

GRE Equal-Cost Multi-Path routing (ECMP)

- Feature history for ECMP, on page 1
- Understanding IP multicast load splitting, on page 2
- What is IP multicast load splitting?, on page 2
- Load splitting versus load balancing, on page 2
- Default behavior for IP multicast when multiple equal-cost paths exist, on page 3
- Methods to load split IP multicast traffic, on page 3
- Relationship with Cisco Express Forwarding (CEF), on page 4
- Restrictions for IP Multicast Load Splitting, on page 4
- How to Load Split IP Multicast Traffic over ECMP, on page 5
- Configuration Examples for Load Splitting IP Multicast Traffic over ECMP, on page 8

Feature history for ECMP

This table provides release and platform support information for the features explained in this module.

These features are available in all the releases subsequent to the one they were introduced in, unless noted otherwise.

Release	Feature name and description	Supported platform
Cisco IOS XE 26.2.1ea	IP multicast load splitting: IP Multicast Load Splitting feature support has been introduced.	Cisco C9550 Series Smart Switches

Release	Feature name and description	Supported platform
Cisco IOS XE 17.18.1	IP multicast load splitting: IP Multicast Load Splitting is a mechanism that distributes (*, *G*) (any-source multicast) and (*S*, *G*) (source-specific multicast) traffic streams across multiple equal-cost Reverse Path Forwarding (RPF) paths. This distribution is achieved by making a pseudorandom decision for each traffic flow, aiming to spread the multicast traffic load across available paths. This process is collectively referred to as equal-cost multipath (ECMP) multicast load splitting	Cisco C9350 Series Smart Switches Cisco C9610 Series Smart Switches

Understanding IP multicast load splitting

This chapter provides an overview of IP Multicast Load Splitting across Equal-Cost Paths, explaining its purpose, how it differs from load balancing, and the various methods available. This feature is introduced on 9350 and 9610 platforms and is only applicable to IPv4. IPv6 ECMP multicast is not supported.

What is IP multicast load splitting?

IP Multicast Load Splitting is a mechanism that distributes (*, *G*) (any-source multicast) and (*S*, *G*) (source-specific multicast) traffic streams across multiple equal-cost Reverse Path Forwarding (RPF) paths. This distribution is achieved by making a pseudorandom decision for each traffic flow, aiming to spread the multicast traffic load across available paths. This process is collectively referred to as equal-cost multipath (ECMP) multicast load splitting.

Load splitting versus load balancing

It is important to understand that load splitting and load balancing are not synonymous.

- **Load Splitting**: Provides a means to randomly distribute traffic streams across multiple equal-cost RPF paths. It attempts to distribute an equal amount of traffic flows on each available RPF path by making a pseudorandom decision, rather than by counting flows. This method results in better load sharing in networks with many traffic streams utilizing similar bandwidth.
- **Load Balancing:** Implies a more even and predictable distribution of traffic. Load splitting does not necessarily result in a perfectly balanced IP multicast traffic load on equal-cost RPF paths, especially if there are only a few (*S*, *G*) or (*, *G*) states.

Default behavior for IP multicast when multiple equal-cost paths exist

By default, for Protocol Independent Multicast (PIM-SM), Source Specific Multicast (PIM-SSM), and bidirectional PIM (bidir-PIM) groups, if multiple equal-cost paths are available, RPF for IPv4 multicast traffic is based on the PIM neighbor with the highest IP address. This is known as the *highest PIM neighbor behavior*. This default behavior does not perform any form of ECMP load splitting.

Methods to load split IP multicast traffic

You can configure ECMP multicast load splitting using the `ip multicast multipath` command with different options. All methods, except the default, enable some form of ECMP multicast load splitting.

ECMP Multicast Load Splitting Based on Source Address (S-hash algorithm)

- **Purpose:** Enables the RPF interface for each (*, *G*) or (*S*, *G*) state to be selected among available equal-cost paths based on the source address (for (*S*, *G*)) or Rendezvous Point (RP) address (for (*, *G*)).
- **Predictability:** This algorithm is predictable because it uses no randomization in calculating the hash value.
- **Polarization:** This method is subject to *polarization*, a problem where routers in some topologies may not effectively utilize all available paths for load splitting.

ECMP Multicast Load Splitting Based on Source and Group Address (basic S-G-hash algorithm)

- **Purpose:** Uses a simple hash based on both the source and group addresses. This provides more flexible support for load splitting, especially for devices sending many streams or broadcasting many channels (e.g., IPTV servers).
- **Predictability:** Predictable due to the absence of randomization.
- **Polarization:** Also subject to *polarization*.



Note ECMP multicast load splitting does not support bidirectional PIM groups.

