# BGP-4 in Vanguard Routers

# Table of Contents

# Introduction to BGP

Border Gateway Protocol BGP is the primary routing protocol used between Autonomous Systems (AS's) in the Internet to exchange routing information. It makes it possible for Internet Service Providers (ISPs) to connect to each other and for end users to connect to ISP's. The primary function of a BGP speaker is to exchange network reachability information with another BGP speaker. This information includes historical data to show the path in the form of AS numbers to reach networks. This historical data prevents routing loops.

BGP has proven itself to be a very scalable and stable protocol with some internet core routers today supporting routing tables in excess of 200,000 networks. Stability is provided because the designers of BGP used TCP as its transport protocol. It is flexible because it supports complex routing distribution policies with mechanisms to signal preferred and redundant paths to another BGP speaker.

Like any protocol BGP has terminology that must be defined to understand a detailed technical discussion of that protocol.

## *BGP terminology*

### AS (Autonomous system):

In the Internet, an Autonomous system (AS) is a collection of IP networks and routers under the control of one entity (or sometimes more) that presents a common routing policy to the Internet. The assignment of AS numbers is under the control of the IANA which also control IP Address distribution.

- A unique AS number (or ASN) is allocated to each AS for use in BGP Routing. With BGP, AS numbers are important because the ASN uniquely identifies each network on the internet.
- ASN Numbers are assigned by the Internet Assigned Numbers Authority (IANA)
  - The first are public AS numbers and range from 1 to 64511.
  - The second range, from 64512 to 65534, are known as private numbers, and can only be used internally within an organization.

### AS connection:

Two AS's are said to be connected when both a Physical (Shared Data Link Subnet) and BGP connection (BGP Session) exists between them.

### BGP Speaker:

A system running BGP. It need not be a router.

**BGP Neighbor/Peer:** A pair of BGP speakers exchanging Inter-AS routing Information

- Internal Peers: A pair of BGP speakers within the same AS. .Internal BGP Peers do not need to be directly connected.
- External Peers: A pair of BGP speakers in different AS's. External BGP Peers need to be directly connected

### BGP Session:

A TCP session established between 2 BGP peers for the purpose of exchanging routing information.

### NLRI:

This stands for Network Layer Reachability Information. This is the address or address prefix whose route is being distributed in a BGP Update.

### Aggregation:

In BGP terminology this means summarization of multiple NLRI's into one.

### Routing Protocol:

The purpose of a routing protocol is to determine the "best" route to each destination and to distribute routing information among the systems on a network

Routing protocols are divided into two general groups: *interior* and *exterior* protocols

- Interior protocols (IGP) are used to exchange routing information within an *autonomous system* (AS). Common IGP protocols are RIP and OSPF.
- Exterior protocols (EGP) are used to exchange routing information between autonomous systems. They are not aware of the finer level of topology of the network on a link-by-link basis within the AS. BGP is an EGP routing protocol.

### Policy:

Policy support within BGP is what makes this protocol flexible and powerful. Policies are used to determine which routes and imported into BGP and which routes are exported both to other BGP speakers and within the BGP speaker to an IGP protocol. Policies can also be used to attach or modify BGP attributes to a NLRI.

### Attributes:

Attributes are the metrics that BGP uses within a BGP Update to tell the receiving router how the route was created, which path it took and which next hop IP address to use. There are other attributes that help influence routing decisions and some attributes that make implementation of policies simpler.

## *How does BGP work?*

Now that we have looked at some of the terminology that BGP uses lets look at how and why BGP was designed and how BGP works.

BGP was designed specifically with the internet in mind. The need was for a protocol that was very powerful, scalable, reliable, and flexible. It was also determined that BGP would need to support a very extensive set of metrics to make complex routing decisions.

To make BGP scalable response to changes in topology were given a back seat. IGP protocols such as OSPF are very quick to discover changes in topology and to signal these changes to neighbors. This quick response is because OSPF was designed to flood changes immediately. OSPF because of its quick response is limited in scaling to huge routing domains and to the number of routers in an area. BGP was designed to be less responsive to changes and to process these changes on timed updates and to limit the amount of change in an update. You can imagine with something the size of the Internet that routing changes are constantly occurring. Routers would be overwhelmed processing routing changes rather than routing packets.

To make BGP reliable the designers selected TCP as the reliable transport rather than designing one from scratch. This allowed a jump start because TCP had years of refinement and development in its use as reliable transport protocol for most IP applications. Using TCP frees BGP up from tasks such as resending messages that are not acknowledged.

To make BGP flexible the designers gave it a support for inbound and outbound filters called policies and gave it an expanded set of metrics called attributes. These policies and attributes allow the Network administrator or designer of routing policy within an AS to block, redistribute or summarize routing information as he/she sees fit. He/she is also able to change or introduce attributes to steer traffic along preferred paths to and from their AS.

## *BGP Typical Uses*

BGP was designed as the Internet routing protocol but it has found other uses as well. Below are the most likely scenarios you would find Vanguard Networks routers using BGP.

### Customer or Small ISP is connected to a single ISP

This scenario a customer does not generally use BGP to the ISP routers. Static routes and default gateway are more commonly used to connect the customer. The exception is where the customer has multiple connections and wants to use BGP attributes to favor a higher bandwidth connection yet provide backup. In the case where the customer uses BGP it most likely uses Private AS number and the ISP strips the Private ASN before redistributing the customer's networks to the Internet.

A small service provider would more often use this connectivity as the ISP buying internet connectivity from larger ISP's. This small ISP will most likely want to obtain their own AS number and obtain their own address space.

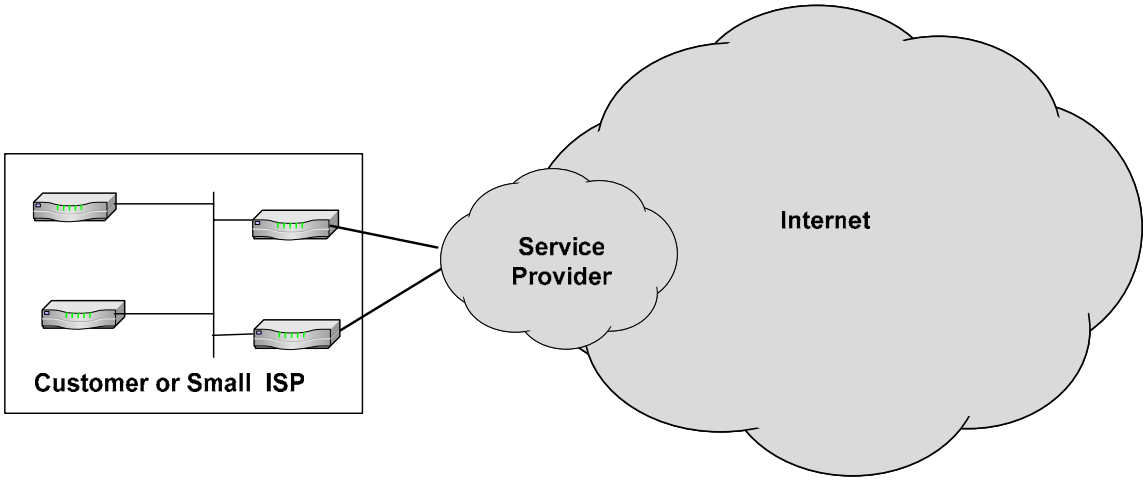


**Figure 1.1 Single or multiple connections to a single ISP**

**When to use BGP to the service provider.**

- Customer is multi-homed to the same ISP
- Customer needs dynamic routing protocol to detect connection failures
  *Note: The customer can use private AS number to the ISP for these connections. (AS numbers 64512-65535) The selection of these numbers must be coordinated with the ISP.*
- Small ISP needs to originate their address space in the Internet

Static routing to the ISP is simpler in other cases.

### Customer is connected to more than one ISP

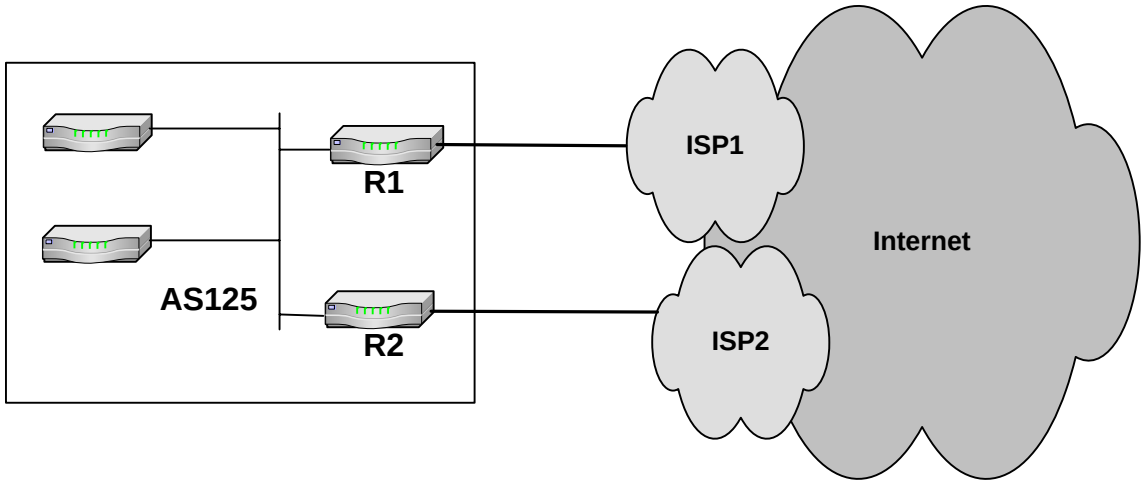


**Figure 1.2 Customer connected to 2 ISPs**

Most common use of this application is the have the ISP send BGP default route or aggregated routes rather than full Internet routing tables. If default route is sent the customer can chose policies to either favor 1 ISP and use the other for backup or to load balance the traffic between the 2 ISP's. (OSPF can be used as the interior routing protocol to load balance traffic between the egress routers to the ISP's) Customer must create Policies to prevent the ISP from using the customer as Transit AS for other traffic.

- Customer must have its own officially assigned AS number and IP address space.
- Customer announces its own official IP networks to both ISP's
- Both ISP's will forward either all routes received from the Internet or send default route to the customer.
- The customer should avoid forwarding any route received from one ISP to the other.
  - If proper policies are not enforced the customer could become transit AS between the 2 ISP's

This connection for a large customer will give them true redundancy and not be dependant on any one ISP for connectivity.

## Service Provider MPLS VPN Network (RFC2547) connecting customer enterprise.

In this application the customer's enterprise network is connected by Service Provider MPLS/BGP-VPN network. (RFC 2547 BGP/MPLS VPN). Unlike a Traditional Frame Relay or ATM network connecting the customer's enterprise where FR or ATM is provisioned end to end. This network uses Frame Relay or ATM as an access protocol to the providers PE Routers. The Provider switches the traffic at the IP layer using MPLS tag switching. BGP is used for the provider to maintain a routing table (VRF) of the customer's enterprise network. This type of offering is rapidly replacing traditional frame relay offerings as headquarters and branch routers need a single PVC to reach all locations. Provider offerings also support QOS across these networks where traditional Frame Relay does not support QOS within the provider cloud.

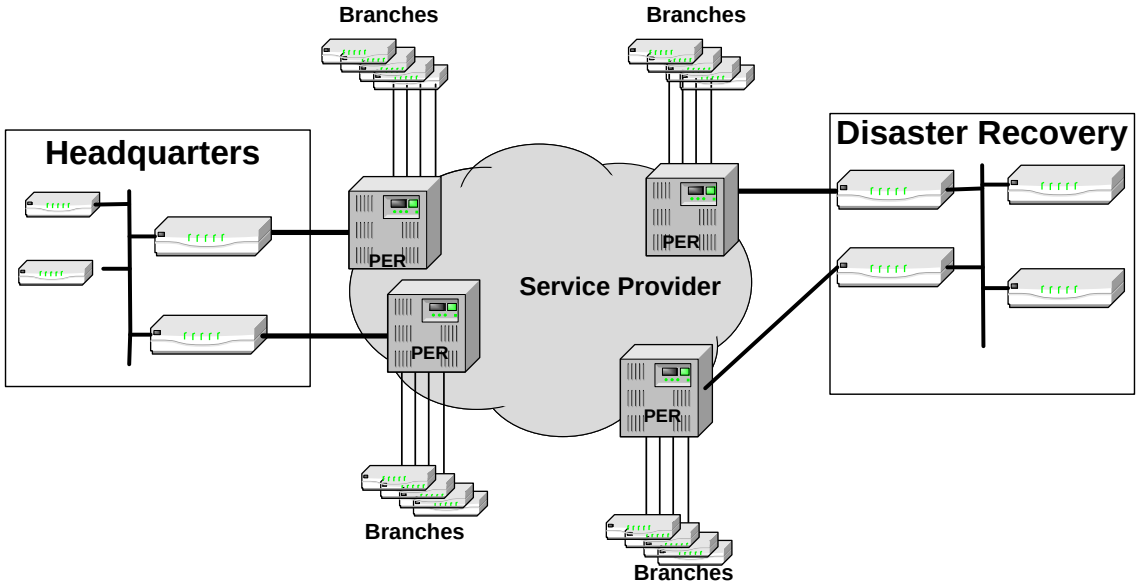


**Figure 1.3 Customer connected to ISP MPLS VPN Network**

Refer to figure 1.3, in this network the Service Provider is providing VPN services separating the customer network from other customers. The provider maintains a VRF (Virtual Routing Forwader) table for each customer. BGP is needed for redundancy at HQ and Disaster recovery site. BGP or static routes can be used at branches. If branches are using backup via ISDN, DSL or Dial, IP can use the absence of BGP routes to trigger backup. This gives higher reliability than dependence on the A bit in the Frame Relay header to trigger backup. With Frame Relay the virtual circuit is no longer a point to point logical connection through the network but only exists between the customers edge router (CE) to the providers edge router (PE), Frame Relay connectivity no longer equates with IP connectivity.

IP-enabled offerings such as AT&T's IP-Enabled Frame Relay fall into the IP-enabled service category in that they allow you to use an existing frame relay access link to tap into a connectionless, Multi-protocol Label Switching (MPLS)-based IP backbone. The primary benefit is that achieving mesh connectivity within your company's VPN requires just a single access permanent virtual circuit (PVC) from each remote site. This type of network will become more popular as more customers are investigating disaster recovery scenarios. The higher cost of the IP enabled PVC starts to balance out when measured against the need for multiple PVC's.

- BGP is needed at headquarters if the headquarters is multi-homed into provider network to provide redundant BGP paths.
- BGP is needed at remotes if customer needs dynamic routing protocol to detect connection failures. (IP triggered backup)
- BGP is needed at all locations if replacing traditional 2 PVC branch FR connections. (Primary and backup PVC)
- Private AS numbers are used by the Service Provider for the customer connections. (Assigned by provider)

## BGP Connecting multiple customer AS's running IGP.

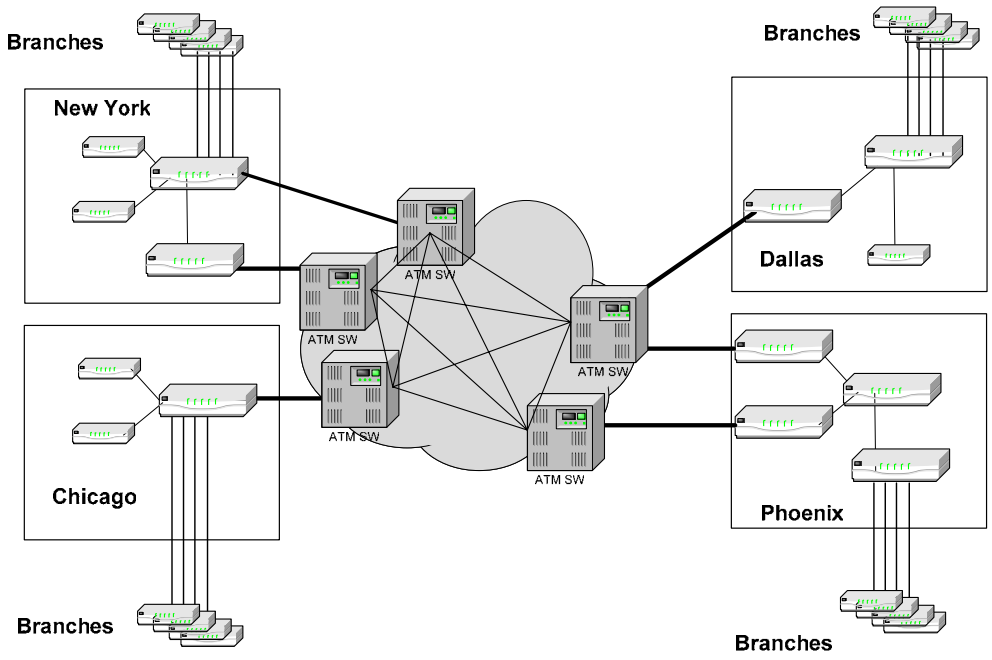


**Figure 1.4 Large Company with BGP Backbone**

Refer to Figure 1.4. In this application the company gives control for a routing domain at the regional level. (New York, Dallas, Chicago and Phoenix) New York and Chicago may be running different OSPF domains while Dallas and Phoenix may be running RIPV2. The backbone routers are meshed and running BGP. The advantage to this company is the flexibility that BGP has with the policies that can be used at backbone routers and the isolation from problems effecting one part of the network having impact on another part.

This type network can also be used by a financial service provider keeping customers divided into separate AS's. BGP provides much easier policy filtering to keep networks separate.

## BGP Design considerations

BGP was designed to perform well in, inter-domain routing applications, inter-networks that require huge routing tables and routing that requires very complex policies.

To achieve these results design tradeoffs were made. Since scalability was the top priority it was necessary to allow for slower convergence than with other routing protocols.

TCP Transport was selected to keep BGP simple and stable but the tradeoff is higher CPU usage because updates must use a separate session with each BGP Peer.

## BGP performs three main Functions

1. Exchange of Network Reachability Information
    - Exchange of Network Reachability Information is the primary function of BGP
2. Policy Enforcement:
    - Policies are a set of rules which represent the AS's routing preference and constraints put on external traffic.
        - Path Selection: Ability to prefer a particular path.
        - Advertisement Control: Ability to control Advertisements
    - Reception Control:
        - Ability to control input of routes into BGP routing table
3. Interaction with IGP
    - Depending on the AS role (Stub, Transit) interaction with routing protocol is required to redistribute the BGP routes within the AS or for the AS to provide transit routing for BGP

## Roles of the BGP Protocol to perform its functions

The BGP protocol uses TCP using well known port number 179. There are 4 different BGP message types.

1. Open Message
2. Keep Alive Message

3. Update Message
4. Notify Message

These messages are used by the 4 different major functions of the BGP protocol.

- Opening and confirming a BGP session with its neighbor
  - BGP Open Messages are used to establish connection with peers.
- Maintaining a BGP connection
  - BGP Keep-alive Messages are used to maintain BGP connections between peers.
- Sending and receiving reachability information
  - BGP Update Messages are used for exchanging reachability information with peers (new and withdrawn routes)
- Notifying Errors
  - BGP Notification Messages are used to inform peer of any error condition or to explicitly close connection.

## BGP Topology Support

Based on Topology and connectivity to other AS's BGP provides support for the following type of AS's

- Stub AS: An AS that has a single connection to ISP or to another AS.

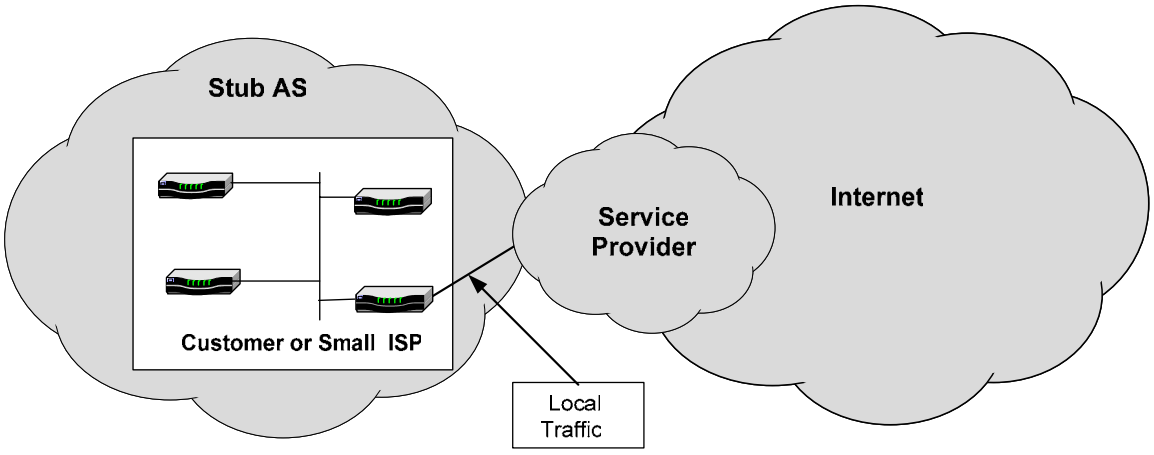


**Figure 1.5 Stub AS**

- Multi-home AS: Has connections to more than one other AS but will not carry transit traffic

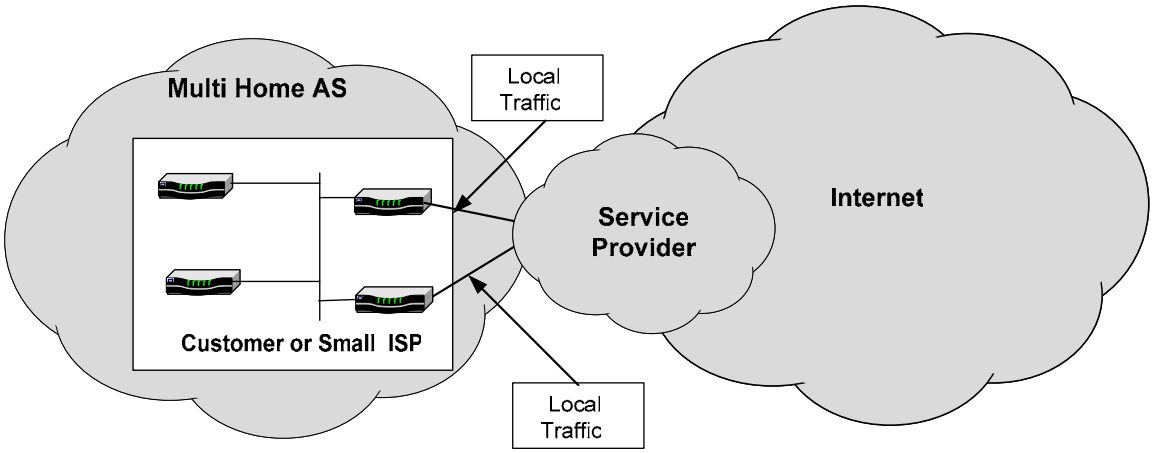


**Figure 1.6 Multi-Homed AS**

- Transit AS: Has connections to more than one other AS and will carry local and transit traffic.

