Transitioning to BGP

ISP Workshops

Last updated 28th September 2020

Acknowledgements

- This material originated from the Cisco ISP/IXP Workshop Programme developed by Philip Smith & Barry Greene
- Use of these materials is encouraged as long as the source is fully acknowledged and this notice remains in place
- Bug fixes and improvements are welcomed
  - Please email *workshop (at) bgp4all.com*

Philip Smith

Scaling the network

How to get out of carrying all prefixes in IGP

Why use BGP rather than IGP?

- IGP has Limitations:
  - The more routing information in the network
    - Periodic updates/flooding “overload”
    - Long convergence times
    - Affects the core first
  - Policy definition
    - Not easy to do

Preparing the Network

- We want to deploy BGP now…
- Because BGP will be used an ASN is required
- If not multihoming, a private ASN is sufficient
- If multihoming to different ISPs is intended in the near future, a public ASN should be obtained:
  - Either go to upstream ISP who is a registry member

    or
  - Apply to the RIR yourself for a one off assignment

    or
  - Ask an ISP who is a registry member

    or
  - Join the RIR and get your own IP address allocation too (this option strongly recommended)!

Preparing the Network

- Will look at two examples of BGP deployment:
  - Example One: network uses only static routes
  - Example Two: network is currently running an IGP

Preparing the Network
Example One

- The network is not running any BGP at the moment
  - single statically routed connection to upstream ISP
- The network is not running any IGP at all
  - Static default and routes through the network to do “routing”

Preparing the Network
First Step: IGP

- Decide on an IGP: OSPF or IS-IS ☺
  - See the ISIS vs OSPF presentation
- Assign loopback interfaces and /32 address to each router which will run the IGP
  - Loopback is used for OSPF and BGP router id anchor
  - Used for IBGP and route origination
- Deploy IGP (e.g. OSPF)
  - IGP can be deployed with NO IMPACT on the existing static routing
  - e.g. OSPF distance might be 110; static distance is 1
  - Smallest distance wins

Preparing the Network
IGP (cont)

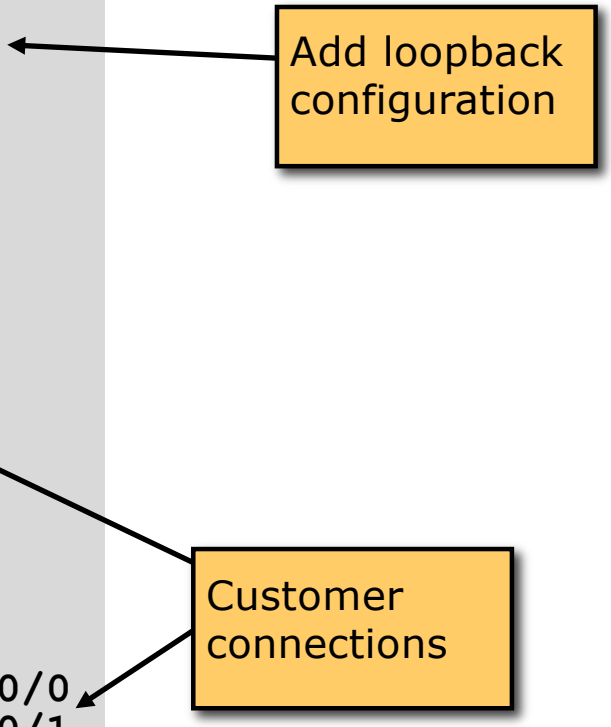
- Be prudent deploying IGP – keep the Link State Database Lean!
  - Router loopbacks go in IGP
  - WAN point to point links go in IGP
  - (In fact, any link where IGP dynamic routing will be run should go into IGP)
  - Summarise on area/level boundaries (if possible) – i.e. think about your IGP address plan

Preparing the Network
IGP (cont)

- Routes which don't go into the IGP include:
  - Dynamic assignment pools (DSL/Cable/Dial)
  - Customer point to point link addressing
    - (using next-hop-self in IBGP ensures that these do NOT need to be in IGP)
  - Static/Hosting LANs
  - Customer assigned address space
  - Anything else not listed in the previous slide

