Multihoming Strategies

for

Internet Connectivity

(For Internal Use Only)

Praveen Akkiraju
Consulting Engineering

Cisco Systems, Inc.
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

Table of Contents

READ ME

Multihoming in the Internet is a complex topic given the number of variables to be considered and their interactions. In order to provide a clear understanding of each of the variables and their interactions, a modular format has been used. Each variable such as Internal Routing, IP addressing issues, and Policy Control tools are organized as separate sections where they are explained in detail. The interactions are studied as a part of the Multihoming Case Studies. The case studies contain detailed configuration examples based on the Cisco IOS™ to provide a practical outlook to the examples.

To get the most out of this document the reader must satisfy the following prerequisites :

Cisco Systems, Inc.
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

- Good grasp of IP : Addressing Issues, Routing fundamentals
- Clear Understanding of Classless InterDomain Routing (CIDR) : Route Aggregation etc.
- Familiarity with BGP-4 : Attributes, Policy Control : Communities, Route Maps etc.
- Familiarity with OSPF : Basics, Route redistribution, Metric Types etc.
- Basic understanding of the Internet Architecture
- Understanding of topics addressed in the White Paper on Inter Domain Routing - CCO
- Familiarity with the Cisco IOS™ configuration language

This document is organized as follows:

- Module 1 talks defines Multihoming and discusses the challenges to be addressed.
- Module 2 lists out all the terms used in this document.
- Module 3 is a general discussion on the key challenges with Multihoming a network.
- Modules 4 through 7 discuss specific areas such as IP address structure, Route advertisements to and from upstream providers.
- Module 8 is a detailed discussion on the policy control mechanism popular in the Internet today.
- Module 9 and 10 are Case studies of Multihoming configurations. These discussions refer heavily from the the earlier modules in order to focus on the unique problems with each configuration.
- Module 12 discusses a few protection mechanism to safeguard the router from routing mishaps. The reader is encouraged to skim over all Modules in order to get the best out of this paper.

I call this a “Living Document”, in the sense that it will evolve over time as Multihoming strategies evolve and newer tools become available. Given the complexity of the topic, this document does not claim to cover all aspects of Multihoming, but I will gladly incorporate any relevant new information that is brought to my attention. Suggestions & Comments should be sent to spa@cisco.com.

Acknowledgments to my fellow Telco-cons’ers Dave O’Leary, Tony Bates, Robert Craig for inputs to this document via. e-m and white board conversations. Thanks to Yakov Rekhter for providing the sanity check, Adriana Vascan for combing through the configs. and providing valuable feedback and Barry Greene for the editorial changes to the doc. Also TIA to all the would be reviewers and their anticipated feedback.

Cisco Systems, Inc.
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

Module 1.0 Introduction

1.1. Definition. Simply stated, a network is said to be Multihomed when it has more than one path to the global Internet via. one or more upstream providers. Some of the common motivations and requirements of Multihomed customers in no particular order are as follows :

Reliability : Ensure connectivity to the Internet in the event of a link failure or an outage in the upstream ISP.

Proximity Routing : Traffic flows statistics for a network may indicate large traffic exchanges with destinations within a particular ISP. In order to provide better response time, the network may choose to connect to that particular ISP. This may result in the network having more than 1 paths to the Internet.

Load Sharing : Effective utilization of all existing Internet links to distribute traffic load.

Symmetric Routing : Ensure that traffic flows to and from the network enter and exit the network via. the same path to the Internet. This maybe an application requirement.

1.2 Cisco Router Specifics Since this document contains configuration examples using the Cisco router, few of the key characteristics of the Cisco IOS™ are listed below to facilitate better understanding :

a] The router prefers routes with more specific routes over aggregated routes, i.e. a route with a /32 mask is preferred over a route with a /24 mask. Standard policy in the Internet.

b] In making routing decisions, the router has a pre-configured list of routing protocol preferences known as Administrative Distances to determine priority. Smaller administrative distances are preferred, exceptions are configurable. A few distances are :

Static Routes	: 0
EBGP	: 20
EIGRP	: 90
OSPF	: 110
Int. IS-IS	:
IBGP	: 200

(Refer the Cisco Configuration Manual for a complete listing)

c] Equal Cost Route Load Balancing : For all IGP's Cisco IOS™ maintains upto 6 equal cost paths. For BGP with Cisco IOS™ 11.2 and upwards the number of equal cost paths is a configurable option using the "BGP Multipath" Feature. When Equal cost paths exist, the Cisco IOS™ load balances traffic depending on the switching type :

Cisco Systems, Inc.
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

Process Switching : Per Packet Load Balancing
Fast/Autonomous/Optimum Switching : Per Destination/Per Session Load Balancing
Distributed Switching:
Cisco Express Forwarding :

Module 2.0 Terminology

Understanding of the language is important. Here are terms used to explain the concepts behind multihomed Internet connections.

- *The "Network"* : At various places in this document reference will be made to the "network", which in the context of this document refers to the Multihomed network.

- *The CIDR notation* : In referring to IP routes the CIDR notation of prefix lengths is used rather than the cumbersome <Address> <Mask> format. Refer to [6] for details. Ex: 200.200.32.0 255.255.224.0 is represented as 200.200.32/19

- *Route/Prefix* : The terms route & prefix maybe be used interchangeably to refer to IP addresses contained within route advertisements. Ex: Route/Prefix - 200.200.32/19

- *Full Routing* : A pseudonym for Full Internet routing tables

- *Internet Service Providers (ISP)* : Commercial networks whose primary mission is to provide access to the global Internet. ISP's are of various sizes and serve different customer bases ranging from other ISP's to residential customers.

- *Network Service Provider (NSP)* : Internet Service Providers (ISP) who operate at the default free zone of the Internet. They carry mostly inter-ISP traffic which is not sourced from or destined to their network. Their customers typically are small/medium ISP's and customer networks. Examples of NSP's are iMCI, Sprintlink and UUnet among others.

- *Access Providers* : Small to Medium tier ISP's who provide Internet access to corporate networks and the SOHO segments. They are in turn connected to NSP's for access to the global Internet. The important aspect of this relationship is that in some case the Access Provider address space is allocated to them from the upstream NSP's address block.

- *Upstream Providers* : In the context of this document, "Upstream providers" refers to an ISP/NSP which provide Internet access to the Multihomed network.

- *Network Access Point (NAP)* : In the US, the Internet topology includes NAP's which are neutral meeting points for ISP's to exchange customer prefix routing information.

Cisco Systems, Inc.
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

- *Peering Arrangements* : These are essentially agreements negotiated between ISP's who want to establish a basis for information exchange at the NAPs' or via. private interconnects. It defines the rules of engagement for both parties.

- *Interior Gateway Protocols (IGP)* : While most ISP's use BGP4 to communicate amongst themselves, the routing inside an ISP's network is controlled by IGP's. In general the most popular IGP's in use are IS-IS and OSPF.

- *Border Routers* : In this document "Border routers" refers to the router which connect the Multihomed network to upstream providers.

- *Internal Routers* : Routers within a network which have no external connectivity.

- *Routing Domain* : In the context of this document, an ISP's routing domain consists of its internal addresses and the addresses belonging to customers who depend on the ISP for connectivity to the Internet.

- *Closest Exit Routing* : The routing scheme whereby routers within the Multihomed network route traffic to external destination to their closest border router using the IGP.

- *Proximity Routing* : The Multihomed network ensures that packets to an external destination are routed to the upstream provider with the shortest path to the destination.

- *Symmetric Routing* : The requirement for packets belonging to a session to exit and enter the Multihomed network via. the same path through the same upstream provider.

- *Local Routing* : Routing information from an upstream provider which includes all the routing information from an upstream provider including the provider's other customers

Module 3.0 General Discussion

3.1. Routing Decisions Cisco routers always give precedence to routes with more specific over aggregated prefixes. When making routing decisions, availability of more specific information allows for the computation of optimal routes to destinations. However, there is a trade off between the requirement for optimal decision making with specific prefix information and the resources (system and management) required to carry the more specific routes.

Default Routes only : This allows the Multihomed network to avoid maintaining detailed routing information in routers. The default is injected into the IGP and outbound traffic will drain to the closest border router that injected the defaults. BGP peering to the provider routers is not a requirement. Sub-optimal routing is a definite possibility.

Partial Routing : In this instance the providers leaks local routes along with a default from both links, this allows the Multihomed network to compute optimal routes to destinations within the provider's routing domain. Default is used to route to destinations not advertised by

Cisco Systems, Inc.
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

the provider. BGP peering with the provider is required along with configuration of policy options. Sub-optimal routing cannot be totally eliminated.

Full Routing : The providers announce full internet routes to the Multihomed network. This allows for optimal routing to all destinations. BGP peering is required and configuration complexity increases to manage and utilize the full routing information. This option will require significant resource allocations and is used in ISP networks where optimal routing is a requirement and by NSP's at higher layers of the Internet.

3.2. Resource Allocation. Depending on the level of routing information to be exchanged the Multihomed network must provision sufficient CPU, memory , bandwidth and administrative resources.

a] *CPU* : Handling large amounts of dynamic routing information implies keeping pace with routing changes in the Internet this requiring appropriate CPU power. Generally,

Cisco 2500 : To be used to connect stub networks which require minimal dynamic routing information and most cases just take a default route. Different configurations available.

Cisco 4500 : With a fast CPU & adequate memory, it is an ideal customer aggregation or border router for small to medium sized networks.

Cisco 7200 : Used as Customer aggregation router where a higher port densities are required. Also used as a Border router in Small to Medium sized ISP's and networks.

Cisco 7500 : Internet Backbone router when fully loaded with memory can handle multiple BGP peer connections with full routing. Has enough CPU to keep up with dynamic changes and efficiently switch traffic on high speed interface like 100BaseT and OC3's.

b] *Memory* : Typical BGP memory usage in the router is about 110 bytes per path, and 370 bytes per route. Depending on the number of routes in the routing table and the number of active BGP neighbors, memory configurations of the router vary. Shown below are snapshots of "sh ip bgp" from an Internet router with full routing tables :

```
Router#sh ip bgp summary
BGP table version is 78732, main routing table version 78732
46382 network entries (46699/92980 paths) using 8172292 bytes of
memory
4237 BGP path attribute entries using 626052 bytes of memory
12884 BGP route-map cache entries using 206144 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory
Doing the math is left as an exercise for the reader.
```

c] *Bandwidth* : Again depending on the number of route prefixes to be exchanged between neighbors sufficient bandwidth must be provisioned between BGP neighbors.

Cisco Systems, Inc.
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

Consider the case where about 40,000 prefixes need to be exchanged between neighbors, at 50 bytes a prefix in a packet, this amounts to 2 Mbytes worth of data. On a 64 Kbps link the initial neighbor routing exchange in the order of 2 Mbytes could take a long time.

d] *Administrative overhead* : Complex configurations imply higher administrative effort.

3.3. Important Considerations

The decision to Multihome opens up a broad range of issues to be considered, a few of these issues are listed below, with more detailed discussions in later sections.

a] *Should the network be Multihomed to one provider or multiple providers ?*

This is depended on a lot of factors including size of the upstream provider, proximity routing considerations, address allocation issues and business aspects of the relationship.

b] *IP address allocation and Prefix Advertisement policies.*

Does the network own its current address space or is it allocated by the present upstream provider. If the multi-provider option is chosen will the network request additional address space from the second provider's address block. Related to this decision is the route advertisement policy. Depending on the addressing scheme and the desired benefit from Multihoming, routing advertisement policy ranges from trivial to very complex.

c] *Routing protocol configurations and Administrative overhead.*

Depending on the network's route and traffic policies , the BGP4 routing configurations to upstream providers vary in complexity, ditto for IGP configurations within networks. More complex the configurations and polices, the higher the administrative overhead.

d] *Load Sharing between multiple upstream provider connections.*

Once the network has multiple exit points to the Internet, there is a motivation to utilize all the links. There are different means of load sharing between the multiple upstream providers and requires manipulation of the networks IGP routing. It should also be noted that the network only controls load sharing of outbound traffic. When Multihomed to a single upstream inbound load sharing can be controlled by means of explicit arrangements or clever routing configurations [4]. In the multi-provider case a measure of inbound load sharing maybe achieved by means of complex BGP configurations.

e] *Symmetric Routing Requirements.*

Cisco Systems, Inc.
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

Once the basic Multihoming begins to work, depending on the applications in use symmetric routing could become an essential requirement. Without symmetric routing, packets originated by a certain application could use one path to exit the network and return via. another path, the effect on the application being inconsistent delays in sending and receiving packets causing session time-outs. This is a particularly difficult problem to solve due to the fact that the return path of a packet through the Internet is beyond the control of the originating Autonomous System or Network.

Module 4.0 IP Addressing Allocation Issues

This is a general discussion on specific issues with different IP address allocation schemes, and related route advertisement issues. The IP addresses in the network come from various sources, this is critical because it affects the route advertisements to the upstream providers.

