

## Description

Routes covered by a resilient equal-cost multi-path (RECMC) prefix are types of routes that make use of hardware tables dedicated for equal-cost multi-path (ECMP) routing. Resilient ECMP deduping is a new feature wherein the switch will reactively attempt to reduce the number of ECMP hardware table entries allocated by forcing routes that share the same set of next hops but point to different hardware table entries to point to the same hardware table entry when hardware resource utilization is high. Forcing RECMC routes to change the hardware table entry that they point to may potentially cause a traffic flow disruption for any existing flows going over that route. The deduping process will attempt to minimize the amount of potential traffic loss caused.

Each route needs to allocate hardware table entries in the ASIC, which contains forwarding information for the route, such as what its next hops are, what egress links each next hop uses, etc. The network device uses these hardware table entries when making forwarding decisions for a packet that is meant for a certain route.. These ECMP hardware tables are limited in size and can fill up quickly if there are a large number of these hardware table entries allocated. One option to alleviate the usage of these hardware tables is to force RECMC routes to share hardware table entries.

There is already an existing feature for RECMC routes to get them to point to the same hardware table entry if they share the same set of next hops and the ordering of the next hops is the same. However, RECMC routes may end up sharing the same set of next hops, but the nexthop ordering may be different between them, and therefore the routes end up occupying different hardware table entries in the ASIC. RECMC routing has a property wherein the current ordering of next hops for a given route is influenced by its previous orderings. The ordering between the routes can differ because these routes may have had a different set of next hops at some previous time before they finally converged onto the same set of next hops.

When the ECMP hardware resource usage crosses the high threshold, the deduping process begins, and it lasts until the ECMP hardware resource usage falls below the low threshold. Users interested in modifying the thresholds can view the configuration commands listed below.

## Platform compatibility

- DCS-7020, DCS-7280 R/R2/R3, DCS-7500 R/R2/R3, DCS-7800R3 series

## Configuration

This feature is enabled by default. All resource optimization features can be enabled or disabled with the following commands. As of 4.25.2F there is no support to individually enable or disable certain hardware resource optimization features.

The following command can be used to disable all hardware resource optimization features:

```
(config)#ip hardware fib next-hop resource optimization disabled
```

To re-enable the all hardware resource optimization features after disabling them, the following command can be used:

```
(config)#no ip hardware fib next-hop resource optimization disabled
```

The following command is used to configure the thresholds for starting and stopping the optimization:

```
(config)#ip hardware fib next-hop resource optimization thresholds low 20 high 80
```

#### Notes:

- The value specified for the threshold represents the percentage of resource utilization, and is an integer between 0 and 100.
- Setting the high threshold to 80 indicates that optimization starts when the resource utilization is above 80%. The default value of this threshold is 90.
- Setting the low threshold to 20 indicates that optimization stops when the resource utilization is below 20%. The default value of this threshold is 85.

## Show commands

The following command is used to show the statistics of this feature:

```
switch#show ip hardware fib summary
Fib summary
-----
Adjacency sharing: disabled
BFD peer event: enabled
Deletion Delay: 0
Protect default route: disabled
PBR: supported
URPF: supported
ICMP unreachable: enabled
Max Ale ECMP: 600
UCMP weight deviation: 0.0
Maximum number of routes: 0
Fib compression: disabled
Resource optimization for adjacency programming: enabled
Adjacency resource optimization thresholds: low 20, high 80
```

The last two lines of the output shows whether the feature is enabled and what are the corresponding threshold values for starting and stopping the optimization process.

The following command is used to show the utilization of the hardware resources:

```
switch#show hardware capacity
```

The above command lists all the hardware resources on track on the device. An example of the multi-level hierarchy ECMP resources is show as follows:

```
Forwarding Resources Usage
```

Table	Feature	Chip	Used	Used	Free	Com
mitted	Best Case	High	Entries	(%)	Entri	
es	Entries	Max	Watermark			
-----						
Entries						
-----						
ECMP				0	0%	40
95	0	4095	0			
ECMP	Mpls			0	0%	4095
	0	4095	0			
ECMP	Routing			0	0%	4095
	0	4095	0			
ECMP	VxlanOverlay			0	0%	4095
	0	4095	0			
ECMP	VxlanTunnel			0	0%	3891
	0	3891	0			

## Limitations

- In this mechanism, optimization of a sub-optimal ECMP route requires releasing and reallocating hardware resources for the route. Therefore the process may increase overall convergence time for route programming. It may not be desirable to always start the optimization when the hardware resource is sufficient. The threshold value for starting the optimization should be adjusted based on the route scale of the network.

- The deduping of ECMP hardware resources may cause potential traffic flow disruption for traffic flows going over RECOMP routes with changing hardware table entries. While the deduping process tries to minimize the amount of traffic flow disruption, it is still sometimes inevitable.
- RECOMP hardware table entries can only be deduped to other RECOMP hardware table entries that share the same set of nexthops. This puts a limit to the amount of RECOMP hardware table entries that can be reduced to the number of RECOMP hardware table entries with unique nexthop sets.

## Resources

- Link to the Resilient ECMP TOI: [Resilient ECMP](#)
- Link to the hardware resource optimization for route programming TOI: [Hardware resource optimization for route programming](#)