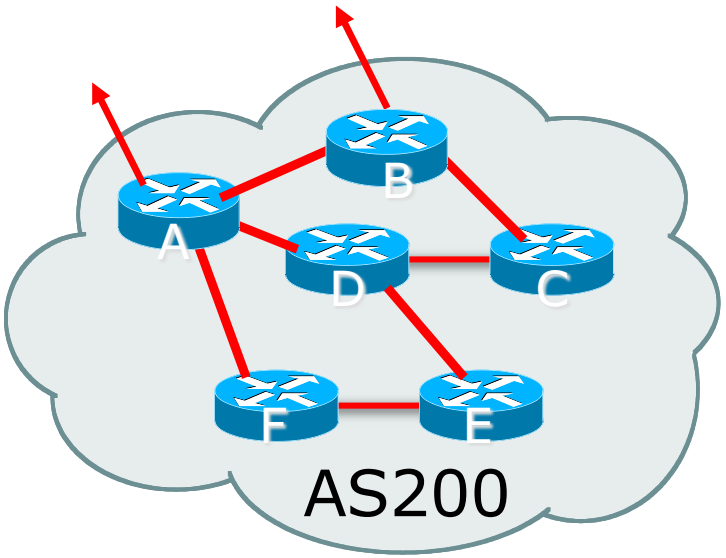
Preparing the Network
OSPF

```
interface loopback 0
 ip address 121.10.255.1 255.255.255.255
 ip ospf 100 area 0
!
interface Ethernet 0/0
 ip address 121.10.2.1 255.255.255.240
 ip ospf 100 area 0
!
interface serial 0/0
 ip address 121.10.0.1 255.255.255.252
!
interface serial 0/1
 ip address 121.10.0.5 255.255.255.252
!
router ospf 100
 passive-interface default
 no passive-interface Ethernet 0/0
!
ip route 121.10.24.0 255.255.252.0 serial 0/0
ip route 121.10.28.0 255.255.254.0 serial 0/1
```



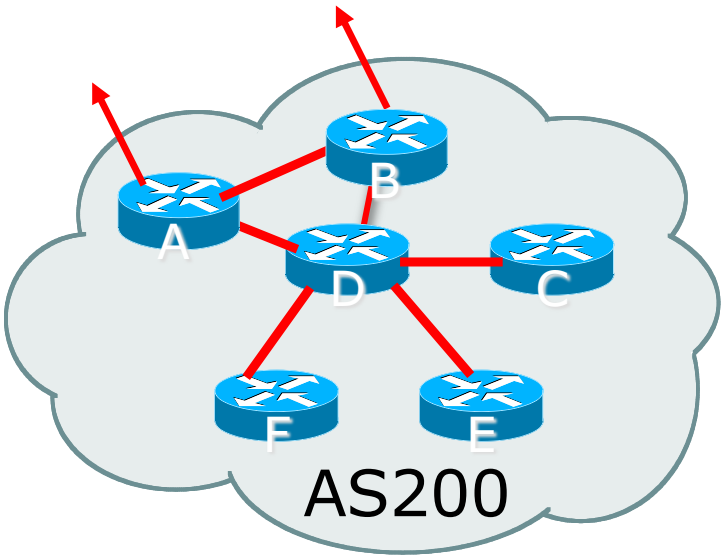
Preparing the Network
Second Step: IBGP

- Second step is to configure the local network to use IBGP
- IBGP can run on
  - all routers, or
  - a subset of routers, or
  - just on the upstream edge
- IBGP must run on all routers which are in the transit path between external connections



Preparing the Network
Second Step: IBGP (Transit Path)

- IBGP must run on all routers which are in the transit path between external connections
- Routers C, E and F are not in the transit path
  - Static routes or IGP will suffice
- Router D is in the transit path
  - Will need to be in IBGP mesh, otherwise routing loops will result



Preparing the Network Layers

- Typical SP networks have three layers:
  - Core – the backbone, usually the transit path
  - Distribution – the middle, PoP aggregation layer
  - Aggregation – the edge, the devices connecting customers

Preparing the Network Aggregation Layer

- IBGP is optional
  - Many ISPs run IBGP here, either partial routing (more common) or full routing (less common)
  - Full routing is not needed unless customers want full table
  - Partial routing is cheaper/easier, might usually consist of internal prefixes and, optionally, external prefixes to aid external load balancing
    - Communities and peer-groups make this administratively easy
- Many aggregation devices can't run IBGP
  - Static routes from distribution devices for address pools
  - IGP for best exit

Preparing the Network Distribution Layer

- Usually runs IBGP
  - Partial or full routing (as with aggregation layer)
- But does not have to run IBGP
  - IGP is then used to carry customer prefixes (does not scale)
  - IGP is used to determine nearest exit
- Networks which plan to grow large should deploy IBGP from day one
  - Migration at a later date is extra work
  - No extra overhead in deploying IBGP, indeed IGP benefits

Preparing the Network
Core Layer

- Core of network is usually the transit path
- IBGP necessary between core devices
  - Full routes or partial routes:
    - Transit ISPs carry full routes in core
    - Edge ISPs carry partial routes only
- Core layer includes AS border routers

Preparing the Network IBGP Implementation

- Decide on:
  - Best IBGP policy
    - Will it be full routes everywhere, or partial, or some mix?
  - IBGP scaling technique
    - Community policy?
    - Route-reflectors?
    - Techniques such as peer groups and peer templates?

