BGP Advanced Routing in SonicOS

Document Scope

This document provides an overview of SonicWALL's implmenetation of Border Gateway protocol (BGP), how BGP operates, and how to configure BGP for your network.

This document contains the following sections:

Feature Overview

The following sections provide an overview of BGP:

What is BGP?

BGP is a large-scale routing protocol used to communicate routing information between Autonomous Systems (ASs), which are well-defined, separately administered network domains. BGP support allows for SonicWALL security appliances to replace a traditional BGP router on the edge of a network's AS. The current SonicWALL implementation of BGP is most appropriate for "single-provider / singly-homed" environments, where the network uses one ISP as their Internet provider and has a single connection to that provider. SonicWALL BGP is also capable of supporting "single-provider / multi-homed" environments, where the network uses a single ISP but has a small number of separate routes to the provider. BGP is configured through the SonicOS Command Line Interface (CLI).

Background Information

Routing protocols are not just packets transmitted over a network, but comprise all the mechanisms by which individual routers, and groups of routers, discover, organize, and communicate network topologies. Routing protocols use distributed algorithms that depend on each participant following the protocol as it is specified, and are most useful when routes within a network domain dynamically change as links between network nodes change state.

Routing protocols typically interact with two databases:

- Routing Information Base (RIB) - Used to store all the route information required by the routing protocols themselves.
- Forward Information Base (FIB) - Used for actual packet forwarding.

The best routes chosen from the RIB are used to populate the FIB. Both the RIB and FIB change dynamically as routing updates are received by each routing protocol, or connectivity on the device changes.

There are two basic classes of routing protocols:

- **Interior Gateway Protocols (IGPs)** - Interior Gateway Protocols are routing protocols designed to communicate routes within the networks that exist inside of an AS. There are two generations of IGPs. The first generation consists of distance-vector protocols. The second generation consists of link-state protocols. The distance-vector protocols are relatively simple, but have issues when scaled to a large number of routers. The link-state protocols are more complex, but have better scaling capability. The existing distance-vector protocols are Interior Gateway Routing Protocol (IGRP), Enhanced Interior Gateway Routing Protocol (EIGRP), Routing Information Protocol (RIP), and RIPv2, an enhanced version of RIP. IGRP and EIGRP are proprietary Cisco protocols. The link-state protocols currently in use are Open Shortest Path First (OSPF) and the little-used Intermediate System to Intermediate System (IS-IS) protocol.

SonicOS supports OSPFv2 and RIPv1/v2 protocols, the two most common routing Interior Gateway Protocols, allowing our customers to use our products in their IGP networks and avoid the additional cost of a separate traditional router.

- **Exterior Gateway Protocols (EGPs)** - The standard, ubiquitous Exterior Gateway Protocol is BGP (BGP4, to be exact). BGP is large-scale routing protocol that communicates routing information and policy between well-defined network domains called Autonomous Systems (ASs). An Autonomous System is a separately administered network domain, independent of other Autonomous Systems. BGP is used to convey routes and route policy between Autonomous Systems. ISPs commonly use BGP to convey routes and route policy with their customers as well as with other ISPs.

  Each Autonomous System has a 16-bit number assigned. Like IP addresses, an AS number may be public or private. Public AS numbers are a limited resource and are provisioned based on a number of factors. ISP customers with large networks multi-homed to two or more ISPs usually have a public AS, whereas smaller customers will be given a private AS administered by their ISP provider.

  As our products evolve in support of enterprise-level requirements, some customers may want to place our products on the edge of their AS in place of a traditional BGP router. To support these topologies, BGP has been added beginning in SonicOS 5.6.5.

Autonomous Systems

Each Autonomous System has a 16-bit number assigned. Like IP addresses, an AS number may be public or private. Public AS numbers are a limited resource and are provisioned based on a number of factors. ISP customers with large networks multi-homed to two or more ISPs usually have a public AS, whereas smaller customers will be given a private AS administered by their ISP provider.

Types of BGP Topologies

BGP is a very flexible and complex routing protocol. As such, BGP routers may be placed in a large variety of topology settings, such as Internet core routers, intermediary ISP routers, ISP Customer Premises Equipment (CPE), or routers in small private BGP networks. The number of BGP routes required for different topologies varies from greater than 300,000 for core routers, to 0 for ISP customers that use a single ISP and use default routing for all destinations outside of their AS. ISP customers are often required to run BGP from their edge router (the CPE) to the ISP regardless of the number of routes they receive from the ISP. This allows ISP customers to control which networks to advertise to the outside world. There's always the fear that a customer will advertise a network, or network aggregate, not owned by the customer, black-holing Internet traffic to those networks. In reality, ISP providers are careful to filter invalid advertisements from their customers (one of BGP's strengths), so this rarely happens.

There are three basic scales of BGP networks:

- **Single-Provider / Singly-Homed** - The network receives a single route (singly-homed) from a single ISP (single-provider). The number of routes an ISP customer receives from its ISP depends on the nature of its AS. An ISP customer that uses only one ISP as their Internet provider, and has a single connection to that provider (single-provider / singly-homed) has no need to receive any routes - all traffic destined outside of the AS will go to their ISP. These customers may still advertise some or all of their inside network to the ISP.
- **Single-Provider / Multi-Homed** - The network receives multiple routes (multi-homed) from a single ISP (single-provider). ISP customers that use a single ISP, but have multiple connections to their ISP may only receive the default route (0.0.0.0/0) at each ISP gateway. If an ISP connection goes down, the advertised default route sent from the connected CPE router to internal routers would be withdrawn, and Internet traffic would then flow to a CPE router that has connectivity to the ISP. The customer's inside network would also be advertised to the ISP at each CPE router gateway, allowing the ISP to use alternate paths should a particular connection to a customer go down.

- **Multi-Provider / Multi-Homed** - ISP customers that use more than one ISP (multi-provider / multi-homed) have one or more separate gateway routers for each ISP. In this case, the customer's AS must be a public AS, and may either be a transit or non-transit AS. A transit AS will receive and forward traffic from one ISP destined for a network reachable through another ISP (the traffic destination is not in the customer's AS). A non-transit AS should only receive traffic destined for its AS - all other traffic would be dropped. BGP routers in a transit AS would often receive a large portion (in many cases, all) of the full BGP route table from each ISP.

