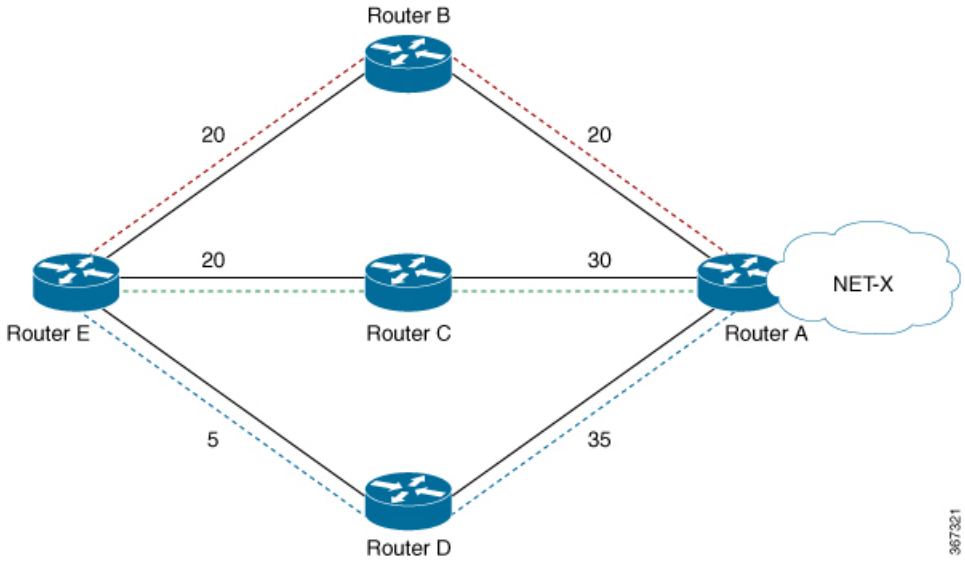
Implementing UCMP

The unequal cost multipath (UCMP) load-balancing provides the capability to load balance traffic proportionally across multiple paths, with different cost. Generally, higher bandwidth paths have lower Interior Gateway Protocol (IGP) metrics configured, so that they form the shortest IGP paths.

With the UCMP load-balancing enabled, protocols can use even lower bandwidth paths or higher cost paths for traffic, and can install these paths to the forwarding information base (FIB). These protocols still install multiple paths to the same destination in FIB, but each path will have a 'load metric/weight' associated with it. FIB uses this load metric/weight to decide the amount of traffic that needs to be sent on a higher bandwidth path and the amount of traffic that needs to be sent on a lower bandwidth path.

In the following example, there are 3 paths to get to Network X as follows:

Figure 1: Topology for UCMP



Paths	Cost from Router E to Net -X
E-B-A	40
E-C-A	50
E-D-A	40

IGP selects the lowest path links, i.e E-B-A and E-D-A. The path E-C-A is not considered for load balancing because of higher cost. The lowest path link E-D (5) is not a tie breaker, as the end to end cost to the Network X is considered.

More than 32 ECMP and UCMP paths are not supported for these features:

- LI
- GRE
- BVI
- NetFlow
- Satellite
- MCAST
- SPAN
- PWHE
- ABF
- P2MP
- MVPN
- VPLS
- L2TPv3
- LISP
- VIDMON
- PBR

- ECMP vs. UCMP Load Balancing, on page 2
- Managing Traffic with ECMP FEC in Network Environments, on page 3
- UCMP Minimum Integer Ratio, on page 3
- Configuring OSPF UCMP with Cost, on page 4
- Configuring OSPF UCMP with Cost Only for Certain Prefixes, on page 6
- Configuring IS-IS With Weight, on page 9
- Configuring IS-IS With Metric, on page 10
- Configuring BGP With Weights, on page 11
- Configuring TE Tunnel With Weights, on page 13
- Policy-Based Tunnel Selection, on page 13
- Interior Gateway Protocol (IGP) Destination-based Load Balancing (DLB), on page 26

ECMP vs. UCMP Load Balancing

Load balancing is a forwarding mechanism that distributes traffic over multiple links based on certain parameters. Equal Cost Multi Path (ECMP) is a forwarding mechanism for routing packets along multiple paths of equal cost with the goal to achieve almost equally distributed link load sharing. This significantly impacts a router's next-hop (path) decision.

In ECMP, it is assumed that all links available are of similar speed which inherently means that the hash values that are computed are equally shared over the multiple paths available.

For instance, if we have two paths available, the buckets (which in the end identify the links to be chosen) will be assigned in a 50% / 50% loadsharing. This can be problematic when one path is say a 10G link and the other link is a 1G link. In this case, you probably want to assign a (near) 90/10 type deviation, but considering that BGP is not bandwidth aware, the 10G path is still chosen 50% of the time as much as the 1G link. In this scenario, not all paths are of equal cost path.

What UCMP does in this case is apply a *weight* to a path which means that we are giving more hash buckets to one path that has a higher weight. The weight applied is *static* in the sense that it is derived by the DMZ bandwidth extended community either assigned to a peer or as configured via the Route Policy Language (RPL) route manipulation functionality.

In general, a routing protocol decides a best path to a destination based on a metric. This metric is generally driven by the bandwidth of the circuit. When we have 3 paths available, say 1G/10G/100G, routing protocols generally discard the 1G/10G paths available. In defined cases, one may want to spread the load over the circuits based on the load they can carry. In this example, one may want to distribute traffic in a 1%/10%/89% fashion over the 1G/10G/100G paths available.

Managing Traffic with ECMP FEC in Network Environments

In a network environment that uses multiple paths, the ECMP Forwarding Equivalence Class (FEC) helps balance traffic across paths that have the same cost. However, when the ECMP FEC reaches the Red threshold, it indicates that resources are depleted. At this point, new ECMP FEC will not be allocated to any newly advertised prefixes. Instead, it assigns a regular FEC, provided it is not in an Out of Resources (OOR) state.

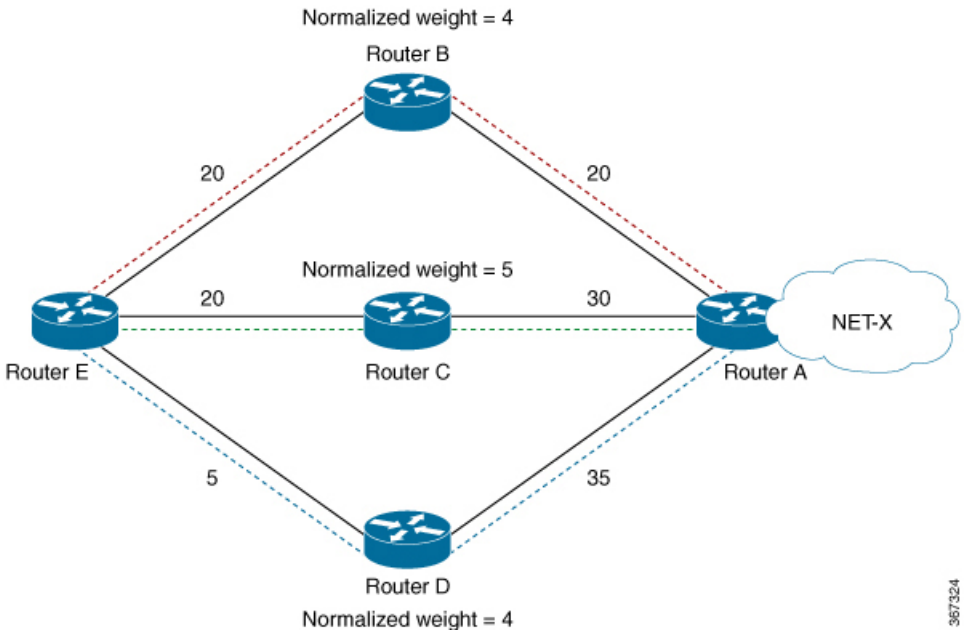
A notable limitation you may encounter is that after the ECMP FEC reaches the Red threshold, and despite stopping the allocation of new ECMP FEC, some in-transit ECMP FEC processes might still get programmed. These processes can run into resource unavailability, resulting in no resource errors, which can ultimately lead to traffic being dropped.

