGP Hygiene at IXPs

BGP Security - Deny Own Prefixes

- Cisco Configuration

```
router bgp 64512
 address-family ipv4
  neighbor 193.242.111.X prefix-list pl-ipv4-own-prefix-range

ip prefix-list pl-ipv4-own-prefix-range seq 5 deny 192.0.2.0/24 le 32
ip prefix-list pl-ipv4-own-prefix-range seq 10 permit 0.0.0.0/0 le 32
```

INEX

BGP Hygiene at IXPs

BGP Security - Deny Own Prefixes

- Juniper Configuration

```
set policy-options prefix-list pl-ipv4-own-address-block 192.0.2.0/24
edit policy-options policy-statement export-ebgp-ipv4-own-prefix-range
set term 10-deny-local from prefix-list-filter pl-ipv4-own-address-block orlonger
set term 10-deny-local then reject
set term 99-accept then accept
top
set protocols bgp group <group> export export-ebgp-ipv4-own-prefix-range
```

BGP Hygiene at IXPs

BGP Security - Deny Own Prefixes

- The default admin distance for eBGP routes on IOS and XR is lower than for IGPs
- I.e. eBGP routes take preference to IGP routes
- Important to change this for handling multihomed customers and to limit the effects of external prefix hijacking
- Cisco Configuration:

```
router bgp 64512
 address-family ipv4
  distance bgp 200 200 200
```