Why Use BGP?

- Even if you are not a large network on the internet, BGP is the standard for multi-homing, load-balancing, and redundancy:
  - Single-provider / Singly-homed – Not typically a strong candidate for BGP, but may still use it to advertise networks to the ISP. Singly-homed networks are not eligible for a public AS from RIRs.
  - Single-provider / Multi-homed – Common to follow RFC2270 suggestion to use a single private AS (64512 to 65535) to get the benefit of BGP while preserving public ASN.
  - Multi-provider / Multi-homed – Highly redundant, typically with dedicated routers to each ISP. Requires public ASN. Large memory footprint
- Route summarization makes routing scalable.

How Does BGP Work?

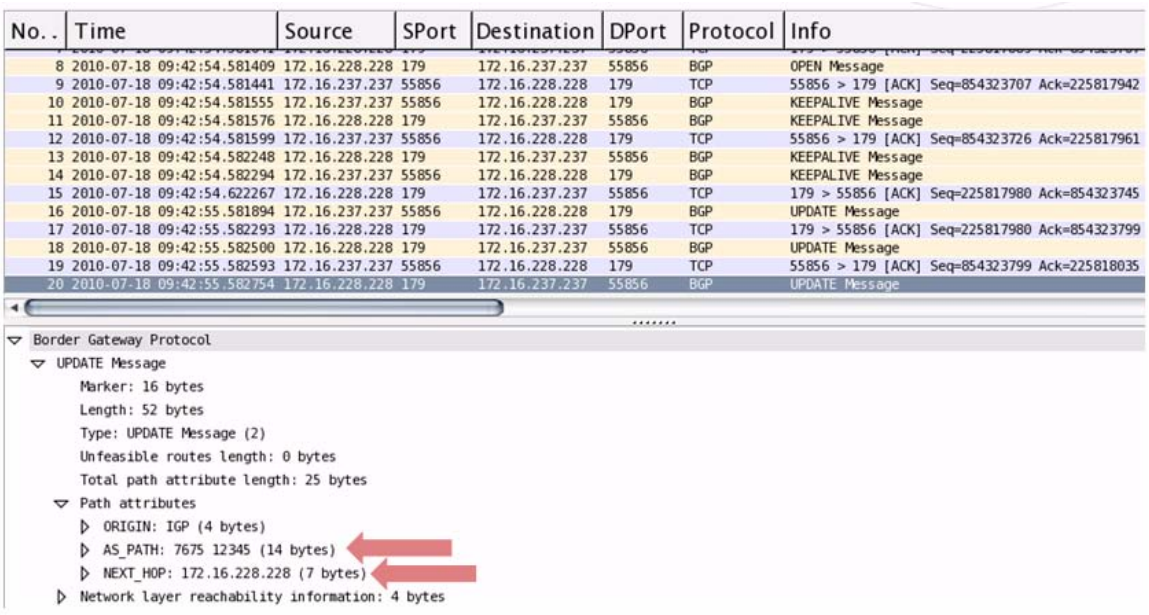
BGP uses TCP port 179 for communication. BGP is considered a path-vector protocol, containing end-to-end path descriptions for destinations. BGP neighbors can either be internal (iBGP) or external (eBGP):

- iBGP – Neighbor is in the same AS.
- eBGP – Neighbor is in a different AS.

Paths are advertised in UPDATE messages that are tagged with various path attributes. AS_PATH and NEXT_HOP are the two most important attributes that describe the path of a route in a BGP update message.

- AS_PATH: Indicates the ASs that the route is traveling from and two. In the example below, the AS_PATH is from AS 7675 to AS 12345. For internal BGP, the AS_PATH specifies the same AS for both the source and destination.

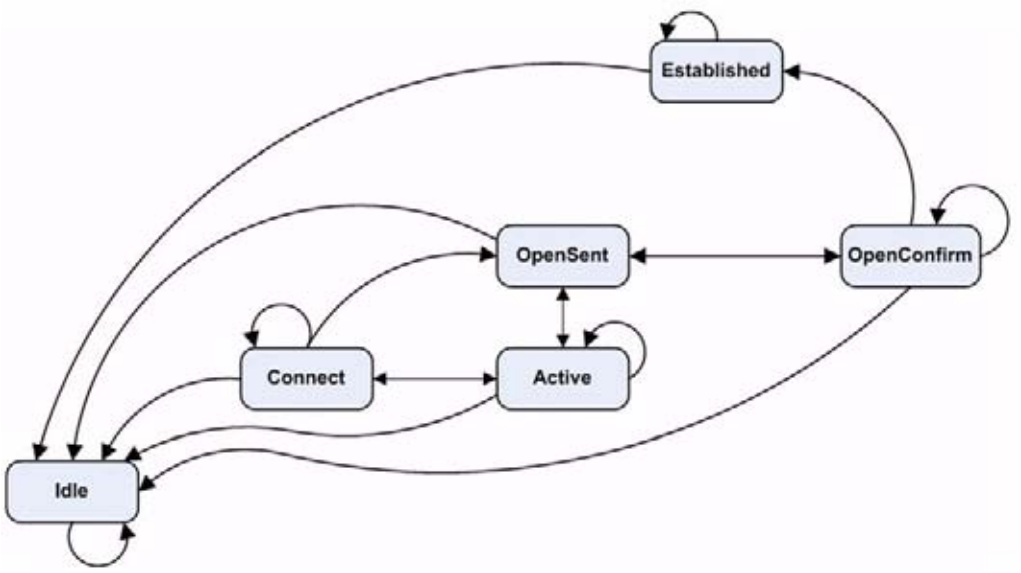
- NEXT_HOP: Indicates the IP address of the next router the path travels to. Paths advertised across AS boundaries inherit the NEXT_HOP address of the boundary router. BGP relies on interior routing protocols to reach NEXT_HOP addresses.



BGP Finite State Machine

RFC 1771, which defines BGP, describes the operation of BGP in terms of the following state machine. The table following the diagram provides additional information on the various states.