ECMP Multicast Load Splitting Based on Source, Group, and Next-Hop Address (next-hop-based S-G-hash algorithm)

- **Purpose:** Employs a more complex hash based on source, group, and the next-hop address.

- **Predictability:** Predictable as it does not use randomization.
- **Polarization: Not subject to polarization**, making it more robust. This is achieved by introducing the actual next-hop IP address of PIM neighbors into the calculation, causing hash results to differ for each device.
- **Stability:** Increases stability of RPF paths during path failures, as only states using the failed path reconverge.
- **IPv6 ECMP:** This is the same algorithm used for IPv6 ECMP multicast load splitting.

Relationship with Cisco Express Forwarding (CEF)

The principles of ECMP multicast load splitting are similar to the default per-flow load splitting in IPv4 CEF or load splitting used with EtherChannels. CEF unicast load splitting also uses methods that do not exhibit polarization.

Restrictions for IP Multicast Load Splitting

Consider these restrictions before configuring IP Multicast Load Splitting.

- **Default Behavior:** By default, multicast traffic is not load split across multiple equal-cost paths. Instead, it flows from the RPF neighbor with the highest IP address if multiple neighbors have the same metric.
- **Load Splitting vs. Load Balancing:** The `ip multicast multipath` command performs load *splitting* , not load *balancing* . This means traffic from a single source will use only one path, even if that traffic significantly outweighs traffic from other sources.
- **PIM Neighbor IP Address Uniqueness:** The `ip multicast multipath` command does not support configurations where the same PIM neighbor IP address is reachable through multiple equal-cost paths. This situation can occur if unnumbered interfaces are used. You must use different IP addresses for all interfaces when configuring this command.
- **Consistent Configuration:** Because the `ip multicast multipath` command changes how an RPF neighbor is selected, it must be configured consistently on all devices in a redundant topology to avoid routing loops.
- **Polarization (S-hash and basic S-G-hash):** The S-hash and basic S-G-hash algorithms are subject to polarization, which can prevent effective utilization of all available paths in certain topologies. The next-hop-based S-G-hash algorithm avoids this.
- **PIM Assert Process:** For PIM-SM and PIM-SSM, the operation of ECMP multicast load splitting can only be guaranteed when all downstream devices on a LAN are consistently configured Cisco devices. Inconsistent configurations or third-party devices can lead to the PIM assert process overriding the load splitting.
- **BGP Multipath Requirement:** To leverage ECMP multicast load splitting for BGP-learned prefixes, BGP multipath must be enabled, as BGP only installs a single path by default.
- **Static Mroute Limitations:** You cannot use static mroutes to configure equal-cost paths for load splitting in IPv4 multicast if you need to configure one static mroute per prefix, due to software limitations. Workarounds using recursive route lookups do not apply to equal-cost multipath routing.

How to Load Split IP Multicast Traffic over ECMP

This section provides step-by-step instructions for enabling ECMP multicast load splitting on your device.

Prerequisites for IP Multicast Load Splitting

IP Multicast Enabled: IP multicast must be enabled on the device. Refer to the *Configuring Basic IP Multicast* module of the *IP Multicast Routing Configuration Guide* for instructions.

Multiple Paths: You must have multiple equal-cost paths available to the source or Rendezvous Point (RP).

Verification: Use the `show ip route ip-address` command (with the IP address of the source or RP) to validate that multiple paths are present in the routing table. ECMP multicast load splitting takes effect only when multiple paths are shown.

Example:

```
Device# show ip route 203.0.113.3
Routing entry for 203.0.113.3/32
  Known via "ospf 10", distance 110, metric 21, type intra area
  Last update from 198.51.100.2 on Ethernet1/2, 00:00:16 ago
  Routing Descriptor Blocks:
    198.51.100.2, from 203.0.113.3, 00:00:16 ago, via Ethernet1/2
      Route metric is 21, traffic share count is 1
  * 192.0.2.2, from 203.0.113.3, 00:04:42 ago, via Ethernet0/2
      Route metric is 21, traffic share count is 1
```

- **Sufficient Sources (for S-hash):** For ECMP multicast load splitting based on source address (S-hash), you must have an adequate number of sources (at least more than two) to see the benefits.
- **PIM-SM SPT Forwarding:** When using PIM-SM with shortest path tree (SPT) forwarding, the T-bit must be set for the forwarding of all (*S*, *G*) states.
- **RPF Validation:** It is best practice to use the `show ip rpf <source-address> [group-address]` command to validate whether sources can take advantage of IP multicast multipath capabilities before configuration.
- **BGP Multipath:** If using BGP, configure multipath using the `maximum-paths` command, as BGP does not install multiple equal-cost paths by default.

Enabling ECMP Multicast Load Splitting Based on Source Address

This procedure enables ECMP multicast load splitting of multicast traffic based on source address (using the S-hash algorithm) to take advantage of multiple paths through the network.