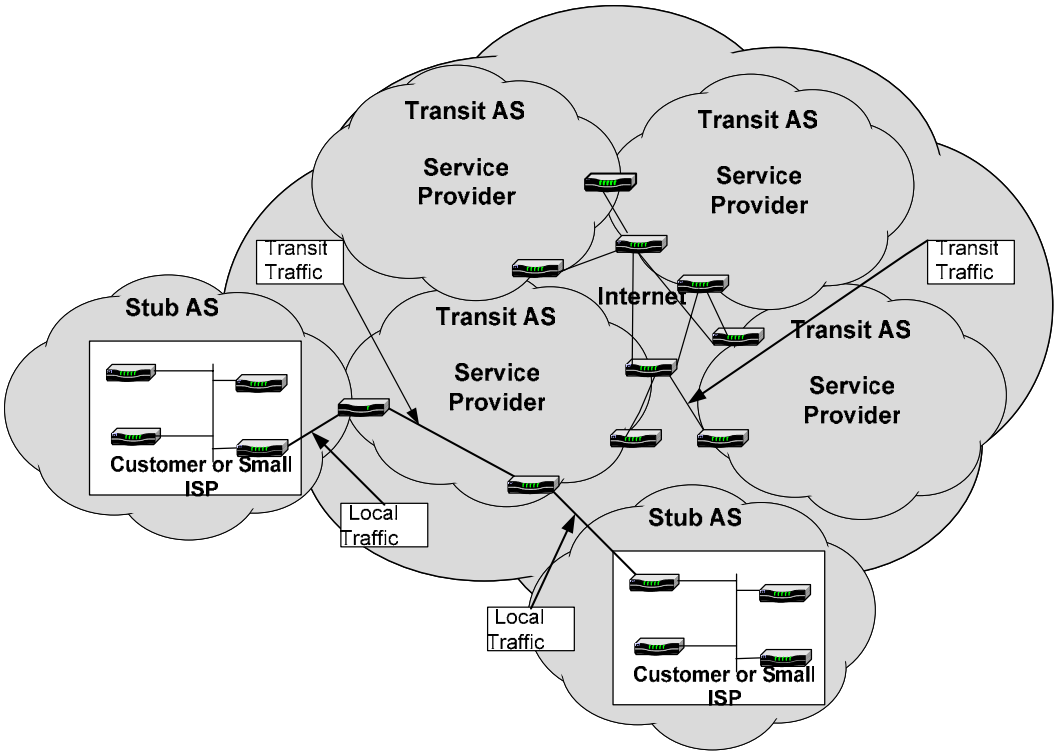


**Figure 1.7 Transit AS**

### BGP provides support for complex AS topologies

BGP supports very complex topologies with topology within the AS being hidden from neighbor AS's. BGP categorizes the AS traffic as either Local Traffic that is traffic that either originates or terminates in the local AS and Transit Traffic that is traffic that neither originates or terminates in the local AS. A stub AS is connected to only one other AS and will only handle local traffic. A multi-homed AS connects to more than one other AS and can either handle local or transit traffic but restricts traffic to local traffic only with policies. Multi-homed AS's are private companies with their own ASN connecting to multiple service providers. Transit AS's are service providers and handle both local and transit traffic. Routing policies in AS's with more than one connection to other AS's determine the types of traffic they handle. We will discuss policies in much greater detail later.

## BGP Protocol and Functions

BGP is a Peer to Peer Routing Protocol. Unlike other routing protocols such as OSPF and RIP all BGP Peers must be statically configured because BGP cannot dynamically acquire neighbors. The Peer Configuration must be done on both sides and the BGP Peer must be on a directly connected or a reachable network. BGP protocol runs on top of TCP using TCP well known port 179. Both routers will attempt to connect to the configured BGP peer using TCP port 179. The Source Address of the connect attempt will be matched against a list of configured neighbors if the connects from both Peers succeed one will be torn down allowing only a single TCP session to remain. In some cases the BGP Peer is configured for Passive Mode and will wait for its Peer to initiate the TCP connection.


**Figure 2.1 BGP TCP Session**

After a TCP session is established BGP uses 4 message types to communicate with its Peer. The Type field is 1 octet of the BGP header.

### *BGP Type Codes:*

The following type codes are defined:

1 - OPEN
2 - UPDATE
3 - NOTIFICATION
4 – KEEPALIVE

## Open Message

After TCP session in connected the first message sent by each side is an Open
Open messages contain:

- BGP Version number
- My Autonomous System Number
- Hold Time
- BGP Identifier
- Optional Parameters

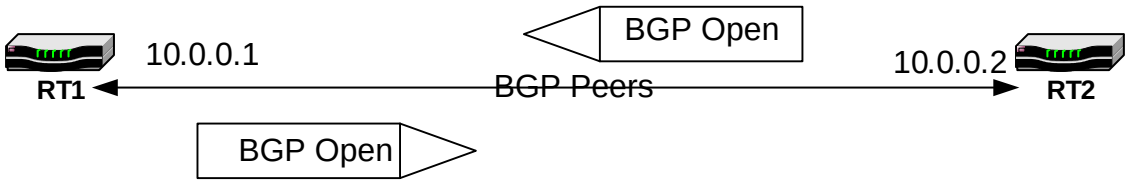


**Figure 2.2 BGP Open Message**

## Update Message

Update Messages are used to send or withdraw routes.

- Update messages can contain:
  - Withdrawn Routes
  - Path Attributes
  - Network Layer Reachability Information (NLRI)

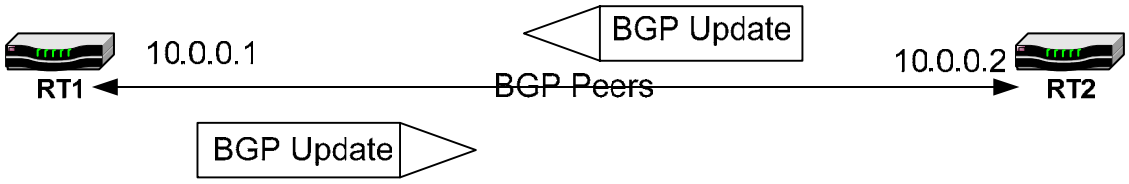


**Figure 2.3 BGP Update Message**

## Notify Message

A Notification message is sent when an error condition is detected. The BGP connection is closed immediately after sending it.

**Notification messages contain:**

- Error code
- Error sub-code

- Data

The following Error Codes have been defined:

| Error Code | Symbolic Name |
|---|---|
| 1 | Message Header Error |
| 2 | OPEN Message Error |
| 3 | UPDATE Message Error |
| 4 | Hold Timer Expired |
| 5 | Finite State Machine Error |
| 6 | Cease |

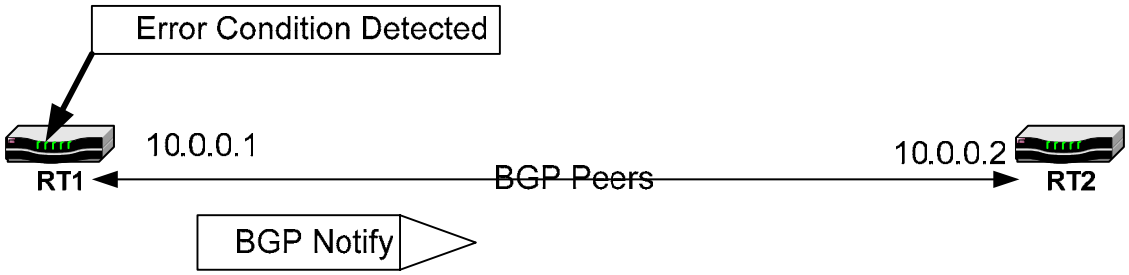


**Figure 2.4 BGP Notify Message**

## Keep Alive Message

KeepAlive messages are sent periodically to ensure the validity of the connection

- KeepAlive message consists of only a message header and has a length of 19 octets.

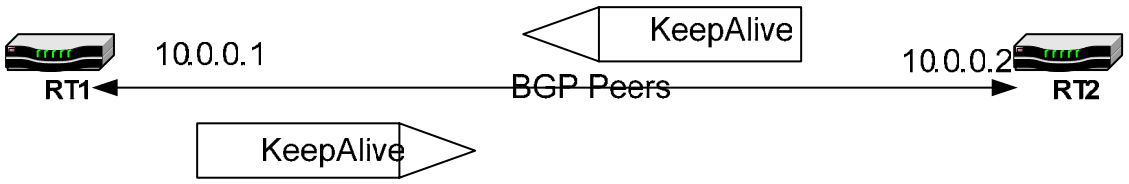


**Figure 2.5 BGP KeepAlive Message**

## *BGP Peer Finite State Machine*

A Finite State Machine (FSM) is maintained for each Peer session.

### BGP Peer States:

1 - Idle (BGP refuses all incoming calls waiting for start event)
2 - Connect (BGP waits for TCP connection to be completed)
3 - Active (In this state BGP is trying to acquire a peer by initiating a TCP connection)
4 - OpenSent (Open sent to Peer. Waiting for open from Peer)
5 - OpenConfirm (Open received from Peer waiting for Keepalive)
6 - Established (Exchange Update, Keepalive or Notification messages)

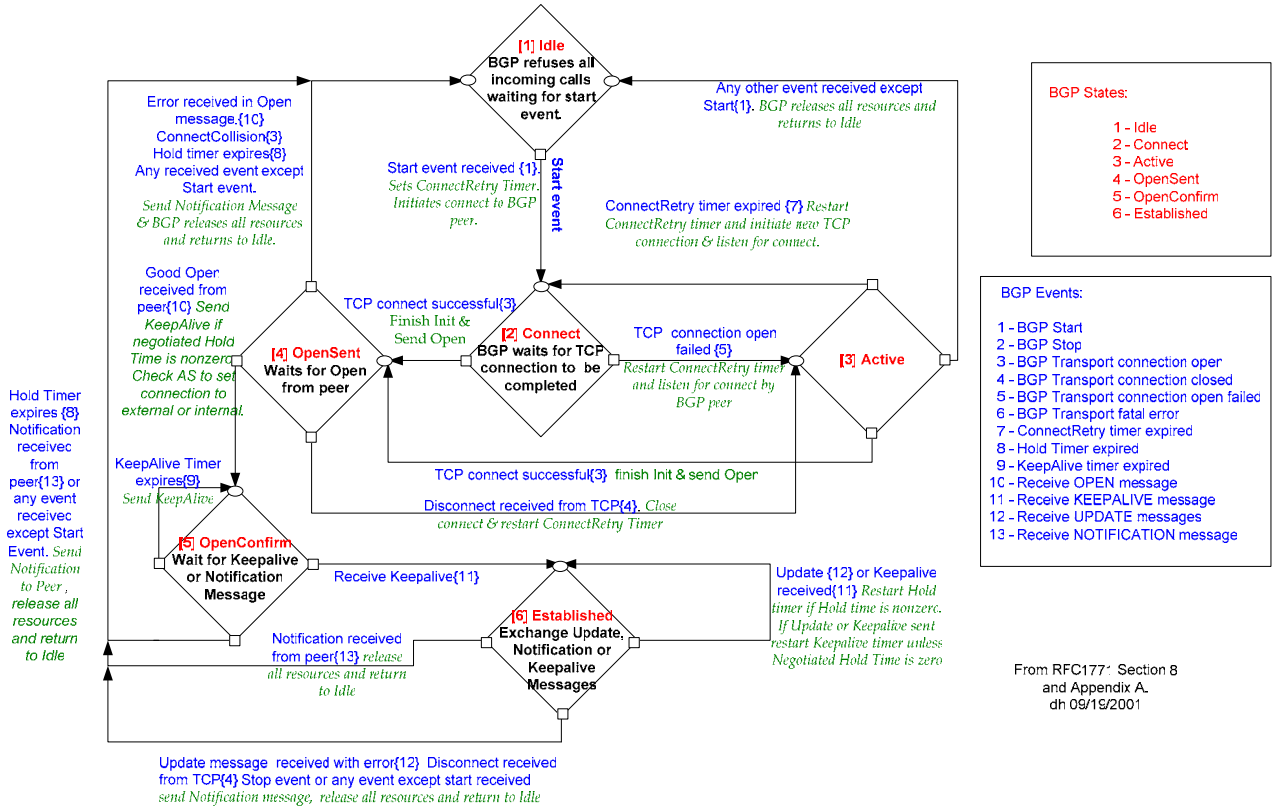


**Figure 2.6 BGP Peer FSM**

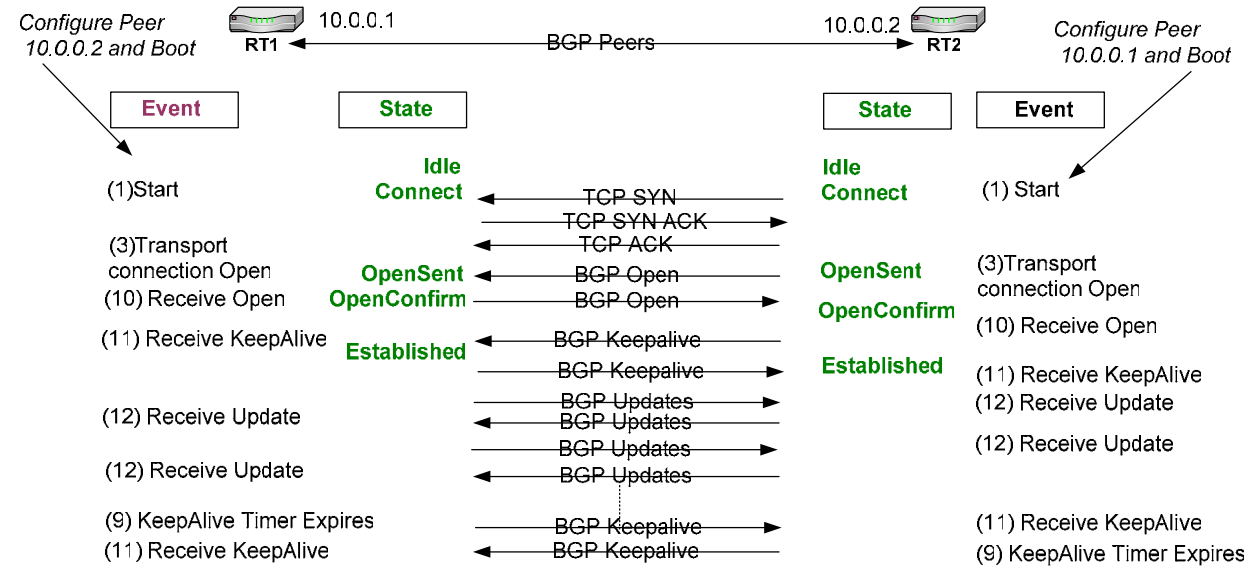


**Figure 2.7 Peer Session establishment**

The BGP Peer FSM will initially be in the Idle state waiting for a Start event. This will initiate the TCP session with the Peer. A normal 3 way handshake (SYN, SYN ACK, and ACK) occurs for the TCP session establishment during this time the Peer FSM is in the Connect State. Once this session is established a Transport Open Event occurs causing the Peer session to send a BGP Open Message and enter the Open Sent state. In this state it waits for an Open message from the Peer, once a good Open message is received it will enter the OpenConfirm State and send Keepalive messages to the Peer as well as waiting for a KeepAlive Message from the Peer. When the KeepAlive is received the Peer session enters the Established state. This is the desired state for BGP to start sending and receiving routing information. In this state Update messages are sent to either advertise or withdraw routes. KeepAlive messages are sent each time the keep alive timer expires to keep the session active. It remains in this state until one side or the other terminates the session. This is done with a Notify message. When a Notify is received the session is immediately terminated. Notifies can be the result of normal termination or error condition. The reason for the Notify is signaled in a code that is part of the message.

## *BGP Attributes*

Routes that are learned via BGP updates have associated fields that are used to pick the best route when multiple Paths exist to a particular destination from different Peers. These fields are referred to as BGP Path attributes. It is necessary for the user to understand of how BGP path attributes influence the route selection process to be able to

design robust networks. This section describes the attributes that BGP uses in the route selection process.

We are familiar with the metrics used in other routing protocols such as RIP and OSPF. These metrics are represented by a number with the larger the value the less desirable the route is. In BGP route metrics are called Path Attributes and are expanded beyond a single metric to decide the best route.

Path Attributes are categorized as:

- Well Known Attributes:
  Well Known Attributes must be supported by all BGP compliant implementations
- Optional Attributes:
  Optional Attributes are supported by some BGP compliant implementations but are not required

Well Known Path Attributes are divided into 2 classes

- Mandatory: Must be present in all update messages
- Discretionary: Optional attributes that may or may not be present but must be used if present.

*All Well Known Path Attributes are always propagated to other Peers*

## Well Known Mandatory Attributes

- Origin: Specifies origin of BGP route
  - IGP: Route was originated in IGP
  - EGP: Route was originated in EGP (Obsolete, used when route came from EGP)
  - Incomplete: Route redistributed into BGP from other than IGP source (Static Route)
- AS-Path: Sequence of AS Numbers through which the network is accessible.
- Next-Hop: IP Address of next hop router

## Well Known Discretionary Attributes

- Local Preference
  - Used to maintain consistent routing policy within the AS
- Atomic Aggregate
  - Used to tell Peer AS that a route has been summarized and that some information that was in originated routes may have been lost when updates were summarized into a single entry

## Optional Attributes

Optional Path Attributes are divided into 2 classes:

- Transitive: Must be propagated if the transitive flag is set regardless of if this router supports this attribute

- Non-transitive: Discarded if not recognized

**Transitive Optional attributes**

- Aggregator: IP address and AS number of the router that performed route aggregation
- Communities: Tag that can be used in downstream AS for filtering or route selection process

**Non-transitive Optional attributes**

- Multi_Exit_Disc (MED): Where multiple connection points exist with neighbor AS MED can be used to inform neighbor AS of preferred link. The path with lower MED will be preferred.

## BGP Origin Attribute

This attribute tell where the route came from at the source that originated the route.

- IGP: Route was originated in IGP for example the route came from RIP or OSPF in the IGP routing tables.
- EGP: Route was originated in EGP. This is obsolete since EGP routing protocol is no longer used.
- Incomplete: The route was redistributed into BGP from other than IGP source such as Static Route.

The purpose of the Origin attribute is for the DOP (Degree Of Preference) algorithm to be able to make the best routing decision when presented with the same route that have different origins. The path with IGP origin type would be selected first, then EGP and then Incomplete.

## BGP AS-Path Attribute

The AS-Path attribute shows the entire path of AS's traveled from origin of the route to the last sender. The attribute is empty when the local route is inserted in the BGP routing table. The senders AS number is pre-pended to the AS-Path when a routing update crosses an AS boundary. This is the boundary between Peers in different AS's. The receiver of the routing update uses the AS-Path to determine through which AS's the routing information has passed. If the AS that receives the routing information detects its own AS number in the AS-Path it will ignore the update to prevent routing loops

## Example of AS_Path loop prevention

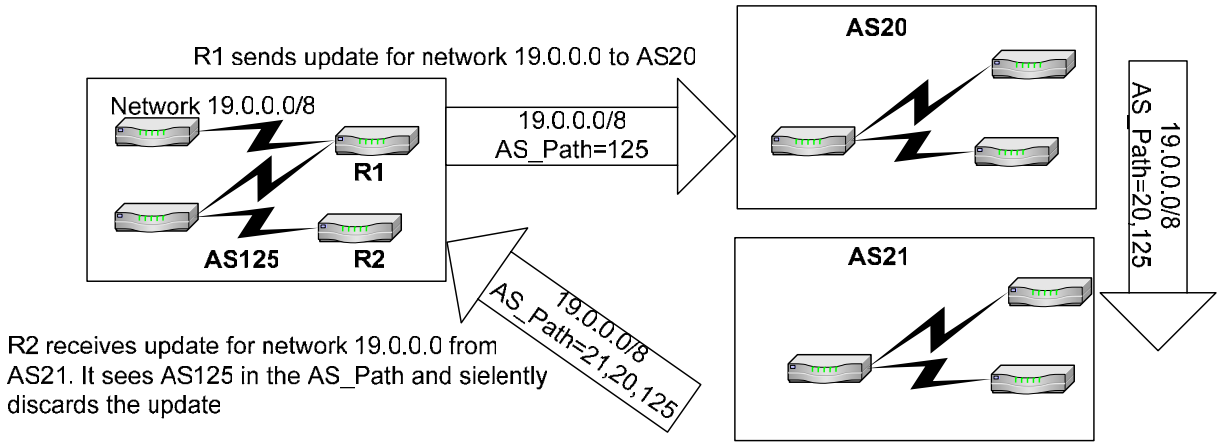


**Figure 2.8 AS-Path Loop prevention**

Figure 2.8 shows how loop prevention works. R1 in AS125 sends update to AS 20 with the AS path equal to 125. This means that the route originates in AS 125. AS 20 propagates this update to AS 21 prepending its AS to the AS-Path AS-Path= 20, 125. AS 21 propagates this update to R2 in AS125 prepending its AS number to the AS-Path. AS-Path= 21, 20, 125. When R2 receives this update it discovers its own AS number in the path and discards the update.

## BGP Next-Hop Attribute

This attribute indicates the next-hop IP address that the receiving router should use for packet forwarding. This is usually set to the address of the sending BGP router but it can be set to a third-party BGP routers IP address for optimal routing when peer and third-party router are in the same subnet.

