

## Chapter 29

# BGP Overview

The Border Gateway Protocol (BGP) is an exterior gateway protocol (EGP) that is used to exchange routing information among routers in different autonomous systems (ASs). BGP routing information includes the complete route to each destination. BGP uses the routing information to maintain a database of network reachability information, which it exchanges with other BGP systems. BGP uses the network reachability information to construct a graph of AS connectivity, thus allowing BGP to remove routing loops and enforce policy decisions at the AS level.

Multiprotocol BGP (MBGP) extensions enable BGP to support Internet Protocol version 6 (IPv6). MBGP defines the attributes `MP_REACH_NLRI` and `MP_UNREACH_NLRI`, which are used to carry IPv6 reachability information. Network layer reachability information (NLRI) update messages carry IPv6 address prefixes of feasible routes.

BGP allows for policy-based routing. You can use routing policies to choose among multiple paths to a destination and to control the redistribution of routing information.

BGP uses the Transmission Control Protocol (TCP) as its transport protocol, using port 179 for establishing connections. Running over a reliable transport protocol eliminates the need for BGP to implement update fragmentation, retransmission, acknowledgment, and sequencing.

The JUNOS routing protocol software supports BGP version 4. This version of BGP adds support for classless interdomain routing (CIDR), which eliminates the concept of network classes. Instead of assuming which bits of an address represent the network by looking at the first octet, CIDR allows you to explicitly specify the number of bits in the network address, thus providing a means to decrease the size of the routing tables. BGP version 4 also supports aggregation of routes, including the aggregation of AS paths.

This chapter discusses the following topics that provide background information about BGP:

BGP Standards on page 456

Autonomous Systems on page 457

AS Paths and Attributes on page 457

External and Internal BGP on page 457

BGP Routes on page 458

BGP Messages on page 459

## BGP Standards

---

The JUNOS software supports BGP version 4 and several extensions to the protocol, which are defined in the following documents:

RFC 1771, *A Border Gateway Protocol 4 (BGP-4)*

RFC 1772, *Application of the Border Gateway Protocol in the Internet*

RFC 1965, *Autonomous System Confederations for BGP*

RFC 1966, *BGP Route Reflection: An Alternative to Full-Mesh IBGP*

RFC 1997, *BGP Communities Attribute*

RFC 2270, *Using a Dedicated AS for Sites Homed to a Single Provider*

RFC 2283, *Multiprotocol Extensions for BGP-4*

RFC 2385, *Protection of BGP Sessions via the TCP MD5 Signature Option*

RFC 2439, *BGP Route Flap Damping*

RFC 2545, *Use of BGP-4 Multiprotocol Extensions for IPv6 Inter-Domain Routing*

RFC 2796, *BGP Route Reflection*

RFC 2858, *Multiprotocol Extensions for BGP-4*

RFC 3065, *Autonomous System Confederations for BGP*

RFC 3107, *Carrying Label Information in BGP-4*

Internet draft draft-ietf-ppvpn-rfc2547bis-00.txt, *BGP/MPLS VPNs* (expires January 2002)

Internet draft draft-ietf-idr-cap-neg-01.txt, *Capabilities Negotiation with BGP4* (expires February 1998)

Internet draft draft-kato-bgp-ipv6-link-local-00.txt, *BGP4+ Peering Using IPv6 Link-local Address* (expires April 2002)

Internet draft draft-ietf-idr-restart-06.txt, *Graceful Restart Mechanism for BGP* (expires July 2003)

Internet draft draft-ietf-mpls-bgp-mpls-restart-03.txt, *Graceful Restart Mechanism for BGP with MPLS* (expires August 2004)

Internet draft draft-ietf-ngtrans-bgp-tunnel-04.txt, *Connecting IPv6 Islands across IPv4 Clouds with BGP* (only MP-BGP over IPv4 Approach) (expires July 2002)

To access Internet Requests for Comments (RFCs) and drafts, go to the Internet Engineering Task Force (IETF) Web site at <http://www.ietf.org>.

## Autonomous Systems

---

An *autonomous system* (AS) is a set of routers that are under a single technical administration and normally use a single interior gateway protocol and a common set of metrics to propagate routing information within the set of routers. To other ASs, an AS appears to have a single, coherent interior routing plan and presents a consistent picture of what destinations are reachable through it.

## AS Paths and Attributes

---

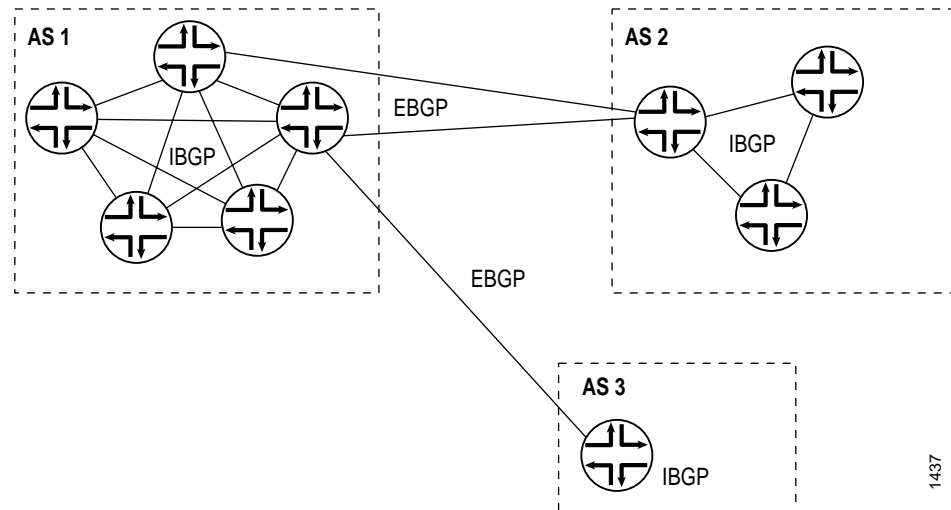
The routing information that BGP systems exchange includes the complete route to each destination, as well as additional information about the route. The route to each destination is called the *AS path*, and the additional route information is included in *path attributes*. BGP uses the AS path and the path attributes to completely determine the network topology. Once BGP understands the topology, it can detect and eliminate routing loops and select among groups of routes to enforce administrative preferences and routing policy decisions.