Procedure

Step 1 Enable privileged EXEC mode.

Example:

```
Device> enable
```

Enter your password if prompted.

Step 2 Enter global configuration mode.

Example:

```
Device# configure terminal
```

Step 3 Enable ECMP multicast load splitting based on source address.

Example:

```
Device(config)# ip multicast multipath
```

This command enables the S-hash algorithm.

Note
This command must be configured consistently on all devices in a redundant topology to avoid looping. It performs load splitting, not load balancing, and traffic from a source will use only one path.

Step 4 Repeat Step 3 on all relevant devices in your redundant topology.

Step 5 Exit global configuration mode.

Example:

```
Device(config)# exit
```

Step 6 (Optional) Verify RPF selection.

Example:

```
Device# show ip rpf <source-address> [group-address]
```

Use this command to ensure IP multicast traffic is being properly load split.

Step 7 (Optional) Verify IP routing table.

Example:

```
Device# show ip route <ip-address>
```

Use this command to confirm multiple paths are available to a source or RP for ECMP multicast load splitting.

Enabling ECMP Multicast Load Splitting Based on Source and Group Address

This procedure enables ECMP multicast load splitting of multicast traffic based on source and group address (using the basic S-G-hash algorithm). This algorithm provides more flexible support for load splitting than the S-hash, particularly for devices sending many streams (e.g., IPTV servers). It is predictable but subject to polarization.

Procedure

Step 1 Enable privileged EXEC mode.

Example:

```
Device> enable
```

Enter your password if prompted.

Step 2 Enter global configuration mode.

Example:

```
Device# configure terminal
```

Step 3 Enable ECMP multicast load splitting based on source and group address.

Example:

```
Device(config)# ip multicast multipath s-g-hash basic
```

This command enables the basic S-G-hash algorithm.

Note
This command must be configured consistently on all devices in a redundant topology to avoid looping.

Step 4 Repeat Step 3 on all relevant devices in your redundant topology.

Step 5 Exit global configuration mode.

Example:

```
Device(config)# exit
```

Step 6 (Optional) Verify RPF selection.

Example:

```
Device# show ip rpf <source-address> [group-address]
```

Use this command to ensure IP multicast traffic is being properly load split.

Step 7 (Optional) Verify IP routing table.

Example:

```
Device# show ip route <ip-address>
```

Use this command to confirm multiple paths are available to a source or RP for ECMP multicast load splitting.

Enabling ECMP Multicast Load Splitting Based on Source Group and Next-Hop Address

This procedure enables ECMP multicast load splitting of multicast traffic based on source, group, and next-hop address (using the next-hop-based S-G-hash algorithm). This algorithm provides the most flexible support for ECMP multicast load splitting and eliminates the polarization problem. It is the recommended method for most deployments.

Procedure

Step 1 Enable privileged EXEC mode.

Example:

```
Device> enable
```

Enter your password if prompted.

Step 2 Enter global configuration mode.

Example:

```
Device# configure terminal
```

Step 3 Enable ECMP multicast load splitting based on source, group, and next-hop address.

Example:

```
Device(config)# ip multicast multipath s-g-hash next-hop-based
```

This command enables the next-hop-based S-G-hash algorithm.

Note
This command must be configured consistently on all devices in a redundant topology to avoid looping.

Step 4 Repeat Step 3 on all relevant devices in your redundant topology.

Step 5 Exit global configuration mode.

Example:

```
Device(config)# end
```

Step 6 (Optional) Verify RPF selection.

Example:

```
Device# show ip rpf <source-address> [group-address]
```

Use this command to ensure IP multicast traffic is being properly load split.

Step 7 (Optional) Verify IP routing table.

Example:

```
Device# show ip route <ip-address>
```

Use this command to confirm multiple paths are available to a source or RP for ECMP multicast load splitting.

Configuration Examples for Load Splitting IP Multicast Traffic over ECMP

Example: Enable ECMP Multicast Load Splitting Based on Source Address

The following example shows how to enable ECMP multicast load splitting on a router based on source address using the S-hash algorithm:

```
Device(config)# ip multicast multipath
```

Example: Enable ECMP Multicast Load Splitting Based on Source and Group Address

The following example shows how to enable ECMP multicast load splitting on a router based on source and group address using the basic S-G-hash algorithm:

```
Device(config)# ip multicast multipath s-g-hash basic
```

Example: Enable ECMP Multicast Load Splitting Based on Source Group and Next-Hop Address

The following example shows how to enable ECMP multicast load splitting on a router based on source, group, and next-hop address using the next-hop-based S-G-hash algorithm:

```
Device(config)# ip multicast multipath s-g-hash next-hop-based
```