Refer to Figure 2.9. This example shows the normal next-hop processing: Router R3 in AS20 sends an update to R1 in AS125 with the next hop set to its own IP address. (10.0.0.2) R2 in AS125 sends this update to R4 in AS21 and modifies the next hop attribute setting the next hop set to its own IP address. (10.0.1.1)

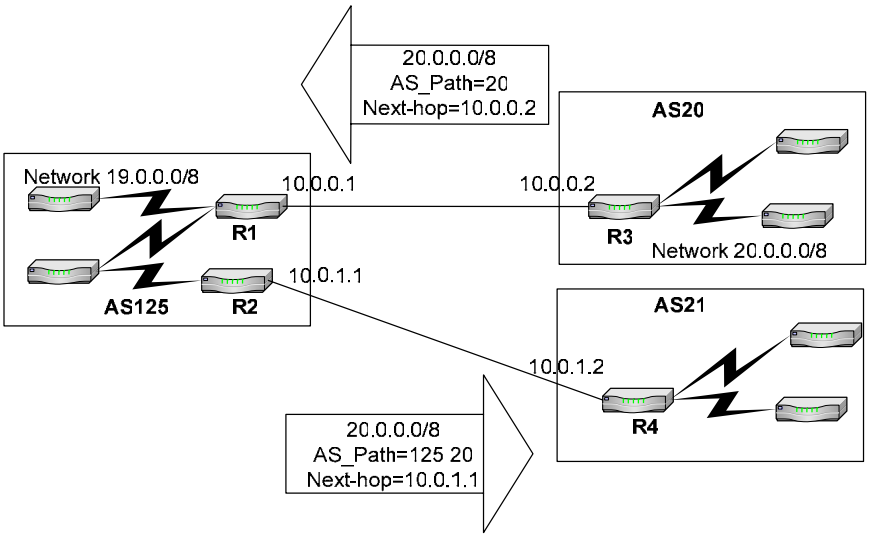


**Figure 2.9 Next Hop Attribute**

However If the peer router is on the same subnet as the current next-hop the next-hop address is not changed.
Refer to Figure 2.10. In the example R1 receives an update from AS20 with the next hop set to 10.0.0.2. When passing the update to its peer in AS21 it sees that the receiving peer is in the same subnet as the next hop in the update. It leaves the next hop unchanged to prevent local routing of packets from AS21 to AS20. Local routing is when a router must forward a packet through the same interface that it was received

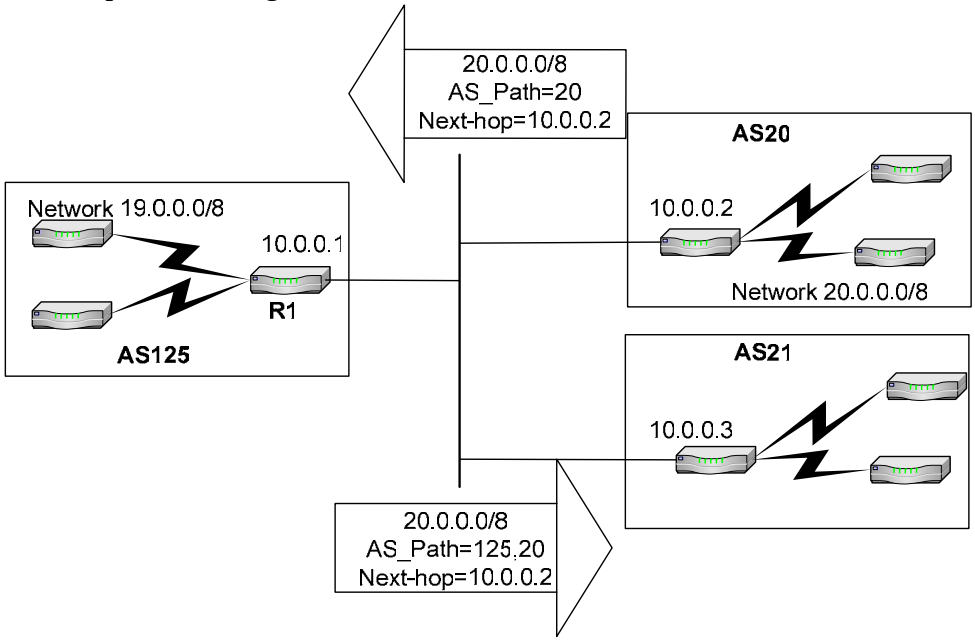


**Figure 2.10 Next Hop Attribute on common subnet**

This rule for process the Next Hop attribute produces special problem for NBMA partial mesh networks. See Figure 2.11 below. Although in the same subnet R2 cannot directly access R3. Both routers must go through the hub router R1 to access the other. If BGP next hop rule was used on this network R3 would be unable to forward packets to 10.0.0.2. To avoid this problem on NBMA LAN-View of the WAN addressed partial meshed networks. The BGP Peer Connection should not use the directly connected NBMA IP addresses. Instead use the Internal address as the BGP Peer address and enable BGP Peer "Parameter Indirect BGP Peering=Allowed" or use the Policy filter to set Next-hop not allowed. When Next-Hop Not Allowed is used the algorithm processes the outbound update as though this were not a common subnet and the next hop on the update R1 to R3 would be 10.0.0.1.

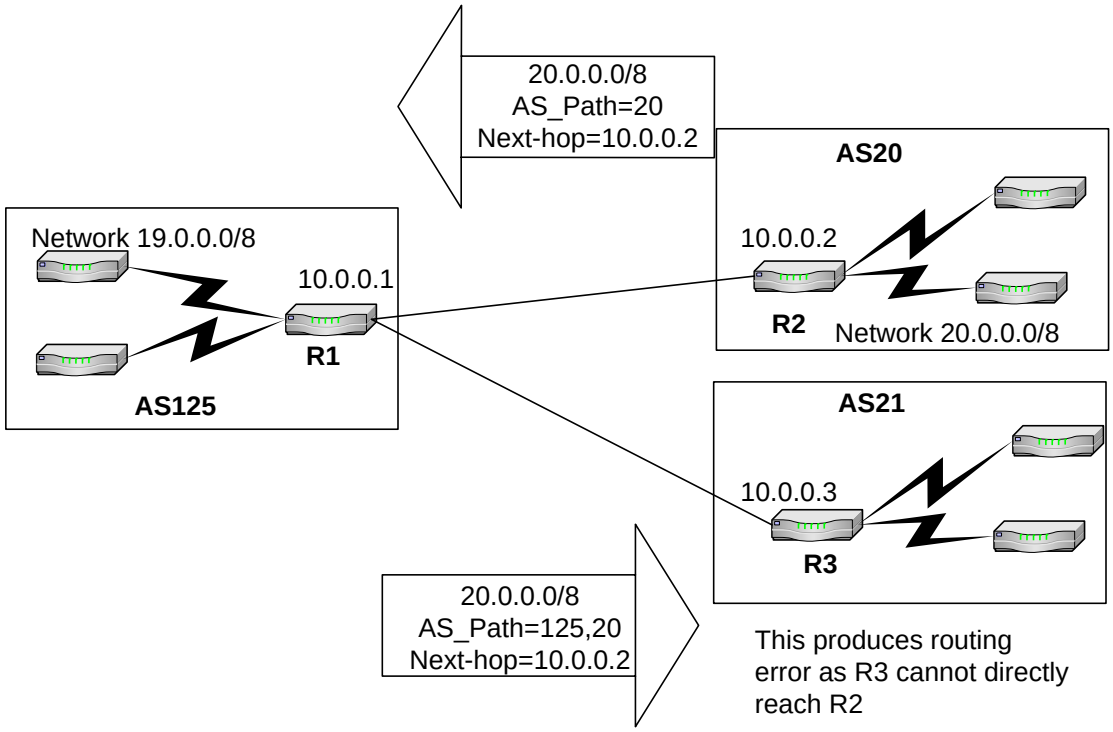


**Figure 2.11 Next Hop Attribute on NBMA common subnet**

## MED and Local Preference Attributes

MED and Local Preference are 2 attributes that are used to influence routing decisions in neighboring routers. MED is used to influence a router in a different AS and Local Preference is used to influence a routing decision in a router in my AS.

- MED is used to tell routers in the Neighbor AS the path that you prefer that they use.

- Local Preference is used to tell routers in your own AS the path that you prefer that they use.

Please refer to Figure 2.12. MED and Local Preference are 2 attributes that can be configured as Peer parameters for AS10 Peers that will influence the routing so that the high speed link is preferred over the low speed link. MED is configured in R1's Peer parameters with R3 to a value of 1. MED is configured in R2's Peer parameters with R4 to a value higher than 1. The lower value of MED on all updates received from R1 will indicate the preferred path. Local Preference is also configured in R1's Peer parameters with R3 to a value higher than 0. All updates that R1 propagates to internal peers that were learned from R3 will have this local preference attribute and will indicated the preferred path.

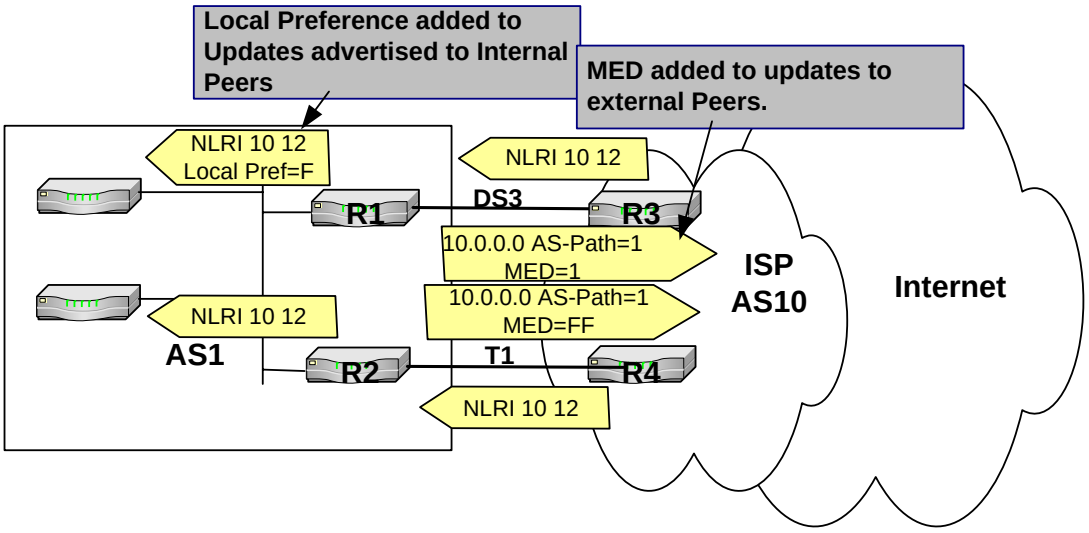


**Figure 2.12 MED and Local Preference Attributes**

## Extra AS Path Pre-pending

Another technique for influencing a neighbor AS to prefer one path over another is to use the peer parameter Number of Extra AS prepends. If AS path is prepended several times on one path and not on another the shorter path will be preferred. The difference is that MED will not be passed to a third AS where AS path prepending can influence AS's many hops away.

Please refer to Figure 2.13. In this example we see AS1 Multi-Homed into 2 ISP's AS10 is the preferred provider with AS20 only being used as backup for email and some high priority Business to Business applications. R1 attaches local preference policy to routes learned through AS10 high bandwidth path so all the AS1 routers will prefer this higher

bandwidth path. Extra AS Prepends are used for all routes advertised to AS20 so all Internet providers should prefer that high bandwidth path through AS10.

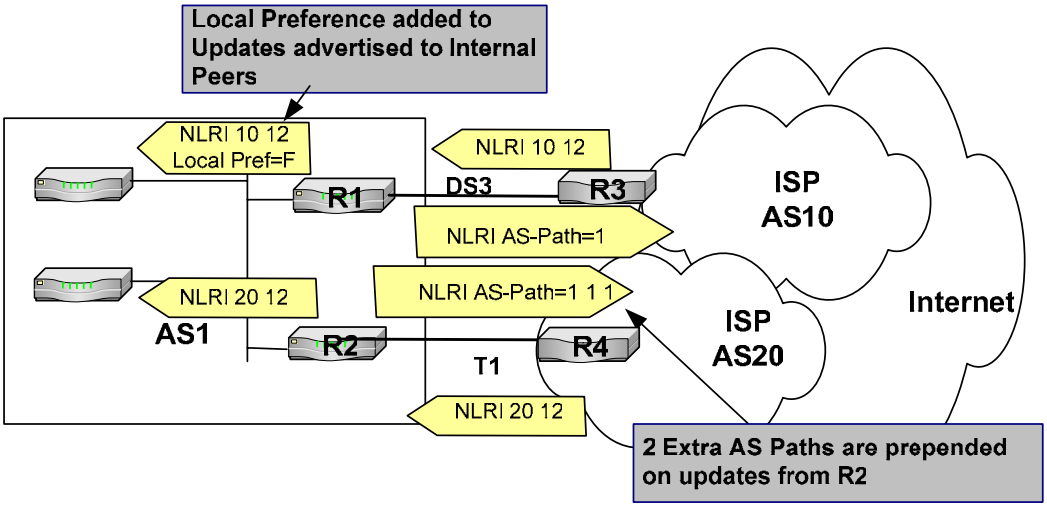


**Figure 2.13 Extra AS Prepending**

## Community Attribute

The Community Attribute is an Optional Transitive Attribute

- Set of 4 Octet Values
  - First 2 Bytes equal AS number
  - 2nd 2 bytes equal Community Value

A NLRI can have more than one Community Attribute. The BGP Speaker receiving the NLRI can act on one, some or all Community Attributes. Community Attributes do not alter the DOP of a route.

Community Attributes are used as flags to mark a set of routes to perform a common function such as:

- Add Local Preference Policy (Inbound)
- Add MED (Outbound)
- Add AS_Path Prepend (Outbound)
- Aggregate Routes

## Degree of Routing Preference (DOP)

Now that we have learned about the BGP attributes that influence routing decisions lets look at how the receiving router uses these attributes to select the best path. When a router processes a NLRI in an update it receives from the peer and it has another NLRI already in its routing database. It looks at the attributes in an order of preference and compares them to the other NLRI's it has in its database.

BGP in Vanguard Routers uses the following rules of preference when selecting the best path for a destination.

- Weight is a special attribute used to modify the DOP in local router only and is not passed in updates. However Weight is the first attribute looked at in DOP Processing. The Path is selected with the highest Weight is selected before other DOP processing is considered.
- Next the path is selected with the largest Local Preference.
- Next the path is selected that was originated by BGP running on this router.
- Next the path is selected with the shortest AS-Path Length
- Next the path is selected with the lowest origin type where IGP is lower than EGP and EGP is lower than Incomplete
- Next the path is selected with the lowest MED
- Next the path is selected for updates learnt from EBGP over IBGP
- Next the path is selected with the lowest BGP Neighbor ID IP Address

## *Differences between EBGP and IBGP*

Next we should look at the different rules that BGP uses when processing updates from a Peer within the same AS and when processing updates from a Peer in a different AS. When BGP Peers are in the same AS the protocol running between them is called Internal BGP (IBGP). When BGP Peers are in different AS's the protocol running between them is called External BGP (EBGP).

- IBGP: BGP session running between peers within the same AS
- EBGP: BGP session running between peers in different AS's

**Differences EBGP and IBGP**

- No AS-path is pre-pended in IBGP sessions. AS-Path is only pre-pended at EBGP border.
- BGP attributes are not changed in IBGP sessions (next-hop remains the EBGP next-hop)
- Route selection process will prefer EBGP route over IBGP route when AS-Path is the same. (equivalent route)
- Local preference attribute is sent only between IBGP peers. It will be removed at EBGP border.

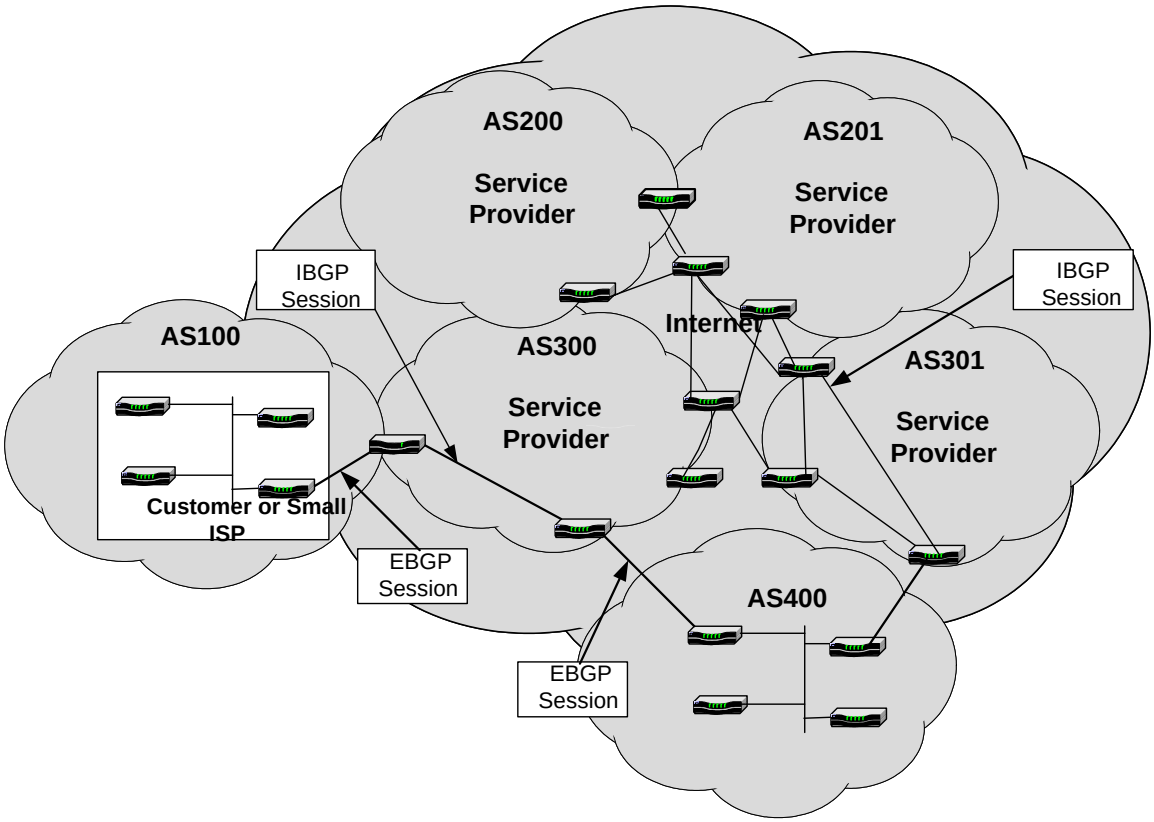


**Figure 2.14 IBGP and EBGP Sessions**

## IBGP and EBGP Loop detection

EBGP and IBGP also detect routing loops differently. IBGP uses a split horizon rule to prevent routing loops. That means IBGP split horizon prohibits any information learned in a IBGP session to be forwarded in another IBGP session.

EBGP uses AS-Path Loop detection to prevent loops. We have looked at this in a previous example. See Figure 2.8. If a NLRI received in a EBGP session contains the receivers AS number it is discarded.

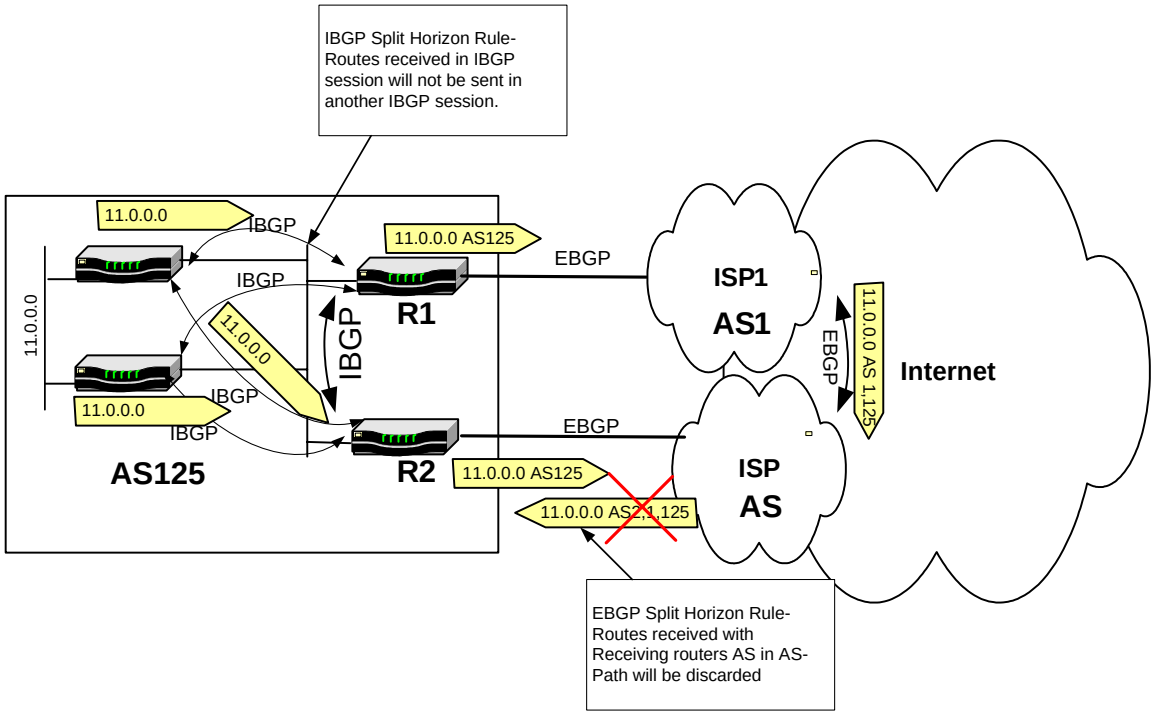


**Figure 2.15 IBGP and EBGP Split Horizon**

Note in figure 2.14 because of the IBGP split horizon rule that full mesh is required for peering sessions with other BGP speakers within the same AS. In very large AS's Route Reflectors and Confederations are used to overcome this requirement. Vanguard Networks routers support up to 128 peer sessions but does not support Route Reflectors and Confederations.

## IBGP Processing of MED

MED attribute received from neighbor AS will be propagated in IBGP sessions but will be stripped before being passed to a third AS in a EBGP session.