Figure 1 BGP Finite State Machine



State	Description
Idle	Waiting for Start event, after establishing new BGP session or resetting an existing session. In the event of errors, falls back to the Idle state. After a Start event, BGP initializes, resets connect retry timer, initiates TCP transport connection, and listens for connections
Connect	Once the TCP layer is up, transition to OpenSent, and send OPEN. If no TCP, transition to Active. If the connect retry timer expires, remain in Connect, reset the timer, and initiate a transport connection. Otherwise, transition back to Idle.
Active	Try to establish TCP connection with peer. If successful, transition to OpenSent and send OPEN. If connect retry expires, restart the timer and fall back to the Connect state. Also actively listen for connection by another peer. Go back to Idle in case of other events. • Connect to Active flapping indicates a TCP transport problem, e.g. TCP retransmissions or unreachability of a peer.
OpenSent	Waiting for OPEN message from peer. Validate on receipt. On validation failure, send NOTIFICATION and go to Idle. On success, send KEEPALIVE and reset the keepalive timer. Negotiate hold time, smaller value wins. If zero, hold timer and keepalive timer are not restarted.
OpenConfirm	Wait for KEEPALIVE or NOTIFICATION. If KEEPALIVE is received, transition to Established. If UPDATE or KEEPALIVE is received, restart the hold timer (unless the negotiated hold time is zero). If NOTIFICATION is received, transition to Idle. • Periodic KEEPALIVE messages are sent. If TCP layer breaks, transition to Idle. If an error occurs, send a NOTIFICATION with error code, transition to Idle.
Established	Session up, exchange updates with peers. If a NOTIFICATION is received, transition to Idle. Updates are checked for errors. On error, send NOTIFICATION, and transition to Idle. In case of hold time expiration, disconnect TCP.

BGP Messages

BGP communication includes the following types of messages

- **Open** – The first message between BGP peers after TCP session establishment. Contains the necessary information to establish a peering session, e.g. ASN, hold time, and capabilities such as multi-product extensions and route-refresh.
- **Update** – These messages contain path information, such as route announcements or withdrawals.
- **Keepalive** – Periodic messages to keep TCP layer up, and to advertise liveliness.
- **Notification** – A request to terminate the BGP session. Non-fatal notifications contain the error code "cease". Subcodes provide further detail:

Subcode	Description
1 – Maximum number of prefixes reached	The configured "neighbor maximum-prefix" value was exceeded
2 – Administratively shutdown	Session was administratively shutdown
3 – Peer unconfigured	Peer configuration has been removed
4 – Administratively reset	Session was administratively reset
5 – Connection rejected	Rejection (sometimes temporary) of BGP session
6 – Other configuration change	Session was administratively reset for some reason

- **Route-refresh** – A request for the peer to resend its routes.

BGP Attributes

BGP update messages can include the following attributes:

Value	Code
1	ORIGIN
2	AS_PATH
3	NEXT_HOP
4	MULTI_EXIT_DISC
5	LOCAL_PREF
6	ATOMIC_AGGREGATE
7	AGGREGATOR
8	COMMUNITY
9	ORIGINATOR_ID
10	CLUSTER_LIST
11	DPA
12	ADVERTISER (Historic)
13	RCID_PATH / CLUSTER_ID (Historic)
14	MP_REACH_NLRI
15	MP_UNREACH_NLRI
16	EXTENDED COMMUNITIES

Value	Code
17	AS4_PATH
18	AS4_AGGREGATOR
19	SAFI Specific Attribute (SSA) (deprecated)
20	Connector Attribute (deprecated)
21	AS_PATHLIMIT (deprecated)
22	PMSI_TUNNEL
23	Tunnel Encapsulation Attribute
24	Traffic Engineering
25	IPv6 Address Specific Extended Community
26	AIGP (TEMPORARY - expires 2011-02-23)
27-254	Unassigned
255	Reserved for development

For more information on BGP attributes, see:
http://www.iana.org/assignments/bgp-parameters/bgp-parameters.xml

Caveats

- **Scale** - Currently, SonicOS supports from 512 to 2,048 policy-based routes (PBRs). This is not sufficient for full or even partial routing tables. The number of routes that exist in the RIB may be greater than the number installed into PBR (which is the FIB). This occurs when multiple competing routes have been received through the routing protocols. For each case in which the RIB contains competing routes to a particular network destination, only one of these routes is chosen to be installed in the FIB.

  Currently, our implementation is most appropriate for the single-provider / singly-homed customers. Single-provider / multi-homed installations may also be appropriate when either the default route is being received from the ISP, or a very small number of ISP-specific routes are received by the customer. The latter allows inside routers to take the optimal path to destinations outside of the AS, but still within the ISP's network domain (this is called partial-routes).
- **Load balancing** - There is currently no multi-path support in SonicOS or Zebos (the 'maximum-paths' capability). This precludes load-balancing without splitting networks.
- **Loopback -** There is currently no loopback interface support.
- **NAT -** BGP is for routing. It does not co-exist well with NAT.
- **VPN updates** - BGP updates over VPN are not currently working.
- **Asymmetric paths** - Stateful firewall will not currently handle asymmetric paths, especially not across multiple firewalls.

Licensing BGP

Licensing for BGP Advanced Routing is included with the following SonicWALL NSA E-Class appliances, when they are registered:

- SonicWALL NSA E8500

- SonicWALL NSA E7500
- SonicWALL NSA E6500
- SonicWALL NSA E5500

To activate these licenses, register each appliance on MySonicWALL. Even when deployed in a High Availability pair, each unit must be individually registered to activate the licenses.

When available, a SonicOS Expanded License can be purchased for the following SonicWALL NSA appliances to activate BGP Advanced Routing:

- SonicWALL NSA 5000
- SonicWALL NSA 4500
- SonicWALL NSA 3500
- SonicWALL NSA 2400
- SonicWALL NSA 240
- SonicWALL TZ 210
- SonicWALL TZ 200

There is no Free Trial available for the BGP Routing Protocol feature.

