

Transitioning to BGP

ISP/IXP Workshops



These materials are licensed under the Creative Commons Attribution-NonCommercial 4.0 International license (<http://creativecommons.org/licenses/by-nc/4.0/>)

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

BGP Videos

- ❑ NSRC has produced a library of BGP presentations, recorded on video, for the whole community to use
 - <https://learn.nsrc.org/bgp>

The screenshot displays the NSRC (Network Startup Resource Center) website. The top navigation bar includes links for Home, About, BGP for All (highlighted), perfSONAR, ScienceDMZ, FedIdM, and Contact Us, along with a search bar. The main content area is divided into three columns. The left column features a 'BGP for All' section with a description of BGP and a 'Video Topics' list with color-coded buttons for BGP for All, perfSONAR, ScienceDMZ, and FedIdM. The middle column, titled 'Introduction to Routing', lists various topics such as Internet Routing, Routing Protocols, and IS-IS Levels. The right column, titled 'Introduction to BGP', lists topics like Introduction to Border Gateway Protocol and Transit and Peering. A large video player is embedded in the right column, showing a video titled 'BGP for All' with a play button and a 'Watch on YouTube' link. Below the video player, there are sections for 'BGP Case Studies' and 'Communities'.

NSRC Network Startup Resource Center

Home About **BGP for All** perfSONAR ScienceDMZ FedIdM Contact Us Search

BGP for All

Border Gateway Protocol (BGP) is the primary routing protocol used to transfer data and information on the Internet or autonomous systems. BGP is a Path Vector Protocol which maintains paths to different hosts, networks and gateway routers and determines the routing decision based on rules, filtering, weight and community.

Understanding the myriad options for routing can produce efficiencies for institutions and create opportunities for research and education networks to collaborate.

Video Topics

- BGP for All
- perfSONAR
- ScienceDMZ
- FedIdM

Introduction to Routing

- Internet Routing
- Routing Protocols
- Introduction to IS-IS UPDATED
- IS-IS Levels
- IS-IS Adjacencies
- Best Configuration Practices for IS-IS on Cisco IOS
- IS-IS Authentication, Default Routes and IPv6
- Introduction to OSPF
- OSPF Areas
- OSPF Adjacencies
- Best Configuration Practices for OSPF on Cisco IOS
- OSPF Authentication, Default Routes and IPv6
- Comparing OSPF and IS-IS
- Choosing between OSPF and IS-IS
- Migrating from OSPF to IS-IS
- Migration Plan
- Finalizing Migration

Introduction to BGP

- Introduction to Border Gateway Protocol
- Transit and Peering
- Autonomous Systems UPDATED
- How BGP works
- Supporting Multiple Protocols
- IBGP versus EBG
- Setting up EBG
- Setting up IBGP

BGP Case Studies

- Peering Priorities NEW
- Transit Provider Peering at an IXP NEW
- Customer Multihomed between two IXPs NEW
- Traffic Engineering for an ISP connected to two IXes NEW
- Traffic Engineering for an ISP with two interfaces on one IX LAN NEW
- Traffic Engineering and CDNs NEW

Communities

- Communities: RFC 1998 Traffic Engineering
- Communities: Simplifying Traffic Engineering
- How to Apply Communities to Originated Routes
- How to Use Communities for Service Identification

Background

- ❑ With the growth of Research & Education institutions and Enterprises seeking redundant Internet connectivity solutions, multihoming is the only scalable technical solution
- ❑ This presentation gives guidelines on how to migrate a network which has only a single upstream provider to one that is (or will be) multihomed between two (or more) upstream providers

Scaling the network



How to get out of carrying all local prefixes
in the IGP

Why use BGP rather than IGP?