UCMP Minimum Integer Ratio

The UCMP Minimum Integer Ratio feature saves hardware resources when programming UCMP, by using optimized number of buckets.

To calculate the UCMP minimum integer ratio, find the greatest common divisor (GCD) and divide all the calculated normalized weights.

In the following Figure, we have three configured weights 40, 50, and 40, with GCD as 10. To calculate the normalized weight, divide the configured weight by GCD. In this example, we need to divide 40 by 10, 50 by 10, and 40 by 10, which is 4, 5, and 4 respectively. Therefore 4, 5, and 4 are the new normalized weights.



New normalized weights are: 40/10 = 4, 50/10 = 5, and 40/10 = 4

If GCD is 1, then Normalized Weight = (Path weight/Total weight) * Maximum bucket size

Configuring OSPF UCMP with Cost

The following example shows how to configure UCMP under OSPF default vrf with IPv4. The same can be done with IPv6.

R5 is advertising its two loopbacks with IP addresses 10.0.0.1 and 192.168.0.1 into OSPF. R1 learns these loopbacks through three directly connected next hops, R2, R3, and R4.

The metrics are as follows:

- Total link metric between R1 and R5 via R2 is 40.
- Total link metric between R1 and R5 via R3 is 50
- Total link metric between R1 and R5 via R4 is 40

Before

Based on the metrics, R1 installs two ECMP paths via R2 and R4 only. Path via R5 is not installed due to higher cost.

```
Router #show route 10.0.0.1/32

Routing entry for 10.0.0.1/32
  Known via "ospf 0", distance 110, metric 41, type intra area
  Installed Aug  3 19:09:21.399 for 03:57:18
  Routing Descriptor Blocks
    10.1.2.2, from 10.0.0.1, via GigabitEthernet0/0/0/0
      Route metric is 41
    10.1.4.2, from 10.0.0.1, via GigabitEthernet0/0/0/2
```

```
      Route metric is 41
  No advertising protos.
```

```
Router #show route 192.168.0.1/32
```

```
Routing entry for 192.168.0.1/32
  Known via "ospf 0", distance 110, metric 41, type intra area
  Installed Aug  3 23:06:05.258 for 00:00:41
  Routing Descriptor Blocks
    10.1.2.2, from 10.0.0.1, via GigabitEthernet0/0/0/0
      Route metric is 41
    10.1.4.2, from 10.0.0.1, via GigabitEthernet0/0/0/2
      Route metric is 41
```

Configuration

```
router ospf Proc1
ucmp
```

Example

```
Router (config)#router ospf 0
Router (config-ospf)#ucmp
Router (config-ospf)#commit
```

Verification

Verify the installation of three paths by R1, to reach each of the R5 loopback IP addresses.

```
Router #show route 10.0.0.1/32
```

```
Routing entry for 10.0.0.1/32
  Known via "ospf 0", distance 110, metric 41, type intra area
  Installed Aug  3 23:15:13.495 for 00:02:01
  Routing Descriptor Blocks
    10.1.2.2, from 10.0.0.1, via GigabitEthernet0/0/0/0
      Route metric is 41, Wt is 4294967295
    10.1.4.2, from 10.0.0.1, via GigabitEthernet0/0/0/2
      Route metric is 41, Wt is 4294967295
    10.1.3.2, from 10.0.0.1, via GigabitEthernet0/0/0/1
      Route metric is 41, Wt is 3452816845
  No advertising protos.
```

```
Router #show ospf routes 10.0.0.1/32
```

```
Topology Table for ospf 0 with ID 172.16.0.1:

Codes: O - Intra area, O IA - Inter area
       O E1 - External type 1, O E2 - External type 2
        O N1 - NSSA external type 1, O N2 - NSSA external type 2

O    10.0.0.1/32, metric 41
       10.1.4.2, from 10.0.0.1, via GigabitEthernet0/0/0/2, ifIndex 8, path-id 1
       10.1.2.2, from 10.0.0.1, via GigabitEthernet0/0/0/0, ifIndex 6, path-id 2
       10.1.3.2, from 10.0.0.1, via GigabitEthernet0/0/0/1, path-id 3, (UCMP, metric 51)
```

```
Router #show route 192.168.0.1/32
```

```
Routing entry for 192.168.0.1/32
  Known via "ospf 0", distance 110, metric 41, type intra area
  Installed Aug  3 23:15:13.495 for 00:02:06
  Routing Descriptor Blocks
    10.1.2.2, from 10.0.0.1, via GigabitEthernet0/0/0/0
      Route metric is 41, Wt is 4294967295
    10.1.4.2, from 10.0.0.1, via GigabitEthernet0/0/0/2
```

```
      Route metric is 41, Wt is 4294967295
    10.1.3.2, from 10.0.0.1, via GigabitEthernet0/0/0/1
      Route metric is 41, Wt is 3452816845
  No advertising protos.
```

```
Router #show ospf routes 192.168.0.1/32

Topology Table for ospf 0 with ID 172.16.0.1:

Codes: O - Intra area, O IA - Inter area
       O E1 - External type 1, O E2 - External type 2
       O N1 - NSSA external type 1, O N2 - NSSA external type 2

O    192.168.0.1/32, metric 41
       10.1.4.2, from 10.0.0.1, via GigabitEthernet0/0/0/2, ifIndex 8, path-id 1
       10.1.2.2, from 10.0.0.1, via GigabitEthernet0/0/0/0, ifIndex 6, path-id 2
       10.1.3.2, from 10.0.0.1, via GigabitEthernet0/0/0/1, path-id 3, (UCMP, metric 51)
```

Configuring OSPF UCMP with Cost Only for Certain Prefixes

The following example shows how to configure UCMP under OSPF default VRF with IPv4 only for prefixes that match an access list. Ensure to follow the same procedure with IPv6.

R5 is advertising its two loopbacks with IP addresses 10.0.0.1 and 192.168.0.1 into OSPF. R1 learns these loopbacks via three directly connected next hops, R2, R3, and R4.

The metrics are as follows:

- Total link metric between R1 and R5 via R2 is 40
- Total link metric between R1 and R5 via R3 is 50
- Total link metric between R1 and R5 via R4 is 40

Before

Based on the metrics, R1 installs two ECMP paths via R2 and R4 only. Path via R5 is not installed due to higher cost.

```
Router #show route 10.0.0.1/32

Routing entry for 10.0.0.1/32
  Known via "ospf 0", distance 110, metric 41, type intra area
  Installed Aug  3 19:09:21.399 for 03:57:18
  Routing Descriptor Blocks
    10.1.2.2, from 10.0.0.1, via GigabitEthernet0/0/0/0
      Route metric is 41
    10.1.4.2, from 10.0.0.1, via GigabitEthernet0/0/0/2
      Route metric is 41
  No advertising protos.
```

```
Router #show route 192.168.0.1/32

Routing entry for 192.168.0.1/32
  Known via "ospf 0", distance 110, metric 41, type intra area
  Installed Aug  3 23:06:05.258 for 00:00:41
  Routing Descriptor Blocks
    10.1.2.2, from 10.0.0.1, via GigabitEthernet0/0/0/0
      Route metric is 41
    10.1.4.2, from 10.0.0.1, via GigabitEthernet0/0/0/2
      Route metric is 41
```

Configuration

```
router ospf Proc1
ucmp prefix-list Acl1
```

Note OSPF UCMP prefix-list configuration should refer to an access list, not a prefix-set.