As a rule, a Multihomed network cannot utilize a provider connection unless the provider advertises the network's routes to the rest of the Internet. This is irrespective of whether the addresses belong to the provider block or not.

4.1. Provider Derived Address Space

Most major providers own blocks of IP addresses out of which they allocated addresses space to downstream networks. This is the simplest of all addressing schemes. Consider the Multihomed network is assigned a contiguous block of /24 addresses ranging from 200.200.16.0/24 to 200.200.31.0/24 from the provider block of 200.200.0.0/16.

a] *Aggregation of the address block at the border routers.*

```
router bgp 100
neighbor 131.108.1.1 remote-as 1
:
aggregate-address 200.200.16.0 255.255.240.0 summary-only
```

This will summarize the /24 prefixes into a /20 prefix when advertising to other BGP peers. The more specific networks will be suppressed by the "summary-only" option. The provider would then further aggregate the 200.200.16.0/20 into the 200.200.0.0/16 block before advertising it to the Internet.

b] *Proxy aggregation by the upstream provider*

This scheme is implemented when the provider seeks detailed information about the internal addressing of the Multihomed network. The Multihomed network simply advertises all the /24 addresses into the upstream provider network. The upstream provider will them aggregate the /24's along with other components of the overall /16 address block when advertising into the Internet.

4.2. Provider Independent Address Space

Cisco Systems, Inc.
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

The Multihomed network owns its address space and is outside the provider address block. The Multihomed customer will advertise an aggregate, say 210.210.16.0/20, to the upstream provider. However, this causes unique problems in the Internet routing tables when the upstream provider advertises routes to other ISP's. Since the Multihomed network address space does not fit into the providers address block, the provider will have to leak the /20 from the Multihomed network out into the Internet along with its own 200.200/16 prefix. This increases the size of Internet routing tables. Another more serious problem occurs when providers with aggressive filters may filter prefixes greater than a /19, at their network boundary routers, causing reachability problems to the /20 prefixes.

4.3. Address space belongs to the other upstream provider

This considerations here are identical to those in Section 5.2. Here the Multihomed networks address block may have been assigned to it by Provider A. When the network Multihomes to Provider B, issues detailed in the section above will have to considered.

4.4. Address Space allocated from both providers

In certain instances, the Multihomed network is allocated addresses out of both provider blocks. This is the most complex and messy case of addressing. Both the providers will have to deal with addresses which do not fit within their address block. The best the network can do is to aggregate its addresses to reduce the amount extra information to be carried by each upstream provider.

Module 5.0 Receiving Route Updates from Upstream Providers

5.1. Default Routing

A default route can be generated in two ways. Either the upstream provider supplies a default route via. BGP or a static default is configured on the Border router pointing to the upstream provider interface. The default is then redistributed into the IGP to propagate it to the rest of the internal routers in the network. It is most often used when connecting stub networks. Multihomed stub network with default routing will have multiple defaults originated by multiple border routers. In such cases, IGP metrics can be used to influence the choice of default gateway. As long as the default origination is tied to the existence of the provider path, this is a clean routing setup.

This option is the simplest in terms of configuration and resources for both the provider and the network. The system resources required are minimal and the routing is simple. Both the networks are isolated from each other routing dynamics. However, the simplicity contributes to sub-optimal routing as the network only sees a single default representing the entire Internet. This could be problem is scenarios such as the one below

BR-1/BR-2 :

Cisco Systems, Inc.
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

```
ip route 0.0.0.0 0.0.0.0 serial 1/0

router ospf 100
:
default-information originate metric 2 metric-type 1
```

In Fig. 2, IR-1, within the Multihomed network receives two defaults from BR-1 and BR-2. Based on the IGP metrics IR-1 picks the BR-1 as its exit to the Internet. Now suppose IR-1 needs to send a packet to Network 220.220.220.0 which belongs to Upstream Provider B who is directly connected to BR-2. IR-1 uses its default route and sends the packet to BR-1 and Provider A. The packet path is clearly sub-optimal.

It is important to note that in certain network situations optimal routing may not be achieved due to the complexity involved. This is a widely recognized fact in the Internet today.

5.2 Full Routing

It has the advantage of ensuring optimal routing given the detailed information available to the network. Assume both Provider A and Provider B carry full Internet routing tables, the Multihomed network may chose to receive full routing from both the providers via. BGP. It is important to understand the basic requirements to support this option :

System Resources : The Border routers need to have enough memory to hold these routes. The presence of large number of dynamic routes also implies the routers must have enough CPU to keep up with changes in the routing information from every nook & corner of the Internet. Sufficient bandwidth must also be provisioned between the Border routers and the provider routers.

Administrative Resources : Sufficiently skilled manpower is required to administer networks which carry full routing tables since it is possible that operator errors may affect the stability of the Internet. The routing at the borders also affect the internal network and situations like unplanned propagation of full Internet routing to the internal routers has the potential to cause severe network outages. Planning is critical to safe operation.

Routing Setup : If full routes are to accepted, there needs to be a clear strategy on how to use that information to optimize routing. This involves choosing the right IGP, the amount of Internet routing to be leaked to the internal network etc. Examples of optimal routing with full routes are given in following sections.

Protective Mechanisms : As mention before presence of the full routing table implies that the routers will see every ripple in the Internet, this may have adverse consequences on the stability of the Multihomed network. Mechanisms such as Route Dampening must be deployed at the borders to insulate the network from constant route flaps. This and other good security precautions are discussed in Section 11.

5.3 Partial Routing

Cisco Systems, Inc.
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

This strategy is the happy middle ground between the previous two options. The goal is to ensure optimal routing to networks belonging to directly connected providers, as it can be argued that despite carrying full routes it not always possible to ensure optimal routing to non-directly connected networks. This is the key concept.

Provider A leaks all the routes that are originated within AS 1 (including customer routes) along with a default route, ditto for Provider B. Configurations for the provider routers are :

Provider A Router :

```
ip as-path access-list 102 permit ^1$
ip as-path access-list 103 deny .*
!
route-map advt_routes permit 10
match as-path 102
route-map advt_routes permit 20
match as-path 103
!
router bgp 1
neighbor < Border Router 1 > remote-as 100
neighbor < Border Router 1 > route-map advt_routes out
neighbor < Border Router 1 > default-information originate
```

In the above configuration, the AS Path filter ensures that only routes originated by AS 1 are permitted, routes with any other AS Path are denied. This will also include the default route 0.0.0.0 generated by the provider to the multihomed network. This policy is then applied to all outbound routing updates sent to the AS 100 peer. A similar configuration can be implemented for B. The default-information originate selectively generates a default to Border Router A.

BR-1 :

```
ip as-path access-list 102 permit ^1$
ip as-path access-list 103 deny .*

route-map rcvd_routes permit 10
match as-path 102
route-map rcvd_routes permit 20
match as-path 103

router bgp 100
neighbor < Provider Router A > remote-as 1
neighbor < Provider Router A > route-map rcvd_routes in
:
!
router ospf 100
:
redistribute bgp 100 distribute-list 100 metric 2 metric-type 1
default-information originate metric 2 metric-type 1
```

The route map applied to incoming updates from the AS 1 peer is double protection in case the route-map on the provider side malfunctions. BGP routes can be redistributed into OSPF

Cisco Systems, Inc.
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

in a controlled manner using the distribute list. The list can be configured to permit all the BGP routes, only the default or a any other combination of routes.

The advantages of this approach are obvious. Optimal routing to directly connected provider networks is ensured. Defaults from both borders provides access to the rest of the Internet and guarantees full connectivity in case of link failures to either provider. System resources are less of a problem, though administrative resources are still required.

The disadvantage is the complexity and coordination required to implement and manage this configuration. Like in the full routing option, routing strategy within the network must be planned before implementation to ensure smooth packet flow. Some of the precautions such as incoming route-maps and route dampening must also be implemented

Ultimately when deciding on an incoming routing strategy the needs of the network must be balanced with the requirements of each option. In some cases the strategy may have to be modified as more information on traffic flows becomes available as operations mature. Also note that it is possible to use different inbound routing strategies to different upstream providers, these options will be discussed in more detailed in Section 9 and 10.

Module 6.0 Internal Routing Strategies

The primary goal of internal routing is to get the packets within the network to the borders. The border routers connect to the upstream provider routers using BGP or static configurations, they serve as gateways to the Internet for the rest of the network. Internal routing maybe qualified with specific routing policies to control outbound traffic flow, ex. Load Sharing. Discussed below are three Internal routing strategies.

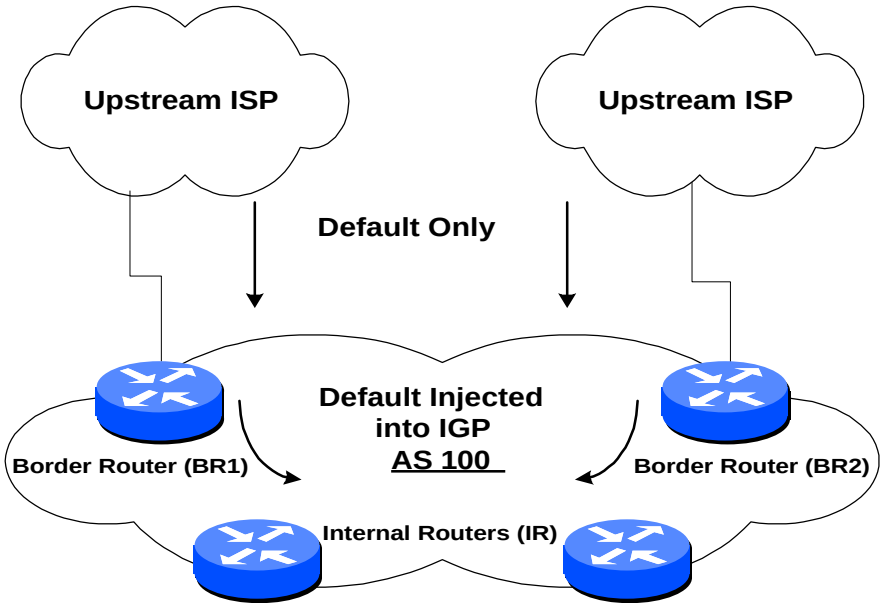
6.1. IGP routing with Default injection at the borders

This option is ideal for small to medium size networks which have an internal routing protocol such as OSPF or EIGRP already running to provide connectivity within the network. Here both Border routers are identical from an internal routing perspective, this is because both inject identical default routes. The relevant topology is shown in Fig. 1.

The border routers participate in external routing to the upstream providers as well as internal routing to the rest of the network. A default can be injected into the IGP running within the network at the border routers. IGP routing updates propagate the Default route to the rest of the internal network with the Border Routers as the source. In instances where there are 2 border routers, an internal router may receive both the Default routes, the choice is then made based on IGP metrics, Ex: Link Costs in OSPF

Cisco Systems, Inc.
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

Figure 1



It should be noted that left as is, this setup automatically provides basic Load Sharing between the two borders. Also known as Closest Exit routing, with default injection being identical, internal routers pick the closest border router based on IGP routing metrics of the advertised Default route. More controlled traffic flow can be achieved by manipulating the IGP metrics of the Default routes. OSPF is used as the IGP throughout.

BR-1 :

```
!# Injection of Default into OSPF
ip route 0.0.0.0 0.0.0.0 serial 0
!
router ospf 100
:
default-information originate metric 2 metric-type 1
:
```

The static default route is tied to the serial line connecting the upstream provider. Hence when there is a link failure to the provider, BR-1 stops originating the default. The “default-information-originate” command under OSPF will ensure OSPF propagates a default route to its neighbors only if it detects a route to 0.0.0.0 in the routing table.

6.2 IGP routing with redistribution of external routes at the borders

The border router establishes BGP peering with the provider routers and participates in dynamic routing. This strategy is ideal in instances where there is need for Proximity routing, i.e. traffic to certain destinations need to be routed to a specific provider router. Closest exit routing is no longer acceptable since it may result in sub-optimal routing. Here BGP handles external routing and the IGP is responsible for all internal routing.