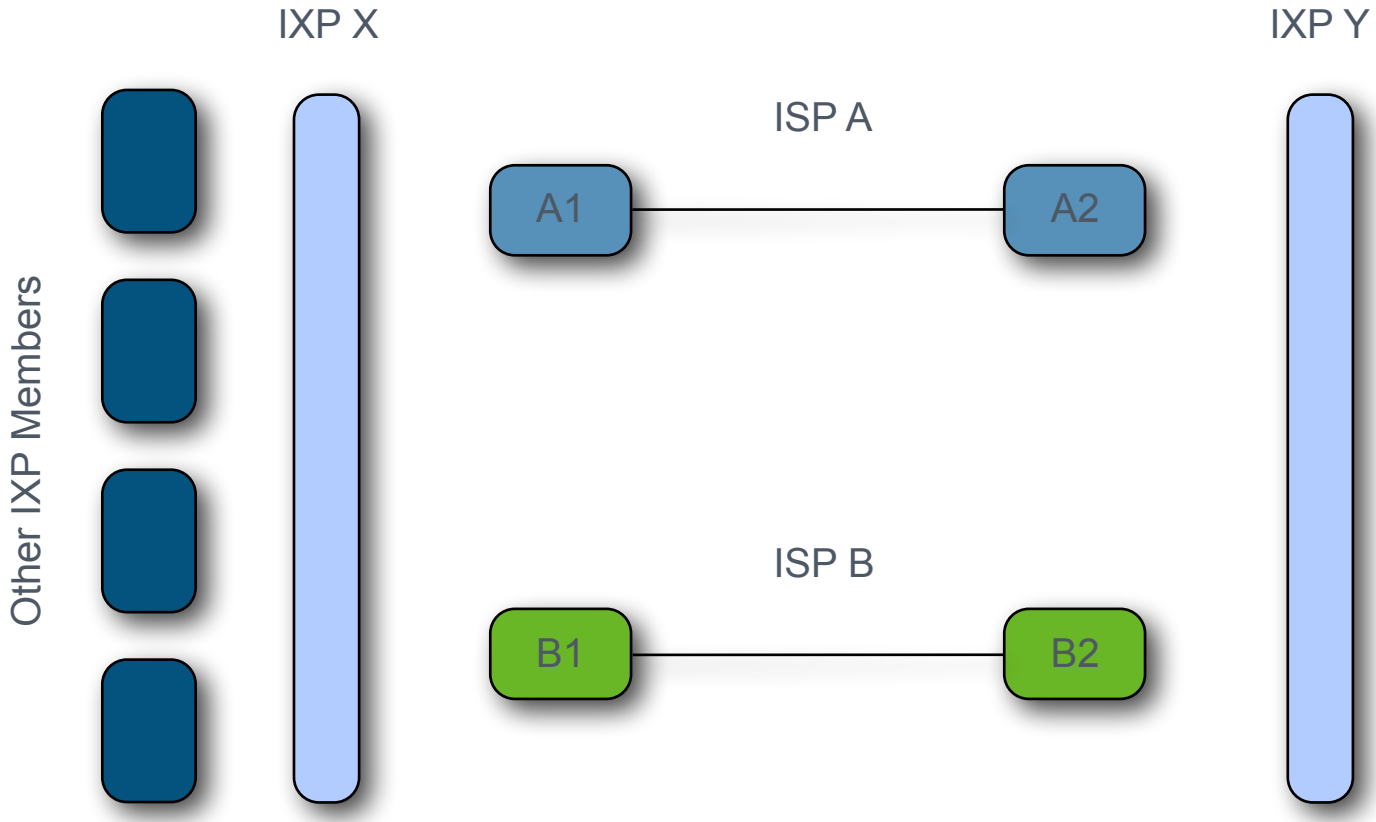
BGP Hygiene at IXPs

BGP Security - Check Next Hop IP Address

- eBGP allows you to set the BGP next-hop address to be any address on the IXP
- At an IXP, you can configure next-hop IP to be any address on IXP
- Recommend checking peer address = next hop address
- Documented in RFC 7948 “Internet Exchange BGP Route Server Operations"

BGP Hygiene at IXPs

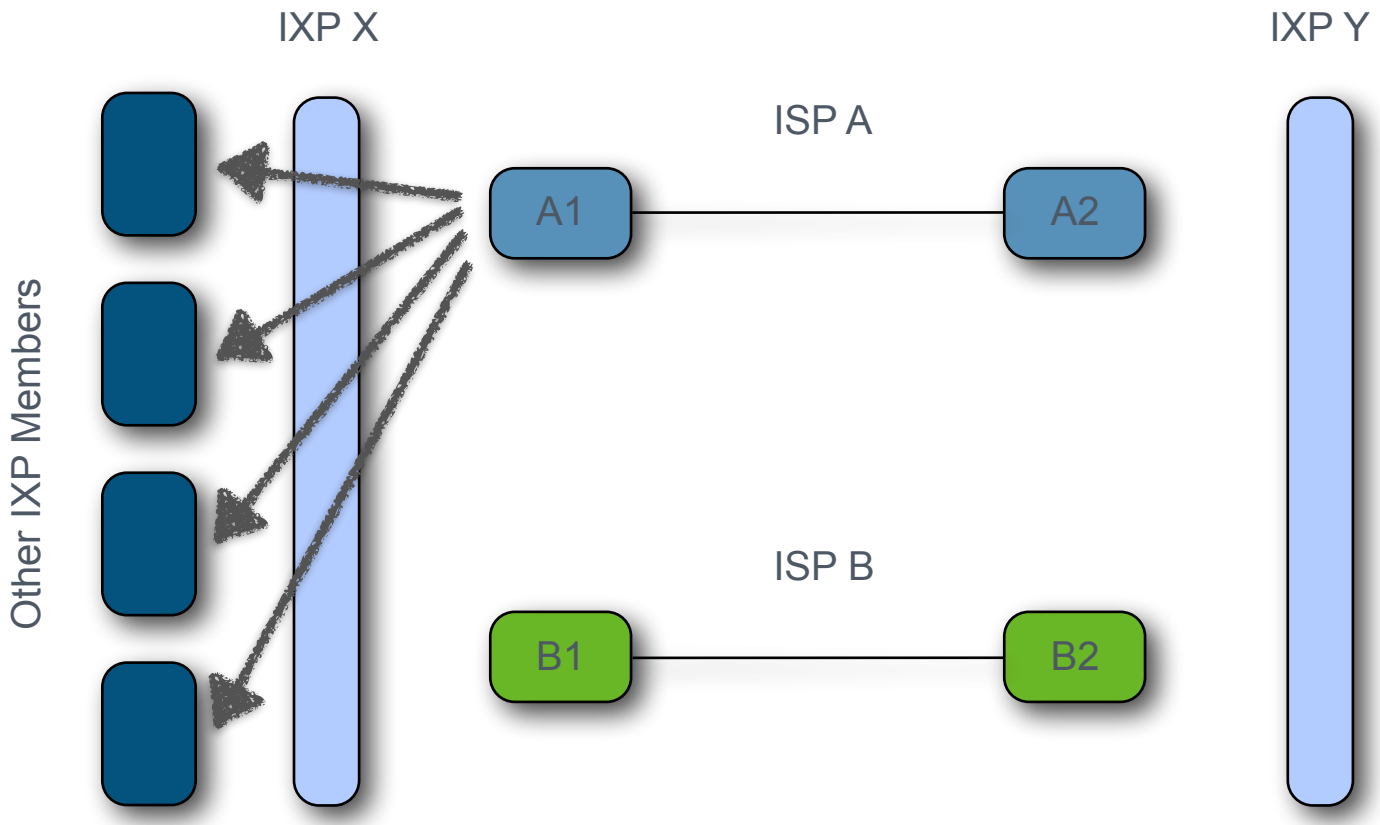
BGP Security - Check Next Hop IP Address