Configuring BGP

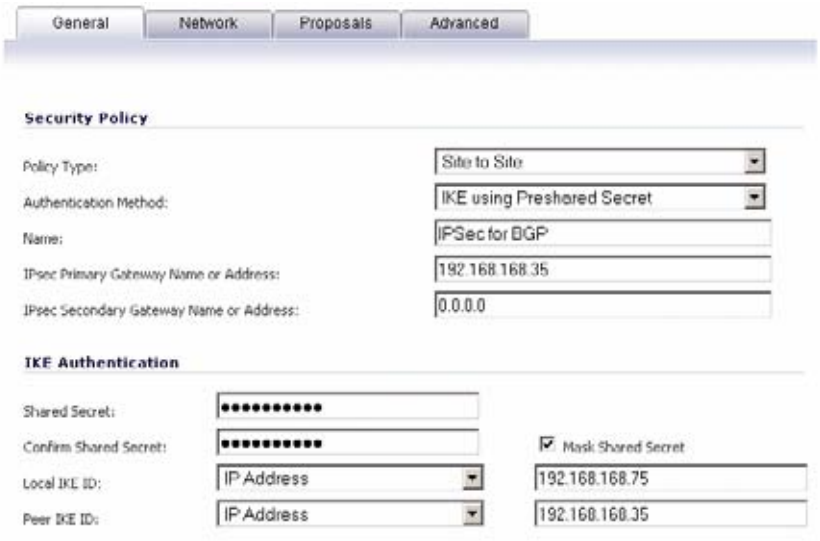
The following sections describe how to configure BGP Advanced Routing for SonicOS:

IPSec Configuration for BGP

BGP transmits packets in the clear. Therefore for strong security, SonicWALL recommends configuring an IPSec tunnel to use for BGP sessions. The configurations of the IPSec tunnel and of BGP are independent of each other. The IPSec tunnel is configured completely within the VPN configuration section of the SonicOS GUI, while BGP is enabled on the **Network > Routing** page and then configured on the SonicOS Command Line Interface. When configuring BGP over IPSec, first configure the IPSec tunnel and verify connectivity over the tunnel before configuring BGP.

The following procedure shows a sample IPSec configuration between a SonicWALL and a remote BGP peer, where the SonicWALL is configured for 192.168.168.75/24 on the X0 network and the remote peer is configured for 192.168.168.35/24 on the X0 network.

1. Navigate to the **VPN > Settings** page and click the **Add** button under the VPN Policies section. The VPN Policies window displays.



2. In the **Policy Type** pulldown menu, make sure that **Site to Site** is selected.

Note A site-to-site VPN tunnel must be used for BGP over IPSec. Tunnel interfaces will not work for BGP.

3. Select the desired **Authentication Method**. In this example, we are using **IKE using Preshared Secret**.
4. Enter a **Name** for the VPN policy.
5. In the **IPsec Primary Gateway Name or Address** field, enter the IP address of the remote peer (for this example it is 192.168.168.35).
6. In the **IPsec Secondary Gateway Name or Address** field, enter 0.0.0.0.
7. Enter a **Shared Secret** and confirm it.
8. In the **Local IKE ID** field, enter the IP address of the SonicWALL (for this example it is 192.168.168.75)
9. In the **Peer IKE ID** field, enter the IP address of the remote peer (192.168.168.35).

10. Click on the **Network** tab.

11. For the local network, select **X0 IP** from the **Choose local network from list** pulldown menu.
12. For the remote network, select the remote peer's IP address from the **Choose destination network from list** pulldown menu, which is 192.168.168.35 for this example. If the remote IP address is not listed, select **Create new address object** to create an address object for the IP address.
13. Click on the **Proposals** tab. You can either use the default IPSec proposals or customize them as you see fit.
14. Click on the **Advanced** tab.
15. Check the **Enable Keep Alive** checkbox.
16. Click **OK**.

The VPN policy is now configured on the SonicWALL appliance. Now complete the corresponding IPSec configuration on the remote peer. When that is complete, return to the **VPN > Settings** page and check the **Enable** checkbox for the VPN policy to initiate the IPSec tunnel.

Use the ping diagnostic on the SonicWall to ping the BGP peer IP address and use Wireshark to ensure that the request and response are being encapsulated in ESP packets.

Note As configured in this example, routed traffic will not go through the IPSEC tunnel used for BGP. That traffic is sent and received in the clear, which is most likely the desired behavior since the goal is to secure BGP, not all the routed network traffic.

For more detailed information on configuring IPSec, see the VPN chapters in the SonicOS Enhanced Administrator's Guide.

Basic BGP Configuration

To configure BGP on a SonicWALL security appliance, perform the following tasks:

1. On the SonicOS GUI, navigate to the **Network > Routing** page.
2. In the **Routing Mode** pulldown menu, select **Advanced Routing**.

Note The actual BGP configuration is performed using the SonicOS command line interface (CLI). For detailed information on how to connect to the SonicOS CLI, see the *SonicOS Command-Line Interface Guide* at: http://www.sonicwall.com/us/support/230_3623.html

3. Log in to the SonicOS CLI through the console interface.
4. Enter configuration mode by typing the **configure** command.
5. Enter the BGP CLI by typing the **route ars-bgp** command. You will now see the following prompt:

```
ZebOS version 7.7.0 IPIRouter 7/2009
ARS BGP>
```

6. You are now in BGP Non-Config Mode. Type **?** to see a list of non-config commands.
7. Type **show running-config** to see the current BGP running configuration.
8. To enter BGP Configuration Mode, type the **configure terminal** command. Type **?** to see a list of configuration commands.
9. When you have completed your configuration, type the **write file** command. If the unit is part of an High Availability pair or cluster, the configuration changes will be automatically conveyed to the other unit or units.

BGP Path Selection Process

The following attributes can be used to configure the BGP path selection process.

Attribute	Description
Weight	Prefer routes learned from neighbors with the highest weight set. Only relevant to the local router.
Local Preference	Administratively prefer routes learned from a neighbor. Shared with the whole AS.
Network or **Aggregate** paths	Prefer paths that were locally originated from the **network** and **aggregate-address** commands.
AS_PATH	Prefer the path with the shortest AS_PATH.
Origin	Prefer the path with the lowest origin type (as advertised in UPDATE messages): IGP < EGP < Incomplete.
Multi Exit Discriminator (MED)	Provides path preference information to neighbors for paths into originating AS.
Recency	Prefer the most recently received path.
Router ID	Prefer the path from the router with the lower router ID.