Cisco Systems, Inc.
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

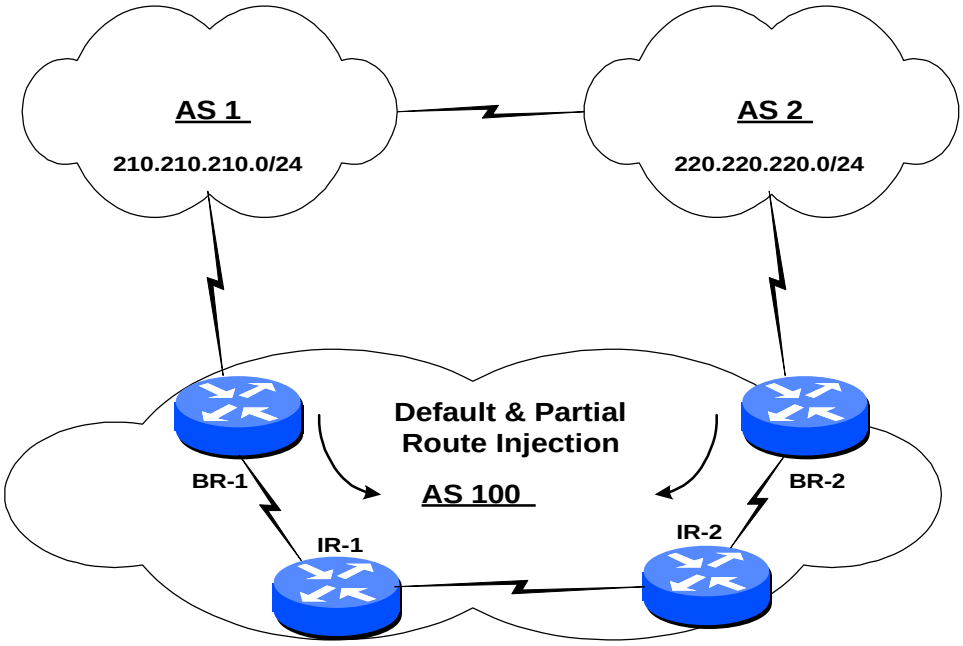


Figure 2

Consider a situation (Fig.2) where the network exchanges a lot of traffic to networks 210.210.210/24 (customer of AS1) and 220.220.220.0/24 (customer of AS2). In the interest of optimal routing all outbound traffic to 210.210.210/24 should be routed to Border A and to Border B for 220.220.220/24. The simplest way to achieve this is to do the following :

- At Border A : Inject 210.210.210/24 and a generic Default route
- At Border B : Inject 220.220.220/24 and a generic Default route

The above configuration will ensure Proximity routing by routing traffic as follows :

Traffic to 210.210.210/24 from anywhere within the network will always be routed to Border A since it advertises the more specific route. More importantly, if connectivity to AS1 is lost, the Default route from Border B will ensure that the traffic to 210.210.210/24 is routed to its destination, albeit using a longer route.

BR-1 :

```
ip route 0.0.0.0 0.0.0.0 Serial 0/0
!
router ospf 100
:
redistribute bgp 10 metric 2 distribute-list 1 in
default-information originate metric 2 metric-type 1
:
!
access-list 1 permit 210.210.210.0 0.0.0.255
access-list 1 deny any any
!
```

Cisco Systems, Inc.
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

The default route origination is identical to that described in Section 4.1. The routing information from BGP is redistributed into OSPF via. an access-list that permits only the /24's (210.210.210.0 in case of BR-1) of interest while denying entry to all other routes.

In such network scenarios the two border routers may run IBGP among themselves. This configuration is useful in situations where complex routing policy needs to be implemented at the borders. One critical requirement of this type of configuration is that the IBGP neighbors may need to be directly connected in order to satisfy BGP synchronization requirements. The alternative to direct connectivity is to redistribute all the BGP routes into the IGP at the borders. More detailed discussion of BGP Synchronization can be found in [1].

NOTE : Be very careful with the distribute-list protecting OSPF from BGP routes. Accidental mis-configuration of the distribute-list could result in BGP redistributing a large number of routes into OSPF. This could result in a catastrophic OSPF failure.

6.3 IBGP as the IGP (a.k.a Pervasive BGP)

This a complex strategy used only in medium to large size ISP's. Such networks are connected to multiple ISP's, receiving significant amounts of routing information from each, with complex internal routing policy requirements. Such a topology is shown Fig.3

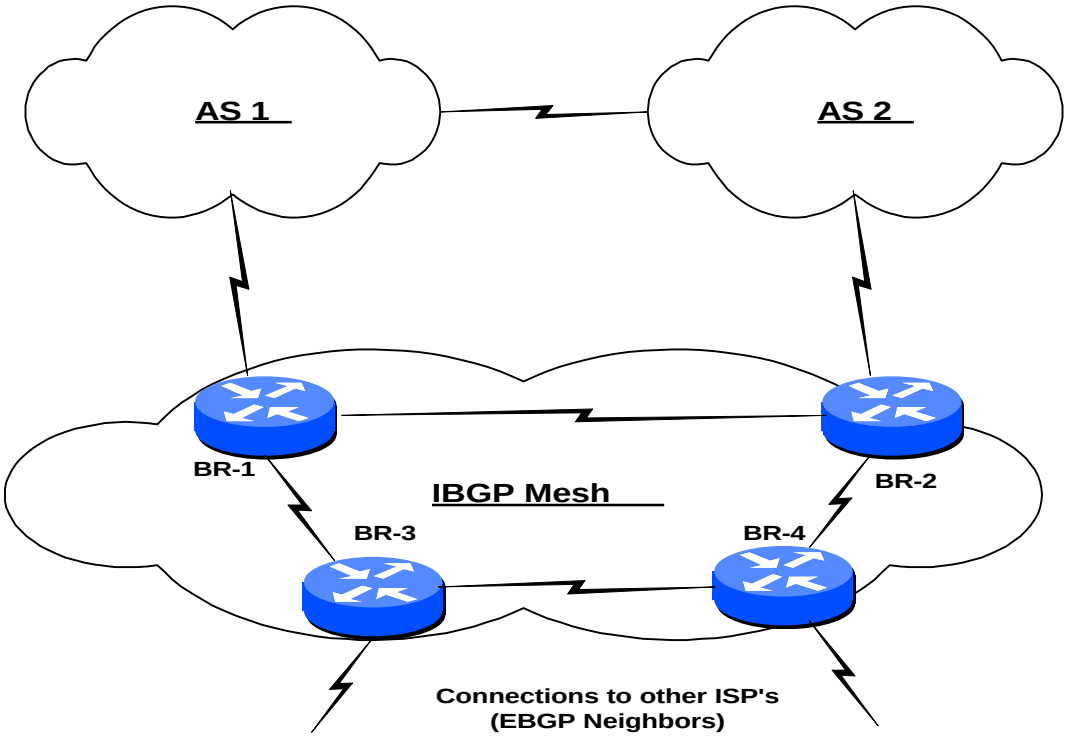


Figure 3

B3R-1 :

```
router bgp 100
neighbor < Router X > remote-as 1
neighbor < Router X > remote-as 1 route-map external_policy out
neighbor internal peer-group
```


Cisco Systems, Inc.
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

```
neighbor internal remote-as 100
neighbor internal route-map internal_policy out
neighbor < Router B > peer-group internal
neighbor < Router C > peer-group internal
```

Here, Router A is EBGP peered with Router X in AS 1, and is IBGP peered with all the other BGP speakers in AS 100 via. a peer-group. Since every router in the AS is a BGP speaker, BGP can be used as the primary routing protocol within the AS. Note, the peer-group is a router performance optimizer and a configuration simplifier. Refer [2] for details.

IBGP by virtue of design offers a variety of policy control tools which can be utilized in this scenario. Full IBGP peering mesh is required among all the IBGP routers. Pervasive BGP if used is most suitable for medium scale ISP's which provide connectivity to other AS's and have an IBGP mesh which is manageable. This option is rarely used in practice.

6.4. Miscellaneous Notes

In most where BGP peering to upstream providers takes place, the Border routers are IBGP peers of each other. This is a required configuration and improves routing decisions.

In networks with large IBGP meshes, scalability of the mesh, and the abilities of the IGP to handle large amounts of routing information are key design considerations in designing Internal routing. BGP Route Reflectors ([2],[5]) is the most commonly used strategy to scale large IBGP meshes. The IGP of choice for NSP's handling full Internet routing tables at the Default Free zone is IS-IS, and by far the most popular IGP is OSPF used by ISP's of all sizes. For more detailed discussions on scaling IBGP Meshes refer to [1].

Module 7.0 Advertising Route Updates to Upstream Providers

Route advertisements to upstream providers determines how the traffic flows back into the network. Generally, the more detailed information upstream providers have the better they are able to redirect traffic to enter the network at the closet entry point.

To revise the contents of Section 5, the origins of the Multihomed network's address space are of interest. If the address space is derived from a providers address block, they can be combined with the rest of the block when the provider announced its routing updates to the Internet. This also enables provider to do proxy aggregation. In cases where the Multihomed network's address block is outside the provider's block, the provider has to carry the network's address space separate from the rest of his block. For example, if the network's addresses space 200.200.16.0/24 to 200.200.31.0/24 (16 Class C's) is assigned out of Provider A's block. Provider B's route advertisements consist of two components :

- The address block assigned to it 210.210.0.0/16
- The Multihomed network's address space 200.200.16.0/20

Cisco Systems, Inc.
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

7.1 Aggregate Only

Multihomed network will advertise the aggregate of its address block to both upstream providers. This will work well for the provider who owns the block since the customer advertise can be further aggregated when advertising to the rest of the Internet. However, provider B will not be able to aggregate the customers block and will be required to carry the more specific route within his network as discussed above. Aggregation of the network address space is configured at the Border routers and shown below :

Border Router :

```
router bgp 100
neighbor 131.108.10.11 remote-as 1
:
aggregate-address 200.200.16.0 255.255.240.0 summary-only
```

Note: For the aggregate address to enter the BGP table, there must be at least one more specific route in the BGP table. Refer [1] for more details.

In certain instances the provider will do the aggregation for the network, also known as proxy aggregation. When aggregating the upstream provider must ensure that all the components of the aggregate are contained within its administrative domain. This situation calls for specialized configurations to ensure connectivity to all the components of the /16 inside and outside the provider AS. Specifically, the provider may have to leak the “holes” of the non-provider /24’s back into its AS and to all its customers to ensure reachability. See detailed discussion in Section 11.

7.2 No Aggregation

As stated above, advertising more specific information to the upstream provider could result in better routing decisions on inbound traffic through the provider into the network. This is applicable in a scenario where the network Multihomes to a single provider at different locations and when multihoming to multiple providers as well when the enterprise has provider independent address space. Consider the scenario in Fig. 4 below :

Cisco Systems, Inc.
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

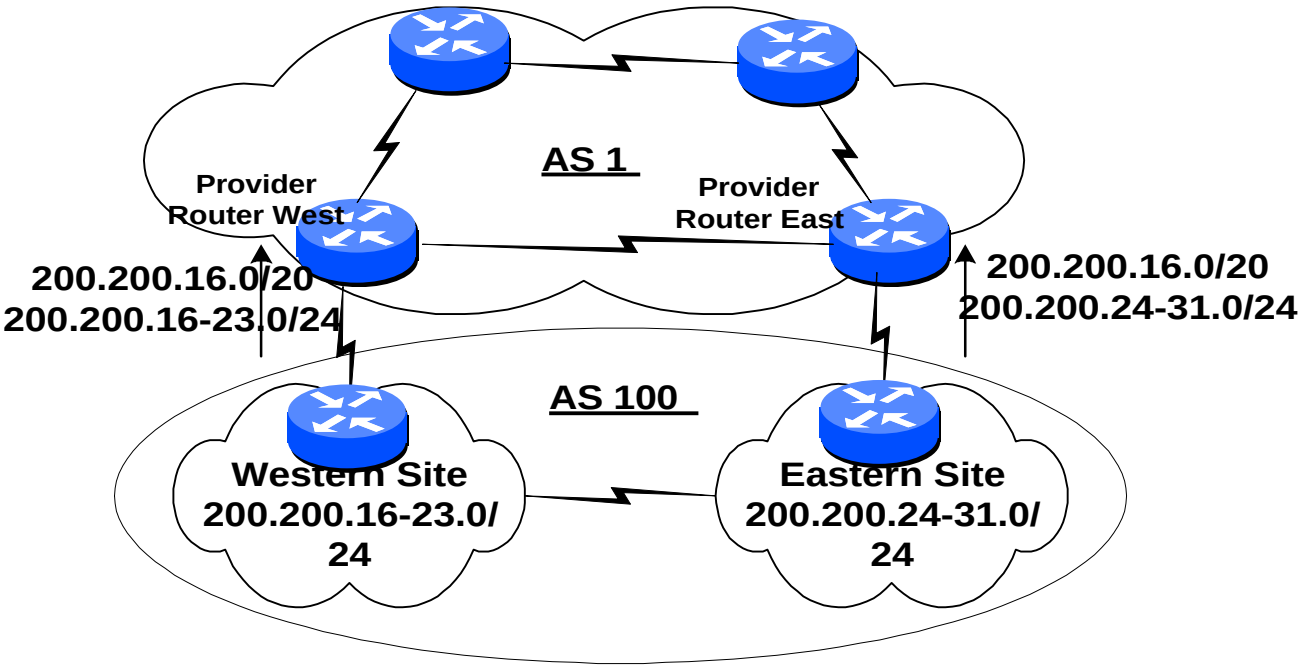


Figure 4

Here's the network Multihomes to a single provider at two geographically distant locations. Assume the Western portion of the network has been assigned the address space 200.200.16-23/24 and the Eastern region is assigned 200.200.24-31/24.