Example

Install UCMP only for prefix 192.168.0.1/32. Other prefixes will not be affected by the UCMP change.

```
Router (config)#ipv4 access-list foo
Router (config-ipv4-acl)#10 permit ipv4 host 192.168.0.1 any

Router (config)#router ospf 0
Router (config-ospf)#ucmp prefix-list foo
Router (config-ospf)#commit
```

Verification

Verify that R1 installs three paths to reach R5 loopback IP address 192.168.0.1/32.

```
Router #show route 192.168.0.1/32

Routing entry for 192.168.0.1/32
  Known via "ospf 0", distance 110, metric 41, type intra area
  Installed Aug  3 23:25:11.997 for 00:15:56
  Routing Descriptor Blocks
    10.1.2.2, from 10.0.0.1, via GigabitEthernet0/0/0/0
      Route metric is 41, Wt is 4294967295
    10.1.4.2, from 10.0.0.1, via GigabitEthernet0/0/0/2
      Route metric is 41, Wt is 4294967295
    10.1.3.2, from 10.0.0.1, via GigabitEthernet0/0/0/1
      Route metric is 41, Wt is 3452816845
  No advertising protos.

Router #show ospf routes 192.168.0.1/32

Topology Table for ospf 0 with ID 172.16.0.1:

Codes: O - Intra area, O IA - Inter area
       O E1 - External type 1, O E2 - External type 2
       O N1 - NSSA external type 1, O N2 - NSSA external type 2

O    192.168.0.1/32, metric 41
       10.1.4.2, from 10.0.0.1, via GigabitEthernet0/0/0/2, ifIndex 8, path-id 1
       10.1.2.2, from 10.0.0.1, via GigabitEthernet0/0/0/0, ifIndex 6, path-id 2
       10.1.3.2, from 10.0.0.1, via GigabitEthernet0/0/0/1, path-id 3, (UCMP, metric 51)

Router #show cef 192.168.0.1/32

192.168.0.1/32, version 186, internal 0x1000001 0x10 (ptr 0xe2ecc50) [1], 0x400 (0xe489c30),
 0x0 (0x0)
Updated Aug  3 23:43:10.261
 remote adjacency to GigabitEthernet0/0/0/0
Prefix Len 32, traffic index 0, precedence n/a, priority 1
  gateway array (0xe2f2060) reference count 2, flags 0x0, source rib (7), 0 backups
                [3 type 3 flags 0x8401 (0xe3a2ca8) ext 0x0 (0x0)]
  LW-LDI[type=3, refc=1, ptr=0xe489c30, sh-ldi=0xe3a2ca8]
  gateway array update type-time 1 Aug  3 23:43:10.261
LDI Update time Aug  3 23:43:10.261
LW-LDI-TS Aug  3 23:43:10.261
```

```
   via 10.1.2.2/32, GigabitEthernet0/0/0/0, 6 dependencies, weight 4294967295, class 0
[flags 0x0]
    path-idx 0 NHID 0x0 [0xf4270b0 0x0]
    next hop 10.1.2.2/32
    remote adjacency
   via 10.1.4.2/32, GigabitEthernet0/0/0/2, 6 dependencies, weight 4294967295, class 0
[flags 0x0]
    path-idx 1 NHID 0x0 [0xf4271e0 0x0]
    next hop 10.1.4.2/32
    remote adjacency
   via 10.1.3.2/32, GigabitEthernet0/0/0/1, 6 dependencies, weight 3452816845, class 0
[flags 0x0]
    path-idx 2 NHID 0x0 [0xf427148 0x0]
    next hop 10.1.3.2/32
    remote adjacency

    Weight distribution:
    slot 0, weight 4294967295, normalized_weight 11, class 0
    slot 1, weight 4294967295, normalized_weight 11, class 0
    slot 2, weight 3452816845, normalized_weight 9, class 0
   Load distribution: 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 (refcount
 3)

    Hash  OK  Interface                 Address
    0     Y   GigabitEthernet0/0/0/0    remote
    1     Y   GigabitEthernet0/0/0/0    remote
    2     Y   GigabitEthernet0/0/0/0    remote
    3     Y   GigabitEthernet0/0/0/0    remote
    4     Y   GigabitEthernet0/0/0/0    remote
    5     Y   GigabitEthernet0/0/0/0    remote
    6     Y   GigabitEthernet0/0/0/0    remote
    7     Y   GigabitEthernet0/0/0/0    remote
    8     Y   GigabitEthernet0/0/0/0    remote
    9     Y   GigabitEthernet0/0/0/0    remote
    10    Y   GigabitEthernet0/0/0/0    remote
    11    Y   GigabitEthernet0/0/0/2    remote
    12    Y   GigabitEthernet0/0/0/2    remote
    13    Y   GigabitEthernet0/0/0/2    remote
    14    Y   GigabitEthernet0/0/0/2    remote
    15    Y   GigabitEthernet0/0/0/2    remote
    16    Y   GigabitEthernet0/0/0/2    remote
    17    Y   GigabitEthernet0/0/0/2    remote
    18    Y   GigabitEthernet0/0/0/2    remote
    19    Y   GigabitEthernet0/0/0/2    remote
    20    Y   GigabitEthernet0/0/0/2    remote
    21    Y   GigabitEthernet0/0/0/2    remote
    22    Y   GigabitEthernet0/0/0/1    remote
    23    Y   GigabitEthernet0/0/0/1    remote
    24    Y   GigabitEthernet0/0/0/1    remote
    25    Y   GigabitEthernet0/0/0/1    remote
    26    Y   GigabitEthernet0/0/0/1    remote
    27    Y   GigabitEthernet0/0/0/1    remote
    28    Y   GigabitEthernet0/0/0/1    remote
    29    Y   GigabitEthernet0/0/0/1    remote
    30    Y   GigabitEthernet0/0/0/1    remote
```

Verify the installation of only two paths by R1, to reach R5 loopback IP address 10.0.0.1/32.

```
Router #show route 10.0.0.1/32

Routing entry for 10.0.0.1/32
  Known via "ospf 0", distance 110, metric 41, type intra area
  Installed Aug  3 23:25:11.897 for 00:16:11
  Routing Descriptor Blocks
    10.1.2.2, from 10.0.0.1, via GigabitEthernet0/0/0/0
```

```
      Route metric is 41, Wt is 4294967295
    10.1.4.2, from 10.0.0.1, via GigabitEthernet0/0/0/2
      Route metric is 41, Wt is 4294967295
  No advertising protos.
```

```
Router #show ospf routes 10.0.0.1/32

Topology Table for ospf 0 with ID 172.16.0.1:

Codes: O - Intra area, O IA - Inter area
       O E1 - External type 1, O E2 - External type 2
       O N1 - NSSA external type 1, O N2 - NSSA external type 2

O    10.0.0.1/32, metric 41
       10.1.4.2, from 10.0.0.1, via GigabitEthernet0/0/0/2, ifIndex 8, path-id 1
       10.1.2.2, from 10.0.0.1, via GigabitEthernet0/0/0/0, ifIndex 6, path-id 2
```

```
Router #show cef 10.0.0.1/32

10.0.0.1/32, version 199, internal 0x1000001 0x10 (ptr 0xe2ed3e8) [1], 0x400 (0xe4899f0),
0x0 (0x0)
Updated Aug  3 23:46:12.770
 remote adjacency to GigabitEthernet0/0/0/0
Prefix Len 32, traffic index 0, precedence n/a, priority 1
  gateway array (0xe2f1838) reference count 1, flags 0x0, source rib (7), 0 backups
                [2 type 3 flags 0x8401 (0xe3a2948) ext 0x0 (0x0)]
  LW-LDI[type=3, refc=1, ptr=0xe4899f0, sh-ldi=0xe3a2948]
  gateway array update type-time 1 Aug  3 23:46:12.770
LDI Update time Aug  3 23:46:12.779
LW-LDI-TS Aug  3 23:46:12.779
   via 10.1.2.2/32, GigabitEthernet0/0/0/0, 8 dependencies, weight 4294967295, class 0
[flags 0x0]
    path-idx 0 NHID 0x0 [0xf4270b0 0x0]
    next hop 10.1.2.2/32
    remote adjacency
   via 10.1.4.2/32, GigabitEthernet0/0/0/2, 8 dependencies, weight 4294967295, class 0
[flags 0x0]
    path-idx 1 NHID 0x0 [0xf4271e0 0x0]
    next hop 10.1.4.2/32
    remote adjacency

    Weight distribution:
    slot 0, weight 4294967295, normalized_weight 1, class 0
    slot 1, weight 4294967295, normalized_weight 1, class 0
    Load distribution: 0 1 (refcount 2)

    Hash  OK  Interface                 Address
    0     Y   GigabitEthernet0/0/0/0    remote
    1     Y   GigabitEthernet0/0/0/2    remote
```