INEX

BGP Hygiene at IXPs

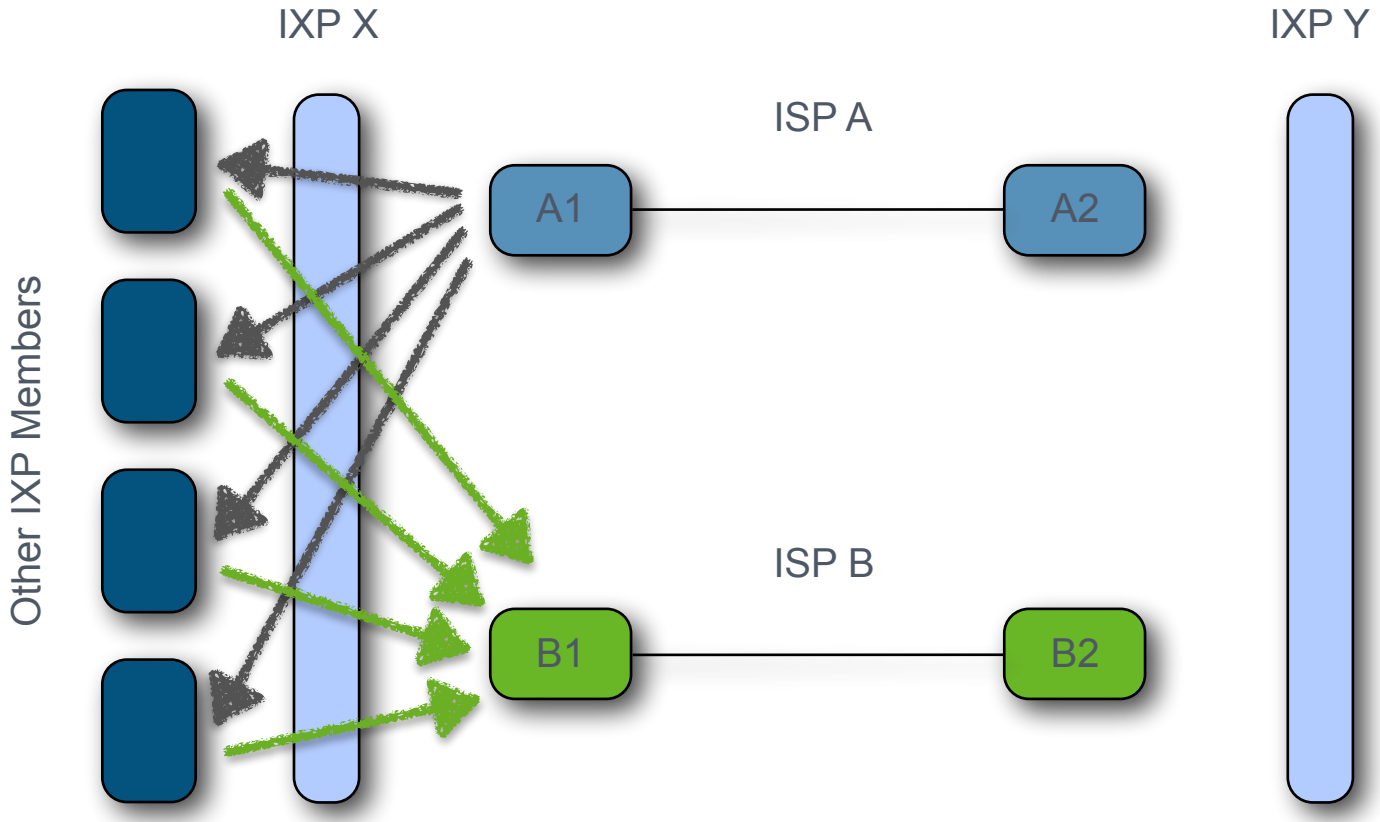
BGP Security - Check Next Hop IP Address