If an aggregate 200.200.16/20 which represents all the 16 /24's were to be advertised into the provider network, there is no way for provider routers to distinguish between a Western address and an Eastern address within the Multihomed network. Sub-optimal routing could occur as provider routers hand off packets to the nearest Border router, i.e. packets to the Western site maybe handed off by the provider to Eastern Border router. Now, consider the case where the Border-W router was able to leak 200.200.16-23/24 and an aggregate 200.200.16/20 and Border-E leaks the aggregate and 200.200.24-31/24. The provider routers can make an informed decision on where to hand off traffic since they have more information about the Multihomed network. Configurations look like

Western Border Router :

```
router bgp 100
neighbor < Provider Router-West > remote-as 1
:
aggregate-address 200.200.16.0 255.255.240.0 suppress-map advt_routes

access-list 101 permit 200.200.24.0 0.0.0.255 255.255.255.0 0.0.0.0
:                              :
access-list 101 permit 200.200.31.0 0.0.0.255 255.255.255.0 0.0.0.0

route-map advt_routes permit 10
match ip address 101
```

Cisco Systems, Inc.
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

The suppress map is used to selectively leak the West coast part of the aggregate. In the above example all the /24's from 200.200.16.0 to 200.200.23.0 are leaked to AS 1, while the /24's from 200.200.24.0 to 200.200.31.0 are suppressed. Because AS 1 now has specific information for these 8 /24's coming in via. the Western Border router, it will prefer the Western router to send packet to the 8 /24's. Ditto for the Eastern border where the specifics for 200.200.24.0 to 200.200.200.31 are leaked.

However, things become much more complicated in a Mutli-provider scenario due to the requirement for inter-provider coordination, making it a less attractive option.

Module 8.0. Traffic Flow Control Tools

Common motivations for network operators to apply routing policy within their network is to optimize traffic flow for a given network design (more traffic on high bandwidth pipes, closet exit routing etc.), to reflect economic realities (specific agreements with peers or customers, making use of all available paths etc.), improve operating efficiencies.

Traffic in this context can be classified into two distinct flows :

Outbound Traffic : Defined as the traffic originating within the Multihomed network destined to networks external to the networks AS. The flow of outbound traffic is under the control of the Multihomed network. The required policy, known here as Internal policy can be enforced by manipulating internal routing protocols (Section 4.0).

Inbound Traffic : This is defined as the traffic originating outside the Mutlihomed network destined to networks within the network's AS. By default, the network has no control over how this traffic is handed to it by the upstream provider. However, using some of the techniques described below, a measure of control can be exercised over this traffic. In some case the Multihomed network may negotiate with the provider to influence the path of Inbound traffic.

Typically, when there exist multiple routes to reach a destination, the network operator has the option of influencing the flow of traffic by manipulating the attributes of the routes. In an IGP context this could mean modifying the metric or cost and in a BGP context tweaking the BGP attribute values to influence route selection. In this section a few of the popular traffic flow control schemes are discussed.

8.1. IGP Metric Manipulation

Within a network, routers may not be BGP speakers and depend primarily on their IGP. In making routing decisions IGP's depend on metric computations. These metrics range from simple interface costs for OSPF to composite metrics as in the case of EIGRP. Hence, when presented with multiple IGP paths to a destination, the best path decision can be controlled by manipulating the metrics of each path. Detailed below are a few configuration examples with OSPF as the IGP

Cisco Systems, Inc.
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

At Redistribution. Consider the redistribution of BGP routes to network 131.108.0.0 into OSPF at two border routers. The configurations are as follows :

Border A	Border B
router ospf 100	router ospf 100
:	:
redistribute bgp 1 metric 2	redistribute bgp 1 metric 4
:	:

Hence, a router within the OSPF network will always chose Border A as the path to get to network 131.108.0.0. It is important to note that by default all redistributed routes are marked as External Type 2 routes, whereby only the external distances are of significance in making routing decisions. If the network operator wants each router within the network to make its routing decision to reflect internal costs to its closest border router ..

Border A	Border B
router ospf 100	router ospf 100
:	:
redistribute bgp 1 metric 2 metric-type 1	redistribute bgp 1 metric 2 metric-type 1

Here, by definition External type 1 routes metric are the sum of the external cost as well as internal interface costs along the path. Hence, the internal topology is reflected in the metric and routers can make closet exit decisions. Refer [7].

The same techniques can be applied to static defaults redistributed into OSPF for destinations which rely on the default route.

8.2. Traffic flow control using BGP

One of the key advantages of BGP is the policy control flexibility it offers to the network operators. BGP's route selection depends on a decision making algorithm which takes into account a combination of attribute values. For reference the algorithm is as follows :

BGP Decision Making Algorithm :

- Is the NEXT HOP reachable
- Is the NEXT HOP synchronized
- Prefer path with the largest weight
- Prefer the route with the highest LOCAL PREFERENCE
- Prefer path originated by BGP process on this router
- Prefer the route with the shortest AS PATH
- Prefer path with lowest origin
- Prefer the route with a smallest MED

Cisco Systems, Inc.
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

- Prefer the external path over the internal path
- Prefer the route with the best IGP metric
- Prefer the path with the lowest BGP router ID

Hence by manipulating any of the attributes mentioned above the BGP route selection process maybe influenced to control traffic flows. Providers when using BGP policy tools have 2 levels of granularity : AS based polices or Prefix based policies. Prefix based polices give providers better control and are more desirable they are complicated and unmanageable due to the number of prefixes in the routing tables. AS Path based polices get applied to an aggregate all prefixes belonging to an AS, but due to customized configurations required for each AS, scalability is a problem. Use of AS based or Prefix based policy depends on the problem being solved, there is no right or wrong method.

8.2.1. Using LOCAL PREFERENCE

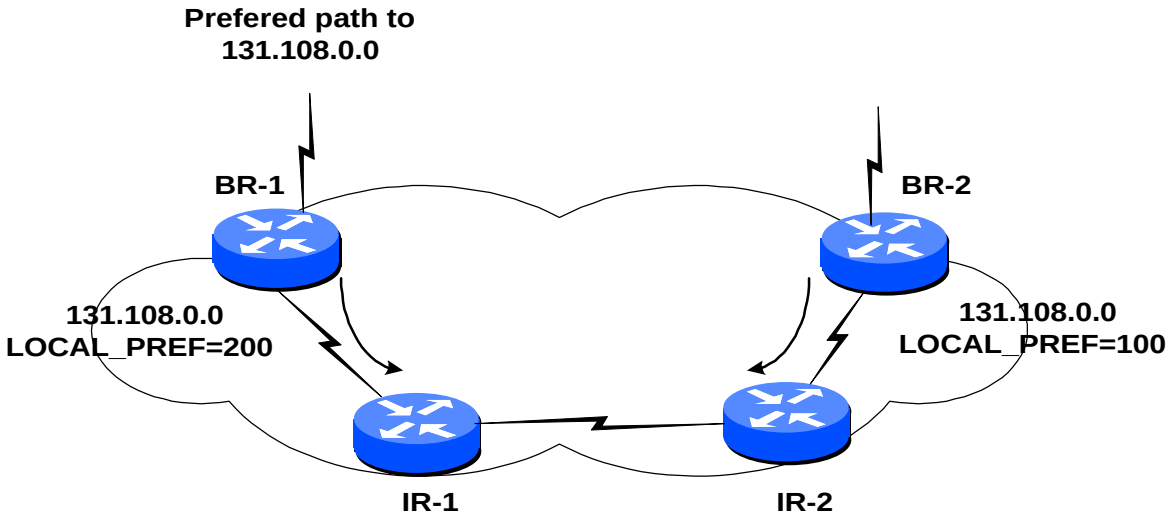


Figure 5

LOCAL_PREF allows internal peers to implement outbound traffic flow polices. Consider the network in Fig. 5, assume that all the routers in the AS are IBGP peers. The IR routers may receive BGP advertisements for the same destination from both BR-1 & BR-2. If the stated policy is to always route all traffic destined to network 131.108.0.0 via. BR-1, the configurations would be

BR-1 :

```
router bgp 1
:
neighbor internal peer-group
neighbor internal remote-as 1
neighbor internal route-map internal_policy out
neighbor < Router BR-2 > peer-group internal
neighbor < Router IR-1> peer-group internal
neighbor < Router IR-2> peer-group internal
:

access-list 10 permit 131.108.0.0 0.0.0.0
access-list 10 deny any any
```

Cisco Systems, Inc.
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

```
route-map internal_policy permit 10
match ip address 10
set local_preference 200
```

In the above example all the Internal BGP peers are part of a Peer Group called Internal since Router A has identical outbound policies for all of them which is to assign a LOCAL_PREF of 200 to the route 131.108.0.0. If Router B is allowed to originate the same route with a default LOCAL_PREF of 100, Router C will always pick Router A as the preferred path to destination 131.108.0.0 as per the algorithm.

8.2.2. Using MED's

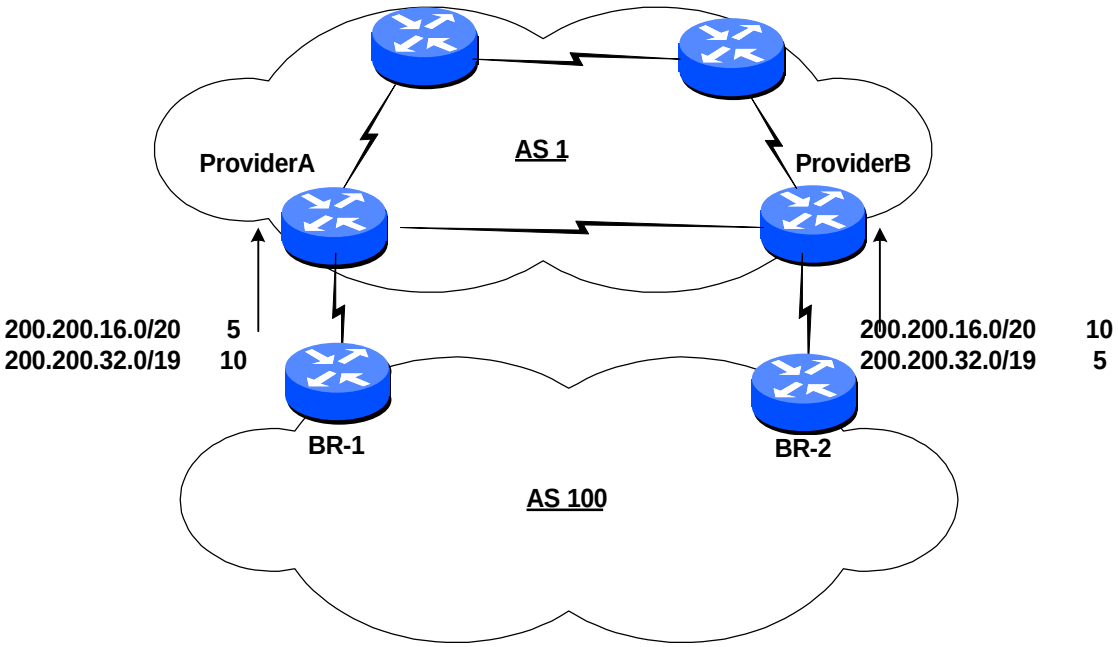


Figure 6

With MED's routers in 1 AS can communicate their inbound traffic policy to router in another AS. It is important to note that MED's are used to communicate policy to directly connected AS's, they cannot be transmitted beyond 1 AS hop. Consider Fig. 6, Routers BR-1 & BR-2 belong to AS 100 which in this example is Multihomed to AS 1 at two points. MED's set by the BR's will not be communicated to other external peers of AS 1.

Using IGP metric manipulation, AS 100 may prefer to route traffic from network 200.200.16.0/20 via. BR-1 to Router A and traffic to 200.200.32.0/19 via. BR-2 to Router B. In order to achieve Symmetric routing, AS 100 can use MED's to inform AS 1 how it would like to receive traffic destined for the 2 sets of routes. The configuration is shown below :

BR-1 :

```
router bgp 100
neighbor < Router A > remote-as 1
neighbor < Router A > remote-as 1 route-map incoming_policy out
:
aggregate-address 200.200.16.0 255.255.240.0 summary-only
```

```
aggregate-address 200.200.32.0 255.255.224.0 summary-only

access-list 101 permit 200.200.16.0 0.0.15.255 255.255.240.0 0.0.0.0
access-list 101 deny any any

access-list 102 permit 200.200.32.0 0.0.31.255 255.255.224.0 0.0.0.0
access-list 102 deny any any

route-map incoming_policy permit 10
match ip address 101
set metric 5
route-map incoming_policy permit 20
match ip address 102
set metric 10
```

at Router Y, the exact same configuration would be implemented with the exception that the route 200.200.16.0/20 would be assigned a MED of 10 and the /19 a MED of 5.