Preparing the Network IBGP Implementation

- Then deploy IBGP:
  - Step 1: Introduce IBGP mesh on chosen routers
    - make sure that IBGP distance is greater than IGP distance (it usually is)
  - Step 2: Install “customer” prefixes into IBGP

    Check! Does the network still work?
  - Step 3: Carefully remove the static routing for the prefixes now in IGP and IBGP

    Check! Does the network still work?
  - Step 4: Deployment of EBGP follows

Preparing the Network IBGP Implementation

Install "customer" prefixes into IBGP?

- Customer assigned address space
  - Network statement/static route combination
  - Use unique community to identify customer assignments
- Customer facing point-to-point links
  - Redistribute connected through filters which only permit point-to-point link addresses to enter IBGP
  - Use a unique community to identify point-to-point link addresses (these are only required for your monitoring system)
- Dynamic assignment pools & local LANs
  - Simple network statement will do this
  - Use unique community to identify these networks

Preparing the Network
IBGP Implementation

Carefully remove static routes?

- Work on one router at a time:
  - Check that static route for a particular destination is also learned by the IBGP
  - If so, remove it
  - If not, establish why and fix the problem
  - (Remember to look in the RIB, not the FIB!)
- Then the next router, until the whole PoP is done
- Then the next PoP, and so on until the network is now dependent on the IGP and IBGP you have deployed

Preparing the Network Completion

- Previous steps are NOT flag day steps
  - Each can be carried out during different maintenance periods, for example:
  - Step One on Week One
  - Step Two on Week Two
  - Step Three on Week Three
  - And so on
  - And with proper planning will have NO customer visible impact at all

Preparing the Network Configuration – Before BGP

```
interface loopback 0
 ip address 121.10.255.1 255.255.255.255
 ip ospf 100 area 0
!
interface ethernet 0/0 ! ISP backbone
 ip address 121.10.1.1 255.255.255.240
 ip ospf 100 area 0
!
interface serial 0/0   ! Customer
 ip address 121.10.0.1 255.255.255.252
!
router ospf 100
 passive-interface default
 no passive-interface ethernet 0/0
!
ip route 121.10.24.0 255.255.252.0 serial 0/0
```

Add loopback configuration if not already there

Preparing the Network Configuration – Steps 1 & 2

```
! interface and OSPF configuration unchanged
!
router bgp 100
 redistribute connected subnets route-map point-to-point
 neighbor 121.10.1.2 remote-as 100
 neighbor 121.10.1.2 next-hop-self
 ...
 network 121.10.24.0 mask 255.255.252.0
 distance bgp 200 200 200
!
ip route 121.10.24.0 255.255.252.0 serial 0/0
!
route-map point-to-point permit 5
 match ip address 1
 set community 100:1
!
access-list 1 permit 121.10.0.0 0.0.255.255
```

Add BGP and related configuration in red

Preparing the Network
Example Two

- The network is not running any BGP at the moment
  - single statically routed connection to upstream ISP
- The network is running an IGP though
  - All internal routing information is in the IGP
  - By IGP, OSPF or ISIS is assumed

Preparing the Network
IGP

- If not already done, assign loopback interfaces (with /32 addresses) to each router which is running the IGP
  - Loopback is used for OSPF and BGP router id anchor
  - Used for IBGP and route origination
- Ensure that the loopback /32s are appearing in the IGP

Preparing the Network
IBGP

- Go through the IBGP decision process as in Example One
- Decide full or partial, and the extent of the IBGP reach in the network

Preparing the Network
IBGP Implementation

- Then deploy IBGP:
  - Step 1: Introduce IBGP mesh on chosen routers
    - make sure that IBGP distance is greater than IGP distance (it usually is)
  - Step 2: Install “customer” prefixes into IBGP

    Check! Does the network still work?
  - Step 3: Reduce BGP distance to be less than the IGP
    - (so that IBGP routes take priority)
  - Step 4: Carefully remove the “customer” prefixes from the IGP

    Check! Does the network still work?
  - Step 5: Restore BGP distance to be greater than IGP
  - Step 6: Deployment of EBGP follows

Preparing the Network IBGP Implementation

Install "customer" prefixes into IBGP?