INEX

BGP Hygiene at IXPs

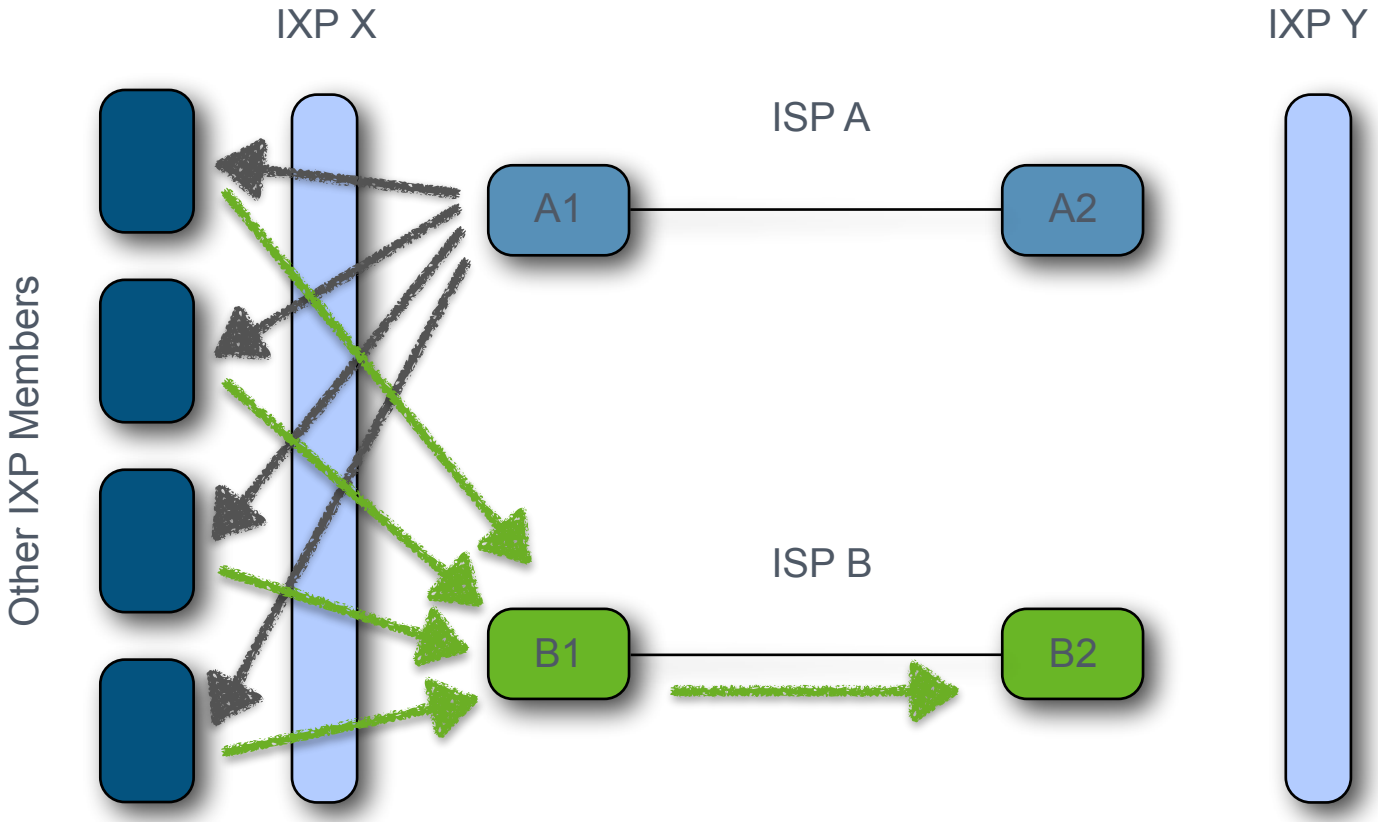
BGP Security - Check Next Hop IP Address