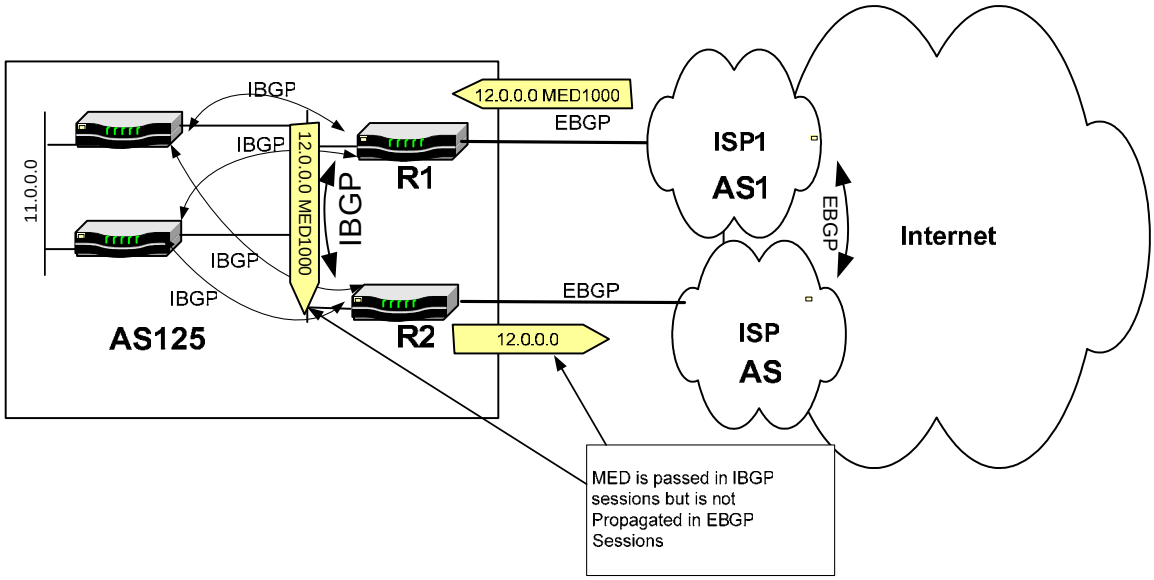


**Figure 2.16 IBGP and EBGP MED Processing**

## IBGP Path Processing in Transit AS

The AS-Path processing in IBGP sessions is also different than between EBGP Peers. AS-Path is only updated at AS Boundaries. So the AS path remains unchanged throughout the transit AS.

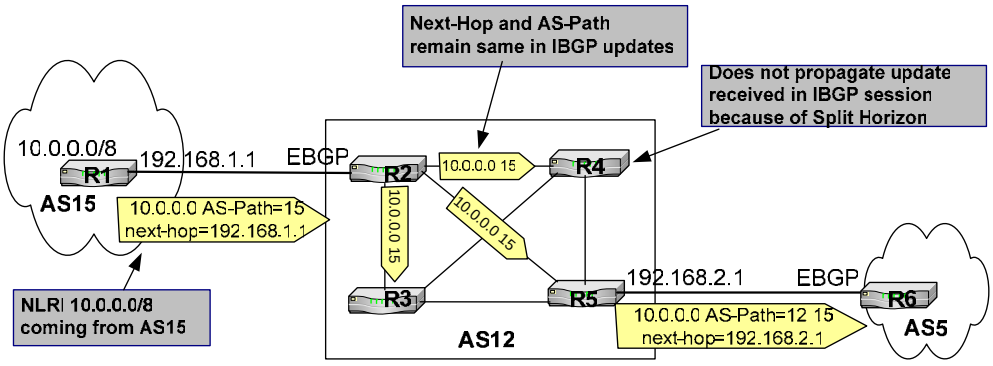


**Figure 2.17 IBGP AS-Path Processing**

Refer to figure 2.17. R2 receives NLRI 10.0.0.0/8 from R1 the AS-Path= 15. When R2 sends this NLRI to R4, R5 and R3 the AS-Path is unchanged because these are IBGP sessions. When R5 sends this NLRI to R6 its AS number (12) is prepended to the AS-Path because this is a EBGP session.

Mandatory Attributes Origin, AS-Path and Next-Hop are carried in all routing updates. As-Path is Pre-pended only when it crosses the EBGP boundary.

All routers within an AS must make a consistent decision about which exit point to use for a NLRI because of this the next hop field of the update remains unchanged as it is propagated within the AS. (The Next-Hop on update received by R2 is R1's IP address. (192.168.1.1) These attributes are propagated unchanged by R2 to R3, R4 and R5.) Because of the IGBP Split Horizon rule R4 and R3 will not forward updates received in IBGP session to another IBGP session. R5 will prepend its AS number when update is propagated to R6. IBGP is dependant on Interior Routing protocol such as Static routes, RIP or OSPF to resolve routing to the external next-hop.

## IP Routing in Transit AS

Because the Next-Hop attribute is not updated in IBGP sessions BGP is dependant on OSPF to find the correct routing through the Transit AS

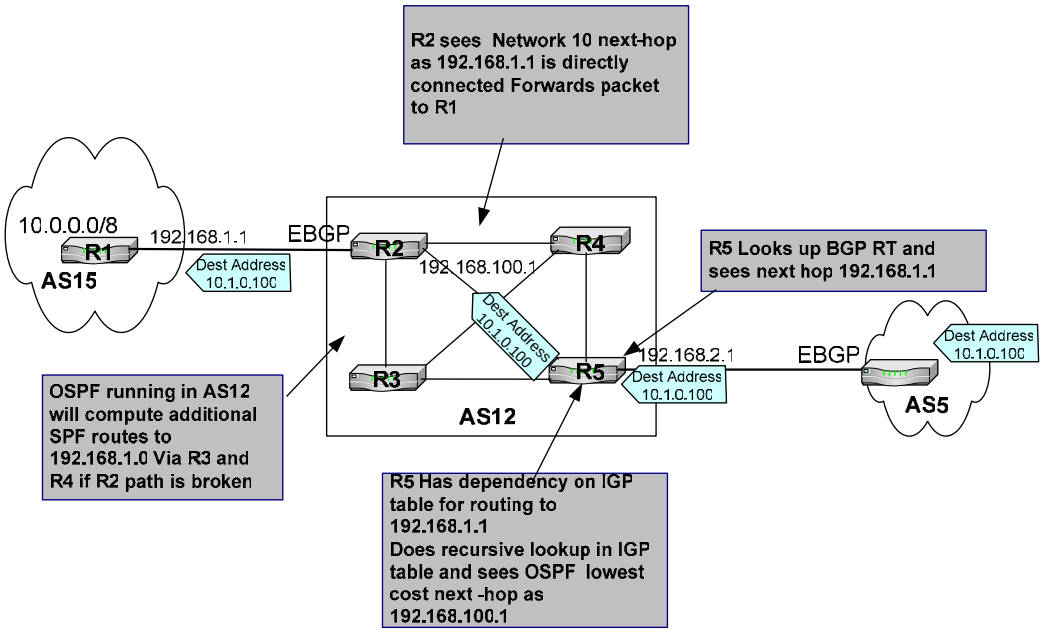


**Figure 2.18 IP routing in Transit AS**

This is best illustrated by at the routing through the network in figure 2.18 a packet is received by the AS5 router on the right with Destination Address of network 10.0.0.x because of its routing table it would forward the packet to R5.
R5 finds BGP route for network 10.0.0.0 via R1's Next hop address 192.168.1.1. Because R5 does not have a direct connection to Network 192.168.1.0/30, R5 must do recursive lookup for network 192.168.1.0/30 and would forward the packet to R2.
R2 finds BGP route for network 10.0.0.0 via R1's Next hop address 192.168.1.1 R2 is directly connected to 192.168.1.0/30 and forwards the packet to R1

It is not necessary to redistribute the BGP into the IGP routing protocol. The IGP supports BGP to find best path through the AS to the exit point and to support redundant paths through the AS. If static routes were used and path of R2<->R5 were unavailable the routing would fail unless higher cost floating static were configured.

## *BGP Policies*

Policies are a set of rules that determine the AS's routing preferences and constraints put on external or internal traffic.
BGP can provide for the following type of policies:

- Path Selection: Provides the AS the capability to prefer a particular path to a destination.
- Advertisement Control: Provides the AS the capability to control the advertisement of BGP learned routes to adjacent AS's.
- Reception Control: Provides the AS the capability of controlling the reception of advertisements it receives from the other BGP speaker in an adjacent AS.

The support for policies to determine which routes are advertised from a peer, which routes are accepted from a peer and which routes are redistributed to and from the IGP routing domain is what makes BGP very powerful and flexible. The architecture of BGP was designed to easily implement routing policies by dividing the Routing Information Bases (RIB) into 3 distinct parts.

- Adj-RIBs-In: The Adj-RIBs-In store routing information that has been learned from inbound UPDATE messages. Their contents represent routes that are available as an input to the decision process.
- Loc-RIB: The Loc-RIB contains the local routing information that the BGP speaker has selected by applying its local policies to the routing information contained in its Adj-RIBs-In.
- Adj-RIBs-Out: The Adj-RIBs-Out store the information that the local BGP speaker has selected for advertisement to its peers.

.

BGP Inbound policies determine what information is accepted into the Loc-RIB. BGP Outbound policies determine what routing information from the Loc-RIB will be put into the Adj-RIB's Out to be advertised to Peers.

## Vanguard Networks routers Policy support

The following is supported in BGP policies in Vanguard routers:

- **AS-Path Filtering Policies**

    - Support for Inbound and Outbound BGP policy and Importing BGP into OSPF and RIP
- **Net Prefix Control Policies**
    - Support for Inbound and Outbound BGP policy and Importing BGP into OSPF and RIP
- **Policy Scope**
    - Peer Scope and Peer List allow policy to apply to specific or groups of peers
- **Policy Action**
    - Deny or Permit: Can be applied to Inbound or Outbound BGP routes
- **Default Policies**
    - Defines default Inbound, Outbound BGP policies for Peers and Default IGP Import Policy

## BGP Inbound Policy

The following previously discussed policy parameters apply to Inbound Policies

- AS-Path Policy
- Net-Prefix Policy
- Policy Scope
- Policy Action
- Default Policy

Additional Inbound Policy Parameters

- Community Profile: Criteria matching specific Community Profile
- Path Weight- Allows the local BGP Speaker to associate weight/preference to route matching above criteria
- Local Preference Policy- Allows the local BGP speaker to set the value of Local Preference to route based on match of above criteria

## BGP Outbound Policy

The following previously discussed policy parameters apply to Outbound Policies

- AS-Path Policy
- Net-Prefix Policy
- Policy Scope
- Policy Action
- Default Policy

Additional Outbound Policy Parameters

- Community Profile: Criteria matching specific Community Profile
- MED Policy- Allows the local BGP Speaker to associate MED attribute to route advertised matching above criteria
- AS Prepends Policy- Allows the local BGP speaker to advertise the route based on match of above criteria with additional Local AS numbers prepended to AS-Path attribute

## BGP Routing Policy Example

There are many ways to use policies to balance bandwidth usage, increase security or hide networks to those outside the AS. Routing policy determines how the BGP speaker determines which routes it receives from and sends to BGP peers or other routing protocols. Routing policy consists of filtering routes, accepting certain routes, accepting and modifying other routes, and rejecting some routes. Think of routing policy as a way the BGP speaker controls the flow of routes into and out of the system.

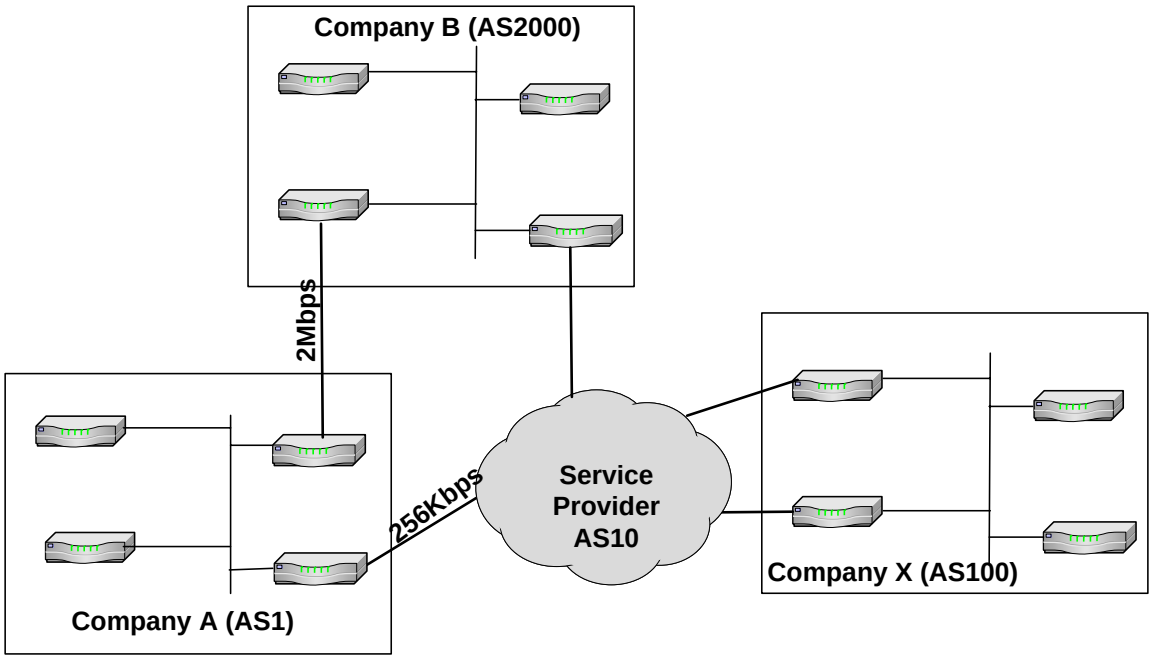


**Figure 2.19 Routing Policy example**

Refer to figure 2.19 Company B has high bandwidth links both to Company A and to the Internet through the Service provider. To insure that it does not become a backup path to the internet for Company A in case of link failure Company B can enforce policies in BGP to only advertise routes to Company A that originated in AS2000 and would not advertise any AS1 routes to the service provider. It is simple to do in BGP by using specified AS-Path in routes advertised to Company A and the Service Provider.

## Community Selection Profile

The NLRI selected by previous Inbound and Outbound Policies can be further filtered by the community attribute.
Community is 4 byte field.

- Bits 0-15 used for AS Number
- Bits 16-31 used for Community Value

Configure BGP Community Profile:

- AS Number: 1-65534, *, MY_AS

Specify the AS Number part of a Community Tag.

- MY_AS, the default value, will be translated to AS Number configured in BGP Global Parameter.
- Community Value: 1-65535, *

  Specify the value part of a Community Tag.

Using a wildcard, (*) in AS Number and Community Value is used for delete community profile only.

## Community Policy Parameters

Community Policy will match, delete or append using community profile numbers as criteria. Match Community Profiles apply to Inbound Policies only. Delete Community Profiles and Append Community Profiles apply to Outbound Policies.
The range of values is:

- None, Community Profile #
- Well-known community attributes
  - NO_EXPORT (0xFFFFFF01)*
  - NO_ADVERTISE (0xFFFFFF02)*
  - NO_EXPORT_SUBCONFED (0xFFFFFF03)*
  - NO_PEER (0xFFFFFF04)**

*Note:*RFC 1997 **RFC 3765*

More than one profile can be used in a Policy. Example: 1-3, 32, NO_EXPORT

## Well-known Communities

The following communities have global significance and their operations shall be implemented in any community-attribute-aware BGP speaker.

- NO_EXPORT (0xFFFFFF01) All routes received carrying a communities attribute containing this value MUST NOT be advertised outside a BGP confederation boundary (a stand-alone autonomous system that is not part of a confederation should be considered a confederation itself).
- NO_ADVERTISE (0xFFFFFF02) All routes received carrying a communities attribute containing this value MUST NOT be advertised to other BGP peers.
- NO_EXPORT_SUBCONFED (0xFFFFFF03) All routes received carrying a community attribute containing this value MUST NOT be advertised to external BGP peers (this includes peers in other members autonomous systems inside a BGP confederation).
- NOPEER, allows an origin AS to specify that a route with this attribute need not be advertised across bilateral peer connections.

## Community Example

Create Local Preference Based on Community

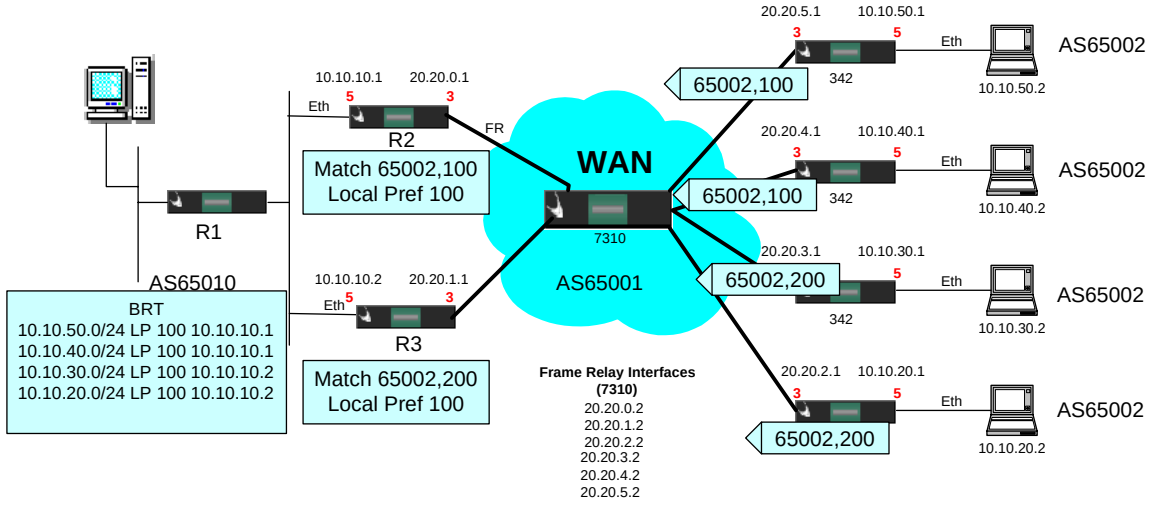


**Figure 2.20 Community example**

**Remote Routers:**

Step 1: Create BGP Profile #1 My_AS, 100 or My_AS, 200
Step 2: Create Outbound Append Community Policy to use Profile 1

This appends Community attribute 65002,100 or 65002,200 to advertised NLRI's

**R2**

Create Inbound Policy to set Local Precedence to 100 for all NLRI with match to Community 65002,100

**R3**

Create Inbound Policy to set Local Precedence to 100 for all NLRI with match to Community 65002,200

**R1**

- R1 automatically computes DOP for routing to 10.10.50.0, 10.10.40.0 first via R2
- R1 automatically computes DOP for routing to 10.10.30.0, 10.10.20.0 first via R3

## BGP Aggregation

The Internet routing tables are huge in some cases Internet core routers support over 200K routes. Without employing aggregation wherever it can be employed the size would be unmanageable. CIDR is employed by BGP to reduce the number of routes in BGP internetworking domain. BGP Aggregation creates aggregate route if the next hop and MED of component routes are equal.

BGP Aggregation can help to:

- Reduce the number of routes in BGP routing domain.

- Implement different traffic engineering strategy. Most suitable data paths can be designated to different traffic streams.
- Support gradual migration when changed from one ISP to another. Versatile control of component routes and route advertisement can help the gradual migration of IP addresses.

### Internet Aggregation Example

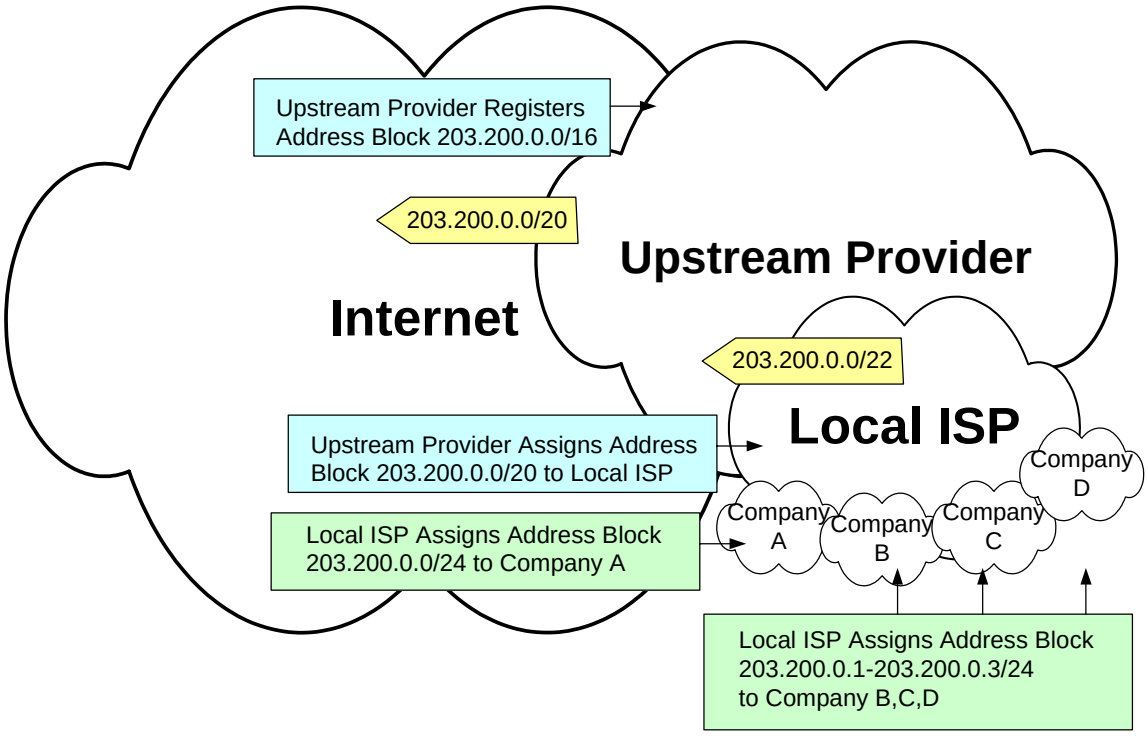


**Figure 2.21**

**Internet Aggregation example**

The Upstream Provider Registers CIDR address Block 203.200.0.0/16 with Regional Addressing authority. This provider assigns part of the block (203.200.0.0/20) to a Local ISP. The Local ISP assigns 203.200.0.0/24 to Company A, 203.200.1.0/24 to Company B, 203.200.2.0/24 to Company C and 203.200.3.0/24 to Company D. It is in the control of the Local ISP and Upstream Provider on what portion of the assigned address space to advertise. The Local ISP may wish aggregate Company A, B C and D as 203.200.0.0/22 and aggregate additional CIDR Blocks as they are assigned. Likewise the Upstream Provider may wish to advertise only assigned CIDR Blocks. It is current best practices for Major Internet providers to only advertise 20 bit CIDR blocks to the Internet. However there are still many thousand of 24 bit CIDR blocks still in internet routing tables.