Configuring IS-IS With Weight

The following example shows the IS-IS weight configuration with IPv4. The same can be done for IPv6, with or without SR.

```
CPU0:router(config)# router isis 1
RP/0/RSP0/CPU0:router(config-isis)# interface GigabitEthernet0/3/0/8
RP/0/RSP0/CPU0:router(config-isis-if)# address-family ipv4 unicast
RP/0/RSP0/CPU0:router(config-isis-if-af)# weight 200
RP/0/RSP0/CPU0:router(config-isis)# interface GigabitEthernet0/3/0/9
RP/0/RSP0/CPU0:router(config-isis-if)# address-family ipv4 unicast
RP/0/RSP0/CPU0:router(config-isis-if-af)# weight 300
```

Verification

The following example verifies CEF entry. Then, for two paths with weights of 200 and 300 respectively, and GCD of 100; the expected normalized weights are 2 and 3.

```
Router# show cef ipv4 97.0.0.0 detail

97.0.0.0/24, version 537, internal 0x1000001 0x0 (ptr 0x71bcaee0) [1], 0x0 (0x71b98870),
0x0 (0x0)
 Updated Oct 16 06:34:46.197
 remote adjacency to GigabitEthernet0/3/0/8
 Prefix Len 24, traffic index 0, precedence n/a, priority 2
  gateway array (0x71a6de10) reference count 13, flags 0x0, source rib (7), 0 backups
                [14 type 3 flags 0x8401 (0x71b02d90) ext 0x0 (0x0)]
  LW-LDI[type=3, refc=1, ptr=0x71b98870, sh-ldi=0x71b02d90]
  gateway array update type-time 1 Oct 16 06:34:46.196
 LDI Update time Oct 16 06:34:46.197
 LW-LDI-TS Oct 16 06:34:46.197
   via 1.0.0.2/32, GigabitEthernet0/3/0/8, 4 dependencies, weight 200, class 0 [flags 0x0]
    path-idx 0 NHID 0x0 [0x7244d2a4 0x0]
    next hop 1.0.0.2/32
    remote adjacency
   via 2.0.0.2/32, GigabitEthernet0/3/0/9, 4 dependencies, weight 300, class 0 [flags 0x0]
    path-idx 1 NHID 0x0 [0x7244d2f8 0x0]
    next hop 2.0.0.2/32
    remote adjacency

    Weight distribution:
    slot 0, weight 200, normalized_weight 2, class 0
    slot 1, weight 300, normalized_weight 3, class 0

    Load distribution: 0 1 0 1 1 (refcount 14)

    Hash  OK  Interface                 Address
    0     Y   GigabitEthernet0/3/0/8    remote
    1     Y   GigabitEthernet0/3/0/9    remote
    2     Y   GigabitEthernet0/3/0/8    remote
    3     Y   GigabitEthernet0/3/0/9    remote
    4     Y   GigabitEthernet0/3/0/9    remote
```

Configuring IS-IS With Metric

The following example shows IS-IS metric configuration with IPv4. The same can be done with IPv6.

```
Router# enable
RP/0/RSP0/CPU0:router(config)# router isis 1
RP/0/RSP0/CPU0:router(config-isis)# interface GigabitEthernet0/3/0/8
RP/0/RSP0/CPU0:router(config-isis-if)# address-family ipv4 unicast
RP/0/RSP0/CPU0:router(config-isis-if-af)# metric 1
RP/0/RSP0/CPU0:router(config-isis)# interface GigabitEthernet0/3/0/9
RP/0/RSP0/CPU0:router(config-isis-if)# address-family ipv4 unicast
RP/0/RSP0/CPU0:router(config-isis-if-af)# metric 100
```

Verification

The following example verifies CEF entry, and checks for the two paths with metric values of 1 and 100, respectively. In this example, the best path route metric is 21 and the UCMP path route metric is 120. Therefore, the calculation is as follows:

The best path route metric, 21 = (1 configured + 20 added by IS-IS), weight 0xFFFFFFFF (4294967295)

The UCMP path route metric, 120 = (100 + 20), weight = (21/120) * 4294967295 = 751619276

GCD is one. So Normalized Weight is:

(4294967295 * 64) / (4294967295 + 751619276) = 54

(751619276 * 64) / (4294967295 + 751619276) = 9

```
Router# show cef ipv4 97.0.0.0 detail


97.0.0.0/24, version 773, internal 0x1000001 0x0 (ptr 0x71bcaee0) [1], 0x0 (0x71b98870),
0x0 (0x0)
 Updated Oct 16 06:36:08.632
 remote adjacency to GigabitEthernet0/3/0/8
 Prefix Len 24, traffic index 0, precedence n/a, priority 2
  gateway array (0x71a6d9d0) reference count 2, flags 0x0, source rib (7), 0 backups
                [3 type 3 flags 0x8401 (0x71b02b90) ext 0x0 (0x0)]
  LW-LDI [type=3, refc=1, ptr=0x71b98870, sh-ldi=0x71b02b90]
  gateway array update type-time 1 Oct 16 06:36:08.632
 LDI Update time Oct 16 06:36:08.632
 LW-LDI-TS Oct 16 06:36:08.632
   via 1.0.0.2/32, GigabitEthernet0/3/0/8, 14 dependencies, weight 4294967295, class 0
[flags 0x0]
    path-idx 0 NHID 0x0 [0x7244d2a4 0x0]
    next hop 1.0.0.2/32
    remote adjacency
  via 2.0.0.2/32, GigabitEthernet0/3/0/9, 14 dependencies, weight 751619276, class 0 [flags
 0x0]
    path-idx 1 NHID 0x0 [0x7244d2f8 0x0]
    next hop 2.0.0.2/32
    remote adjacency


    Weight distribution:
    slot 0, weight 4294967295, normalized_weight 54, class 0
    slot 1, weight 751619276, normalized_weight 9, class 0
```

Configuring BGP With Weights

The following example shows BGP configuration with weights.

```
RP/0/RSP0/CPU0:router(config)# route-policy BW1
RP/0/RSP0/CPU0:router(config-rpl)# set extcommunity bandwidth (2906:45750000)
RP/0/RSP0/CPU0:router(config-rpl)# end-policy
RP/0/RSP0/CPU0:router(config)# !
RP/0/RSP0/CPU0:router(config)# route-policy BW2
RP/0/RSP0/CPU0:router(config-rpl)# set extcommunity bandwidth (2906:47250000)
RP/0/RSP0/CPU0:router(config-rpl)# end-policy
RP/0/RSP0/CPU0:router(config)# !
RP/0/RSP0/CPU0:router(config)# route-policy pass-all
RP/0/RSP0/CPU0:router(config-rpl)# pass
RP/0/RSP0/CPU0:router(config-rpl)# end-policy
RP/0/RSP0/CPU0:router(config)# !
RP/0/RSP0/CPU0:router(config)# router bgp 1
RP/0/RSP0/CPU0:router(config-bgp)# bgp bestpath as-path multipath-relax
RP/0/RSP0/CPU0:router(config-bgp)# address-family ipv4 unicast
RP/0/RSP0/CPU0:router(config-bgp-af)# maximum-paths eibgp 64
RP/0/RSP0/CPU0:router(config-bgp-af)# !
RP/0/RSP0/CPU0:router(config-bgp-af)# neighbor 1.0.0.2
RP/0/RSP0/CPU0:router(config-bgp-nbr)# remote-as 2
RP/0/RSP0/CPU0:router(config-bgp-nbr)# ebgp-multihop 255
```

```
RP/0/RSP0/CPU0:router(config-bgp-nbr)# dmz-link-bandwidth
RP/0/RSP0/CPU0:router(config-bgp-nbr)# address-family ipv4 unicast
RP/0/RSP0/CPU0:router(config-bgp-nbr-af)# multipath
RP/0/RSP0/CPU0:router(config-bgp-nbr-af)# route-policy BW1 in
RP/0/RSP0/CPU0:router(config-bgp-nbr-af)# route-policy pass-all out
RP/0/RSP0/CPU0:router(config-bgp-nbr-af)# !
RP/0/RSP0/CPU0:router(config-bgp-nbr-af)# neighbor 2.0.0.2
RP/0/RSP0/CPU0:router(config-bgp-nbr)# remote-as 2
RP/0/RSP0/CPU0:router(config-bgp-nbr)# ebgp-multihop 255
RP/0/RSP0/CPU0:router(config-bgp-nbr)# dmz-link-bandwidth
RP/0/RSP0/CPU0:router(config-bgp-nbr)# address-family ipv4 unicast
RP/0/RSP0/CPU0:router(config-bgp-nbr-af)# multipath
RP/0/RSP0/CPU0:router(config-bgp-nbr-af)# route-policy BW2 in
RP/0/RSP0/CPU0:router(config-bgp-nbr-af)# route-policy pass-all out
```