This strategy is ideal for a network which has East & West coast operations and is connected to the same upstream provider at both locations. Also the network may not have East-West connectivity and depends on the providers links for traffic flows. Using MED's informs the provider to route all the East coast bound traffic to the East coast border and vice-versa.

8.2.3 Symmetric Routing and AS PATHs

Consider the scenario in the Fig. 7. This represents one or many situations where Symmetric routing becomes a challenging problem to solve. AS 100 is Multihomed to ISP1 and ISP2, both ISP's are in turn connected to ISP3. The problem is as follows :

All traffic from AS 100 network 200.200.32/19 destined to ISP3 is routed via. ISP2, the challenge is to have all the traffic from ISP3 destined to the /19 in AS 100 come back via. ISP2. AS 100 has no way of controlling ISP3's decision making. There needs to be some way to communicate to ISP3 that the optimum way to route traffic to AS 100 is via. ISP 2. This is a major problem since ISP 3 may have its own set of traffic flow policies. It would be fair to state that total Symmetric routing is impossible to achieve. However one currently used strategy to achieve a measure of symmetric routing is using AS Path Prepend in Cisco routers.

Cisco Systems, Inc.
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

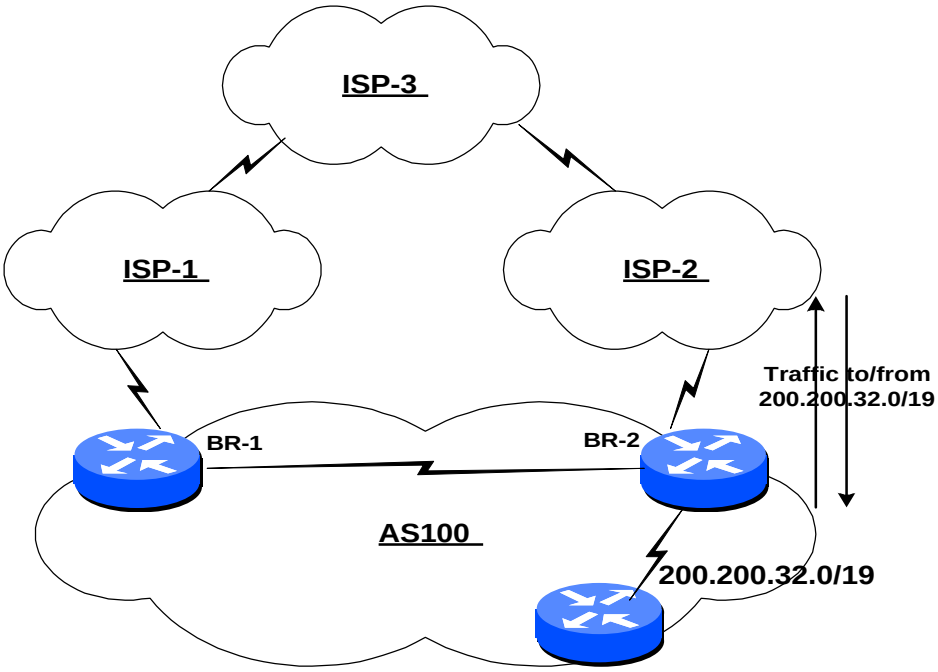


Figure 7

Router BR-2 :

```
router bgp 100
neighbor < ISP 2 Router > remote-as 1 route-map aspath out
neighbor < Router BR-1 > remote-as 100
:
aggregate-address 200.200.16.0 255.255.240.0 summary-only
aggregate-address 200.200.32.0 255.255.224.0 summary-only

access-list 101 permit 200.200.16.0 0.0.15.255 255.255.240.0 0.0.0.0
access-list 101 deny any any

access-list 12 permit 200.200.32.0 0.0.31.255 255.255.224.0 0.0.0.0
access-list 12 deny any any

route-map aspath permit 20
match ip address 102
set as-path prepend 100
```

From ISP3's perspective it now has 2 BGP advertisements to 200.200.32.0/19 :

	AS Path	
From ISP 1 :	1 100 100	ISP1 advertises the same route with a longer
From ISP 2 :	2 100	AS Path than ISP2.

Since BGP prefers shorter AS Paths, ISP will pick the route advertised via. ISP2.

This approach however has the disadvantage of propagating superfluous information in the Internet. It is also important to note that the above case is simplistic . For this scheme to work, the number of prepended AS's must be proportional to the number of connections. Symmetric

Cisco Systems, Inc.
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

routing to directly connected provider networks is achievable with creative configurations and co-ordination with the provider.

8.2.4 Use of Communities with LOCAL PREFERENCE

While Communities themselves do not alter the BGP decision making process, they can be used as flags to mark a set of routes. Upstream provider routers may then use these flags to apply specific routing polices within their network. Communities therefore are a route aggregation tool for policy purposes. Reference [4] illustrates an elegant method for Multihomed customers to use Communities and influence the setting of Local Preference for customer routes within Upstream Provider networks. In general Community based policy control provides prefix level granularity rather than AS based policy.

To summarize this concept, providers establish a mapping between customer configurable Community values and the corresponding LOCAL PREFERENCE values within the providers network. The idea is that customers with specific policies that require the modification of LOCAL_PREF. in the provider network will set the corresponding Community values in their routing updates.

From RFC 1998 :

Example of Community to LOCAL_PREF mappings are :

Category	LOCAL_PREF	Community Values
Customer Routes	100	None
Customer Backup Routes	90	3561:90 (0x0DE9005A)
Other ISP Routes	80	3561:80 (0x0DE90050)
Customer Provided Backup	70	3561:70 (0x0DE90046)

So depending on the type of route (as designated by the customer), the appropriate Community values are set by the customer Border router to its MCI peer. The MCI peer router then assigns the mapped LOCAL_PREF values to the incoming routes from a customer based on their Community values.

This scheme allows ISP's to let customers decide how their routes will be handled within the ISP's network, and simplifies ISP router configurations by reducing customizations.

Example Configuration (from RFC 1998):

Customer Border Router :

```
router bgp 100
neighbor < x.x.x.x > remote-as 3561
neighbor < x.x.x.x > filter-list 20 out
neighbor < x.x.x.x > route-map config-community out
neighbor < x.x.x.x > send-community
!
```


Cisco Systems, Inc.
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

```
! # Match All
ip as-path access-list 1 permit .*
!
! # List of Customer AS's
ip as-path access-list 20 permit ^$
ip as-path access-list 20 permit ^64700_
ip as_path access-list 20 deny .*
!
! # IP access-list matching, customer provided backup
access-list 101 permit ip 192.160.154.0 0.0.0.0 255.255.255.0 0.0.0.0
!
! # Route Map config-community
route-map config-community permit 10
match ip address 101
set community 0x0DE90046                    >>>>> Community Value 3561:70
route-map config-community permit 20
match as-path 1
```

At the ISP Border Router :

```
! #Customer provided Backup
ip community-list 70 permit 0x0DE90046
ip community-list 70 deny
!
! #Other ISP Routes
ip community-list 80 permit 0x0DE90050
ip community-list 80 deny
!
! #Customer Backup routes
ip community-list 90 permit 0x0DE9005A
ip community-list 90 deny
!
! #Route Map applied to all BGP customers
route-map set-customer-local-pref permit 10
match community 70
set local-preference 70
route-map set-customer-local-pref permit 20
match community 80
set local-preference 80
route-map set-customer-local-pref permit 30
match community 90
set local-preference 90
route-map set-customer-local-pref permit 40
match as-path 1
set local-preference 100
```

Hence, all ISP Border routers have the same configuration mapping the Community values to LOCAL-PREF's. The above example is one way to use Communities, in general similar concepts can be used to apply customized polices to a group of routes. Please refer to RFC 1998 for a more detailed discussion of this scheme.

Module 9.0 Multihoming to a Single Provider

This configuration is common with stub customers connected to highly reliable Network Service Provider networks or in situations where strategic relationships exit between the network and providers. These are also common in regions such as Asia and Latin America

Cisco Systems, Inc.
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

where access to multiple Network Service Providers is limited. It is assumed that the Multihomed network gets its addressing space from the Upstream provider block. If this is not the case then as discussed in Section 4.1, the provider will be required to propagate the Multihomed network routes into the Internet along with its assigned address blocks. There are two basic scenarios which are discussed below :

9.1 Single Router Configuration

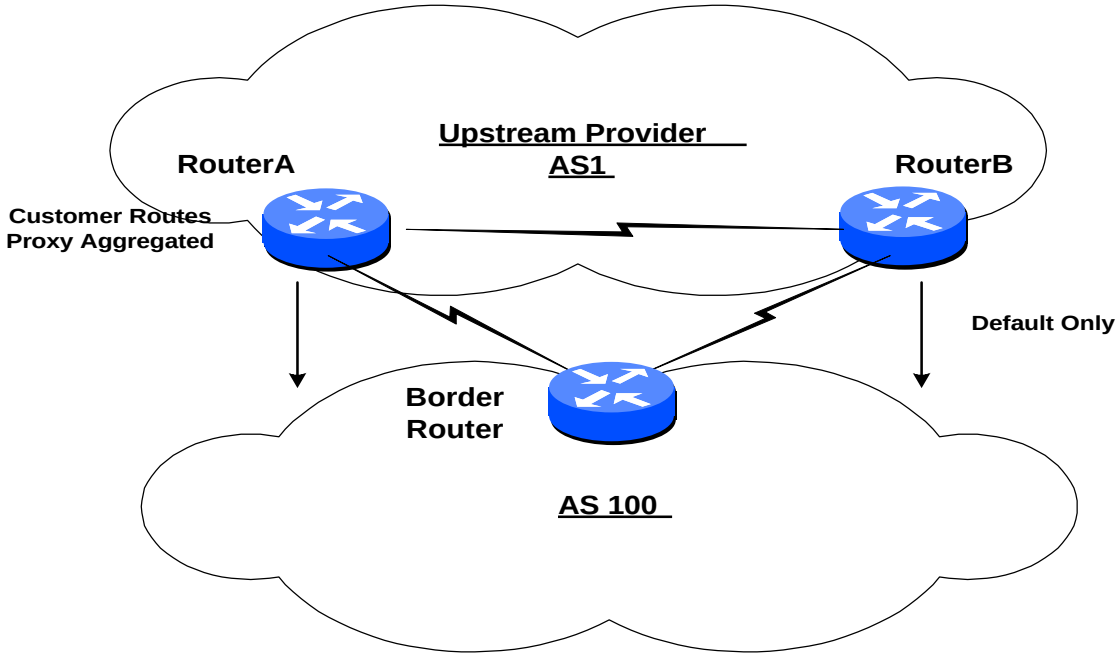


Figure 8

a] *Routing to the Upstream Provider*. The topology is shown in Fig. 8. With both the links from the upstream provider terminate in one Border router, there are two ways to interface with the provider routers

i] Static Default routing. In this configuration, the Border router does not do dynamic BGP routing to the Upstream provider routers. Static defaults are configured on the Border router pointing to the provider routers for outbound traffic. On the provider routers, static routes for the Multihomed network are configured and redistributed into the providers dynamic routing.

Border Router

```
interface s 0/0
!# 1st link to Upstream Provider (AS1)
ip address 131.108.10.10 255.255.255.255
!
interface s 1/0
!# 2nd link to Upstream Provider (AS1)
ip address 131.108.20.20 255.255.255.255
!
!# Static Default Routes
ip route 0.0.0.0 0.0.0.0 serial 0/0
ip route 0.0.0.0 0.0.0.0 serial 1/0
```

Cisco Systems, Inc.
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

```
!
router ospf 100
:
default-information originate metric 2 metric-type 1
```

Here, two defaults each pointing to one of the outbound interfaces are statically configured. The static defaults are introduced into OSPF for propagation to the rest of the network. Since the are tied to specific interfaces, they would only be propagated if the interface is up and running.

Provider Router A

```
interface s 2/0
! Link from Router A to Multihomed Network
ip address 131.108.10.11 255.255.255.255
!
!# Static Route for the Multihomed network's address space (Proxy
Aggregation)
ip route 200.200.16.0 255.255.240.0 serial 1/0
!
router ospf 1
:
redistribute static metric 2

router bgp 1
:
aggregate-address 200.200.16.0 255.255.240.0 summary-only
:
```

In the provider router, a static is configured for the prefixes belonging to the Multihomed network pointing to serial 2/0. Assuming the provider runs BGP on his Border routers, BGP synchronization requirements may require redistribution of the static into the IGP. The route is introduced into BGP by configuring an aggregate.

ii] BGP Peering. When the Border router runs BGP with the upstream provider, dynamic changes in the Multihomed network are conveyed to the provider instantaneously. BGP peering also reduces overhead in terms of maintenance of static defaults and routes in both networks. However, as shown in the configurations below route-maps must be configured to protect against inconsistencies in routing updates from the provider. Here we assume the Multihomed network accepts only default.