In Vanguard routers to make this feature work, the BGP Aggregation in BGP Global parameters menu should be enabled. BGP can aggregate the component routes originated by other BGP speaker if BGP Proxy Aggregation is also enabled.

To create an aggregate route the next hop and MED of component routes should be the same. Compared with aggregate route, the IP address of specific route falls into the range of aggregate. Not all specific routes are the component routes as next hop/MED could be different. All specific routes are grouped based the next hop/MED pair, Aggregation initially choose the group with most member routes. The attribute of the aggregate is generated based on the component routes.

- Attribute of aggregate route can be set automatically or manually.

Advertisement of component routes is controlled by configuration the component routes are not sent by default. They can be advertised by proper configuration. Check outbound policy setting if unsuccessful.

Since aggregate route is created based on the next hop/MED pair, the number of specific route in each group with the same next hop/MED pair can change. The changing will cause the attribute modification of aggregate route and advertisement of aggregate route. If number of component route is less than the number of different next hop/MED pair route group by 2 or more, the attribute will be changed. This is used to alleviate the route flapping of aggregate route.
Aggregate profile contains attribute setting method for aggregate route, either automatic or manual.

## BGP Aggregation Features in VGMS Router

Aggregation Profile Configurables:

- Aggregate Network- Range of IP Addresses
- More Specific component networks- Can include or exclude more specific networks with aggregate
- Proxy Aggregation- Aggregates a route not originated by its own AS
- Attribute of Aggregate route- Admin can change attributes for traffic engineering
- Specific Routes- Configure which specific routes are advertised
- Suppressed Routes- Configure which specific routes are suppressed

**Aggregation Profiles Support:** The BGP aggregation profiles allow the network administrator to control what kinds of routes are aggregated. The network administrator has several levels of controlling whether specific routes are sent out and proxy aggregation is enabled.

- **Configurable aggregate network:** The network administrator is allowed to configure the aggregate network. The aggregate network, which is a range of IP addresses, is used to match to more specific networks. If there is more than one specific network, which falls in the range of the aggregate network, the aggregate network is created and advertised to the other BGP peers.
- **Configurable more specific component networks:** The network administrator is allowed to configure a set of more specific component networks based on which

aggregate network is created. The network administrator can choose a set of more specific component networks or exclude a set of more specific networks for an aggregate network.

- **Configurable proxy aggregation:** Although proxy aggregation is not recommended in the Vanguard router, it is configurable. By configuring this parameter the network administrator has the right to control if a Vanguard router aggregates a route that is not originated by its own AS. Vanguard BGP routers can act as the aggregation proxy of a third party product.
- **Configurable attribute of aggregate route:** The network administrator is allowed to change the attribute value of the aggregate route to implement traffic engineering strategy.
- **Configurable specific routes:** The network administrator is allowed to configure which specific routes are sent out to other AS and which routes are suppressed. By configuring these parameters the users can handle complex network layout.
- **Configurable suppressed route:** The network administrator is allowed to configure which specific routes are suppressed.

**Aggregation and De-aggregation Support in Vanguard routers:** When a new BGP route is added to BGP routing table, the route is compared to the configured profiles to determine if an aggregate route should be formed or not. If an aggregate route is formed then a new aggregate route entry is created and added to routing table. The attribute of a newly formed route is filled accordingly. If a route is withdrawn from the routing table, then a de-aggregation algorithm is applied. If one specific route is corresponding to an aggregate route, then the aggregate route entry is deleted from the routing table.

**Full ATOMIC_AGGREGATE and AGGREGATOR Support:** In ONS versions 6.0-6.2, the Vanguard BGP implementation only supports receiving of ATOMIC_AGGREGATE and AGGREGATOR attribute. With the introduction of BGP aggregation in 6.3, ATOMIC_AGGREGATE and AGGREGATOR attribute is fully supported; that is, Vanguard send ATOMIC_AGGREGATE and AGGREGATOR attributes if needed.

**AS path attributes preserved:** By forming an aggregate route, the originating AS or the AS information before the proxy aggregation could be lost. By using more AS attribute information in the routing table and updates packet, Vanguard can preserve the AS path attribute successfully.

## MPLS VPN Aggregation Example

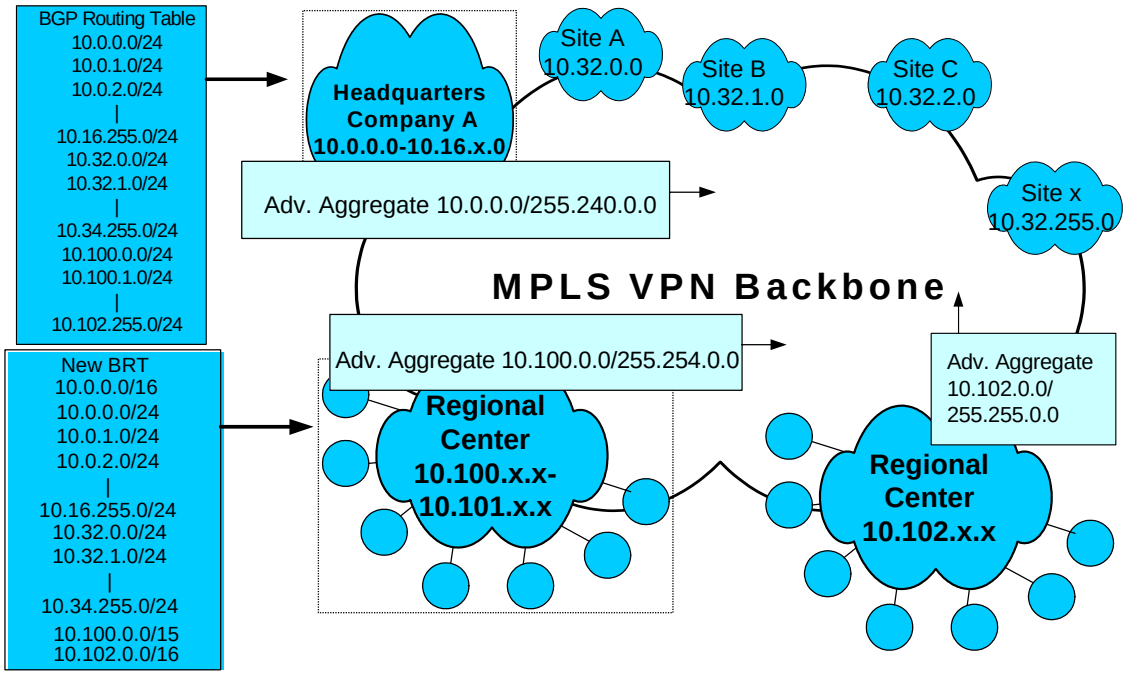


**Figure 2.22 MPLS VPN Aggregation example**

In some cases to keep routing table sizes to a minimum in remote routers it is necessary to aggregate routes in parts of the network. In this example the core network at Headquarters and Regional centers have large numbers of component networks. The Headquarters and Regional Centers aggregate the routes they are sending to the MPLS backbone. This has the effect of greatly reducing the size of remote VPN sites routing tables without implementing inbound policies.

## Aggregation Rules and Routes

Aggregation applies to routes which exist in the BGP routing tables and can be applied if more than one specific route of the aggregate exists in the BGP routing table. Aggregation adds a new level of complexity when deciding what routes to advertise to peers and how to forward packets because of a loss of information associated with route summarization.

The following general rules must be adhered to for aggregation to work properly.

- Longest match - forwarding is done based on the longest match (most specific route) found in the routing table.
- Aggregation can only be performed if (at least) more than one specific route exists in the routing table.
- The IP addresses must be assigned on hierarchical or topological lines for aggregation to have its optimal benefits.
- Destinations which are multi-homed relative to a routing domain can be explicitly announced into that routing domain.

- In order to prevent routing loops, routers that aggregate multiple routes must discard packets which match the aggregate route but do not match any of the explicit routes which make up the aggregate.

The following three situations describe the update options which must be supported.

1. **Aggregate Only, Suppress More Specific:** An aggregate is advertised and all of its specific routes are suppressed. This is usually done when the more specific routes do not offer any extra benefits such as making a better forwarding decision when forwarding traffic. An example of this is a single homed network
2. **Aggregate Plus More Specific Routes:** Situations exist in which it is beneficial to advertise the aggregate and its more specific routes. This usually occurs when a customer is multi-homed to a single provider. The provider would use the more specific routes to make a better decision when sending traffic towards the customer. At the same time the provider can propagate the aggregate only towards the NAP to minimize the number of routes leaked to the Internet. The use of this method allows the provider to balance the load to the customer. This is accomplished by sending different metrics for different routes on each of the links. In a backup situation with a primary/secondary topology, the forwarding of specific routes on the primary causes all traffic to be sent on this link unless the link fails and the routes are withdrawn from the routing domain.
3. **Aggregation With a Subnet of More Specific Routes:** In some situations a subset of more specific routes need to be advertised in addition to the aggregate. This is used to direct certain traffic to a specific AS such as in the case of multi-homed, geographically dispersed networks. This type of configuration allows the administrator to direct traffic to routes closer to the user in very large AS's. In some situations it is required that the attributes of an aggregate be changed.

## ORIGIN Attribute when aggregating routes

If at least one route among routes that are aggregated has ORIGIN with the value INCOMPLETE, then the aggregated route must have the ORIGIN attribute with the value INCOMPLETE. Otherwise, if at least one route among routes that are aggregated has ORIGIN with the value EGP, then the aggregated route must have the origin attribute with the value EGP. In all other cases, the value of the ORIGIN attribute of the aggregated route is INTERNAL (IGP).

## AS_PATH Attribute when aggregating routes

If routes to be aggregated have identical AS_PATH attributes, then the aggregated route has the same AS_PATH attribute as each individual route. When proxy aggregation is being performed (routes being aggregated from different AS's) then the AS_PATH attribute can be

- The AS number of the router doing the aggregation, if the type is AS_SEQUENCE.

- An unordered set of AS's that the aggregate has formed from with the aggregating routers, with its own AS number in the last position

The detailed information that existed in the specific prefixes are lost when summarized in the form of aggregates. The purpose of AS_SET is to try to save the attributes carried in the specific routes. Without AS_SET the aggregate formed is considered to have originated from the AS in which it was formed and all the specific attributes are lost. With the use of the AS_SET, type the specific attribute information is retained.

### Atomic Aggregate Attribute and Aggregator Attribute handling

- **ATOMIC_ AGGREGATE Attribute**
  The ATOMIC_AGGREGATE attribute is used to indicate a loss of AS_PATH information if an aggregate is formed. The sources of the aggregate can have different attributes. If a system propagates an aggregate that causes a loss of AS_PATH information then it is required to attach the ATOMIC_AGGREGATE attribute to that route.
  If a BGP speaker receives an UPDATE with the ATOMIC_AGGREGATE attribute set, it must not remove the attribute from the route when propagating it to other speakers.
- **AGGREGATOR Attribute**
  The AGGREGATOR attribute specifies the autonomous system and the router that has generated the aggregate.

## *BGP Redistribution*

Vanguard routers support BGP redistribution for both IGP routing protocols (OSPF and RIPV2). OSPF redistribution has been supported since ONS version 6.0. RIPV2 redistribution was added in version 6.3. Both work in a similar fashion so we will only cover BGP to RIPV2 in this paper.

### BGP to RIPv2 redistribution major functions

- **BGP to RIPv2 Route Redistribution Policy**
  - Each Policy is a single element of a Policy set
  - A set of single policies form the BGP-to RIP route redistribution
  - If a BGP route satisfies one single policy it meets the criteria
  - The route is imported if the action is Permit
  - If the BGP route satisfies no single policy it will use the default policy.
  - Default policy can be Permit or Deny

  BGP route redistribution policy is used to set the criteria that are used by the route importing process. If a BGP route satisfies the criteria, the route is imported to RIPv2 routing domain. A set of single route redistribution policy forms the router's whole policy of BGP to RIPv2 route redistribution. Each single policy is

an element of the router's whole policy set. Stating that BGP route satisfies the route importing criteria indicates that the BGP route meets the requirement of at least one single policy in the router's whole policy set. If a BGP route satisfies one single policy, then the route satisfies the criteria. This route can be imported into RIPv2 routing domain if the action type of the policy is Permit.
There are two different levels of policy in the router:
Non-default policy
Default policy
Both policies are elements of the router's whole policy set. The default policy exists in the node and works once BGP to RIPv2 route redistribution is enabled. Although you can change the parameters of default policy, it is not required to configure explicitly. The non-default policy is not stored in the router unless you configure it explicitly. If the network administrator wants to enforce more policies in traffic engineering, he must configure non-default policies to implement it. If both non-default policy and default policy exist in a router, non-default policy always takes precedence.
When the user configures the policy, several levels should be provided to the user. The user can choose what kind of routes are selected, based on the IP address and the AS number(s) in the AS path attributes of the route. The user should be able to choose to import the route type, such as, whether a specific route is selected when an aggregate route is there. The user should be able to choose the action on the selected route, either permit the route to enter RIPv2 domain or deny.
If a route is permitted to enter the RIPv2 domain, the user should have ability to set the metric and tag fields of RIPv2 route. The user can choose to set them to a fixed value to control the traffic flow or let the router calculate the value. For the tag field the user even could disable the value setting.

- **BGP to RIPV2 Policy Criteria**
  - IP Address: IP Address and mask determine route to match criteria.
  - Match Type: Exact or range determine if this is specific or range of addresses
  - Originating AS: Can match only routes originated from a specific AS
  - Advertising AS: Can match only routes received from a specific AS
  - Filtering Action: Permit or Deny
    ***The following parameters are configured if Filtering action is permit***
    - Import Specific Route: Import specific route that is part of aggregate
    - RIP Aggregative: Route can be combined with others to form aggregate
    - RIP metric: The metric to use for BGP imported routes 1-15
    - RIP Route Tag: Controls handling of Route Tag Field
    - Route Tag Value: 0, 1 to 65,535 when "RIP Route Tag" is set to Manual.

- **BGP to RIPv2 Route Importing**
  - Searches BGP Routing Table (BRT)
  - Selects Qualified Routes based on Policy Set
  - Translates from BGP to Global RT format
  - Sets the proper flags in Global Routing Table
  - Two importing Methods
  - Batch : All routes in BRT are checked
  - Incremental: Only the modified routes are checked

  BGP to RIPv2 route importing is a key component of BGP to the RIPv2 route redistribution feature. Route importing is responsible for searching the BGP routing table, selecting the qualified route entries, translating the qualified route entries from BGP format, putting them into the global routing table, and setting the proper flags for these imported entries. From the global routing table, RIPv2 can advertise the imported BGP routes in the local AS.
  There are two route importing methods:
    - Batch input
    - Incremental input

  Batch and incremental input methods are supported for the BGP to RIPv2 route redistribution feature. In batch input method all the route entries in BRT are checked. Based on the checking result, the qualified route entries are input in RIPv2 routing table in block-by-block manner until all the qualified routes are done. In each block, many routes are included (each block can contain as many as 24 qualified routes). The reason why batch input method is introduced is that the number of RIPv2 triggered updates can be drastically decreased compared to the incremental input method. In incremental input method only the modified (changed/added/deleted) routes are put into the RIPv2 routing table in one-to-one way until all the qualified routes entries are done. If the BGP routing table is very large, it's not a good idea to go through the whole routing table when only one entry is modified. Normally in a stable network, a route entry is not modified frequently. By utilizing the incremental input method, the router could avoid a lot of resource consumption.
  Not every BGP route is qualified to enter the RIPv2 routing domain. The BGP routes must satisfy a few predefined conditions. First, the BGP routes that were input from the local AS are not eligible to be re-input back to IGP again. That indicates the originating AS of a BGP route is equal to the configured local AS number then this route is not considered in route input process.

To qualify to be imported to RIPv2 domain, a BGP route must satisfy the following conditions:

- Not imported from the local AS
- Meet at least one single policy in the router's whole policy set

- Is an aggregate BGP route or is a specific route and specific route is allowed to be selected
- The action value of the matched policy is Permit

If a BGP route passes all the checks, then the action is used to determine if this route should be permitted to enter RIPv2 routing domain or should be denied. If a route is allowed to enter RIPv2 domain, a flag is set to indicate that a BGP route is input to RIPv2.
To add the imported BGP route through RIPv2 in the GRT, several RIPv2 fields have to be filled according the attributes of BGP routes:

- The IP address field (the destination IP address) is copied directly from the destination IP address field of BGP route.
- The subnet mask is copied directly from the mask field of the BGP route.
- The next hop is copied from the BGP route.

The metric is set according to the user's configuration:

- If the configured value is between 1 and 15, this value is copied to the metric field.

The route tag field is set according to user's configuration:

- If the configuration is Disable or Manual but has invalid value, set the value as 0,
- If the configuration is Manual and has a valid value, set the value as the configured one,
- If the configuration is AdvertisingAS/OriginatingAS_Automatic, set the value as the advertising/originating AS number.
    - If automatic is chosen then the number of the originating AS or of the advertising AS of the BGP route are copied.

The next hop field should be set as:

- IP address of the peer's router interface from which the best path of the BGP route is relayed, if the peer is IBGP peer
- or the IP address of the router interface from which the best path of the BGP route is received for all other cases. When RIPv2 receives these imported BGP routes, RIPv2 should not check from which interface this route came. RIPv2 should set the corresponding flag to indicate that this route was imported from BGP.

## • RIPv2 Aggregation

RIPV2 Aggregation will combine many specific routes into one advertised aggregate route.

- This feature exists in releases prior to 6.3
- New requirements added to aggregation algorithm
- User may select the imported route to not take part of the aggregation process
- If the imported route does take part in route aggregation the aggregate route carries the route tag of the specific route

- This is required for IGP-BGP route synchronization

RIPv2 route aggregation is used to combine many specific routes into an aggregate route entry. RIPv2 route aggregation is an existing feature in Vanguard routers. The introduction of BGP to RIPv2 route redistribution in Release 6.3 and greater, add new requirements to the RIPv2 route aggregation algorithm.
If the user set the imported routes as not eligible to be aggregated in BGP to RIPv2 route redistribution policy configuration, the imported routes do not take part in the route aggregation process. This check should be done when route aggregation is trying to form the aggregate route.
If the imported routes can take part in the aggregation process, the aggregate route carries the route tag of the specific route. This is required because of BGP-IGP route synchronization. The other BGP speaker in the local AS could not see the IGP route if aggregate route is used and BGP-IGP does not recognize the aggregate route. By carrying the route tag in the aggregate route, the other BGP speaker could determine if an aggregate IGP route can be considered as the IGP route or not.

- **RIPv2 Route Advertisements**
  - Normal RIP Request/ Response Packets
  - Imported Routes carry Tag Value
  - Other RIPV2 Routers carry tag in propagated routes
  - Other BGP Routers in AS use tag to determine in route is IGP route or not.

The RIPv2 Advertisement is used to send RIPv2 request and response packets. For the imported BGP routes the RIPv2 advertisement treats them as normal routes except that the tag value is copied into the packet and then the value is installed into the routing table.

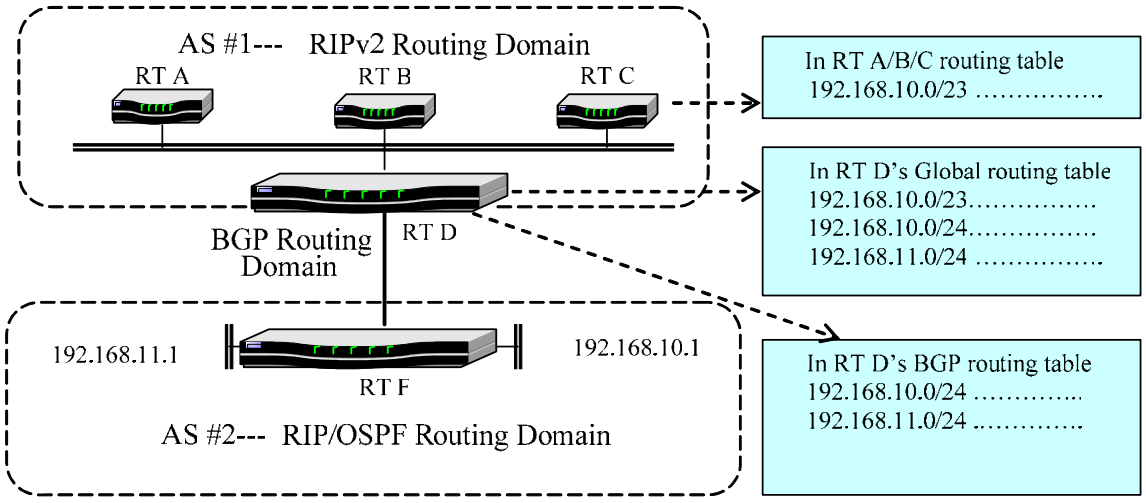


**Figure 2.23 BGP-4 to RIP Redistribution Example**

An example of the redistribution process is shown above. AS #1 and AS #2 are two IGP domains which are using RTD and RTF as their ASBR respectively.
RTF and RTD are using both BGP and IGP. RTF's IGP could be RIP or OSPF or a combination. RTD's IGP is RIPv2. RTA, RTB and RTC are all RIPv2 only routers in AS #1.
If BGP aggregation is not implemented in router RTF, RTF sends a BGP update packet to RTD which includes the 192.168.10.0/24 and 192.168.11.0/24. When RTD receives RTF's BGP update packet it puts 192.168.10.0/24 and 192.168.11.0/24 into its BGP routing table and lets RIPv2 take these two routes into GRT. If route aggregation is enabled in RT D, then another route 192.168.10.0/23 is formed as an aggregate route. RIPv2 advertises these route entries to other routers in the AS, either as 192.168.10.0/24 and 192.168.11.0/24 or 192.168.10.0/23 depending on the configuration in the corresponding interface.