Verification

Step 1: Verify CEF entry:

Via 1.0.0.2: set extcommunity bandwidth (2906:45750000) – Weight = 45750000/125=366000 (125 ratio because baud)

Via 2.0.0.2: set extcommunity bandwidth (2906:47250000) – Weight = 47250000/125=378000

GCD is 6, so norm_weight = 61 and 63. Though 61 + 63 > 64.

Step 2: GCD of weights 61 and 63 is 1. Therefore, Normalised Weight = (Path weight/Total weight) * Maximum bucket size. The maximum bucket size value is 64. Total weight = 61+63 = 124.

norm_weight1 = (61/124) * 64 = 31, norm_weight2 = (63/124) * 64 = 32

You can verify the weight distribution in BGP, using the following command:

```
Router # show cef vrf default ipv4 97.0.0.0 detail


97.0.0.0/24, version 1965, internal 0x5000001 0x0 (ptr 0x71bcb620) [1], 0x0 (0x0), 0x0 (0x0)
 Updated Oct 16 08:15:02.958
 Prefix Len 24, traffic index 0, precedence n/a, priority 4
  gateway array (0x72a5e2f8) reference count 10, flags 0x2010, source rib (7), 0 backups
                [1 type 3 flags 0x48441 (0x71b02cd0) ext 0x0 (0x0)]
  LW-LDI[type=0, refc=0, ptr=0x0, sh-ldi=0x0]
  gateway array update type-time 1 Oct 16 08:15:02.958
 LDI Update time Oct 16 08:15:02.959


    Weight distribution:
    slot 0, weight 366000, normalized_weight 31
    slot 1, weight 378000, normalized_weight 32


  Level 1 - Load distribution: 0 1 0 1 0 1 0

1 0 1 0 1 0 
```

Configuring TE Tunnel With Weights

Verification

In the following example, the weight is distributed among the five TE tunnels.

```
Router# show cef ipv4 97.0.0.0 detail
(…)
   via 200.0.0.1/32, tunnel-te1, 3 dependencies, weight 8, class 0 [flags 0x0]
    path-idx 0 NHID 0x0 [0x7244d2f8 0x0]
    next hop 200.0.0.1/32
    local adjacency
   via 200.0.0.1/32, tunnel-te2, 3 dependencies, weight 4, class 0 [flags 0x0]
    path-idx 1 NHID 0x0 [0x7244e948 0x0]
    next hop 200.0.0.1/32
    local adjacency
   via 200.0.0.1/32, tunnel-te3, 3 dependencies, weight 1, class 0 [flags 0x0]
    path-idx 2 NHID 0x0 [0x7244d544 0x0]
    next hop 200.0.0.1/32
    local adjacency
   via 200.0.0.1/32, tunnel-te4, 3 dependencies, weight 1, class 0 [flags 0x0]
    path-idx 3 NHID 0x0 [0x7244d694 0x0]
    next hop 200.0.0.1/32
    local adjacency
   via 200.0.0.1/32, tunnel-te5, 3 dependencies, weight 1, class 0 [flags 0x0]
    path-idx 4 NHID 0x0 [0x7244d7e4 0x0]
    next hop 200.0.0.1/32
    local adjacency

    Weight distribution:
    slot 0, weight 8, normalized_weight 8, class 0
    slot 1, weight 4, normalized_weight 4, class 0
    slot 2, weight 1, normalized_weight 1, class 0
    slot 3, weight 1, normalized_weight 1, class 0
    slot 4, weight 1, normalized_weight 1, class 0
```

Policy-Based Tunnel Selection

Policy-Based Tunnel Selection (PBTS) provides a mechanism that lets you direct traffic into specific TE tunnels based on different criteria.

PBTS is a special case in UCMP calculation. It uses load share command to configure weight. The UCMP algorithm normalizes each class independently and it uses max_path from PD specific max_tunnels_per_class. UCMP with PBTS can have more total_paths (buckets) than the supported number of paths (buckets) for all Forwarding Classes (FCs).

All other XR platform sets 8 buckets per FC and 64 buckets for all 8 (0-7) FCs. After normalization, the total number buckets do not exceed platform limit.