Border Router

```
router bgp 100
neighbor < Router A > remote-as 1
neighbor < Router A > route-map rcvd_routes in
:
aggregate-address 200.200.16.0 255.255.240.0 summary-only

router ospf 100
:
redistribute bgp 100 metric 2 distribute-list 1
:
```

Cisco Systems, Inc.
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

```
access-list 1 permit 0.0.0.0 0.0.0.0
access-list 1 deny any any
route-map rcvd_routes permit 10
match ip address 1
```

The provider may elect to send the Multihomed network a default route via. BGP which is redistributed into the networks IGP for propagation to the rest of the network. The access-list as highlighted before protects the IGP from accidental injection of large number of external routes which could cause major network outages.

Provider Router

```
router bgp 1
neighbor < Border Router > remote-as 100
neighbor < Border Router > distribute-list 100 in

access-list 100 permit 200.200.16.0 0.0.15.255 255.255.240.0 0.0.0.0
access-list 100 deny any any
```

Distribute-list 10 will permit the Multihomed network to advertise only its /20 prefix. This step is commonly used to protect the ISP's network against misconfigurations in the Border router which may leak unwanted routes into the ISP's network.

b] *IGP Routing*. The Network may use any of the 3 options discussed in Section 4.2. Most often networks run an IGP such as OSPF for internal connectivity. The border routers will redistribute the default routes into the IGP.

c] *Routing Requirements and Policy*. All outbound traffic exits the AS via. one router so the defaults in the IGP will ensure all traffic drains to the border router. Similarly for Inbound traffic no explicit routing policy is required since there is only one entry into the Multihomed network.

d] *Traffic Flow Control.* Load Sharing and Symmetric routing are non-issues given the single router configuration. There is a single point of failure in the Border router and hence Reliability is available only from a provider perspective.

Given that reliability is a key requirement, the presence of a single point of failure defeats the purpose of Multihoming. Depending on the size of the network the Border router must have enough system resources to switching the traffic load to and from the network.

9.2 Multiple Router Configuration

Cisco Systems, Inc.
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

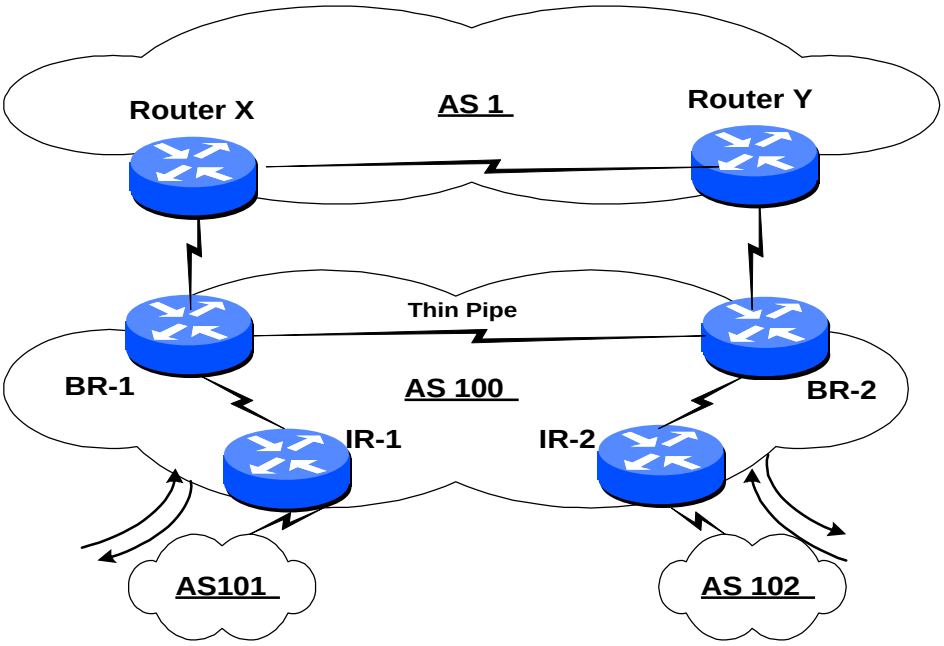


Figure 9

In this topology, the Multihomed network connects to the upstream provider with multiple routers. The Multihomed network co-ordinates with the upstream provider to ensure the desired traffic flow to and from the network. For example, in Fig. 9 the Network may use the providers connectivity to pass traffic between BR-1 and BR-2 and use the Thin pipe for routing updates only or implement Symmetric routing using MED's. The amount of routing to be exchanged will depend on the desired policy control.

a] *Routing to the Upstream Provider* :

- Static defaults are used only if the Multihomed network has no major routing or traffic flow policy requirements to and from the provider. The provider will do closest exit routing to the Multihomed network as more detailed information is unavailable.

- Most Multihomed networks do BGP routing to the provider as it provides policy control and enable them to use the connectivity effectively. Using BGP any of the traffic flow control techniques detailed in Section 8.2 can be applied to control Inbound traffic.

b] *Internal Routing* :

The Network may use any of the 3 options discussed in Section 4.2. Most often networks run an IGP such as OSPF for internal connectivity. The border routers will inject either defaults or a partial set of external routes into the IGP depending on the Internal policy and traffic flow requirements. Example: The internal network may want to see the directly connected provider routes so that it may make closest exit routing decision as will be shown in the configurations below.

In certain cases the Multihomed network is itself an ISP providing connectivity to other networks. Here the Multihomed Network does BGP routing to the other AS and obtains routing information from customer networks which need to be propagated to the upstream

provider. This can be achieved by introducing customer routes into BGP either by redistribution or using the Aggregate address command (Refer [1] for detailed configuration examples). These routes are carried within IBGP to the Border routers where they become a part of the Multihomed network's routing advertisement to the Upstream Provider. A configuration example of such a case is shown.

c] *Routing Requirements and Policy Issues*

Given that the network is Multihomed to a single upstream provider, the logical option is go with the Partial routing option described in Section 4.3.2. All or few of these routes maybe redistributed into the IGP. This will give the network visibility into the providers internal routing structure, allowing for optimal routing decisions to provider networks and networks in the provider's routing domain.

Sometimes the Multihomed network may request Full Internet routes so that it may provide the same to its customers who require all routes. Here Full Routes will be accepted at the Border routers and passed on only to those BGP speakers who are connected to the requesting customers. In most case since carrying Full Routes impact the operations of the network financially and administratively (Section 6.2).

d] *Traffic Flow Control*

Outbound traffic can be manipulated by setting the appropriate IGP metrics as discussed in Section 4.5.1. Inbound traffic policy is enforced using the BGP routing policy options and co-operation with the Upstream Provider. Load Sharing may not always be a requirement. Consider the situation where the network connects to the provider at 2 different sites. The Inbound policy would be to hand off all traffic coming into Site 1 at Border Router 1 and vice versa and the traffic to Site 1 maybe greater than the traffic to/from Site 2. However, the bigger challenge is that of Symmetric Routing. As discussed in Section 4.5.2.3, total Symmetric routing is impossible to achieve. If it is a critical requirement, the AS Path Prepend technique can be used with Site 2 routes appearing longer when advertised by Border 1 and vice versa. For traffic sourced from the directly connected provider the MED or Community based LOCAL_PREF techniques can be used to influence Inbound traffic from the provider.

e] *Configuration Case Study*

Fig. 9 shows a Multihomed network with the following characteristics

- OSPF is used as the IGP, all routers in Area 0
- Full Routing Updates at the Border Routers (BR-1 & BR-2)
- Partial Updates to Internal Router IR-1 and IR-2
- Router IR-1 connects AS 101 and Router IR-2 connects AS 102
- No Load Sharing requirement
- Proximity Routing is a requirement
- Symmetric Routing required to provider destinations from AS 101 & AS 102. All traffic from AS 101 must exit and enter AS 100 via. Router A, and traffic from AS 102 via. Router B.

Cisco Systems, Inc.
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

- Partial Route redistribution from BGP into OSPF. Important not to protect IGP from redistributing Full Internet routes. May cause significant Network Outage.

For Outbound routing therefore, traffic to destinations within the Provider's routing domain will take the best path available as determined by the IGP decision process. To non-provider destinations, Internal routers use the defaults in the IGP to reach the Border routers which have Full routing information to get the packets out. Inbound policy is discussed along with the configurations. Upstream Provider routes are not propagated to IBGP neighbors since the IGP takes precedence while making routing decisions anyway.

BR-1 :

```
! # IGP configuration
!
router ospf 100
network 200.200.200.0 0.0.0.255 area 0
!
! #Selectively leak Provider routes into the IGP for Proximity Routing
redistribute bgp 100 metric 2 metric-type 1 distribute-list 101
!
! #Source a Default for propagation to the the rest of the Internal routers
! Metric Type 1 is used so routing decisions reflect the Internal topology
default-information originate metric 2 metric-type 1
!
router bgp 100
neighbor < Router X > remote-as 1
!
!# Full routes from neighbor, so no Incoming route-map
!# Outgoing route-map ensure only the correct routes are sent out
neighbor < Router X > route-map outgoing_route out
neighbor < Router X > route-map outgoing_policy out
!
!# All internal neighbors who exchange Partial Routes are part of a Peer
Group.
!
neighbor internal peer-group
neighbor internal remote-as 100
neighbor internal route-map partial_routes out
neighbor < Router IR-1 > peer-group internal
neighbor < Router IR-2 > peer-group internal
!
!# Router BR-2 exchanges Full Routes and hence is configured separately.
neighbor < Router BR-2 > remote-as 100
!
!# Static Default tied to the interface connected to the provider
ip route 0.0.0.0 0.0.0.0 serial 1/0
!
!# Access Lists
!# Access List 101 permits Provider routes only
access-list 101 permit < Networks belonging to the Provider >
access-list 101 deny any any
!
!# AS Path filter applied to routing updates going to the Provider
!# ensure only route originating in AS 100, 101 and 102 are sent.
ip as-path access-list 102 permit  ^$
ip as-path access-list 102 permit _101$
ip as-path access-list 102 permit _102$
!
```

Cisco Systems, Inc.
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

```
!# AS Path filter sends only routes originated by AS X (which is a network
directly connected to BR-1. Not shown in Fig.9) to Router IR-1 & IR-2
ip as-path access-list 103 permit _X$
!
!# AS Path filter to select all routes originated from AS 101
ip as-path access-list 104 permit _101$
!
!# AS Path filter to select all routes originated from AS 102
ip as-path access-list 105 permit _102$
!
!# Generic AS Path filter for all other routes
ip as-path access-list 106 deny .*
!
!# Route Maps
!
!# Route Map applies the AS Path filters outbound to the Upstream Provider
route-map outgoing_routes permit 10
match as-path 102
route-map outgoing_routes permit 20
match as-path 106
!
!# Route Map ensures IBGP neighbors C and D are sent Partial Routes only
!# ie. the routes orginated in AS X directly connected to BR-1.
route-map partial_routes permit 10
match as-path 103
route-map partial_routes permit 20
match as-path 106
!
!# Route Map uses MED's to the provider to ensure Symmetric routing for AS
101 & 102
route-map outgoing_policy permit 10
match as-path 104
set metric 5
route-map outgoing_policy permit 20
match as-path 105
set metric 10
```

All AS 101 routes get a MED of 5 assigned out of Router A. In Router B AS 101 routes get a MED of 10 assigned. Therefore routers in AS 1 will always prefer to route packets to AS 101 via. the Router A border, vice versa for AS 102. However, note that this will work only if AS 1 doesn't override the MED with attributes which come before the MED in the BGP decision process, i.e. setting a LOCAL_PREF which contradicts the MED's.