Weight

The weight command assigns a weight value, per address-family, to all routes learned from a neighbor. The route with the highest weight gets preference when the same prefix is learned from more than one peer. The weight is relevant only to the local router.

The weights assigned using the **set weight** command override the weights assigned using this command.

When the weight is set for a peer-group, all members of the peer-group will have the same weight. The command can also be used to assign a different weight to a particular peer-group member.

The following example shows weight configuration:

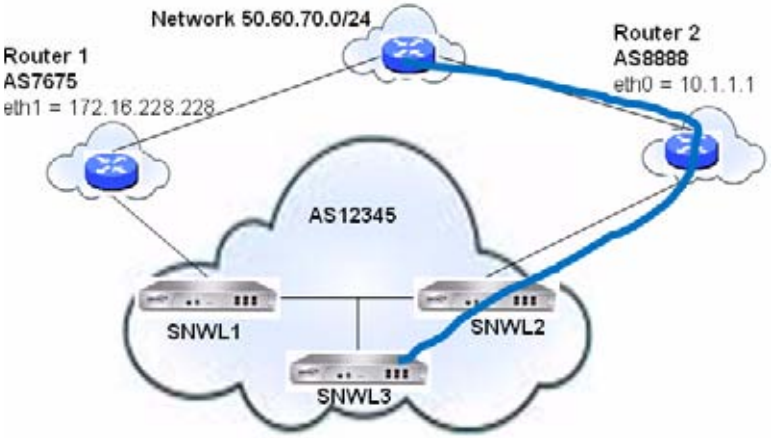
```
router bgp 12345
 neighbor 12.34.5.237 remote-as 12345
 neighbor 12.34.5.237 weight 60

router bgp 12345
```

```
neighbor group1 peer-group
neighbor 12.34.5.237 peer-group group1
neighbor 67.78.9.237 peer-group group1
neighbor group1 weight 60
```

Local Preference

The Local Preference attribute is used to indicate the degree of preference for each external route in an appliance's routing table. The Local Preference attribute is included in all update messages sent to devices in the same AS. Local Preference is not communicated to outside AS. The following figure shows a sample topology illustrating how Local Preference affects routes between neighboring ASs.

Figure 2 BGP Local Preference topology



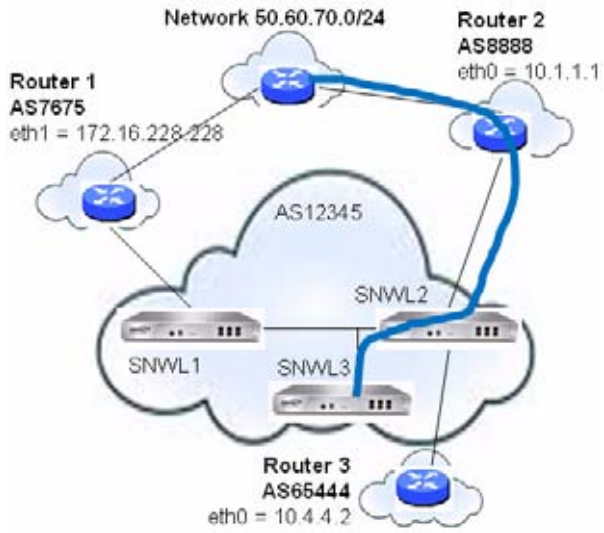
The following BGP configurations are entered on SNWL1 and SNWL2. The higher Local Preference on SNWL2 leads to SNWL2 being the preferred route advertised by AS 12345 (the SonicWALL AS) to outside ASs.

SNWL1 Configuration	SNWL2 Configuration
x0 = 12.34.5.228 x1 = 172.16.228.45 ------------------ router bgp 12345 neighbor 172.16.228.228 remote-as 7675 neighbor 12.34.5.237 remote-as 12345 bgp default local-preference 150	x0 = 12.34.5.237 x1 = 10.1.1.2 ------------------ router bgp 12345 neighbor 10.1.1.1 remote-as 8888 neighbor 12.34.5.228 remote-as 12345 bgp default local-preference 200

Local Preference used with Route Maps

Route Maps are similar to Access Control Lists. They consist of a series of Permit and/or Deny statements that determine how the appliance processes the routes. Route maps are applied to inbound traffic—not outbound traffic. The following diagram shows a sample topology that uses a route map to configure local preference.

Figure 3 BGP Local Preference topology with Route Maps



The following BGP configurations are entered on SNWL1 and SNWL2.

SNWL1 Configuration	SNWL2 Configuration

`x1 = 172.16.228.45`
 	
`------------------`
`router bgp 12345`
` neighbor 172.16.228.228 remote-as 7675`
` neighbor 12.34.5.237 remote-as 12345`
` bgp default local-preference 150`
	`x0 = 12.34.5.237`
	`x1 = 10.1.1.2`
	`x4 = 10.4.4.1`
	`------------------`
	`router bgp 12345`
	` neighbor 10.1.1.1 remote-as 9999`
	` neighbor 10.1.1.1 route-map rmap1 in`
	` neighbor 12.34.5.237 remote-as 12345`
	`....`
	`ip as-path access-list 100 permit ^8888$`
	`...`
	`route-map rmap1 permit 10`
	` match as-path 100`
	` set local-preference 200`
	
	`route-map rmap1 permit 20`
	` set local-preference 150`

The Route Map configured on SNWL2 (rmap1) is configured to apply to inbound routes from neighbor 10.1.1.1. It has two permit conditions:

- route-map rmap1 permit 10: This permit condition matches access list 100 that is configured to permit traffic from AS 8888 and set routes from AS 8888 to a Local Preference of 200.
- route-map rmap1 permit 10: This permit condition sets all other traffic that doesn't match access list 100 (i.e. traffic coming from ASs other than 8888) to a Local Preference of 150.

AS_PATH Prepending

AS_Path Prepending is the practice of adding additional AS numbers at the beginning of a path update. This makes the path for this route longer, and thus decreases its preference.

AS_Path Prepending can be applied on either outbound or inbound paths. AS_Path Prepending may not be honored if it is over-ruled by a neighbor.