Example

The **show cef ipv6** command displays the PBTS class information in the following output.

```
Router# show cef ipv6 97:: detail

97::/64, version 88177, internal 0x1000001 0x0 (ptr 0x980eef7c) [1], 0x0 (0x974366b8), 0xa28
```

```
(0x988842c0)
Updated Mar  7 05:44:46.875

Prefix Len 64, traffic index 0, precedence n/a, priority 2
 gateway array (0x97e54770) reference count 11, flags 0x28, source rib (7), 0 backups
              [12 type 1 flags 0x200401 (0x9799a3f8) ext 0x0 (0x0)]
 LW-LDI[type=1, refc=1, ptr=0x974366b8, sh-ldi=0x9799a3f8]
 gateway array update type-time 4 Mar  7 05:46:11.118
LDI Update time Mar  7 05:46:11.118
LW-LDI-TS Mar  7 05:46:11.118


  via ::ffff:200.0.0.1/128, tunnel-te45, 3 dependencies, weight 1, forward class 6 [flags
0x0]

   path-idx 0 NHID 0x0 [0x97b51978 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}

  via ::ffff:200.0.0.1/128, tunnel-te46, 3 dependencies, weight 1, forward class 6 [flags
0x0]

   path-idx 1 NHID 0x0 [0x97b51648 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}

  via ::ffff:200.0.0.1/128, tunnel-te47, 3 dependencies, weight 1, forward class 6 [flags
0x0]

   path-idx 2 NHID 0x0 [0x97b51c20 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}

  via ::ffff:200.0.0.1/128, tunnel-te48, 3 dependencies, weight 1, forward class 6 [flags
0x0]

   path-idx 3 NHID 0x0 [0x97b52308 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}

  via ::ffff:200.0.0.1/128, tunnel-te49, 3 dependencies, weight 1, forward class 7 [flags
0x0]

   path-idx 4 NHID 0x0 [0x97b518f0 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}

  via ::ffff:200.0.0.1/128, tunnel-te1, 3 dependencies, weight 3, forward class 1 [flags
0x0]

   path-idx 5 NHID 0x0 [0x97b4f338 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
```

```
     labels imposed {ExpNullv6}

   via ::ffff:200.0.0.1/128, tunnel-te2, 3 dependencies, weight 500, forward class 1 [flags
 0x0]

    path-idx 6 NHID 0x0 [0x97b50328 0x0]
    next hop VRF - 'default', table - 0xe0000000
    next hop ::ffff:200.0.0.1/128
    local adjacency
     labels imposed {ExpNullv6}

   via ::ffff:200.0.0.1/128, tunnel-te3, 3 dependencies, weight 1, forward class 1 [flags
0x0]

    path-idx 7 NHID 0x0 [0x97b4ede8 0x0]
    next hop VRF - 'default', table - 0xe0000000
    next hop ::ffff:200.0.0.1/128
    local adjacency
     labels imposed {ExpNullv6}

   via ::ffff:200.0.0.1/128, tunnel-te4, 3 dependencies, weight 1, forward class 1 [flags
0x0]

    path-idx 8 NHID 0x0 [0x97b4eb40 0x0]
    next hop VRF - 'default', table - 0xe0000000
    next hop ::ffff:200.0.0.1/128
    local adjacency
     labels imposed {ExpNullv6}

   via ::ffff:200.0.0.1/128, tunnel-te5, 3 dependencies, weight 1, forward class 1 [flags
0x0]

    path-idx 9 NHID 0x0 [0x97b4fff8 0x0]
    next hop VRF - 'default', table - 0xe0000000
    next hop ::ffff:200.0.0.1/128
    local adjacency
     labels imposed {ExpNullv6}

   via ::ffff:200.0.0.1/128, tunnel-te6, 3 dependencies, weight 1, forward class 1 [flags
0x0]

    path-idx 10 NHID 0x0 [0x97b4f778 0x0]
    next hop VRF - 'default', table - 0xe0000000
    next hop ::ffff:200.0.0.1/128
    local adjacency
     labels imposed {ExpNullv6}

   via ::ffff:200.0.0.1/128, tunnel-te7, 3 dependencies, weight 1, forward class 1 [flags
0x0]

    path-idx 11 NHID 0x0 [0x97b4f118 0x0]
    next hop VRF - 'default', table - 0xe0000000
    next hop ::ffff:200.0.0.1/128
    local adjacency
     labels imposed {ExpNullv6}

   via ::ffff:200.0.0.1/128, tunnel-te8, 3 dependencies, weight 1, forward class 1 [flags
0x0]

    path-idx 12 NHID 0x0 [0x97b4ee70 0x0]
    next hop VRF - 'default', table - 0xe0000000
    next hop ::ffff:200.0.0.1/128
    local adjacency
     labels imposed {ExpNullv6}
```

```
  via ::ffff:200.0.0.1/128, tunnel-te9, 3 dependencies, weight 1, forward class 2 [flags
0x0]

   path-idx 13 NHID 0x0 [0x97b4f090 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}

  via ::ffff:200.0.0.1/128, tunnel-te10, 3 dependencies, weight 1, forward class 2 [flags
0x0]

   path-idx 14 NHID 0x0 [0x97b4f448 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}

  via ::ffff:200.0.0.1/128, tunnel-te11, 3 dependencies, weight 1, forward class 2 [flags
0x0]

   path-idx 15 NHID 0x0 [0x97b4faa8 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}

  via ::ffff:200.0.0.1/128, tunnel-te12, 3 dependencies, weight 1, forward class 2 [flags
0x0]

   path-idx 16 NHID 0x0 [0x97b4f008 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}

  via ::ffff:200.0.0.1/128, tunnel-te13, 3 dependencies, weight 1, forward class 2 [flags
0x0]

   path-idx 17 NHID 0x0 [0x97b50218 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}

  via ::ffff:200.0.0.1/128, tunnel-te14, 3 dependencies, weight 1, forward class 2 [flags
0x0]

   path-idx 18 NHID 0x0 [0x97b4fbb8 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}

  via ::ffff:200.0.0.1/128, tunnel-te15, 3 dependencies, weight 1, forward class 2 [flags
0x0]

   path-idx 19 NHID 0x0 [0x97b4ed60 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}
```

```
  via ::ffff:200.0.0.1/128, tunnel-te16, 3 dependencies, weight 1, forward class 2 [flags
0x0]

   path-idx 20 NHID 0x0 [0x97b4fcc8 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}

  via ::ffff:200.0.0.1/128, tunnel-te17, 3 dependencies, weight 1, forward class 3 [flags
0x0]


   path-idx 21 NHID 0x0 [0x97b50190 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}


  via ::ffff:200.0.0.1/128, tunnel-te18, 3 dependencies, weight 1, forward class 3 [flags
0x0]


   path-idx 22 NHID 0x0 [0x97b4f998 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}


  via ::ffff:200.0.0.1/128, tunnel-te19, 3 dependencies, weight 1, forward class 3 [flags
0x0]


   path-idx 23 NHID 0x0 [0x97b4fee8 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}


  via ::ffff:200.0.0.1/128, tunnel-te20, 3 dependencies, weight 1, forward class 3 [flags
0x0]

   path-idx 24 NHID 0x0 [0x97b505d0 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}


  via ::ffff:200.0.0.1/128, tunnel-te21, 3 dependencies, weight 1, forward class 3 [flags
0x0]

   path-idx 25 NHID 0x0 [0x97b4fc40 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}


  via ::ffff:200.0.0.1/128, tunnel-te22, 3 dependencies, weight 1, forward class 3 [flags
0x0]
```

```
   path-idx 26 NHID 0x0 [0x97b50988 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}


  via ::ffff:200.0.0.1/128, tunnel-te23, 3 dependencies, weight 1, forward class 3 [flags
0x0]

   path-idx 27 NHID 0x0 [0x97b50080 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}


  via ::ffff:200.0.0.1/128, tunnel-te24, 3 dependencies, weight 1, forward class 3 [flags
0x0]

   path-idx 28 NHID 0x0 [0x97b4fd50 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}


  via ::ffff:200.0.0.1/128, tunnel-te25, 3 dependencies, weight 1, forward class 4 [flags
0x0]


   path-idx 29 NHID 0x0 [0x97b503b0 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}


  via ::ffff:200.0.0.1/128, tunnel-te26, 3 dependencies, weight 1, forward class 4 [flags
0x0]


   path-idx 30 NHID 0x0 [0x97b507f0 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}


  via ::ffff:200.0.0.1/128, tunnel-te27, 3 dependencies, weight 1, forward class 4 [flags
0x0]

   path-idx 31 NHID 0x0 [0x97b4ff70 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}


  via ::ffff:200.0.0.1/128, tunnel-te28, 3 dependencies, weight 1, forward class 4 [flags
0x0]
```

```
   path-idx 32 NHID 0x0 [0x97b50548 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}


  via ::ffff:200.0.0.1/128, tunnel-te29, 3 dependencies, weight 1, forward class 4 [flags
0x0]

   path-idx 33 NHID 0x0 [0x97b4fb30 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}

  via ::ffff:200.0.0.1/128, tunnel-te30, 3 dependencies, weight 1, forward class 4 [flags
0x0]

   path-idx 34 NHID 0x0 [0x97b506e0 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}


  via ::ffff:200.0.0.1/128, tunnel-te31, 3 dependencies, weight 1, forward class 4 [flags
0x0]

   path-idx 35 NHID 0x0 [0x97b51208 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}


  via ::ffff:200.0.0.1/128, tunnel-te32, 3 dependencies, weight 1, forward class 4 [flags
0x0]

   path-idx 36 NHID 0x0 [0x97b502a0 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}


  via ::ffff:200.0.0.1/128, tunnel-te33, 3 dependencies, weight 1, forward class 5 [flags
0x0]

   path-idx 37 NHID 0x0 [0x97b514b0 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}


  via ::ffff:200.0.0.1/128, tunnel-te34, 3 dependencies, weight 1, forward class 5 [flags
0x0]

   path-idx 38 NHID 0x0 [0x97b50c30 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}
```

```
  via ::ffff:200.0.0.1/128, tunnel-te35, 3 dependencies, weight 1, forward class 5 [flags
0x0]

   path-idx 39 NHID 0x0 [0x97b50b20 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}


  via ::ffff:200.0.0.1/128, tunnel-te36, 3 dependencies, weight 1, forward class 5 [flags
0x0]

   path-idx 40 NHID 0x0 [0x97b50cb8 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}


  via ::ffff:200.0.0.1/128, tunnel-te37, 3 dependencies, weight 1, forward class 5 [flags
0x0]


   path-idx 41 NHID 0x0 [0x97b51180 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}


  via ::ffff:200.0.0.1/128, tunnel-te38, 3 dependencies, weight 1, forward class 5 [flags
0x0]

   path-idx 42 NHID 0x0 [0x97b51428 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}


  via ::ffff:200.0.0.1/128, tunnel-te39, 3 dependencies, weight 1, forward class 5 [flags
0x0]


   path-idx 43 NHID 0x0 [0x97b51758 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}


  via ::ffff:200.0.0.1/128, tunnel-te40, 3 dependencies, weight 1, forward class 5 [flags
0x0]

   path-idx 44 NHID 0x0 [0x97b520e8 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}
```

```
  via ::ffff:200.0.0.1/128, tunnel-te41, 3 dependencies, weight 1, forward class 6 [flags
0x0]

   path-idx 45 NHID 0x0 [0x97b51538 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}


  via ::ffff:200.0.0.1/128, tunnel-te42, 3 dependencies, weight 1, forward class 6 [flags
0x0]

   path-idx 46 NHID 0x0 [0x97b50dc8 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}


  via ::ffff:200.0.0.1/128, tunnel-te43, 3 dependencies, weight 1, forward class 6 [flags
0x0]

   path-idx 47 NHID 0x0 [0x97b51b10 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}