## *BGP Indirect Peer Load Balancing*

Release 6.5S150 of Vanguard ONS supports Load Balancing between multiple links (8) between 2 Peers. To load balance multiple links a single BGP session is established between the loopback address or internal address of the 2 routers. An Equal Cost Static route is mapped to each peer router loopback address for each Link IP address as Next Hop.

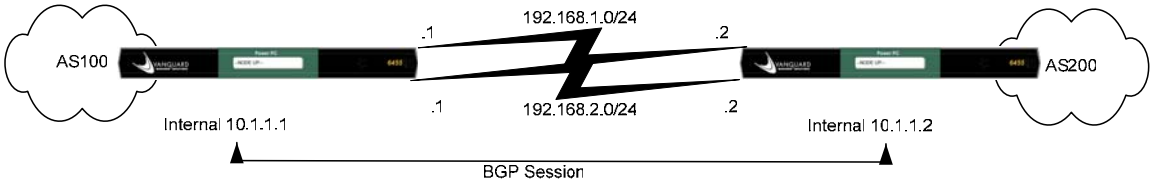


**Figure 2.24 BGP Load Balancing**

Refer to figure 2.23. BGP load balancing is achieved using the 2 T1 paths using equal cost static route load balancing. (Up to 8 paths are supported) A BGP Peer session is configured between the internal addresses in AS100 router (10.1.1.1/32) and AS200 router (10.1.1.2/32). A single BGP path is established between these 2 routers. The next hop for all routes received from AS200 router in the AS100 routers BGP routing table will be 10.1.1.2. The next hop for all routes received from AS100 router in the AS200 routers BGP routing table will be 10.1.1.1.

In the AS100 router 2 static routes for 10.1.1.2/32 will be added with equal metric with next hop being 192.168.1.2 and 192.168.2.2. When a match occurs for BGP routing entry during lookup in the BGP table with next hop of 10.1.1.2 a recursive lookup occurs because 10.1.1.2/32 is an indirect IP address. During the recursive lookup for 10.1.1.2 the next hop will be assigned for the oldest used static entry. This lookup is placed in IP cache and all subsequent packets for this session will follow to the same next hop. This is known as session based load balancing.

## *BGP Passive Peer*

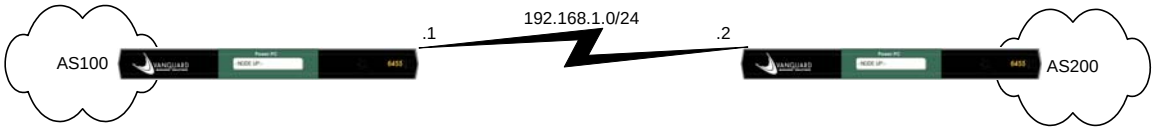


**Figure 2.25 BGP Passive Peer**

During BGP session connection collisions can occur where both Peers attempt BGP Open at the same time. When this Open collision occurs both peers back off and retry the Open. This can cause delay in getting BGP session established and the subsequent exchange of routing information. This feature allows the Peer to be enabled in Passive mode. If configured as Passive the router will not attempt a BGP open but allow the Peer router to initiate the BGP session. If configured for Passive the other router must be configured for Active Mode.

# Routing in IP Enabled BGP/MPLS VPN

The number one application for BGP in Vanguard routers is as the routing protocol in a BGP/MPLS VPN. This service is available from most Service Providers worldwide.

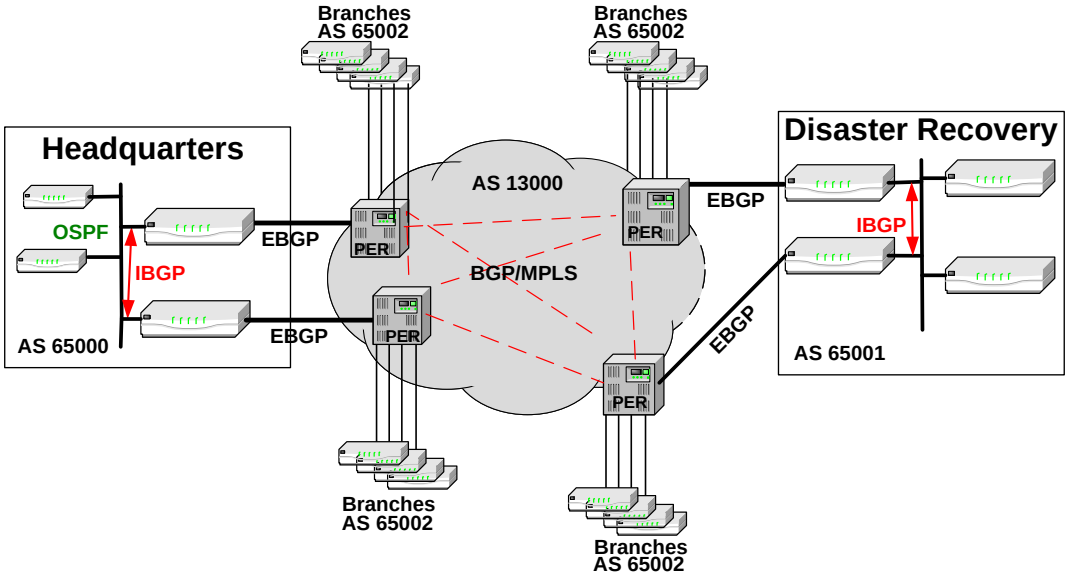


**Figure 3.1 BGP/MPLS VPN**

This is a service provided by Service Providers such as AT&T or Embratel. Customers have been using Frame Relay or ATM VPN's for some time and leasing part of a shared service (bandwidth between switches) from the service provider. In this new offering FR or ATM is connected to the edge of the provider network and the provider switches the traffic based on IP addresses. Usually BGP is required for the branches and all branches use the same AS number and advertise their routes to the PER. The PER overrides the Branch AS number with its own so all routes except for the DR site and the HQ site will have an AS-Path of AS13000. Normally this type of network will only require BGP to run at the branches if the application is making secondary routing decisions or if triggering backup on loss of BGP.

In this network example the application will send packets to the HQ site 10.1.x.x. If the HQ site is down the disaster recovery site will advertise network 10.1.x.x. Branches also may directly access another branch. Branches can implement policies to only accept default BGP route from the network. If the BGP route is not present the branch will use dial backup triggered by a static default gateway. All PER routers will have full routing table for the VPN.

HQ and DR sites are using IBGP between its access routers to understand all BGP exit points from the AS for redundant routing. BGP is redistributed into OSPF for load balancing.

## *Development of MPLS*

Developed by IETF by combining features of:

- Cisco's Tag Switching
- IBM's Aggregate Route-based IP Switching (ARIS)
- IP Switching (Ipsilon)
- IP Navigator (Lucent/Ascend)

MPLS Requirements

- Simple, Efficient high speed forwarding of IP packets
- Scalable to very large networks
- QOS (Routing based on service class)
- Traffic Engineering (Control of Network Bandwidth)

MPLS was developed as an standards based answer to many incompatible switching protocols from several vendors. MPLS has combined features from Cisco, IBM, Ipsilon and Lucent to give a standard way that vendors can exchange tagged frames for high speed switching of IP frames. MPLS is compatible with many layer 2 protocols and in some cases such as Frame Relay and ATM uses the layer 2 header as part of the MPLS label.

**MPLS Basics: The Components**
The MLPS Architecture is fully described in RFC-3031

**Label Switch Routers** (LSR's) are High Speed router devices at the core of the MPLS network that participates in establishing Label Switched Paths (LSP's) using the appropriate label signaling protocol and high speed switching based on established paths.

**Label Edge Routers** (LER's) are devices that operate on the edge of the access network and MPLS network. LER's support multiple ports connected to dissimilar networks. These routers establish LSP's using the label signaling protocol at the ingress and distribute traffic back to the access network at the egress. The LER plays an important role in the assignment and removal of labels.

**Forwarding Equivalence Class** (FEC). A group of packets that share the same requirements for their transport.

**Label**: Identifies the path a packet should traverse. Label assignment may be based on Destination routing, Traffic Engineering, VPN, QOS and Multicast.

MPLS as opposed the regular IP forwarding assign a packet to a particular FEC just once as the packet enters the network. FEC's may be based on service requirement policies or simply on destination address prefixes. Forwarding tables are built called label Information Base (LIB) based on FEC to label bindings. Label values are derived from the underlying data link layer (such as FR or ATM) DLCI's or VPI/VCI can be used directly as labels
**MPLS performs the following functions:**

- Specifies mechanisms to manage traffic flows of various granularities
- MPLS remains independent of Layer 2 and Layer 3 Protocols
- Provides a means to map IP addresses to fixed length labels used by different packet forwarding and switching technologies.
- Interfaces to existing Routing Protocols such as RSVP, BGP and OSPF
- Supports the IP, ATM and Frame-relay Layer 2 protocols

## MPLS Routing

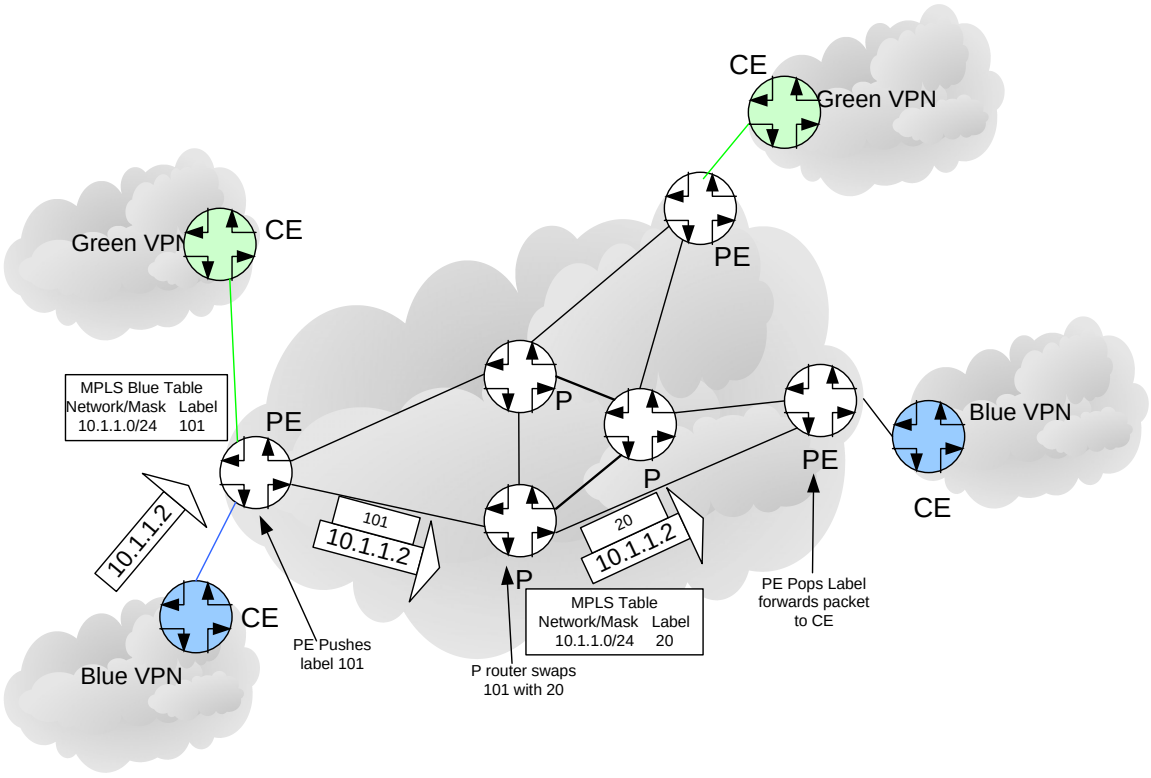


**Figure 3.2 MPLS Routing**

MPLS is based on routers and switches performing Label switching to provide a Label Switch Path (LSP) through the network. When a packet is received with destination address of 10.1.1.2 the Ingress MLPS router looks at its MPLS forwarding table associated with the Blue VPN and assigns the packet to a Forwarding Equivalency Class (FEC). The FEC assignment can be based on a number of factors beside the destination address such as Source Address, BGP next Hop or Diffserv Code Point. It attaches (PUSHES) the MPLS Label between the Layer 2 header and the IP Header that is associated with the packet.

MPLS Labels only have Local significance. Intervening Label Switch Routers (LSR) swap the incoming label with a label defined in their own MPLS forwarding database. When the Egress MLPS router receives the packets it removes (POPS) the label before forwarding the packet based on its BGP forwarding table. The IP header is not examined by LSR's. You may say that MPLS works very similar to how Frame Relay works. The

Frame Relay Access Router (FRAD) takes an IP packet and prepends a FR header (DLCI) on the packet. The frame is sent into the Frame Relay network The Frame Relay switch receiving the packet looks at the DLCI and its forwarding tables. It attaches a new FR header (DLCI) and forwards the packet. The process continues until the frame is received by the remote FRAD. The FR header is stripped and the packet is forwarded with a new layer 2 header for the proper network.

## MPLS Label

Once the LER has been determined the FEC's route a label is inserted in each frame or cell. Typically this label gets appended to the layer 2 header. (Ethernet) If the egress network is based on ATM the label populates the VPI/VCI field. If the Egress network is frame based the label is enclosed in a shim between the data link header and the IP header.

- 20 Bits of Label Information
- 3 Bits Class of Service Information
- 1 Bit Stack Field
- 8 Bit Time to Live Field

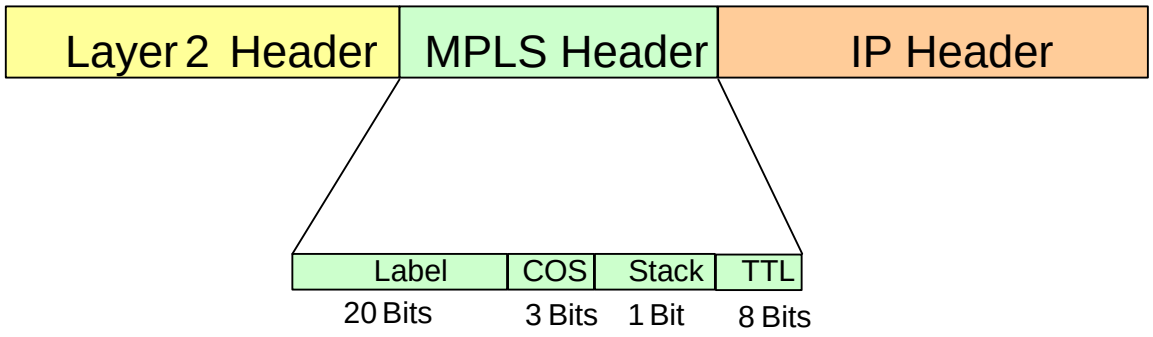


**Figure 3.3 MPLS Label**

## MPLS Label Distribution

MPLS does not require a single method of label distribution. BGP-4 has extensions based on RFC3107 that allows the distribution of labels within the BGP protocol. RSVP also has been enhanced to allow for the support of exchanging MPLS labels. The most common method of label distribution is through Label Distribution Protocol (LDP) because of its support for QOS. IETF has defined Label Distribution Protocol for Unicast IP Packets in RFC-3036. LDP supports both Class of Service (CoS) and QOS. PIM is used for label mapping for multicast packets.

## *RFC2547 Network Components*

In March 1999 a new RFC was introduced by Engineers from Cisco describing a framework for BGP/MPLS VPN's. This new architecture would introduce a new layer 2 VPN that separated customer's networks by running multiple VRF routing tables. An important feature was allowing each customer to maintain their existing addressing plan without requiring Network Address Translation. This allows customers to continue to use RFC1918 Private address space.
10.0.0.0
172.16.0.0-172.32.0.0
192.168.0.0-192.168.255.0

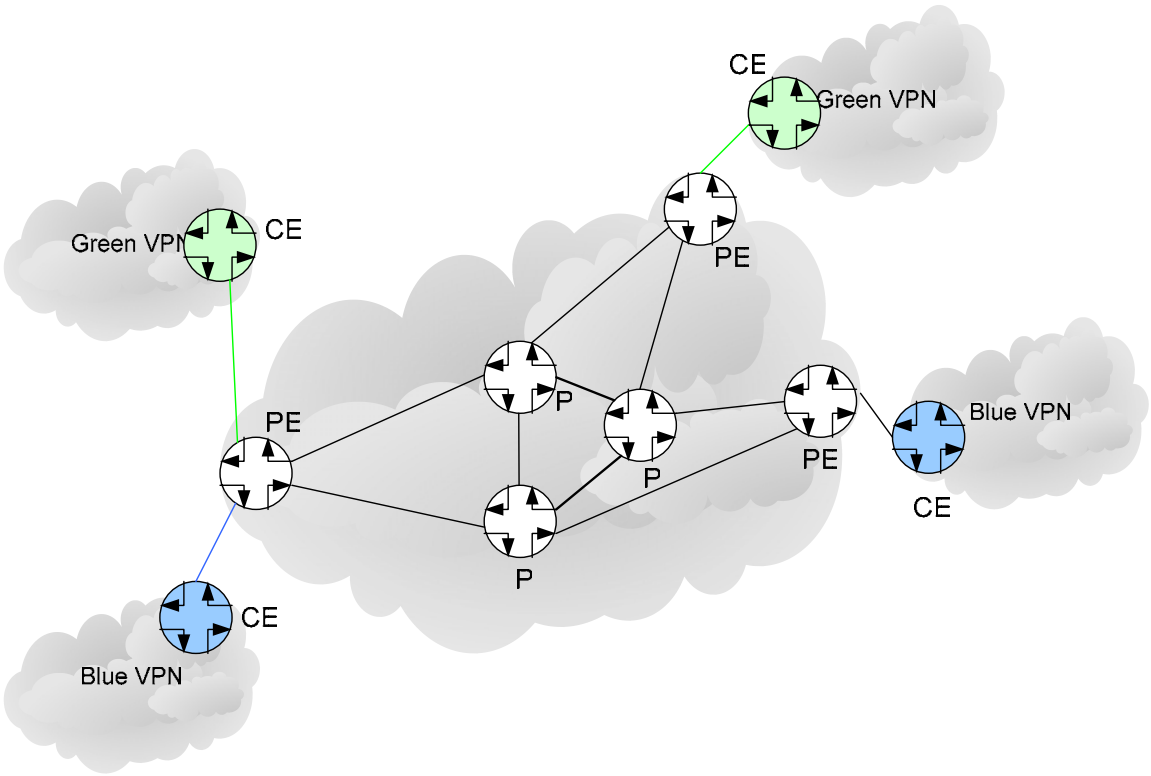


**Figure 3.4 RFC2547 Network Components**

The components in a BGP/MPLS VPN are defined by RFC 2547 as follows:

- **CE:** Customer Edge Device connects Site to Service Providers Network. Can be Host or Layer 2 Switch but typically the CE is an IP router that establishes BGP adjacency with a Provider Edge Router. The CE advertises the sites VPN routes to the PE and learns VPN routes from the PE
- **PE:** Provider Edge Router: PE routers exchange routing information with CE using static routing, RIPv2, OSPF or BGP4. Typically SP only supports Static or

BGP routing with CE. Although the PE maintains a separate forwarding table for each VPN it is only required to maintain VPN routes for VPN's that are directly attached to the PE. After learning VPN routes from the CE a PE exchanges VPN routing information with other PE's using IBGP sessions or BGP route Reflectors. PE routers also Create MPLS tags. Using MPLS to forward VPN data packets across the Service Providers backbone The ingress PE functions as the ingress Label Switch Router (LSR) and the egress PE functions as the egress LSR.

- **P:** A Provider Router (P) is any router in the provider's network that does not directly attach to CE devices. P routers function as MPLS transit LSR's when forwarding VPN data through the provider's backbone between PE routers. P routers are only required to maintain routes to the providers PE routers they are not required to maintain routing information for each specific VPN customer site.

## *BGP/MPLS/VPN Fundamentals*

### Access links (CE to PE):

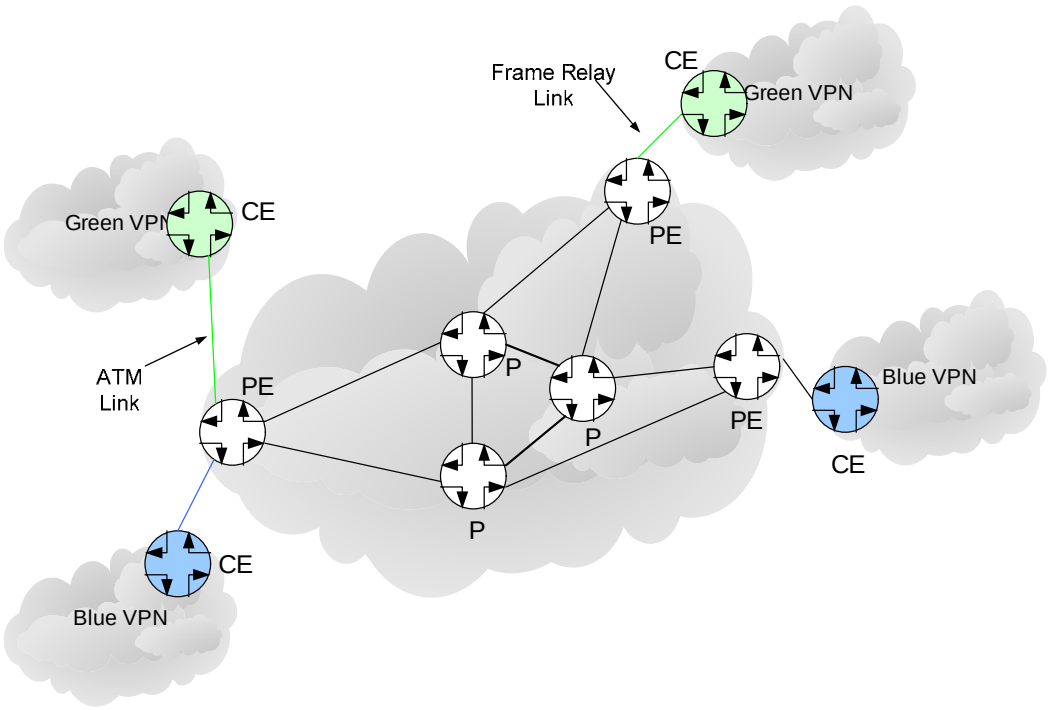


**Figure 3.5 Access Links PE to CE**

The Access links from Customer edge router can be almost any of today's WAN protocols. Common used access protocols are ATM or Ethernet First Mile (EFMA) where high bandwidth is required and Frame Relay or PPP on lower bandwidth links. There is no requirement for CE Routers to exchange routing information as there is in a traditional FR or private link network so routing tables in CE Routers can be very simple. It is very common to filter everything except default route in remote CE routers. This

allows large VPN's to be supported and yet use CE devices that do not have the capacity to support large routing tables.