Outbound Path Configuration	Inbound Path Configuration
router bgp 12345 bgp router-id 10.50.165.233 network 12.34.5.0/24 neighbor 10.50.165.228 remote-as 7675 neighbor 10.50.165.228 route-map long out ! route-map long permit 10 set as-path prepend 12345 12345	router bgp 7675 bgp router-id 10.50.165.228 network 7.6.7.0/24 neighbor 10.50.165.233 remote-as 12345 neighbor 10.50.165.233 route-map prepend in ! route-map prepend permit 10 set as-path prepend 12345 12345

This configuration leads to a route being installed to the neighbor 10.50.165.233 with the AS_Path Prepended as 12345 12345. This can be viewed by entering the **show ip bgp** command.

```
ARS BGP>show ip bgp
BGP table version is 98, local router ID is 10.50.165.228
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal, l -
labeled
              S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*> 12.34.5.0/24     10.50.165.233            0             0 12345 12345 12345 i
*> 7.6.7.0/24       0.0.0.0                       100  32768 i

Total number of prefixes 2
```

Multiple Exit Discriminator (MED)

The **set metric** command can be used in a route map to make paths more or less preferable:

```
router bgp 7675
network 7.6.7.0/24
 neighbor 10.50.165.233 remote-as 12345
 neighbor 10.50.165.233 route-map highmetric out
!
route-map highmetric permit 10
 set metric 300
```

The Multi Exit Discriminator (MED) is an optional attribute that can be used to influence path preference. It is non-transitive, meaning it is configured on a single appliance and not advertised to neighbors in update messages. In this section, we will consider the uses of the **bgp always-compare-med** and **bgp deterministic-med** commands.

bgp always-compare-med command

The **bgp always-compare-med** command allows comparison of the MED values for paths from different ASs for path selection. A path with lower MED is preferred.

As an example, consider the following routes in the BGP table and the **always-compare-med** command is enabled:

```
Route1: as-path 7675, med 300
```

```
Route2: as-path 200, med 200
Route3: as-path 7675, med 250
```

Route2 would be the chosen path because it has the lowest MED.

If the **always-compare-med** command was disabled, MED would not be considered when comparing Route1 and Route2 because they have different AS paths. MED would be compared for only Route1 and Route3.

bgp deterministic-med command

The selected route is also affected by the **bgp deterministic-med** command, which compares MED when choosing among routes advertised by different peers in the same autonomous system.

When the **bgp deterministic-med** command is enabled, routes from the same AS are grouped together, and the best routes of each group are compared. If the BGP table showed:

```
Route1: as-path 200, med 300, internal
Route2: as-path 400, med 200, internal
Route3: as-path 400, med 250, external
```

BGP would have a group of Route1 and a second group of Route2 and Route3 (the same AS).

The best of each group is compared. Route1 is the best of its group because it is the only route from AS 200.

Route1 is compared to the Route2, the best of group AS 400 (the lower MED).

Since the two routes are not from the same AS, the MED is not considered in the comparison. The external BGP route is preferred over the internal BGP route, making Route3 the best route.

BGP Communities

A community is a group of prefixes that share some common property and can be configured with the transitive BGP community attribute. A prefix can have more than one community attribute. Routers can act on one, some or all the attributes. BGP communities can be thought of as a form of tagging. The following is an example of a BGP communities configuration.

```
router bgp 12345
 bgp router-id 10.50.165.233
 network 12.34.5.0/24
 network 23.45.6.0/24
 neighbor 10.50.165.228 remote-as 7675
 neighbor 10.50.165.228 send-community
 neighbor 10.50.165.228 route-map comm out
!
access-list 105 permit 12.34.5.0/24
access-list 110 permit 23.45.6.0/24
!
route-map comm permit 10
 match ip address 105
 set community 7675:300
!
route-map comm permit 20
 match ip address 110
 set community 7675:500
!
router bgp 7675
 bgp router-id 10.50.165.228
 network 7.6.7.0/24
 neighbor 10.50.165.233 remote-as 12345
 neighbor 10.50.165.233 route-map shape in
!
ip community-list 1 permit 7675:300
ip community-list 2 permit 7675:500
!
```

```
route-map shape permit 10
 match community 1
 set local preference 120

route-map shape permit 20
 match community 2
 set local preference 130
```

Synchronization and Auto-Summary

The synchronization setting controls whether the router advertises routes learned from an iBGP neighbor based on the presence of those routes in its IGP. When synchronization is enabled, BGP will only advertise routes that are reachable through OSPF or RIP (the Exterior Gateway Protocols as opposed to BGP, the Exterior Gateway Protocol). Synchronization is a common cause of BGP route advertisement problems.

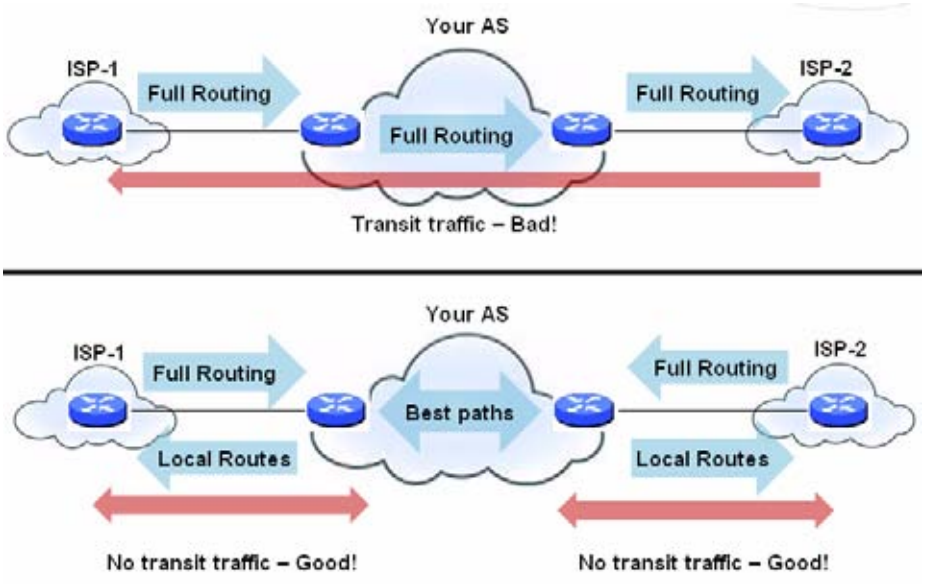
The auto-summary setting controls whether or not routes are advertised classfully. Auto-summary is another common cause of BGP configuration problems

By default, auto-summary and synchronization are disabled on Zebos.