  via ::ffff:200.0.0.1/128, tunnel-te44, 3 dependencies, weight 1, forward class 6 [flags
0x0]

   path-idx 48 NHID 0x0 [0x97b516d0 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}


  via ::ffff:200.0.0.1/128, tunnel-te50, 3 dependencies, weight 1, forward class 7 [flags
0x0]

   path-idx 49 NHID 0x0 [0x97b525b0 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}

  via ::ffff:200.0.0.1/128, tunnel-te51, 3 dependencies, weight 1, forward class 7 [flags
0x0]

   path-idx 50 NHID 0x0 [0x97b52638 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}


  via ::ffff:200.0.0.1/128, tunnel-te52, 3 dependencies, weight 1, forward class 7 [flags
0x0]

   path-idx 51 NHID 0x0 [0x97b51f50 0x0]
   next hop VRF - 'default', table - 0xe0000000
```

```
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}

  via ::ffff:200.0.0.1/128, tunnel-te53, 3 dependencies, weight 1, forward class 7 [flags
0x0]

   path-idx 52 NHID 0x0 [0x97b52060 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}

  via ::ffff:200.0.0.1/128, tunnel-te54, 3 dependencies, weight 1, forward class 7 [flags
0x0]

   path-idx 53 NHID 0x0 [0x97b527d0 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}

  via ::ffff:200.0.0.1/128, tunnel-te55, 3 dependencies, weight 1, forward class 7 [flags
0x0]

   path-idx 54 NHID 0x0 [0x97b52280 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}


  via ::ffff:200.0.0.1/128, tunnel-te56, 3 dependencies, weight 1, forward class 7 [flags
0x0]

   path-idx 55 NHID 0x0 [0x97b52d20 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}

  via ::ffff:200.0.0.1/128, tunnel-te57, 3 dependencies, weight 1, class 0 [flags 0x0]

   path-idx 56 NHID 0x0 [0x97b51ca8 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}

  via ::ffff:200.0.0.1/128, tunnel-te58, 3 dependencies, weight 1, class 0 [flags 0x0]

   path-idx 57 NHID 0x0 [0x97b52858 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}

  via ::ffff:200.0.0.1/128, tunnel-te59, 3 dependencies, weight 1, class 0 [flags 0x0]

   path-idx 58 NHID 0x0 [0x97b52390 0x0]
   next hop VRF - 'default', table - 0xe0000000
   next hop ::ffff:200.0.0.1/128
   local adjacency
    labels imposed {ExpNullv6}
```

```
via ::ffff:200.0.0.1/128, tunnel-te60, 3 dependencies, weight 1, class 0 [flags 0x0]

 path-idx 59 NHID 0x0 [0x97b52a78 0x0]
 next hop VRF - 'default', table - 0xe0000000
 next hop ::ffff:200.0.0.1/128
 local adjacency
  labels imposed {ExpNullv6}

via ::ffff:200.0.0.1/128, tunnel-te61, 3 dependencies, weight 1, class 0 [flags 0x0]

 path-idx 60 NHID 0x0 [0x97b52c10 0x0]
 next hop VRF - 'default', table - 0xe0000000
 next hop ::ffff:200.0.0.1/128
 local adjacency
  labels imposed {ExpNullv6}

via ::ffff:200.0.0.1/128, tunnel-te62, 3 dependencies, weight 1, class 0 [flags 0x0]

 path-idx 61 NHID 0x0 [0x97b52da8 0x0]
 next hop VRF - 'default', table - 0xe0000000
 next hop ::ffff:200.0.0.1/128
 local adjacency
  labels imposed {ExpNullv6}


via ::ffff:200.0.0.1/128, tunnel-te63, 3 dependencies, weight 1, class 0 [flags 0x0]

 path-idx 62 NHID 0x0 [0x97b52c98 0x0]
 next hop VRF - 'default', table - 0xe0000000
 next hop ::ffff:200.0.0.1/128
 local adjacency
  labels imposed {ExpNullv6}


 Weight distribution:

 slot 0, weight 1, normalized_weight 1, class 0
 slot 1, weight 1, normalized_weight 1, class 0
 slot 2, weight 1, normalized_weight 1, class 0
 slot 3, weight 1, normalized_weight 1, class 0
 slot 4, weight 1, normalized_weight 1, class 0
 slot 5, weight 1, normalized_weight 1, class 0
 slot 6, weight 1, normalized_weight 1, class 0
 slot 7, weight 1, normalized_weight 1, forward class 1
 slot 8, weight 3, normalized_weight 1, forward class 1
 slot 9, weight 500, normalized_weight 1, forward class 1
 slot 10, weight 1, normalized_weight 1, forward class 1
 slot 11, weight 1, normalized_weight 1, forward class 1
 slot 12, weight 1, normalized_weight 1, forward class 1
 slot 13, weight 1, normalized_weight 1, forward class 1
 slot 14, weight 1, normalized_weight 1, forward class 1
 slot 15, weight 1, normalized_weight 1, forward class 2
 slot 16, weight 1, normalized_weight 1, forward class 2
 slot 17, weight 1, normalized_weight 1, forward class 2
 slot 18, weight 1, normalized_weight 1, forward class 2
 slot 19, weight 1, normalized_weight 1, forward class 2
 slot 20, weight 1, normalized_weight 1, forward class 2
 slot 21, weight 1, normalized_weight 1, forward class 2
 slot 22, weight 1, normalized_weight 1, forward class 2
 slot 23, weight 1, normalized_weight 1, forward class 3
 slot 24, weight 1, normalized_weight 1, forward class 3
 slot 25, weight 1, normalized_weight 1, forward class 3
 slot 26, weight 1, normalized_weight 1, forward class 3
```

```
   slot 27, weight 1, normalized_weight 1, forward class 3
   slot 28, weight 1, normalized_weight 1, forward class 3
   slot 29, weight 1, normalized_weight 1, forward class 3
   slot 30, weight 1, normalized_weight 1, forward class 3
   slot 31, weight 1, normalized_weight 1, forward class 4
   slot 32, weight 1, normalized_weight 1, forward class 4
   slot 33, weight 1, normalized_weight 1, forward class 4
   slot 34, weight 1, normalized_weight 1, forward class 4
   slot 35, weight 1, normalized_weight 1, forward class 4
   slot 36, weight 1, normalized_weight 1, forward class 4
   slot 37, weight 1, normalized_weight 1, forward class 4
   slot 38, weight 1, normalized_weight 1, forward class 4
   slot 39, weight 1, normalized_weight 1, forward class 5
   slot 40, weight 1, normalized_weight 1, forward class 5
   slot 41, weight 1, normalized_weight 1, forward class 5
   slot 42, weight 1, normalized_weight 1, forward class 5
   slot 43, weight 1, normalized_weight 1, forward class 5
   slot 44, weight 1, normalized_weight 1, forward class 5
   slot 45, weight 1, normalized_weight 1, forward class 5
   slot 46, weight 1, normalized_weight 1, forward class 5
   slot 47, weight 1, normalized_weight 1, forward class 6
   slot 48, weight 1, normalized_weight 1, forward class 6
   slot 49, weight 1, normalized_weight 1, forward class 6
   slot 50, weight 1, normalized_weight 1, forward class 6
   slot 51, weight 1, normalized_weight 1, forward class 6
   slot 52, weight 1, normalized_weight 1, forward class 6
   slot 53, weight 1, normalized_weight 1, forward class 6
   slot 54, weight 1, normalized_weight 1, forward class 6
   slot 55, weight 1, normalized_weight 1, forward class 7
   slot 56, weight 1, normalized_weight 1, forward class 7
   slot 57, weight 1, normalized_weight 1, forward class 7
   slot 58, weight 1, normalized_weight 1, forward class 7
   slot 59, weight 1, normalized_weight 1, forward class 7
   slot 60, weight 1, normalized_weight 1, forward class 7
   slot 61, weight 1, normalized_weight 1, forward class 7
   slot 62, weight 1, normalized_weight 1, forward class 7

   PBTS class information:

     class 0: 7 paths, offset 0

     forward class 1: 8 paths, offset 7
     forward class 2: 8 paths, offset 15
     forward class 3: 8 paths, offset 23
     forward class 4: 8 paths, offset 31
     forward class 5: 8 paths, offset 39
     forward class 6: 8 paths, offset 47
     forward class 7: 8 paths, offset 55

   Load distribution: 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25
26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55
56 57 58 59 60 61 62 (refcount 12)


   Hash  OK  Interface                 Address

   0     Y   tunnel-te57               point2point
   1     Y   tunnel-te58               point2point
   2     Y   tunnel-te59               point2point
   3     Y   tunnel-te60               point2point
   4     Y   tunnel-te61               point2point
   5     Y   tunnel-te62               point2point
   6     Y   tunnel-te63               point2point
   7     Y   tunnel-te8                point2point
```

```
8     Y   tunnel-te1               point2point
9     Y   tunnel-te2               point2point
10    Y   tunnel-te3               point2point


11    Y   tunnel-te4               point2point
12    Y   tunnel-te5               point2point
13    Y   tunnel-te6               point2point
14    Y   tunnel-te7               point2point
15    Y   tunnel-te16              point2point
16    Y   tunnel-te9               point2point
17    Y   tunnel-te10              point2point
18    Y   tunnel-te11              point2point
19    Y   tunnel-te12              point2point
20    Y   tunnel-te13              point2point
21    Y   tunnel-te14              point2point
22    Y   tunnel-te15              point2point
23    Y   tunnel-te24              point2point
24    Y   tunnel-te17              point2point
25    Y   tunnel-te18              point2point
26    Y   tunnel-te19              point2point
27    Y   tunnel-te20              point2point
28    Y   tunnel-te21              point2point
29    Y   tunnel-te22              point2point
30    Y   tunnel-te23              point2point
31    Y   tunnel-te32              point2point
32    Y   tunnel-te25              point2point
33    Y   tunnel-te26              point2point
34    Y   tunnel-te27              point2point
35    Y   tunnel-te28              point2point
36    Y   tunnel-te29              point2point
37    Y   tunnel-te30              point2point
38    Y   tunnel-te31              point2point
39    Y   tunnel-te40              point2point
40    Y   tunnel-te33              point2point
41    Y   tunnel-te34              point2point
42    Y   tunnel-te35              point2point
43    Y   tunnel-te36              point2point
44    Y   tunnel-te37              point2point
45    Y   tunnel-te38              point2point
46    Y   tunnel-te39              point2point
47    Y   tunnel-te44              point2point
48    Y   tunnel-te45              point2point
49    Y   tunnel-te46              point2point
50    Y   tunnel-te47              point2point
51    Y   tunnel-te48              point2point
52    Y   tunnel-te41              point2point
53    Y   tunnel-te42              point2point
54    Y   tunnel-te43              point2point
55    Y   tunnel-te56              point2point
56    Y   tunnel-te49              point2point
57    Y   tunnel-te50              point2point
58    Y   tunnel-te51              point2point
59    Y   tunnel-te52              point2point
60    Y   tunnel-te53              point2point
61    Y   tunnel-te54              point2point
62    Y   tunnel-te55              point2point
```