## PE's and CE's are BGP routing Peers

- Sites do not directly exchange routing information.
- Very large VPN's can be supported by very simple CE routing table

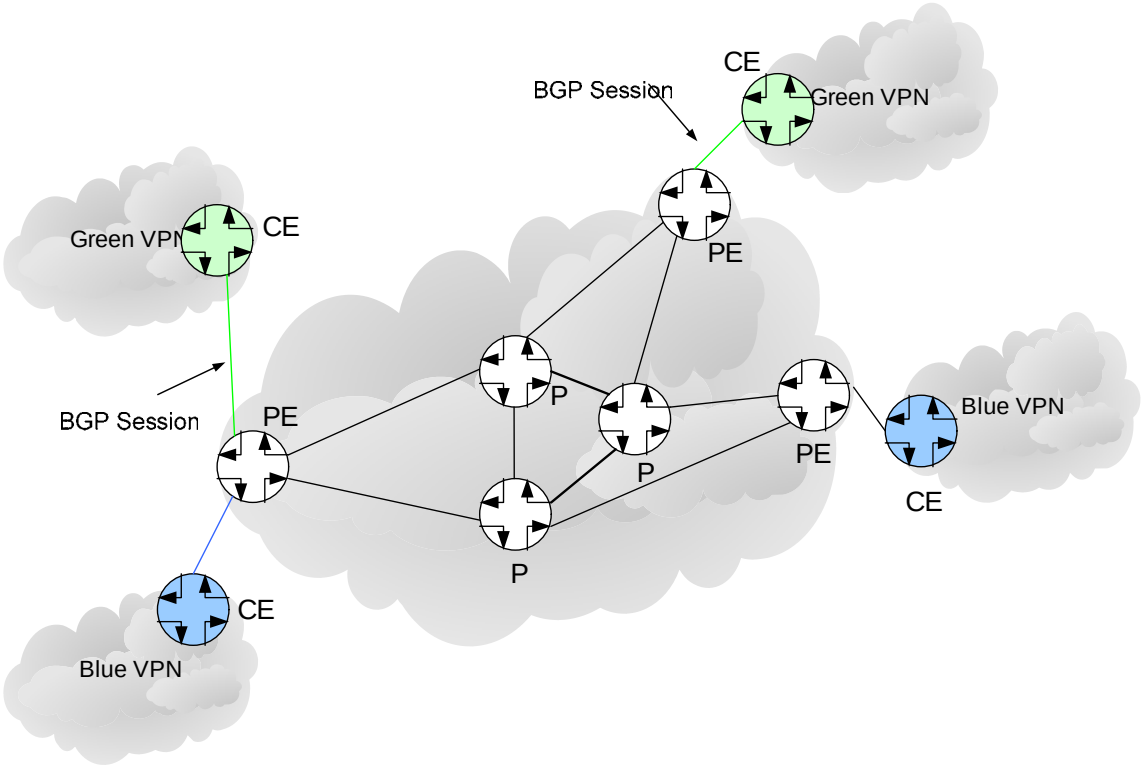


**Figure 3.6 BGP between PE and CE**

There is no direct exchange of routing information between sites as there is in traditional Frame Relay or ATM WAN networks. Routing information is exchanged between CE and their peer PE router across a BGP session.

On many sites with a single connection to the provider a very simple routing table is required such as default route and still maintains any to any routing since the provider maintains a complete routing table of the VPN.

## PE Routers maintain multiple forwarding tables (VRF)

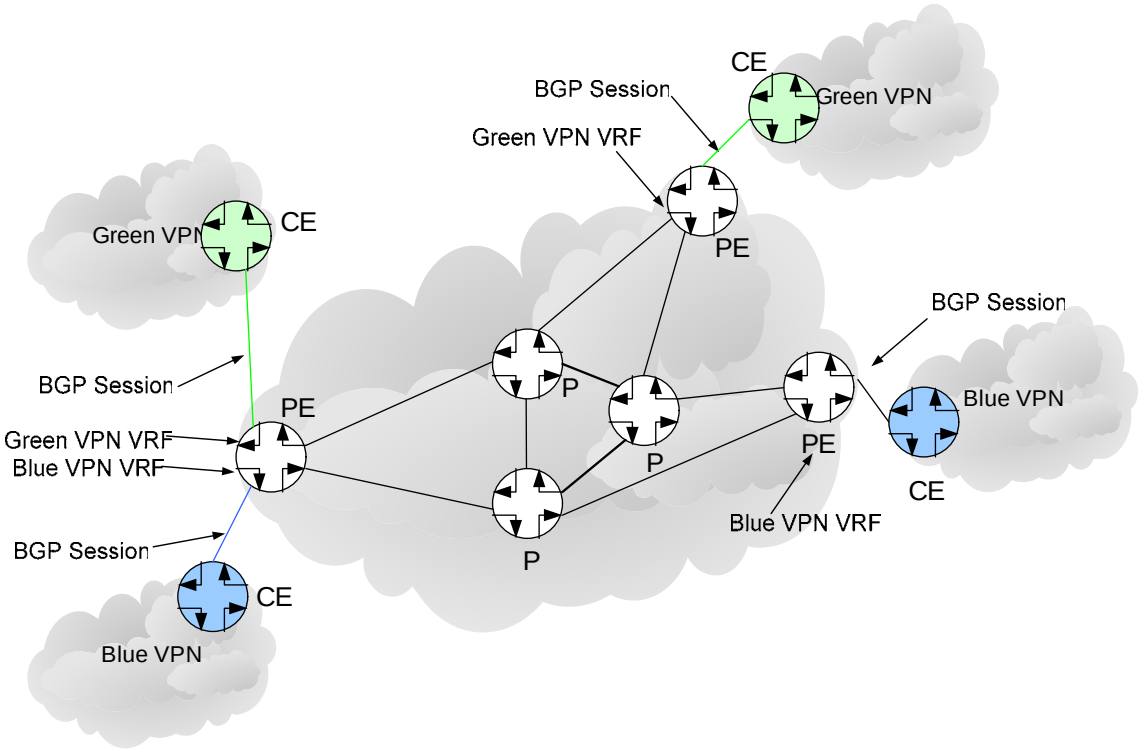


**Figure 3.7 PE Routers maintain multiple VRF'S**

PE's are high end switching devices capable of supporting multiple forwarding tables. These forwarding tables are called VRF's and they maintain one VRF for each VPN supported by that PE router.

PE's are BGP peers with other PE's and P routers. The IP switching within the provider network is through the exchange of MPLS labels with PE's and other P routers.

The Service Provider may provide value added services such as QOS, Web Hosting, NAT, Firewall and Internet access for the VPN. The Customer alternately may elect controlled access to the internet through there own central site firewall and ISP access points.

MPLS is not required at CE device. Multiple Virtual sites can be supported through VLAN tags or multiple PVC's.

## IP Routing in BGP/MPLS VPN

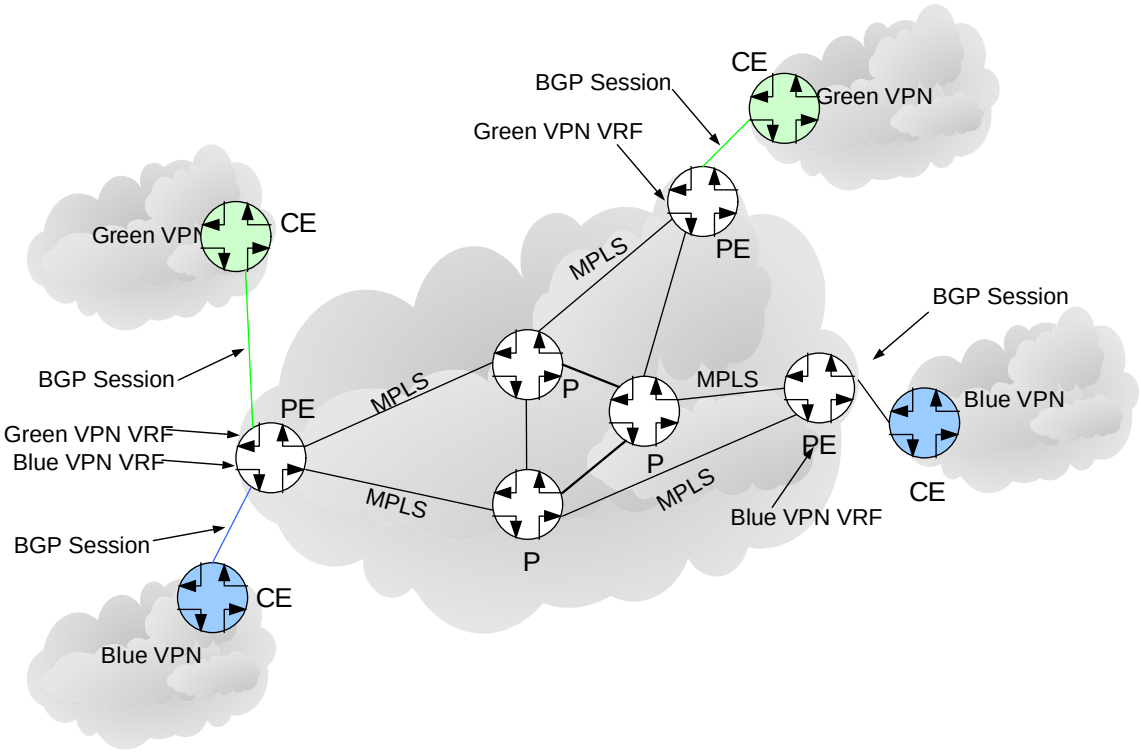


**Figure 3.8 IP Routing in BGP/MPLS VPN**

PE Routers exchange VPN routing information with other PE Routers using BGP protocol. The PE Routers form the IBGP peer relationship or use IBGP route reflector. When the PE router distributes a VPN-IP4 route via BGP it uses its own address as the "BGP next hop". It also assigns and distributes a MPLS label. When the PE processes a received packet that has this label at the top of the stack, the PE will pop the stack and send the packet directly to the CE router from which it learned the route. All backbone forwarding decisions are made with MLPS. The procedure used to define the Label Switched Path (LSP) between BGP peers is left to the Service Provider by the RFC. It may be setup in a pre-established way or it may be setup when a route would need to be installed. It may be a best effort route or a traffic engineered route. There may be a single LSP or there may be several perhaps with different QoS characteristics. All that matters is that some LSP exists between the PER and its BGP next hop. This is similar to what we may see within any transit AS in BGP. The IGP provides the routing within the AS and may provide multiple paths in some cases such as OSPF.
At the edges of the network between the PE and the CE any routing protocol can be used. It is most common for the SP to only allow BGP or Static routing. It is not required that the CE router support MPLS. If CE does support MPLS it must import entire MPLS tag table for its VPN.

## Security

The security is about the same as in other layer 2 VPN's such as Frame Relay or ATM networks. In the absence of mis-configuration or deliberate interconnection VPN's cannot gain access to each other.

Some business requirements may mandate IP-Sec for security. This can be supported the same as any FR network with Tunnel between site CE and central site tunnel terminator. Individual hosts at the VPN sites could also use IP-Sec tunnel through the provider network.

## QOS:

One big advantage of BGP/MPLS VPN's is the support for QOS within the provider network. QOS is not supported in traditional FR and ATM networks. QOS comes at the IP layer using the Diffserv model. This model of QOS is supported by most Service Providers. It is an important component of MPLS/VPN networks that the ingress traffic to the PE router be classified properly for Class of Service handling in the provider network. The CE's role will be to insure which traffic will get priority handling both within the CE and through the provider network. The CE will also act as a QOS domain border router, sometimes being required to classify packets for the trip through the provider network with one QOS value (DSCP) and another value through the customer's network. The role of the PE is to map TOS or DSCP within the IP header of a packet to MPLS QOS. This is part of the process of assigning a packet to a MPLS Forwarding Equivalency Class (FEC).

All Vanguard routers fully support the DiffServ model of QOS defined by RFC2474, RFC2475, RFC2597 and RFC2598. We are going to take four
DiffServ functions and see how they are implemented in a Vanguard router. First we will trace IP packets as they flow through the router and see how they are affected as they pass through the different functions.

First let's take a look at where the packets can come from. IP packets are originated from two general sources; externally arriving in the router from a LAN or WAN connection, or internally such as RIP or SNMP packets. Note that packets that are entering the Vanguard externally may, or may not already have a DSCP assigned.

In traffic classification the Vanguard will apply Multi Field, or MF classification. The Vanguard will check for matching criteria in one, some or all of the following header fields: IP address fields, the protocol ID fields and the protocol port number fields. If a match is found the packet will be given a temporary IP MF classification tag. This tag does not leave the router with the packet, in fact it won't even leave this function, but it will be used to assign the packet the appropriate DSCP. If the packet arrived with a DSCP it maybe changed after passing through traffic classification. With its DSCP the packets next stop is the Traffic Conditioning function.

This function is used to regulate the traffic flow and effect packet discarding if necessary. This function is very complex and the subject of its own lesson. Suffice to say

that packets maybe dropped at this point in the process but those that aren't will go on to the QOS Mapping function.
This function looks at the packets DSCP and assigns it the appropriate PHB tag. The tagged packet is then sent to the last function; Queuing and Scheduling. Of course the packet still has its DSCP.
This function actually has, as the name implies two separate functions. The first is queuing or buffering as it is commonly called. Here all the packets, with the exception of those with an Expedite Forwarding PHB, are buffered prior to being released to the WAN port. In the case of those packets with the EF PHB they in effect skip this function and go directly to the WAN port. This is why the EF PHB is usually reserved for VoIP traffic.
The purpose of the Scheduling part of Queuing and Scheduling is to sort out which queue should send a packet to the WAN port next. This is basically a prioritization function but it is the heart of QOS implementation.
Vanguard supports both Credit based (Normal Delay Services) and Priority based (Low Delay services) Queuing and Scheduling services as well as EF (Minimum Delay services) for Voice.

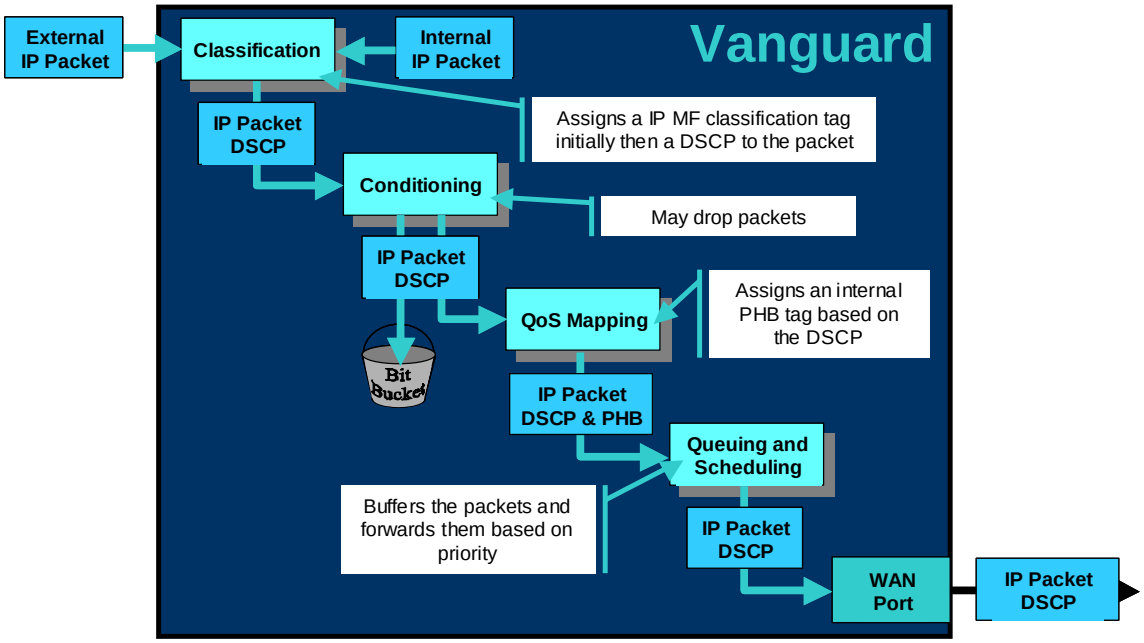


**Figure 3.9 VanguardMS support for QOS**

## Customers Enterprise Network connected by RFC 2547 BGP/MPLS VPN

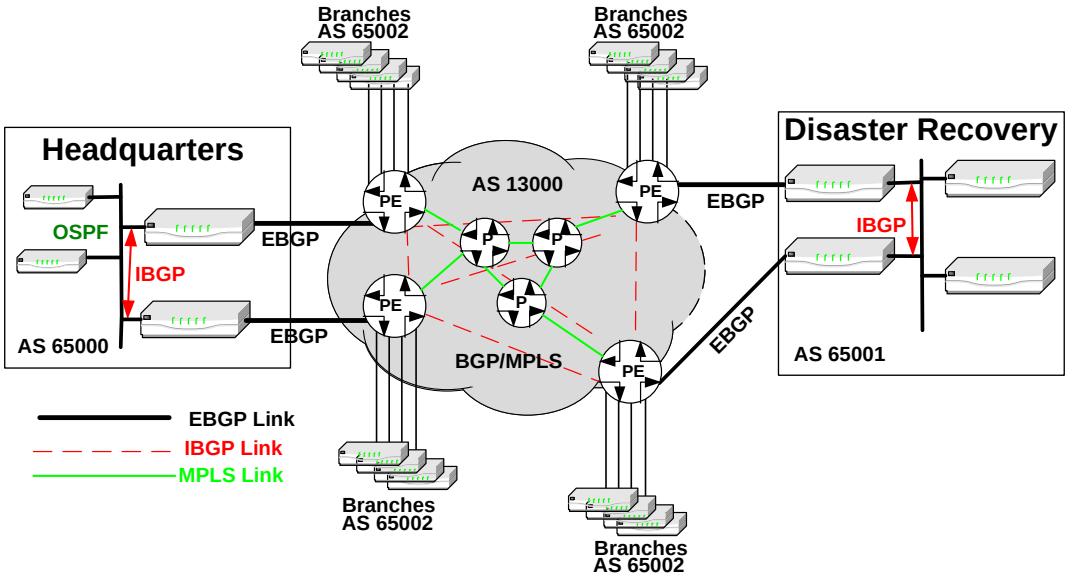


**Figure 3.10 MPLS/BGP VPN**

In this network the Service Provider is providing VPN services separating the customer network from other customers. BGP is needed for redundancy at HQ and Disaster recovery site. BGP or static routes can be used at branches. If branches are using backup via ISDN, DSL or Dial, IP can use the absence of BGP routes to trigger backup. This gives higher reliability than dependence on the Frame Relay A bit to trigger backup. With Frame Relay the PVC is no longer a point to point logical connection through the network but only exists between the customers edge router (CE) to the providers edge router (PE), Frame Relay connectivity no longer equates with IP connectivity. IP-enabled offerings such as AT&T's Network-Based VPN and Verizon IP-VPN services both fall into the IP-enabled service category in that they allow you to use an existing frame relay access link to tap into a connectionless, Multi-protocol Label Switching (MPLS)-based IP backbone. The primary benefit is that achieving mesh connectivity within your company's VPN requires just a single access permanent virtual circuit (PVC) from each remote site. This type of network will become more popular as more customers are investigating disaster recovery scenarios. The higher cost of the IP enabled PVC starts to balance out when measured against multiple PVC's.

## Advantages to MPLS VPN's

- SP can offer Service Level Guarantees (SLA's)

- Many implementations of IPSec VPN's rely on Internet Backbone where ISP's cannot offer SLA's
- SP can provide QOS granularity
- Easier Migration lower equipment cost than IPSec
- The cost and complexity is hidden in the SP network
- Full mesh connectivity with single PVC

As with any offering there are both advantages and disadvantages to MPLS VPN's. Since many of the IPSec VPN's rely on the Providers IP backbone and most often part of the journey for a packet may be across the Internet few SLA's are offered. Where they are offered IPSec looses the advantage of flexibility because all end points must be directly to the VPN providers network.
MPLS VPN's have been engineered with SLA's in mind. The SP can take advantage of the Traffic Engineering features of MPLS.
Unless the tunnel endpoint is Diffserv aware and copies DSCP's to Tunnel header IPSec makes it impossible to differentiate between packets inside the tunnel since all data is encrypted including IP headers of the original packets. Frame Relay and ATM switches do not look at IP headers when traffic is switched through their network. It is impossible to utilize Diffserv QOS in these VPN's. MPLS VPN's utilize a shim header which includes 3 CoS bits. PE routers set the packets priority as it enters the MPLS VPN and QOS is engineered into the network. Most are offering 3 or 4 QOS classes but up to 8 can be supported by the architecture.
MPLS VPN's support the same access protocols (FR/ATM) that is prevalent in corporate networks today. So no replacement is required for CPE equipment. Many times a IPSec VPN implementation will require a forklift upgrade to the network. The complexity of IPSec VPN implementation falls on the customer if they maintain there own CPE routers. The complexity of MPLS VPN remains in the provider network.
The biggest advantage that MPLS VPN's offer is full meshing of the VPN network at the cost of a single PVC. This allows rollout of any to any applications without the cost of a full mesh network or the latency of a hub and spoke network.