- Every 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 IS-IS 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, and often is the basis of the IS-IS NET
 - Used for IBGP and route origination
- ❑ Deploy IGP (e.g. IS-IS)
 - IGP can be deployed with NO IMPACT on the existing static routing
 - e.g. IS-IS distance might be 115; static distance is 1
 - **Lowest 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

IS-IS

```
interface loopback 0
 ip address 100.64.255.1 255.255.255.255
!
interface GigabitEthernet 0/0
 ip address 100.64.2.1 255.255.255.240
 ip router isis ISP
!
interface GigabitEthernet 0/1.100
 ip address 100.64.0.1 255.255.255.252
!
interface GigabitEthernet 0/1.101
 ip address 100.64.0.5 255.255.255.252
!
router isis ISP
 net 49.0000.1000.6425.5001.00
 is-type level-2-only
 metric-style wide
 passive-interface loopback 0
!
ip route 100.64.24.0 255.255.252.0 100.64.0.2
ip route 100.64.28.0 255.255.254.0 100.64.0.6
```

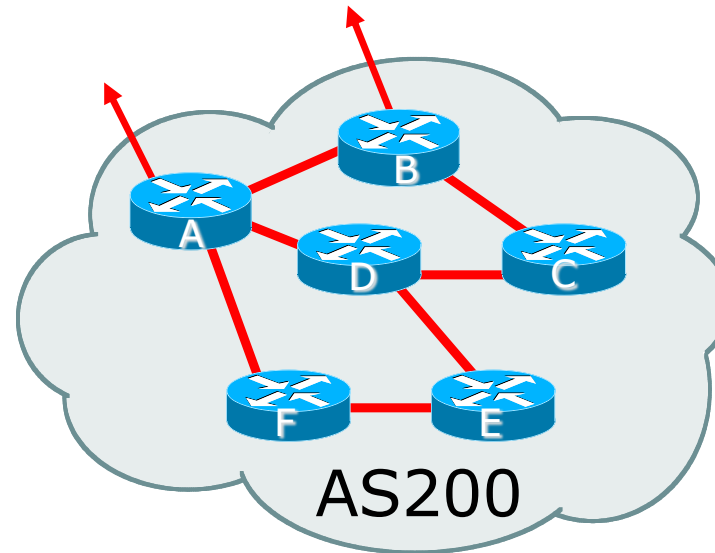
Add loopback
configuration

Customer
connections

Preparing the Network

Second Step: IBGP

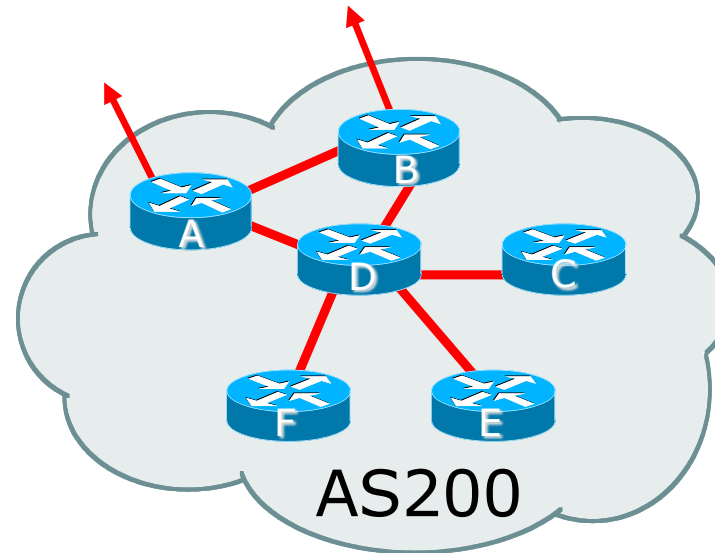
- ❑ Second step is to configure the local network to use IBGP
- ❑ IBGP can run on
 - All routers, or
 - A subset of routers, or
 - Just on the upstream edge
- ❑ IBGP must run on all routers which are in the transit path between external connections



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 100.64.255.1 255.255.255.255
!
interface GigabitEthernet 0/0
 description ISP backbone
 ip address 100.64.1.1 255.255.255.240
 ip router isis ISP
!
interface GigabitEthernet 0/1.100
 description Customer Link
 ip address 100.64.0.1 255.255.255.252
!
router isis ISP
 net 49.0000.1000.6425.5001.00
 is-type level-2-only
 metric-style wide
 passive-interface loopback 0
!
ip route 100.64.24.0 255.255.252.0 100.64.0.2
```

Add loopback configuration if not already there

Preparing the Network Configuration – Steps 1 & 2