Interior Gateway Protocol (IGP) Destination-based Load Balancing (DLB)

Currently, the router supports upto 2K labelled prefixes with Equal Cost Multi Path (ECMP). From Cisco IOS XR Software Release 7.1.1, with the introduction of the Interior Gateway Protocol (IGP) Destination-based Load Balancing (DLB) feature, the router can support higher scale of labelled prefixes.

If this feature is enabled, the software selects one path out of the available multipaths based on a software hash of the destination IP address and then programs the data plane with the selected single path. Thus, the traffic gets distributed on available paths by pre-selection of the path based on the per-prefix hash.

However, if Layer 2 Virtual Private Network (L2VPN) or Border Gateway Protocol (BGP) service uses a IGP DLB prefix for next-hop reachability, then the software removes DLB from that particular prefix and it behaves as if L2VPN or BGP runs on a non-DLB IGP.

Restrictions for IGP DLB

The following are the restrictions for IGP DLB:

- IGP DLB is not enabled in software, by default.
- ECMP resource utilization happens only on enabling BGP or L2 services for the DLB prefix. However disabling BGP or L2 services for the prefix does not release the ECMP resource immediately. ECMP resources get released only when the prefix is re-learned again or with a router reload.
- For the configuration command **hw-module fib dlb level-1 enable** to take effect, you should reload the LC.

Configuring IGP DLB

Without IGP DLB enabled, all prefixes consume ECMP hardware resources.

However, with IGP DLB enabled, DLB IP and labeled prefixes do not consume ECMP resources. Only BGP and L2VPN services over a DLB IGP prefix consume ECMP resources. This enables the router to handle higher prefix scale by the efficient use of the ECMP hardware resources.

The following example shows the ECMP software-based load-balancing configuration:

```
Router# config terminal
Router(config)# hw-module fib dlb level-1 enable
Router(config)# commit
Router(config)# end
Router# reload location all
Proceed with reload? [confirm]
```

Verification

In the below command, resource usage will be less when IGP DLB is enabled:

```
Router# show controllers npu resources ecmpfec location 0/3/CPU0
HW Resource Information
```

```
    Name                            : ecmp_fec
    Asic Type                       : Jericho Plus

NPU-0
OOR Summary
        Estimated Max Entries       : 4096
        Red Threshold               : 95
        Yellow Threshold            : 80
        OOR State                   : Green


Current Usage
        Total In-Use                : 3        (0 %)
        ipnhgroup                   : 3        (0 %)
        ip6nhgroup                  : 0        (0 %)


NPU-1
OOR Summary
        Estimated Max Entries       : 4096
        Red Threshold               : 95
        Yellow Threshold            : 80
        OOR State                   : Green


Current Usage
        Total In-Use                : 3        (0 %)
        ipnhgroup                   : 3        (0 %)
        ip6nhgroup                  : 0        (0 %)


NPU-2
OOR Summary
        Estimated Max Entries       : 4096
        Red Threshold               : 95
        Yellow Threshold            : 80
        OOR State                   : Green


Current Usage
        Total In-Use                : 3        (0 %)
        ipnhgroup                   : 3        (0 %)
        ip6nhgroup                  : 0        (0 %)


NPU-3
OOR Summary
        Estimated Max Entries       : 4096
        Red Threshold               : 95
        Yellow Threshold            : 80
        OOR State                   : Green


Current Usage
        Total In-Use                : 3        (0 %)
        ipnhgroup                   : 3        (0 %)
        ip6nhgroup                  : 0        (0 %)

Router# show cef 209.165.200.225 hardware egress detail location npu0
….
HW Walk:
LEAF:
    PI:0x308bec58b8 PD:0x308bec5958 rev:4910 type: IPV4 (0)
    LEAF location: LEM
```

```
FEC key: 0x20e20000e0b

LWLDI:
    EOS0/1 LDI: ECD_MARKED  SERVICE_MARKED  # shows ECMP resource usage with IGP DLB
    LSP pattern:3
```