- Customer assigned address space
  - Network statement/static route combination
  - Use unique community to identify customer assignments
- Customer facing point-to-point links
  - Redistribute connected through filters which only permit point-to-point link addresses to enter IBGP
  - Use a unique community to identify point-to-point link addresses (these are only required for your monitoring system)
- Dynamic assignment pools & local LANs
  - Simple network statement will do this
  - Use unique community to identify these networks

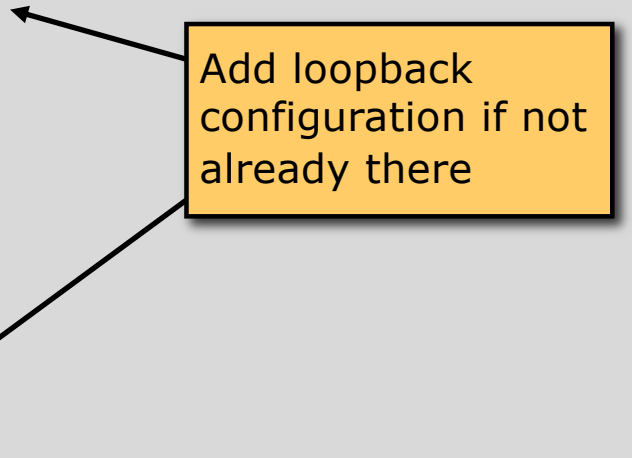
Preparing the Network
IBGP Implementation

Carefully remove "customer" routes from IGP?

- Work on one router at a time:
  - Check that IGP route for a particular destination is also learned by IBGP
  - If so, remove it from the IGP
  - If not, establish why and fix the problem
  - (Remember to look in the RIB, not the FIB!)
- Then the next router, until the whole PoP is done
- Then the next PoP, and so on until the network is now dependent on the IBGP you have deployed

Preparing the Network
Example Two Configuration – Before BGP

```
interface loopback 0
 ip address 121.10.255.1 255.255.255.255
 ip ospf 100 area 0
!
interface serial 0/0
 ip address 121.10.0.1 255.255.255.252
!
interface serial 0/1
 ip address 121.10.0.5 255.255.255.252
!
router ospf 100
 passive-interface loopback 0
 redistribute connected subnets              ! Point-to-point links
 redistribute static subnets                 ! Customer networks
!
ip route 121.10.24.0 255.255.252.0 serial 0/0
ip route 121.10.28.0 255.255.254.0 serial 0/1
```



Preparing the Network
Example Two Configuration – Steps 1 & 2

```
! interface and OSPF configuration unchanged
!
router bgp 100
 redistribute connected subnets route-map point-to-point
 neighbor 121.10.1.2 remote-as 100
 neighbor 121.10.1.2 next-hop-self
 ...
 network 121.10.24.0 mask 255.255.252.0
 network 121.10.28.0 mask 255.255.254.0
 distance bgp 200 200 200
!
ip route 121.10.24.0 255.255.252.0 serial 0/0
ip route 121.10.28.0 255.255.254.0 serial 0/1
!
route-map point-to-point permit 5
 match ip address 1
 set community 100:1
!
access-list 1 permit 121.10.0.0 0.0.255.255
```

Preparing the Network
Example Two Configuration – Steps 3 & 4

```
router ospf 100
 passive-interface default
 no passive-interface ethernet 0/0
!
router bgp 100
 redistribute connected route-map point-to-point
 neighbor 121.10.1.2 remote-as 100
 neighbor 121.10.1.2 next-hop-self
 ...
 network 121.10.24.0 mask 255.255.252.0
 network 121.10.28.0 mask 255.255.254.0
 distance bgp 20 20 20              ! reduced BGP distance
!
ip route 121.10.24.0 255.255.252.0 serial 0/0
ip route 121.10.28.0 255.255.254.0 serial 0/1
!
...etc...
```

OSPF redistribution has been removed, OSPF tidied up

Preparing the Network
Example Two Configuration – Step 5

```
router ospf 100
 passive-interface default
 no passive-interface ethernet 0/0
!
router bgp 100
 redistribute connected route-map point-to-point
 neighbor 121.10.1.2 remote-as 100
 neighbor 121.10.1.2 next-hop-self
 ...
 network 121.10.24.0 mask 255.255.252.0
 network 121.10.28.0 mask 255.255.254.0
 distance bgp 200 200 200 ! BGP distance restored
!
ip route 121.10.24.0 255.255.252.0 serial 0/0
ip route 121.10.28.0 255.255.254.0 serial 0/1
!
...etc...
```

Preparing the Network Completion

- Previous steps are NOT flag day steps
  - Each can be carried out during different maintenance periods, for example:
  - Step One on Week One
  - Step Two on Week Two
  - Step Three on Week Three
  - And so on
  - And with proper planning will have NO customer visible impact at all

Preparing the Network Configuration Summary

- IGP essential networks are in IGP
- Customer networks are now in IBGP
  - IBGP deployed over the backbone
  - Full or Partial or Upstream Edge only
- BGP distance is greater than any IGP
- Now ready to deploy EBGP

Transitioning to BGP

ISP Workshops