Module 10.0 Multihoming to different providers

In this scenario, the Mutlihomed network connects to more than 1 upstream provider, resulting in increased complexity in the routing and traffic flow.

a] Addressing structure of the Multihomed network : As discussed in detail in Section 5 the address space could belong to either provider or the network itself. In either case *the providers will have to carry routes which do not belong to their address block in their advertisements. The Multihomed network cannot utilize a provider connection unless the provider advertises the network's routes to the rest of the Internet.*

b] *Routing Strategies* : There are 4 routing strategies commonly used in today's Internet :

Cisco Systems, Inc.
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

The more detailed the routing information, the lesser the amount of sub-optimal routing

i] Default from both Providers : When the Multihomed network has very minimal policy and traffic flow requirements, this option is the way to go. Both Providers send defaults and the Multihomed network makes routing decision based on its IGP metrics. The resources needed here are minimal and it works fairly well in stub networks. The Multihomed network's routes are proxy aggregated by the Provider or if BGP peering exists dynamic routes or aggregates are sent to the Provider.

ii] Dynamic routes + Default from Provider 1 and default from the other : This strategy is used when the Multihomed network determines that Proximity routing (as defined in Section 2) to certain destinations is a critical requirement. The provider leaks more specific routes for these select destinations, all other destinations, the Multihomed network uses the default and may take a path through either Provider's network. The dynamic routing could consist of Partial routes or the Full Internet routing table depending on the policies. The Multihomed network's advertisements could be carried by proxy aggregation or by dynamic updates via. BGP peering. This will depend on the routing and traffic flow policies of the Multihomed network as discussed in Section 4.5. Sub-optimal routing will occur due to the lack of detailed routing info to all destinations.

iii] Dynamic routes + Defaults from both Providers : This strategy is the most common and efficient for non-stub Multihomed networks and is discussed with configuration examples below. Here, the Multihomed network accepts the Local routes from both Providers along with a Default for the rest of the Internet routes. This is very efficient because, the Network now has the ability to take the best path to destinations within the directly connected Provider domains and reach the rest of the Internet through either provider. This simplifies Outbound traffic flow control. However, Sub-optimal routing is still a possibility, when given the lack of more specific routing information, internal routers (IR) use defaults to reach the closest border. For Inbound traffic control, once again either of the strategies discussed in Section 4.5 maybe applied. Resource requirements are not major, though depending on the routing and traffic policies the configuration could get quite complex.

iv] Full Routes from both Providers : This configuration is used in cases where large NSP's multihome to other ISP's at the higher layers of the Internet or in Networks which have a requirement for Optimum routing at any cost. Optimum routing is possible as the Multihomed network now has visibility to the entire Internet routing topology. With Full routing the Multihomed network must have the suitable IGP, system resources, administrative talent and protective mechanisms. Due to the dynamic nature of the Internet, routers in the Network carrying Full routes will experience routing churn. A detailed discussion of large NSP architectures and policies is beyond the scope of this document, but a simpler Optimum routing example is discussed as an example.

Cisco Systems, Inc.
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

c] *Routing and Traffic Flow Policies* : As mention above, the presence of multiple provider significantly complicates the implementation of the policies. Certain options such as the use of MED's as discussed in Section 8.2.2 cannot be used. In general policies may require Inter-Provider cooperation which is difficult to achieve. Hence, in this setup the Multihomed network must accept lesser amount of Inbound traffic flow control.

10.1 Single Router Configuration

As discussed in Section 9.1, Single router configurations are not recommended since they are single points of failure and defeat the purpose of Multihoming. However, this configuration is described here for the sake of completeness.

a] *Routing to the Provider*

i] Static Routing. With Multiple Upstream provider, static routing is not an option unless the network uses one provider as the primary and the other as the secondary. In such a case the Border router is configured with 2 static routes, one of which is assigned a higher preference. (Note : Since both defaults are static routes, administrative distance must be used to assign a higher distance to the static route pointing towards the secondary provider). When both defaults have the same administrative distance, the router will round robin packets between them. Sample Configuration :

Border Router :

```
ip default-network 0.0.0.0
ip route 0.0.0.0 0.0.0.0 < Provider 1 Router Address >
ip route 0.0.0.0 0.0.0.0 < Provider 2 Router Address > 2
```

The "ip default-network" command will set the Gateway of Last Resort on the router to the Provider addresses configured as the Next Hop to reach the default network 0.0.0.0 which has 2 static routes, one with a higher distance of 2.

ii] BGP Routing. In most instances the Border router BGP peers with the provider routers to exchange dynamic routing information. Outbound routing will exit through the single border router and is taken care of by the IGP. Once the traffic gets to the Border, the router must decide which provider to use for which traffic.

Each provider sends the Border routers its Local routes AND a default. The Border router is able to make efficient decisions on routing to destination within the directly connected providers, to the rest of the destinations the defaults sourced by both providers are used. The defaults are propagated to the rest of the network by injecting them into the IGP.

Fig. 10 shows the topology of such as setup.

Cisco Systems, Inc.
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

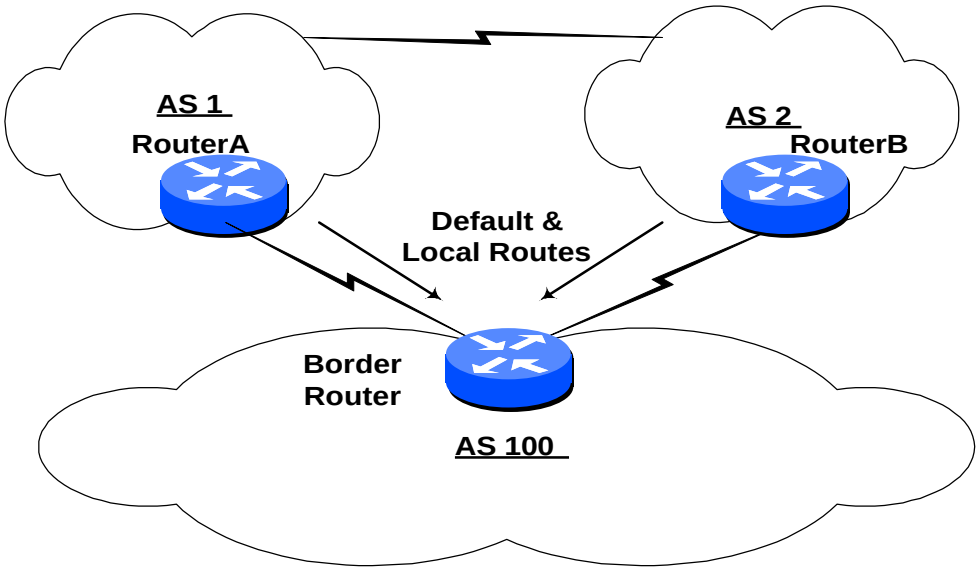


Figure 10

Border Router :

```
router bgp 100
neighbor < Router A > remote-as 1
!# Permit in routes originated by AS 1 and advertise out routes from AS 100
neighbor < Router A > route-map A_routes in
neighbor < Router A > route-map advt_routes out
neighbor < Router B > remote-as 2
!# Permit in routes originated by AS 2 and advertise out routes from AS 100
neighbor < Router B > route-map B_routes in
neighbor < Router B > route-map advt_routes out
!
!# Since both the routing peers provide Default routes, those routes are
injected in the IGP. The default is then propagated to the rest of the
Internal routers.
!
router ospf 100
:
default-information originate
:
!
!# AS Path Filter Lists
!
!# The next 2 filter permit routes originated from AS 1 and 2 respectively.
ip as-path access-list 101 permit _1$
ip as-path access-list 102 permit _2$
!# This filter permits all routes originated by the AS to which this router
belongs (AS100)
ip as_path access-list 103 permit ^$
ip as_path access-list 104 deny .*
!
!# Route Maps
!
route-map A_routes permit 10
match as-path 101
route-map A_routes permit 20
match as-path 104
!
route-map B_routes permit 10
match as-path 102
```

Cisco Systems, Inc.
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

```
route-map B_routes permit 20
match as-path 104
!
route-map advt_routes permit 10
match as-path 103
route-map advt_routes permit 20
match as-path 104
!
```

b] *Internal Routing.* As shown in the example above OSPF can be used as the IGP and default routes redistributed into it at the Borders.

c] *Routing Requirements and Policy Issues.* Most of the issues are part of the discussion in the configuration example given in [a].

d] *Traffic Flow Control.* All traffic will flow through the Border router. Proximity routing is performed due to the presence of detailed Local routing information for each of the providers in the Border router. Traffic to destinations beyond the providers is routed using defaults via. either of the provider. The Multihomed network would have no major Inbound traffic requirements since it has a single Border router.

However, it is important that both providers advertise the networks routes into the Internet whether the routes belong to their address block or not. If one of the providers doesn't advertise the routes to the Internet, inbound traffic to the network cannot be routed via. that provider. This is very important.

10.2 Multiple Router Configuration

Cisco Systems, Inc.
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

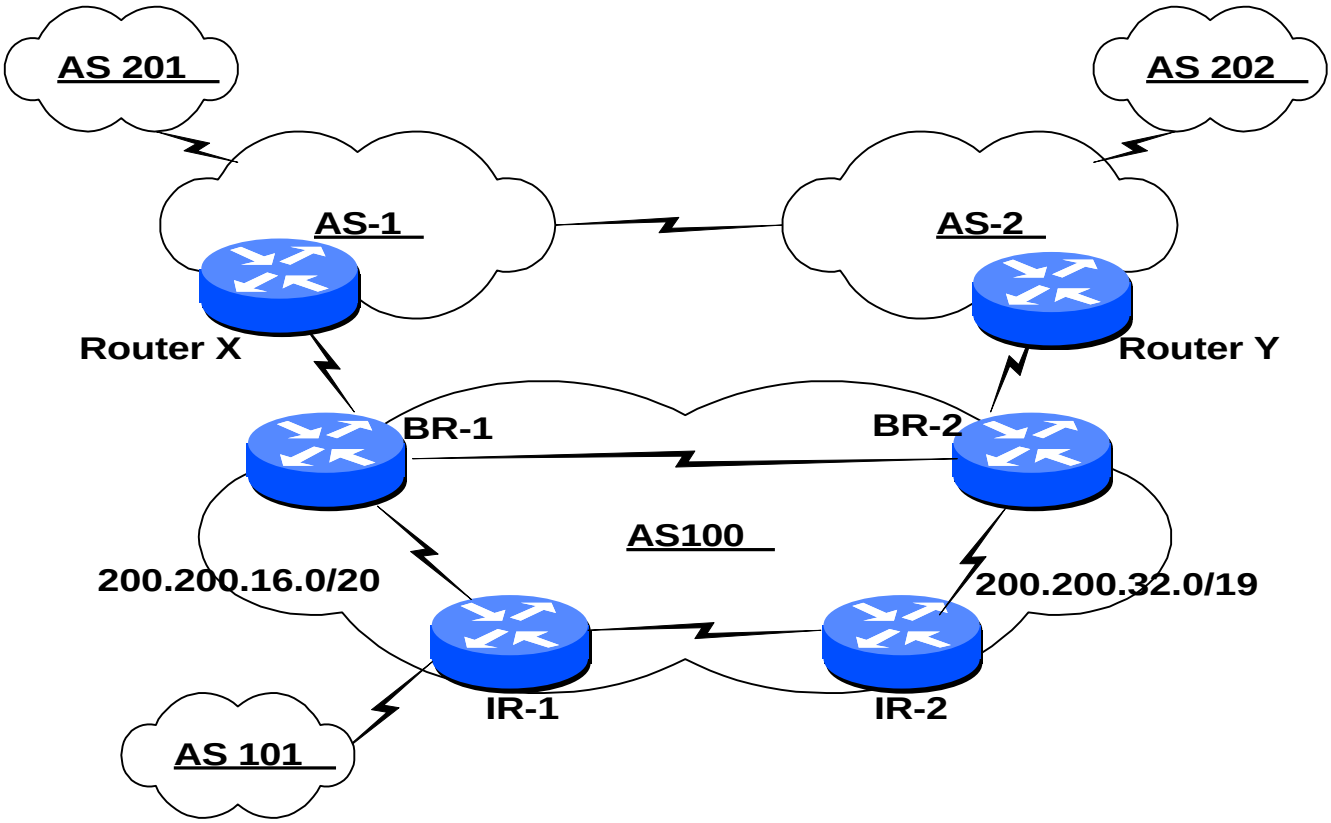


Figure 11

This is perhaps the most common Multihoming configuration in the Internet today and also the most complex. A typical configuration is shown in Fig. 11. The Multihomed network AS 100 connects to two upstream Providers AS 1 and AS2. AS 100 maybe required to Multihome in order for Reliability or it maybe geographically dispersed requiring the network to connect to different providers each site. Let consider each case :

a] Multihoming for Reliability [Case 1] : Here, the Network needs to maintain its connectivity to the Internet in the event of failure of either Provider. Cisco in San Jose is an example. Providers may supply AS 100 with Full routes or Local routes only depending on the level of Optimum routing requirements. Outbound traffic flow can be controlled by managing the BGP and IGP routing within the network. There are no specific Inbound routing requirements other than to try and maintain Symmetric routing for which strategies described in Section 4.5.3 maybe used. The complexity of the configuration and the network is dictated by the policy requirements of the network.

b] Geographical Multihoming [Case 2] : A typical example of this setup is when a company maintains two sites on the East and West Coasts. Economics or connectivity may result in each site connecting to a different provider. This is similar to the scenario studied in Section 4.5.2.2. Outbound policy would be to hand off to the closest Border router and this can be managed via. BGP and IGP interaction. The Inbound policy is more critical, especially if the company does not have an infrastructure to handle significant East-West traffic. In that case, AS 100 will require all traffic to the Western site exit and enter AS 100 via. Provider 1 and vice versa. This is essentially a Symmetric routing requirement which is critical for business

Cisco Systems, Inc.
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

operation. One way to address this requirement is using the strategy discussed in Section 4.5.2.3 to force the Internet to implement the Inbound policy

1] *Routing to the Providers*. In majority of the cases dynamic BGP peering exists between the Multihomed Network and the Upstream Providers. This ensures that the Network gets the maximum flexibility and policy control.