```
! interface and IS-IS configuration unchanged
!
router bgp 100
  redistribute connected subnets route-map point-to-point
  neighbor 100.64.1.2 remote-as 100
  neighbor 100.64.1.2 next-hop-self
  ...
  network 100.64.24.0 mask 255.255.252.0
  distance bgp 200 200 200
!
ip route 100.64.24.0 255.255.252.0 100.64.0.2
!
route-map point-to-point permit 5
  match ip address prefix-list ptp-links
  set community 100:1
!
ip prefix-list ptp-links permit 100.64.0.0/16 ge 30 le 31
!
```

← Add BGP and related configuration in blue

← Allows for /30 and /31 point-to-point addressing

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 IS-IS 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
 - ❑ And often also the basis for the IS-IS NET
 - 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 100.64.255.1 255.255.255.255
!
interface GigabitEthernet 0/1.100
 ip address 100.64.0.1 255.255.255.252
!
interface GigabitEthernet 0/1.101
 ip address 100.64.0.5 255.255.255.252
!
router isis ISP
 net 49.0000.1000.6425.5001.00
 is-type level-2-only
 metric-style wide
 passive-interface loopback 0
 redistribute connected subnets      ! Point-to-point links
 redistribute static subnets         ! Customer networks
!
ip route 100.64.24.0 255.255.252.0 100.64.0.2
ip route 100.64.28.0 255.255.254.0 100.64.0.6
```

Add loopback
configuration if not
already there

Preparing the Network

Example Two Configuration – Steps 1 & 2

```
! interface and IS-IS configuration unchanged
!
router bgp 100
  redistribute connected subnets route-map point-to-point
  neighbor 100.64.1.2 remote-as 100
  neighbor 100.64.1.2 next-hop-self
  ...
  network 100.64.24.0 mask 255.255.252.0
  network 100.64.28.0 mask 255.255.254.0
  distance bgp 200 200 200
!
ip route 100.64.24.0 255.255.252.0 100.64.0.2
ip route 100.64.28.0 255.255.254.0 100.64.0.6
!
route-map point-to-point permit 5
  match ip address prefix-list ptp-links
  set community 100:1
!
ip prefix-list ptp-links permit 100.64.0.0/16 ge 30 le 31
!
```

← Add BGP and related configuration in blue

← Allows for /30 and /31 point-to-point addressing

Preparing the Network

Example Two Configuration – Steps 3 & 4

```
router isis ISP
 net 49.0000.1000.6425.5001.00
 is-type level-2-only
 metric-style wide
 passive-interface loopback 0
!
router bgp 100
 redistribute connected route-map point-to-point
 neighbor 100.64.1.2 remote-as 100
 neighbor 100.64.1.2 next-hop-self
 ...
 network 100.64.24.0 mask 255.255.252.0
 network 100.64.28.0 mask 255.255.254.0
 distance bgp 20 20 20
!
ip route 100.64.24.0 255.255.252.0 100.64.0.2
ip route 100.64.28.0 255.255.254.0 100.64.0.6
!
...etc...
```

← IS-IS redistribution
has been removed,
IS-IS tidied up

← Reduced BGP distance

Preparing the Network

Example Two Configuration – Step 5

```
router isis ISP
 net 49.0000.1000.6425.5001.00
 is-type level-2-only
 metric-style wide
 passive-interface loopback 0
!
router bgp 100
 redistribute connected route-map point-to-point
 neighbor 100.64.1.2 remote-as 100
 neighbor 100.64.1.2 next-hop-self
 ...
 network 100.64.24.0 mask 255.255.252.0
 network 100.64.28.0 mask 255.255.254.0
 distance bgp 200 200 200
!
ip route 100.64.24.0 255.255.252.0 100.64.0.2
ip route 100.64.28.0 255.255.254.0 100.64.0.6
!
...etc...
```

Restored BGP distance



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/IXP Workshops