INEX

BGP Hygiene at IXPs

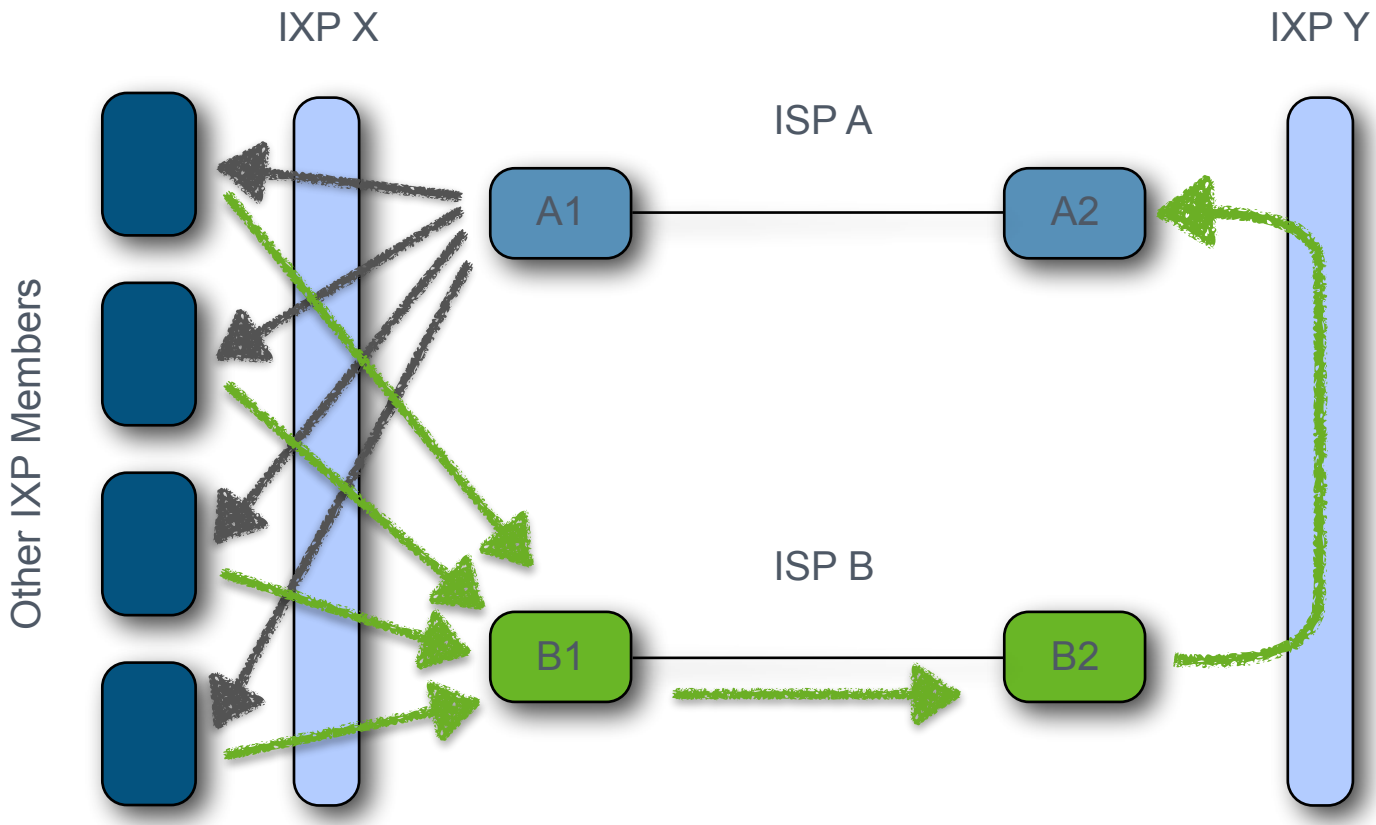
BGP Security - Check Next Hop IP Address



INEX

BGP Hygiene at IXPs

BGP Security - Check Next Hop IP Address



INEX

BGP Hygiene at IXPs

BGP Security - Check Next Hop IP Address

- Cisco Configuration

```
ip prefix-list pl-nh-peer1 seq 5 permit 193.242.111.X/32

route-map ixp-peer1-in permit 10
 match ip next-hop prefix-list pl-nh-peer1
route-map ixp-peer1-in deny 20

router bgp 64512
 address-family ipv4
  neighbor 193.242.111.X route-map ixp-peer1-in in
```

INEX

BGP Hygiene at IXPs

BGP Security - Check Next Hop IP Address

- Juniper Configuration

```
edit policy-options policy-statement import-ixp-peer1-in
set term 10-permit-nhip from next-hop X.Y.Z.W
set term 10-permit-nhip then accept
set term 99-deny then reject
top
set protocols bgp group <group> neighbor X.Y.Z.W import import-ixp-peer1-in
```

BGP Hygiene at IXPs

BGP Security - Input Validation

- Install ACLs on your IXP interface which allows only traffic destined to your own network
- Don’t use uRPF. It will break things.