Preventing an Accidental Transit AS

As we discussed earlier, an AS peer can either be a transit peer (allowing traffic from an outside AS to another outside AS) or a non-transit peer (requiring all traffic to either originate or terminate on its AS). Transit peers will have dramatically larger routing tables. Typically, you will not want to configure a SonicWALL security appliance as a transit peer.

Figure 4 Transit Peers vs. Non-Transit Peers



To prevent your appliance from inadvertently becoming a transit peer, you will want to configure inbound and outbound filters, such as the following:

Outbound Filters

Permit only routes originated from the local AS out

```
ip as-path access-list 1 permit ^$
```

```
router bgp 12345
 bgp router-id 10.50.165.233
 network 12.34.5.0/24
 neighbor 10.50.165.228 remote-as 7675
 neighbor 10.50.165.228 filter-list 1 out
 neighbor 172.1.1.2 remote-as 9999
 neighbor 10.50.165.228 filter list 1 out
```

Permit only owned prefixes out

```
ip prefix-list myPrefixes seq 5 permit 12.34.5.0/24
ip prefix-list myPrefixes seq 10 permit 23.45.6.0/24

router bgp 12345
 bgp router-id 10.50.165.233
 network 12.34.5.0/24
 network 23.45.6.0/24
 neighbor 10.50.165.228 remote-as 7675
 neighbor 172.1.1.2 remote-as 9999
 neighbor 10.50.165.228 prefix-list myPrefixes out
 neighbor 172.1.1.2 prefix-list myPrefixes out
```

Inbound Filters

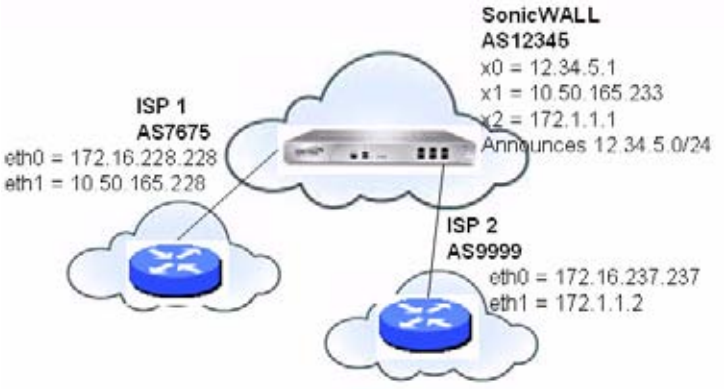
Drop all owned and private inbound prefixes

```
ip prefix-list unwantedPrefixes seq 5 deny 12.34.5.0/24 le 32
ip prefix-list unwantedPrefixes seq 10 deny 23.45.6.0/24 le 32
ip prefix-list unwantedPrefixes seq 20 deny 10.0.0.0/8 le 32
ip prefix-list unwantedPrefixes seq 21 deny 172.16.0.0/12 le 32
ip prefix-list unwantedPrefixes seq 22 deny 192.168.0.0/16 le 32
ip prefix-list unwantedPrefixes seq 30 permit 0.0.0.0/0 le 32

router bgp 12345
 bgp router-id 10.50.165.233
 network 12.34.5.0/24
 network 23.45.6.0/24
 neighbor 10.50.165.228 remote-as 7675
 neighbor 172.1.1.2 remote-as 9999
 neighbor 10.50.165.228 prefix-list unwantedPrefixes in
 neighbor 172.1.1.2 prefix-list unwantedPrefixes in
```

Using Multi-Homed BGP for Load Sharing

The following topology shows an example where a SonicWALL security appliance uses a multi-homed BGP network to load share between two ISPs.

Figure 5 Multi-Homed BGP for Load Sharing Topology



The SonicWALL security appliance is configured as follows:

```
router bgp 12345
 bgp router-id 10.50.165.233
 network 12.34.5.0/24
 neighbor 10.50.165.228 remote-as 7675
 neighbor 10.50.165.228 route-map ISP1 out
 neighbor 172.1.1.2 remote-as 9999
 neighbor 10.50.165.228 route-map ISP2 out
!
route-map ISP1 permit 10
match ip address 1
set weight 100

route-map ISP1 permit 20
match ip address 2

route-map ISP2 permit 10
match ip address 1

route-map ISP2 permit 20
match ip address 2
set weight 100

access-list 1 permit 12.34.5.0/25
access-list 2 deny 12.34.5.0/25
access-list 2 permit any
```

Verifying BGP Configuration

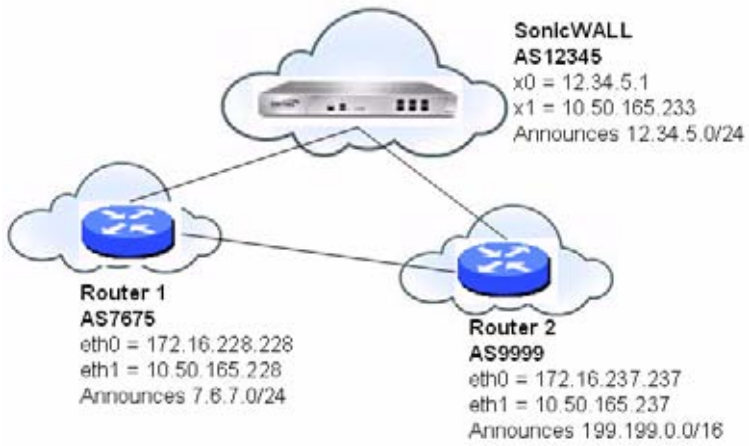
The following sections describe methods to verify a BGP configuration:

- "Viewing BGP FIB and RIB routes" on page 19
- "Configuring BGP Logging" on page 21

Viewing BGP FIB and RIB routes

Figure 6 shows a basic BGP topology where a SonicWALL security appliance is configured for BGP to connect to two routers on two different ASs.