## Disadvantages to MPLS VPN's

- Less security than IPSec VPN's but about the same as FR/ATM
- Greater geographic limitations than IPSec VPN's
- Emerging Standards/ immature technology

There is some concern by some network engineers that traffic is carried unencrypted across a public network, however the use of MPLS labels provides traffic isolation and is as secure as any Layer2 VPN.
MPLS VPN's at this time do not offer the flexibility that IPSec implementations do. With MPLS VPN's you are limited to where the providers network goes while IPSec VPN can be implemented anywhere the Internet goes.

## Critical CE Features

All Service Providers stress that one of the chief advantages of MPLS VPN's is that this service works with present CPE equipment. (MPLS forwarding is not required to be supported by the CE) Most SP's offer routing between CE and PE to be either static or BGP-4. However FRAD and CPE vendors will need to keep in step with SP's to provide the features necessary to stay on SP's preferred CE lists.
Vanguard MS introduced BGP-4 in version 6.0 of ONS to address the BGP/MPLS/VPN market. Vanguard's BGP supports full OSPF to/from BGP and RIP to/from BGP redistribution, BGP Community support, BGP aggregation as well as full inbound outbound policy support.

The following BGP features in are required CE routers in a MPLS VPN:

- Inbound and Outbound Policy support
- BGP to IGP Routing Redistribution
- IGP to BGP Redistribution
- BGP Aggregation
- Community Support

Almost all offerings from Service Providers stress that one of the primary advantages of this service offering over traditional Frame Relay VPN's is the support of QOS within MPLS.
Although the Intserv model (RSVP) has some interaction with MPLS the Diffserv model seems to fit best. CE routers will be used to re-classify packets from the QOS model used within the customer's network to the QOS model used within the SP network. It will be important for the CE to fully implement DiffServ to support traffic shaping, WRED for AF classes, QOS domain management through traffic classification or reclassification. MLPPP will be important access protocols to support demands for higher bandwidth, (higher than T1/E1 but less than T3/E3) These access protocols will also be important for segmentation strategies to support QOS for VoIP and Video.
Vanguard MS has long been supporting the QOS features required for the CE router. It is clear if IP MPLS VPN's become a dominant offering from carriers and Service Providers that packet Voice will go the direction of VoIP rather than VoFR. Multilink and segmentation features as well as QOS will determine if the CE vendor be successful with VoIP. Vanguard has supported full Diffserv QOS since ONS release 5.5. Support for FRF.12 and MLPP have also been available for some time. Vanguard routers are compatible with major PE vendors in the area of MLPPP interleaving and queuing and FRF.12

The following CE QOS Features are required:

- QOS Classification
- QOS Traffic shaping
- QOS Priority Queuing

- QOS Scheduling
- QOS Diffserv: Full Diffserv support will be required in the future to address this market.

Real time applications such as VoIP require Link Segmentation and Interleaving features such as:

- Multi-Link PPP (MLP)
- FRF-12

Multicast support:

- Most service have offerings that support Multicast

### MPLS-VPN Networks and Legacy Protocols

Legacy protocols such as SNA and BSC can be encapsulated in TCP/IP to run across these MPLS/VPN's. Vanguard routers can support a large array of legacy Protocols encapsulated within IP with its SOTCP. Another solution is to continue to keep the legacy applications on traditional Frame Relay PVC's while IP applications run on the MPLS VPN. All major providers that provide Frame Relay access to these networks can provide hybrid access where some Frame Relay PVC's are mapped to traditional Frame Relay and others are mapped to the MPLS VPN. SOTCP gives the customer the option of converting the network to MPLS VPN and taking advantage of the single PVC for any to any connectivity while continuing to support applications requiring legacy protocol support. The support for QOS and CoS makes it very likely that there will be no timing issues with Legacy applications across MPLS VPN Networks.

## BGP Basic Configuration

It is assumed that basic IP configuration is done before configuring BGP. Before you can configure BGP-4 parameters you must configure the following:

- LAN and WAN ports
- LAN Connection Table
- IP Router Interfaces
- IP Router Parameters

From the Main menu BGP configuration can be found at:

Configure->Router->Configure BGP*

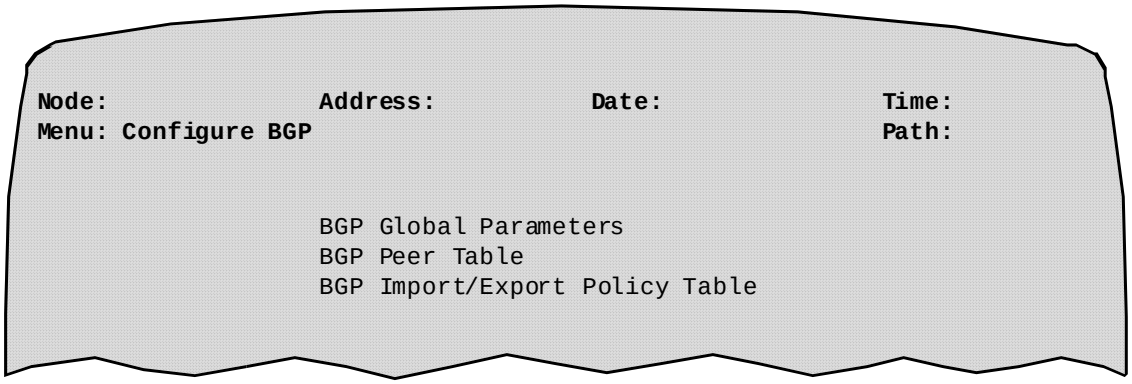


**Figure 4.1 BGP Main Menu**

Basic minimum BGP configuration of BGP Global Parameters:

- Set the BGP parameter to Enable
- Set AS Number parameter to your AS number
- Set BGP Identifier to Internal IP Address (Suggested)
- Set Default IGP Import Policy (OSPF, RIP, Direct, etc.)
- Set Default BGP Inbound Policy to Permit
- Set Default BGP Outbound Policy to Permit

You will need to configure a BGP Peer Table entry for each BGP Peer router you wish to exchange routing information with. The following example is for only 1 connection.

In BGP Peer Table Peer Entry 1:

- Set Peer Control parameter to Enable
- Set Peer AS Number parameter to AS number of Peer router
- Set Peer IP Address List parameter to Peers IP address

## *Simple BGP to OSPF redistribution*

Configure OSPF AS Boundary Routing Parameters

- Set Import BGP Routes to Yes
- Set Default BGP->OSPF Import Policy to Permit

This basic configuration added to running OSPF configuration will Import all BGP routes into OSPF. Most instances the default Import policy will be set to deny and policy filters will determine what BGP routes will be imported into OSPF.

### *Common Startup Problems*

If BGP Peer session does not go to Established state these are some things to check in your configuration.

- BGP is disabled or BGP Peer not configured in other router
- TCP session never gets established. Check to see if you can ping the other router and you are not going through a firewall that may be blocking TCP connection for Port 179.
- BGP Peer is not reachable or cannot Ping Peer. BGP is dependent on basic TCP connection being correct.
- BGP Peer mis-configured (Wrong AS Number). This router needs to configure the BGP Peer with the Peers AS number. The Peer also needs to have a proper BGP configuration. Symptom is the Peer oscillates between active and idle states.
- BGP ID is not a valid IP address in this router.
    - If the BGP ID configured in BGP Global parameters is not correct the status of BGP will be disabled. This can happen if the interface selected as BGP ID is down. Check IP Interface status to see the state of the Interface. It is suggested that Internal Interface is used for BGP ID so this will not happen.

## BGP Statistics

To view BGP statistics go to Main menu and follow the path:

- Main->Status/statistics->Router Stats->BGP Statistics

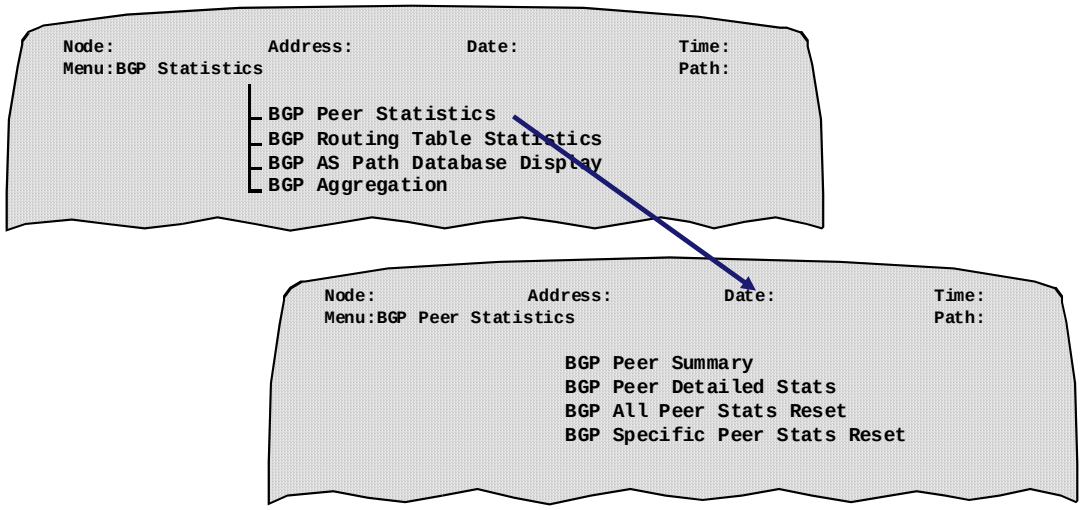


**Figure 4.1 BGP Statistic Menu**

## Peer Summary Statistics

Peer Summary statistics will show overall status of all BGP Peer sessions. A Peer connection should be in the Established state.

```
Node: <Node name>  Address: <node address>    Date: <date>

 BGP Peer Statistics Summary                                  Page 1 of 1

Total Number of Peers: 3

  PeerNo  Peer AS   State        BGP ID          Peer IP Addr     Up Since
                                                                 (DAY:HH:MM:SS)
  1       100       Established  150.1.1.1       1.1.1.1          000:00:48:00
  2       200       Idle         0.0.0.0         0.0.0.0          000:00:00:00
  3       Local_AS  Established  160.1.1.1       20.1.1.1         121:23:56:55


Press any key to continue (ESC to exit)...
```

**Figure 4.2 Peer Summary Statistics**

## Peer Detailed Statistics

Included in Peer detailed statistics are TCP Connection details, Count for BGP packets exchanged, Count of BGP routes exchanged, and Count of BGP Connection Errors.

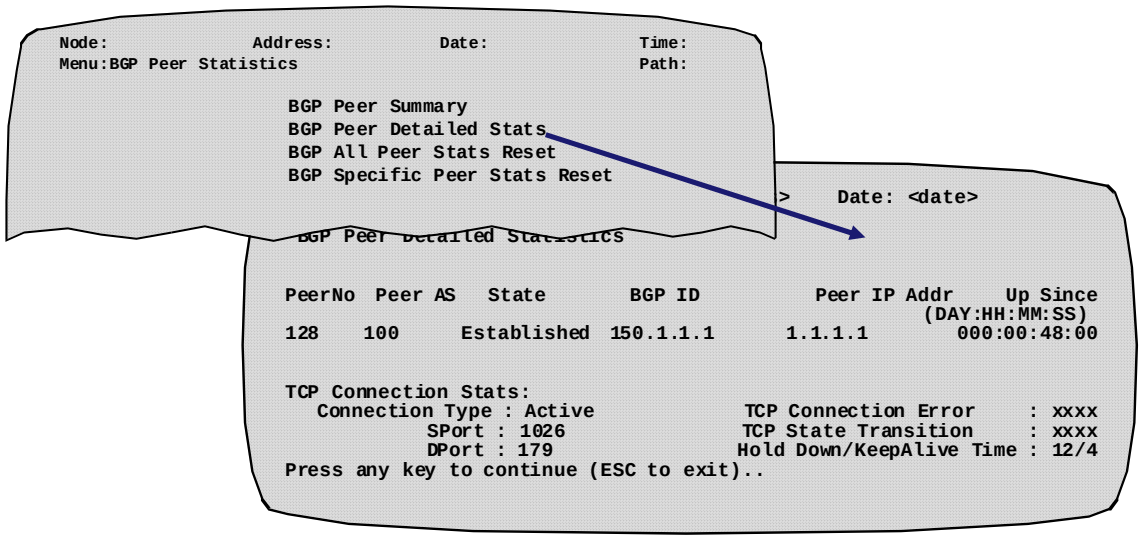


**Figure 4.3 Peer Detailed Statistics**

## BGP Routing Table statistics

BGP Routing table (BRT) statistics show details of BGP routes

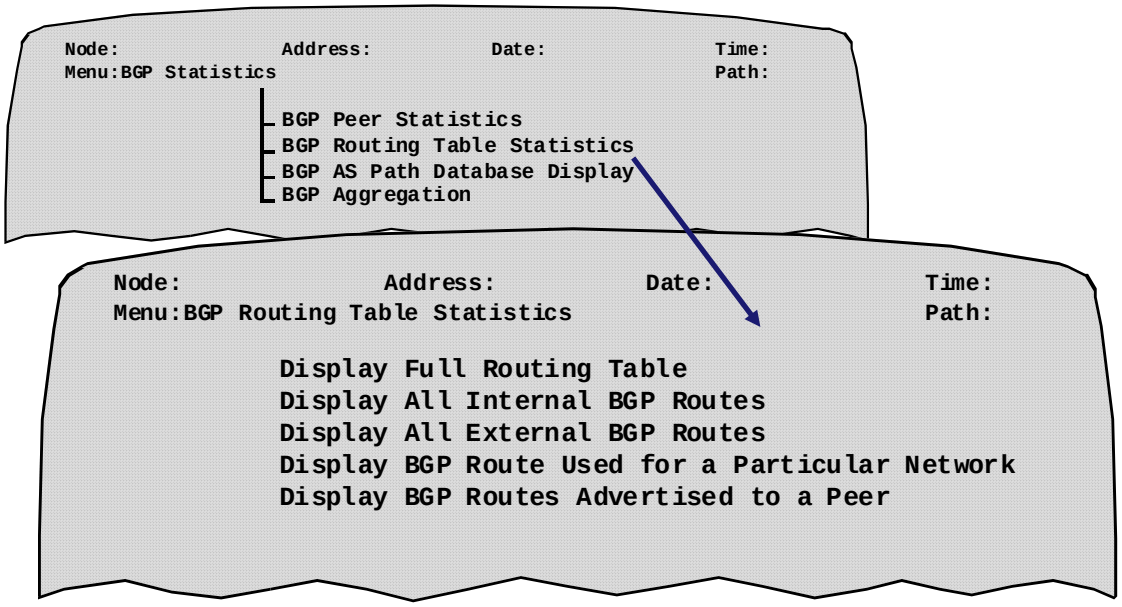


**Figure 4.4 BGP Routing Table Statistics**

BGP Full routing table displays all External and Internal routes.

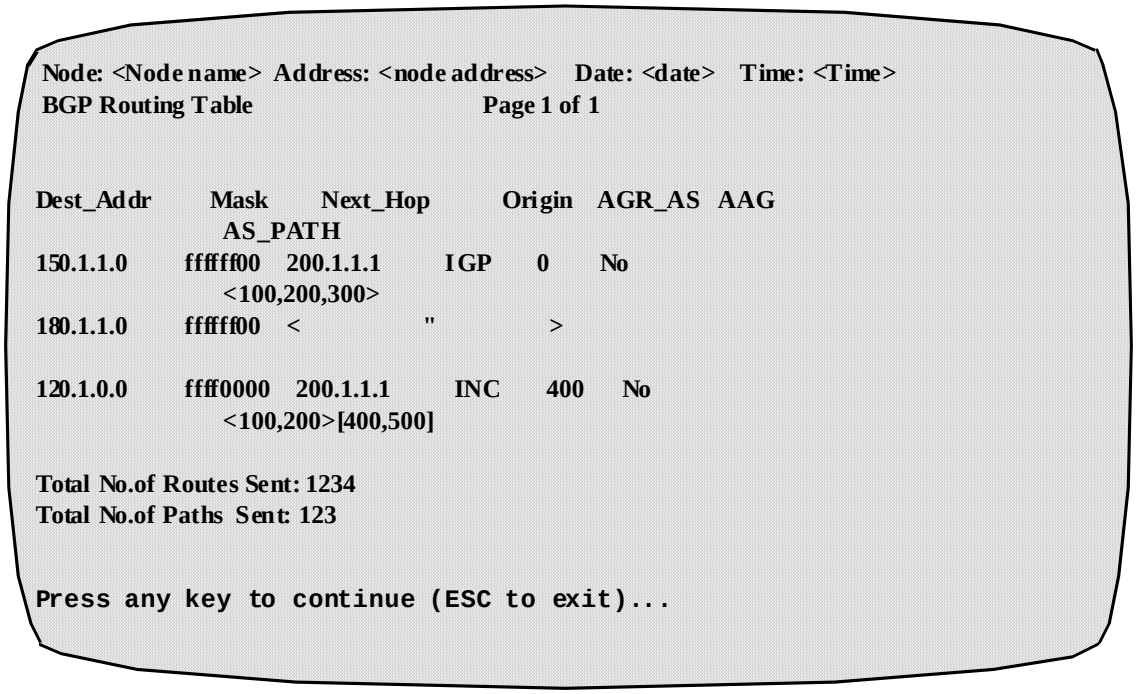


**Figure 4.5 BGP Full Routing Table Statistics**

## Using BGP AS Path Database Display

The BGP AS Path Database contains valuable data for trouble shooting BGP. The BRT will only display the route selected by the DOP algorithm. The BGP AS Path Database contains all Paths.

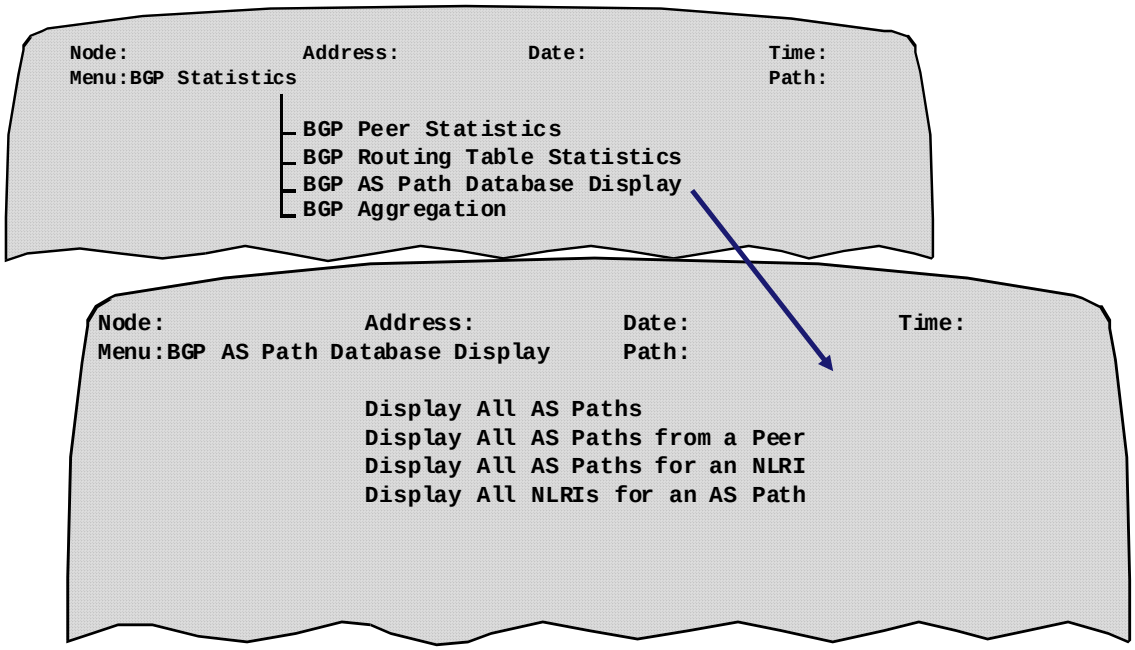


**Figure 4.6 BGP AS Path Database Display Menu**

## Display all Paths for NLRI

A lookup for all paths of a single NLRI can be done with the Display all Paths for NLRI selection. Refer to Figure 4.7. To display all paths for NLRI. 100.1.0.0 /16 Key in NLRI (100.1.0.0) and mask (ffff0000). Here we see there are 2 paths the first with next hop of 2.1.1.1 and a DOP of 50 and a second with a next hop of 10.1.1.1 and a DOP of 25. Only the first path would be displayed in the BRT but if this Path failed the second would take its place.

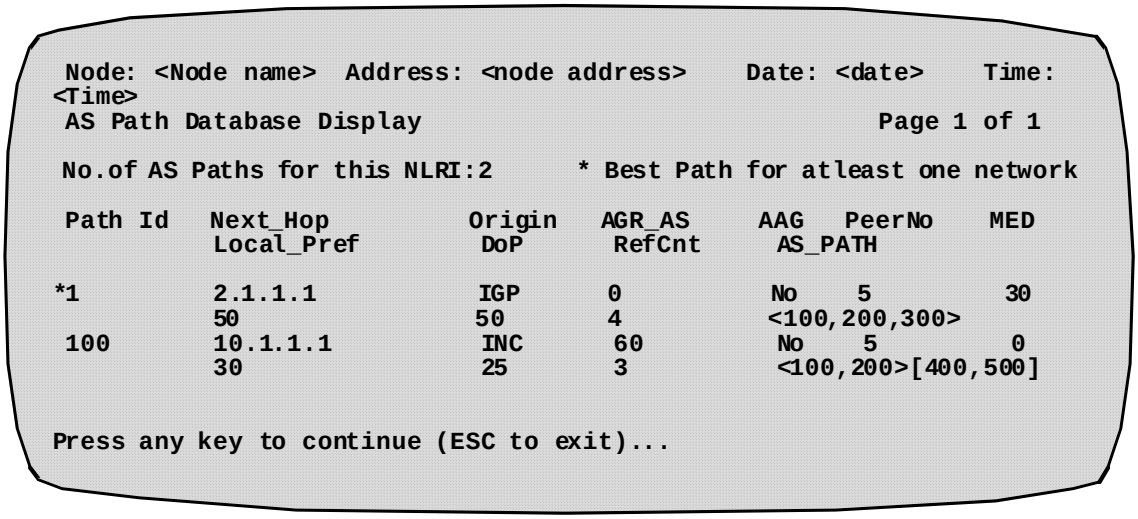


**Figure 4.7 Display all Paths for NLRI**