Cisco – BGP Best Path Selection Algorithm

Table of Contents

BGP Best Path Selection Algorithm

Description

Border Gateway Protocol (BGP) routers typically receive multiple paths to the same destination. The BGP best path algorithm decides which is the best path to install in the IP routing table and to use for forwarding traffic.

Let's begin by assuming that all received paths for a particular prefix are arranged in a list, similar to the output of the **show ip bgp <longer−prefixes>** command. Some paths received by the router aren't considered as candidates for the best path. Such paths typically don't have the valid flag in the output of the **show ip bgp <longer−prefixes>** command. The following is a list of reasons that cause routers to ignore paths.

- Paths marked as "not synchronized" in the **show ip bgp <longer−prefixes>** output. If BGP synchronization is enabled, which it is by default in Cisco IOS® Software, there must be a match for the prefix in the IP routing table in order for an internal (iBGP) path to be considered a valid path. If the matching route is learned from an OSPF neighbor, its OSPF router ID must match the BGP router ID of the iBGP neighbor. Most users prefer to disable synchronization using the **no synchronization** BGP subcommand.

- Paths for which the NEXT_HOP is inaccessible. This is why it's important to have an IGP route to the NEXT_HOP associated with the path.

- Paths from an external (eBGP) neighbor if the local autonomous system (AS) appears in the AS_PATH. Such paths are denied upon ingress into the router, and are not even installed in the BGP routing−information base (RIB). The same applies to any path denied by routing policy implemented via access, prefix, AS_PATH, or community lists, unless you've configured **soft−reconfiguration inbound** for the neighbor.

- If you enabled **bgp enforce−first−as** and the UPDATE doesn't contain the AS of the neighbor as the first AS number in the AS_SEQUENCE, the router sends a notification and closes the session.

- Paths marked as "(received−only)" in the **show ip bgp <longer−prefixes>** output. These paths have been rejected by policy, but have been stored by the router because **soft−reconfiguration inbound** has been configured for the neighbor sending the path.

How the Best Path Algorithm Works

BGP assigns the first valid path as the current best path. It then compares the best path with the next path in list, until it reaches the end of the list of valid paths. Following is a list of rules used to determine the best path:

1. Prefer the path with the largest WEIGHT. **Note**: WEIGHT is a Cisco−specific parameter, local to the router on which it's configured.

2. Prefer the path with the largest LOCAL_PREF.

3. Prefer the path that was locally originated via a **network** or **aggregate** BGP subcommand, or through redistribution from an IGP. Local paths sourced by **network/redistribute** commands are preferred over local aggregates sourced by the **aggregate−address** command.

4. Prefer the path with the shortest AS_PATH. Note the following:

    ♦ This step is skipped if **bgp bestpath as−path ignore** is configured.
    ♦ An AS_SET counts as 1, no matter how many ASs are in the set.
    ♦ The AS_CONFED_SEQUENCE is not included in the AS_PATH length.

5. Prefer the path with the lowest origin type: IGP is lower than EGP, and EGP is lower than INCOMPLETE.

6. Prefer the path with the lowest multi−exit discriminator (MED). Note the following:

    ♦ This comparison is only done if the first (neighboring) AS is the same in the two paths; any confederation sub−ASs are ignored. In other words, MEDs are compared only if the first AS in the AS_SEQUENCE is the same for multiple paths. Any preceding AS_CONFED_SEQUENCE is ignored.
    ♦ If **bgp always−compare−med** is enabled, MEDs are compared for all paths. This option needs to be enabled over the entire AS, otherwise routing loops can occur.
    ♦ If **bgp bestpath med−confed** is enabled, MEDs are compared for all paths that consist only of AS_CONFED_SEQUENCE (paths originated within the local confederation).
    ♦ Paths received from a neighbor with a MED of 4,294,967,295 will have the MED changed to 4,294,967,294 before insertion into the BGP table.
    ♦ Paths received with no MED are assigned a MED of 0, unless **bgp bestpath missing−as−worst** is enabled, in which case they are assigned a MED of 4,294,967,294.
    ♦ The **bgp deterministic med** command can also influence this step as demonstrated in the How BGP Routers Use the Multi−Exit Discriminator for Best Path Selection.

7. Prefer external (eBGP) over internal (iBGP) paths. **Note**: Paths containing AS_CONFED_SEQUENCE are local to the confederation, and therefore treated as internal paths. There is no distinction between Confederation External and Confederation Internal.

8. Prefer the path with the lowest IGP metric to the BGP next hop.

9. If **maximum−paths *n*** is enabled, and there are multiple external or confederation−external paths from the same neighboring AS or sub−AS, BGP inserts up to *n* most recently received paths in the IP routing table. This allows eBGP multipath load sharing. The maximum value of *n* is currently 6. The default value, when this option is disabled, is 1. The oldest received path is marked as the best path in the output of **show ip bgp <longer−prefixes>**, and the equivalent of **next−hop−self** is performed before forwarding this best path to internal peers.

10. When both paths are external, prefer the path that was received first (the oldest one). This step minimizes route−flap, since a newer path won't displace an older one, even if it was the preferred route based on additional decision criteria, as described in steps 11, 12, and 13.

    Skip this step if any of the following is true:

    ♦ The **bgp best path compare−routerid** command is enabled.

> **Note:** This command was introduced in Cisco IOS® Software Releases 12.0.11S, 12.0.11SC, 12.0.11S3, 12.1.3, 12.1.3AA, 12.1.3.T, and 12.1.3.E.

- ♦ The router ID is the same for multiple paths, since the routes were received from the same router.

- ♦ There is no current best path. An example of losing the current best path occurs when the neighbor offering the path goes down.

11. Prefer the route coming from the BGP router with the lowest router ID. The router ID is the highest IP address on the router, with preference given to loopback addresses. It can also be set manually using the **bgp router−id** command. **Note**: If a path contains route−reflector (RR) attributes, the originator ID is substituted for the router ID in the path selection process.

12. If the originator or router ID is the same for multiple paths, prefer the path with the minimum cluster ID length. This will only be present in BGP route−reflector environments. It allows clients to peer with RRs or clients in other clusters. In this scenario, the client must be aware of the RR−specific BGP attribute.

13. Prefer the path coming from the lowest neighbor address. This is the IP address used in the BGP **neighbor** configuration, and corresponds to the remote peer used in the TCP connection with the local router.

Tools Information

For additional resources, refer to Cisco [TAC Tools for Routing Protocol Technologies](#).

Related Information

- **[IP Routing Top Issues](#)**
- **[BGP Support Page](#)**
- **[More BGP Technical Tips](#)**
- **[More Routing Protocol Technical Tips](#)**

Updated: Feb 22, 2002   Document ID: 13753

Cisco – BGP Best Path Selection Algorithm