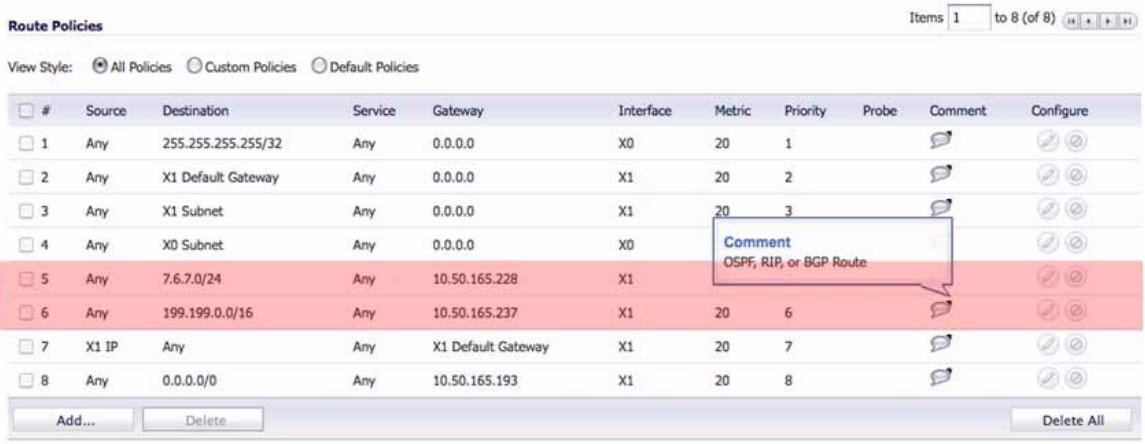
Figure 6 BGP Topology



The routes in the FIB for this network can be viewed either in the SonicOS GUI or by using the CLI.

Viewing FIB routes in the GUI

The BGP routes in the FIB can be viewed on the SonicOS GUI in the Routing Policies table on the **Network > Routing** page.



Viewing FIB Routes in the CLI

To view the FIB routes in the CLI, perform the following commands:

```
NSA 2400> configure
(config[NSA 2400])> route ars-nsm

ZebOS version 7.7.0 IPIRouter 7/2009
ARS NSM>show ip route
Codes: K - kernel, C - connected, S - static, R - RIP, B - BGP
       O - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2
       i - IS-IS, L1 - IS-IS level-1, L2 - IS-IS level-2, ia - IS-IS inter area
       * - candidate default

B       7.6.7.0/24 [20/0] via 10.50.165.228, X1, 05:08:31
B       199.199.0.0/16 [20/0] via 10.50.165.237, X1, 05:08:31
C       10.50.165.192/26 is directly connected, X1
C       127.0.0.0/8 is directly connected, lo0
C       12.34.5.0/24 is directly connected, X0
```

Viewing RIB Routes in the CLI

To view the RIB routes in the CLI, enter the **show ip bgp** command:

```
ARS BGP>show ip bgp
BGP table version is 98, local router ID is 10.50.165.233
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal, l -
labeled
              S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*> 7.6.7.0/24       10.50.165.228            0             0 7675 i
*> 12.34.5.0/24     0.0.0.0                       100  32768 i
*> 199.199.0.0/16   10.50.165.228            0             0 7675 9999 i

Total number of prefixes 3
```

Note The last route is the path to AS9999 that was learned through AS7675.

Configuring BGP Logging

SonicWALL BGP offers a comprehensive selection of debug commands to display log events related to BGP traffic. BGP logging can be configured on the CLI by using the **debug bgp** command followed by of the following keywords:

BGP Debug Keywords	Description
all	Enables all BGP debugging.
dampening	Enables debugging for BGP dampening.
events	Enables debugging for BGP events.
filters	Enables debugging for BGP filters.
fsm	Enables debugging for BGP Finite State Machine (FSM).
keepalives	Enables debugging for BGP keepalives.
nht	Enables debugging for NHT messages.
nsm	Enables debugging for NSM messages.
updates	Enables debugging for inbound/outbound BGP updates.

To disable BGP debugging, enter the "no" form of the command. For example, to disable event debugging, type the **no debug events** command.

BGP log messages can also be viewed on the SonicOS GUI on the **Log > View** page. BGP messages are displayed as part of the **Advanced Routing** category of log messages.

The above message indicates that an update to the outgoing RIB was denied because the router from which the update was received was not directly connected to the appliance.

To allow for BGP peers that are not directly connected, use the **ebgp-multihop** keyword with the **neighbor** command. For example:

```
neighbor 10.50.165.228 ebgp-multihop
```

BGP Terms

ARD – Autonomous Routing Domain – A collection of networks/routers that have a common administrative routing policy.

AS - Autonomous System – An ARD that has been assigned an identifying number, typically running BGP4 at its border router(s).

BGP4 - Border Gateway Protocol 4: The most prevalent EGP.

CIDR – Classless inter-domain routing, enables efficient route advertisement through route aggregation.

CPE – Customer Premise Equipment - The equipment at the edge of a customer's network used to interface with the ISP.

EGP - Exterior Gateway Protocol – Any protocol (in practice, BGP4) used to communicate routing information between Autonomous Systems.

Full-Routes - The entire global BGP route table.

FIB - Forwarding Information Base – Our existing route table, used to find the egress interface and next hop when forwarding packets.

Looking Glass* - A Looking Glass (LG) server is a read-only view of routers of organizations running the LG servers. Typically, publicly accessible looking glass servers are run by ISPs or NOCs.

Multi-Homed - An ISP customer that has multiple connections to one or more ISPs.

Multi-Provider - An ISP customer that uses multiple ISPs to connect to the Internet.

NSM – Network Services Module - The ZebOS component that centralizes the interface to the FIB and RIB. The separate routing protocol daemons interface with the NSM for all RIB updates. NSM alone updates the FIB with best-route information from the RIB.

Partial Routes - A subset of the full BGP route table, usually specific to destinations that are part of an ISP's domain.

RIB - Route Information Base – A run-time database owned by the NSM, and used to store all route information gathered and used by the routing protocols.