## External and Internal BGP

---

BGP supports two types of exchanges of routing information: exchanges between different ASs and exchanges within a single AS. When used between ASs, BGP is called *external BGP* (EBGP) and BGP sessions perform *inter-AS routing*. When used within an AS, BGP is called *internal BGP* (IBGP) and BGP sessions perform *intra-AS routing*. Figure 6 on page 458 illustrates ASs, IBGP, and EBGP.

Figure 6: ASs, External BGP, and Internal BGP



1437

A BGP system shares network reachability information with adjacent BGP systems, which are referred to as *neighbors* or *peers*.

BGP systems are arranged into *groups*. In an internal BGP group, all peers in the group—called *internal peers*—are in the same AS. Internal peers can be anywhere in the local AS and do not have to be directly connected to each other. Internal groups use routes from an IGP to resolve forwarding addresses. They also propagate external routes among all other internal routers running internal BGP, computing the next hop by taking the BGP next hop received with the route and resolving it using information from one of the interior gateway protocols.

In an external BGP group, the peers in the group—called *external peers*—are in different ASs and normally share a subnet. In an external group, the next hop is computed with respect to the interface that is shared between the external peer and the local router.

## BGP Routes

A BGP route consists of the following:

- A destination, described as an IP address prefix.

- Information that describes the path to the destination, including the following:

- AS path, which is a list of numbers of the ASs that a route passes through to reach the local router. The first number in the path is that of the last AS in the path—the AS closest to the local router. The last number in the path is the AS farthest from the local router, which is generally the origin of the path.

- Path attributes, which contain additional information about the AS path that is used in routing policy.

BGP peers advertise routes to each other in update messages.

BGP stores its routes in the JUNOS software routing table. The routing table stores the following information about BGP routes:

- Routing information learned from update messages received from peers

- Local routing information that the BGP system selects by applying local policies to routes received in update messages

- Information that the BGP system selects to advertise to its BGP peers in the update messages it sends

For each prefix in the routing table, the routing protocol process selects a single best path, called the active path. The algorithm for determining the active path is described in “How the Active Route Is Determined” on page 7.

## BGP Messages

---

BGP systems send four types of messages:

- Open

- Update

- Keepalive

- Notification

All BGP messages have the same fixed-size header, which contains a marker field indicating the total length of the message and a type field indicating the message type.

### ***Open Messages***

After a TCP connection is established between two BGP systems, they exchange BGP open messages to create a BGP connection between them. Once the connection is established, the two systems can exchange BGP messages and data traffic.

Open messages consist of the BGP header plus the following fields:

- Version—The current BGP version number is 4.

- Local AS number—You configure this with the `autonomous-system` statement at the [edit routing-options] or [edit logical-routers *logical-router-name* routing-options] hierarchy levels.

- Hold time—Proposed hold-time value. You configure the local hold time with the `BGP hold-time` statement.

BGP identifier—IP address of the BGP system. This address is determined when the system starts up and is the same for every local interface and every BGP peer. You can configure the BGP identifier with the `router-id` statement at the [edit routing-options] or [edit logical-routers *logical-router-name* routing-options] hierarchy levels. By default, BGP uses the IP address of the first interface it finds in the router.

Parameter field length and the parameter itself—These are optional fields.

## Update Messages

BGP systems send update messages to exchange network reachability information. BGP systems use this information to construct a graph that describes the relationships among all known ASs.

Update messages consist of the BGP header plus the following optional fields:

Unfeasible routes length—Length of the field that lists the routes being withdrawn from service because they are no longer deemed reachable

Withdrawn routes—IP address prefixes for the routes being withdrawn from service

Total path attribute length—Length of the field that lists the path attributes for a feasible route to a destination

Path attributes—Properties of the routes, including the path origin, the multiple exit discriminator (MED), the originating system's preference for the route, and information about aggregation, communities, confederations, and route reflection

Network layer reachability information (NLRI)—IP address prefixes of feasible routes being advertised in the update message

## Keepalive Messages

BGP systems exchange keepalive messages to determine whether a link or host has failed or is no longer available. Keepalive messages are exchanged often enough so that the hold timer does not expire. These messages consist only of the BGP header.

## Notification Messages

BGP systems send notification messages when an error condition is detected. After the message is sent, the BGP session and the TCP connection between the BGP systems are closed. Notification messages consist of the BGP header plus the error code and subcode, and data that describes the error.