2] *Internal Routing*. The Network may use any of the 3 options discussed in Section 4.2. Most often networks run an IGP such as OSPF for internal connectivity. The border routers will redistribute either defaults or a partial set of external routes into the IGP depending on the Internal policy and traffic flow requirements. An example of this requirement is when the internal network may want to see the directly connected provider routes so that it may make closest exit routing decision as will be shown in the configurations below.

3] *Routing Requirements and Policy Issues*. As discussed above, the two example cases considered in this document have unique set of policy requirements. Generally for most Outbound traffic the IGP is setup to do Closest Exit routing to the Border routers. Proximity routing is another requirement which can be achieved by the providers leaking at least Local routes in the Multihomed network. For Inbound traffic depending on the particular case under consideration, option discussed in Section 4.5.2 are applicable. More details will be discussed in the configuration examples.

4] *Traffic Flow Control*. Outbound traffic can be manipulated by setting the appropriate IGP metrics as discussed in Section 4.5.1. Inbound traffic policy is enforced using the BGP routing policy options and co-operation with the Upstream Provider. Load Sharing, as mentioned in previous sections may not always be a requirement. Once again the bigger challenge is achieving Symmetric Routing. As discussed in Section 4.5.2.3, total Symmetric routing is impossible to achieve.

C] *Configuration Examples* :

Case 1 : Multihoming for Reliability [Fig. 11]

In discussing this configuration the following assumptions are made :

- The network is an ISP itself. Router IR-1 connects AS 101 to the Internet
- Full Routing from both Providers
- Partial Route redistribution from BGP into OSPF. Important to protect IGP from redistributing Full Internet routes. May cause significant Network Outage.
- OSPF is used as the IGP, all routers in Area 0
- No Load Sharing requirement, but Optimal routing required.
- Proximity Routing is a requirement
- Symmetric Routing is required whenever possible

Router BR-1 : (Identical configuration for Router BR-2)

Cisco Systems, Inc.
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

```
!# IGP Configuration
router ospf 100
network 200.200.200.0 0.0.0.255 area 0
!# Partial BGP routes redistributed into IGP along with the Static Default
redistribute bgp 100 metric 2 metric-type 1 distribute-list 101
default-information originate metric 2 metric-type 1
!
!# BGP Configuration
router bgp 100
neighbor < Router X > remote-as 1
!# The next line only permit advertisement of routes originated by AS 100
neighbor < Router X > route-map advt_routes out
neighbor < Router BR-2 > remote-as 100
!# Internal Policy route-map for application of LOCAL_PREF etc. to specific
routes
neighbor < Router BR-2 > route-map internal_policy out
neighbor < Router IR-1 > remote-as 100
!# IR-1 a non-Border IBGP peer receives only routes within provider domain
neighbor < Router IR-1 > route-map internal_advt out
!
!# Static Default tied to the interface connecting the Provider router
ip route 0.0.0.0 0.0.0.0 serial 1/0
!
!# Access Lists
!
access-list 101 permit < Provider Networks  >
access-list 101 deny any any
!
!# AS Path Filter Lists
!
!# Permit all routes within the provider's routing domain
ip as-path access-list 101 permit < Regular Expression for all routes
within the provider's routing domain >
!
!# This filter permits all routes originated by the AS to which this router
belongs (AS100) and those orginated by AS 101.
ip as_path access-list 102 permit ^$
ip as _path access-list 102 permit _101$
ip as_path access-list 103 deny .*
!
!# Route Maps
!
route-map internal_advt permit 10
match as-path 101
route-map internal_advt permit 20
match as-path 103

route-map advt_routes permit 10
match as-path 102
route-map advt_routes permit 20
match as-path 103
```

The route-map internal_policy is shown there for completeness. In this example we do not a specific Internal policy outside of Closest Exit routing done by the IGP.

Note : In Fig. 11, AS 201 is closer to AS 1 and AS 202 is closer to AS 2. Ideally, the Multihomed Network would want to send packets to AS 201 from any Internal router to Border BR-1. This will ensure Optimal routing. In this example the IGP drains the traffic to

Cisco Systems, Inc.
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

the Closest Border router, since BR-1 & BR-2 have Full Internet tables they are able to determine AS 201 is closer via. AS 1 than via. AS2. Optimal routing is achieved. Router IR-1 which connects customer AS 101 will introduce those routes into BGP and OSPF for transport to the Borders and the Internet as shown.

Case 2 : Geographical Multihoming

For the sake of this discussion we assume the following :

- Assume a non-ISP network with sites in 2 distinct geographical locations (Fig. 11). BR-1 & IR-1 are in the West and BR-2 and IR-2 are in the East.
- The East-West Border routers are connected to each other and are IBGP peers
- A common IGP domain is assumed for the Network.
- Also, the East-West bandwidth is adequate to handle Internal traffic. The policy on External traffic is for all Inbound traffic delivered at the closest E/W Border.
- Local Routes + Default from each Upstream Provider at the Border routers
- Local Routes + Default introduced into OSPF to ensure Closest Exit routing
- OSPF is used as the IGP, all routers in Area 0
- Internal Router need not run BGP
- No Load Sharing requirement, but Optimal routing required.
- Proximity Routing is a requirement
- Symmetric Routing required to and from East and West sites as discussed.

Router BR-1 : (Identical configuration for Router BR-2)

```
!# IGP Configuration
router ospf 100
network 200.200.200.0 Area 0
:
!# Routes in the Provider's local routes are redistributed into the IGP.
redistribute bgp 100 metric 2 metric-type 1 distribute-list 101
default-information originate
!
!# BGP Configuration
router bgp 100
neighbor < Router X > remote-as 1
!# Route-maps to protect advertisements in and out of the router to AS 1
neighbor < Router X > route-map incoming_routes in
neighbor < Router X > route-map outgoing_routes out
neighbor < Router B > remote-as 100
!
!# Access Lists
!
access-list 101 permit < List of Provider 1's Local Routes + Default >
!# West Site Routes
access-list 105 permit 200.200.16.0 0.0.15.255 255.255.240.0 0.0.0.0
!# East Site Routes
access-list 106 permit 200.200.32.0 0.0.31.255 255.255.224.0 0.0.0.0
!
!# AS Path Filters
!
ip as-path access-list 101 permit _1$
!# Where X is any other AS which is within Provider 1's routing domain.
```


Cisco Systems, Inc.
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

```
ip as-path access-list 101 permit _X$
ip as-path access-list 102 deny .*
!
!# Route Maps
!
route-map incoming_routes permit 10
match as-path 101
route-map incoming_routes permit 20
match as-path 102
!
route-map outgoing_routes permit 10
match ip address 106
set as-path prepend 100
route-map outgoing_routes permit 20
match ip address 105
```

Note : Refer to Section 8.2.3 for more details on the AS Path Prepend method.

In Router BR-2, access-list 105 will be matched and prepended with AS 100 to make its AS Path longer as it is advertised out of Router BR-2. The result as explained in Section 8.2.3 is that all traffic to 200.200.16.0 routes will always come in via. Router BR-1 and vice versa. Since each Provider advertises their Local routes at each Border, the Outbound traffic from the Network will always drain to the right Border. Symmetric and Optimal routing. It is important to note that policy and aggregation decisions at upstream provider routers will influence the effectiveness of this approach to traffic flow control. Note that as stated in Section 8.2.3, AS Path prepend can be overruled by upstream providers using Local Pref.

Module 11.0 Precautions & Gotchas

11.1 Double Route-Map protection between peers

As seen in many of the example configurations it is always advisable to for the Mutlihomed Network and the Provider to install protective route-maps to filter each others route advertisements.

Border Router :

```
router bgp 100
neighbor < Router X > remote-as 1
neighbor < Router X > route-map incoming_routes in
neighbor < Router X > route-map outgoing_routes out
:
```

The route-maps are self explanatory, more detailed examples given in Sections 9 and 10.

11.2 Leaking Holes

In certain case, a Network maybe advertising an Aggregate route which cover certain specifics which are outside its routing domain. Example : A Network maybe advertising 200.200.32.0/19, which covers specifics from 200.200.32.0/24 - 200.200.63.0/24. If the networks 200.200.60 - 63/24 are outside its routing domain routing problems result. These /24's are known as "Holes". See Fig. 12. This weighs in as a Gotcha since most operators do not realize the problem till it bites them.

Cisco Systems, Inc.
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

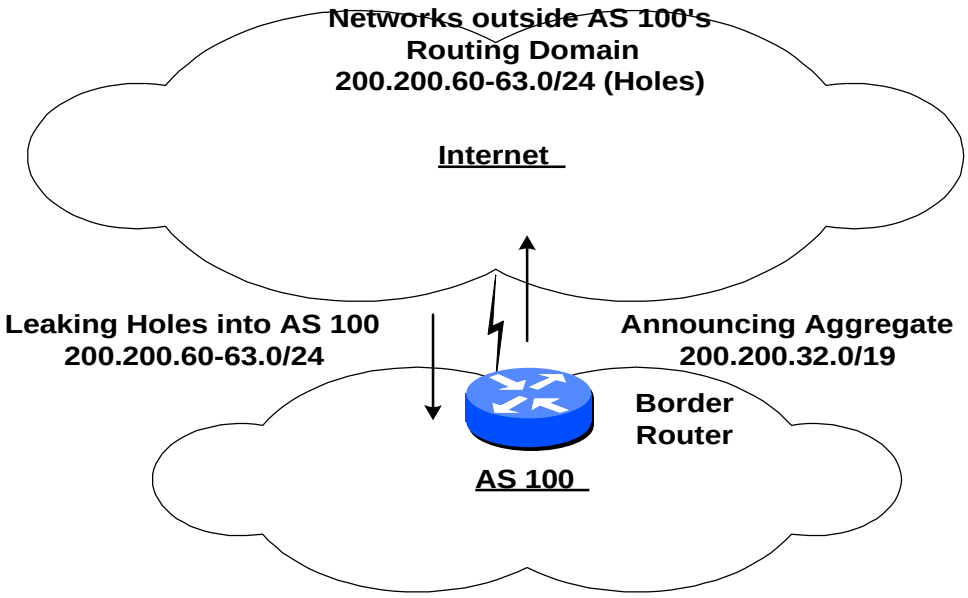


Figure 12

Packets to these 4 /24 will still reach them if more specific routes exist in the Internet routing tables. This is because more specific routes are preferred over aggregates. However, problems may result when routers within the Network try to reach these 4 /24 destinations unless Full Routing is propagated within the Network itself. In the absence of Full Routing, Internal routers get confused over where to send the packets since the Borders generate an Aggregate which covers these routes. Packets get bounced around and eventually time out.

The way to solve this problem is to "Leak the Holes" into the Internal network at the Borders. This will ensure that Internal routers have information on how to get to the "Holes" using more specific routes leaked at the borders.

Note : All this is relevant only if the Network advertises an Aggregate which covers some networks outside its routing domain.

11.3 Route Dampening (Refer [1], [2])

This is a feature in the Cisco IOS whereby repeatedly flapping routes advertised via. EBGP peers are ignored after a couple of flaps. The flapping route is put in the penalty box and is not advertised to other BGP peers of the router till it stabilizes. This feature is now widely installed on most border routers to protect them from constant routing instability.

12.0 Conclusions & References

Multihoming is a complex routing problem which does not lend itself to canned solutions. Every scenario has its own unique routing quirks which affect the Multihoming strategy to be used. As routing in the Internet continues to evolve so will the Multihoming strategies. The above discussion is intended to bring out some of the current problems, strategies and examples in use today in the Internet. The reader is encouraged to use this information as a basis to determine the Multihoming strategy for their particular network.

Cisco Systems, Inc.
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

Some of the recent efforts in the area of Multihoming are listed below :

- draft-bates-multihoming-00.txt : Scaleable support for multi-homed multi-provider connectivity
- NAT based Multihoming : Presented by Yakov Rekhter at Nanog '96 - Ann Arbor.

References

1] Inter Domain Routing White Paper : Sam Halabi - Good Overview of BGP4
2] Cisco Reference Manuals - Discussions on BGP features such as Route Dampening, Route Reflectors etc. and router configuration syntax
3] Multihoming discussions on Nanog & inet-access archives.
4] RFC 1998 : E.Chen et.al : An Application of the BGP Community attribute in Multi-home routing.
5] RFC 1966 : T.Bates et.al. : BGP Route Reflection : An Alternative to Full Mesh IBGP
6] RFC 1519 : V. Fuller et.al. : Classless Inter-Domain Routing : An Address Assignment and Aggregation strategy.
7] RFC 1583 : J. Moy : OSPF - Version 2

Cisco Systems, Inc.
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100