BGP Hygiene at IXPs

BGP Security - IXP Prefix Redistribution

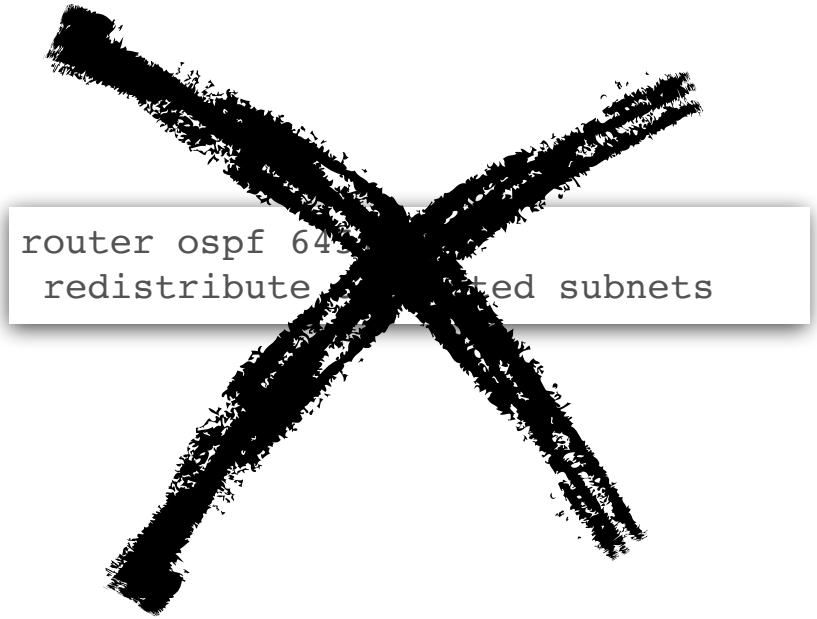
- Cisco Configuration

```
router ospf 64512
 redistribute connected subnets
```

BGP Hygiene at IXPs

BGP Security - IXP Prefix Redistribution

- Cisco Configuration



INEX

BGP Hygiene at IXPs

BGP Security - IXP Prefix Redistribution

- IXP router interfaces are used as attack targets for DDoS
- If there's a DDoS, the IXP needs to be able to stop announcing the IXP prefix
- If you redistribute the peering LAN into your OSPF / ISIS network, the IXP operator cannot control propagation of the prefix on your network
- This can cause member links / routers to break and the IXP operator cannot fix it

BGP Hygiene at IXPs

BGP Security - IXP Prefix Redistribution

- iBGP between loopbacks only
- Set next-hop-self on iBGP peerings
- Cisco: use route-maps to stop IXP LAN from being accepted, or don’t use “redistribute connected subnets”
- Juniper: use policy-statements to reject IXP peering LAN
- Peering LAN prefixes can be announced via route servers

BGP Hygiene at IXPs

BGP Security - Summary

- Do use max prefixes where possible
- Do check next-hop = peer ip address
- Do deny own prefixes
- Do use prefix lists, communities
- Don’t use as-path list or access-lists for filtering prefixes
- Change the default Cisco BGP admin distance
- Don’t redistribute your IXP’s peering LAN addresses into OSPF / ISIS
- Don’t worry too much about MD5
- Read what other Operators have written about handling BGP

INEX

BGP Hygiene at IXPs

BGP Security - Further Reading

- RFC 7454: “BGP Operations and Security” (BCP 194)
- RFC 7948: “Internet Exchange BGP Route Server Operations”
- http://blog.ipspace.net/
- http://www.team-cymru.org/
- Archived talks from: NANOG, RIPE, MENOG, APRICOT, UKNOF, NLNOG.

THANK YOU

Any Questions?