WRED-Explicit Congestion Notification Feature Overview

Currently, the congestion control and avoidance algorithms for Transmission Control Protocol (TCP) are based on the idea that packet loss is an appropriate indication of congestion on networks transmitting data using the best-effort service model. When a network uses the best-effort service model, the network delivers data if it can, without any assurance of reliability, delay bounds, or throughput. However, these algorithms and the best-effort service model are not suited to applications that are sensitive to delay or packet loss (for instance, interactive traffic including Telnet, web-browsing, and transfer of audio and video data). Weighted Random Early Detection (WRED), and by extension, Explicit Congestion Notification (ECN), helps to solve this problem.

RFC 3168, *The Addition of Explicit Congestion Notification (ECN) to IP*, states that with the addition of active queue management (for example, WRED) to the Internet infrastructure, routers are no longer limited to packet loss as an indication of congestion.

WRED Explicit Congestion Notification

Information About WRED Explicit Congestion Notification

Guidelines and Limitations for WRED Explicit Congestion Notification

- Explicit Congestion Notification (ECN) parameters are configurable only at system level.
- Weighted Random Early Detection (WRED) cannot be configured alone on a Quality of Service (QoS) group. ECN is enabled by default.
- You must configure WRED thresholds for 10 G interfaces and 40 G interfaces even when no interfaces are present.
- WRED ECN is not applicable to multicast or broadcast traffic.
- WRED ECN is not supported on Nexus 5000 series switches.

How WRED Works

WRED makes early detection of congestion possible and provides a means for handling multiple classes of traffic. WRED can selectively discard lower priority traffic when the router begins to experience congestion and provide differentiated performance characteristics for different classes of service. It also protects against global synchronization. Global synchronization occurs as waves of congestion crest, only to be followed by periods of time during which the transmission link is not used to capacity. For these reasons, WRED is useful on any output interface or router where congestion is expected to occur.

WRED is implemented at the core routers of a network. Edge routers assign IP precedences to packets as the packets enter the network. With WRED, core routers then use these precedences to determine how to treat different types of traffic. WRED provides separate thresholds and weights for different IP precedences, enabling the network to provide different qualities of service, in regard to packet dropping, for different types of traffic. Standard traffic may be dropped more frequently than premium traffic during periods of congestion.

For more information about WRED, refer to the "Congestion Avoidance Overview" module.

ECN Extends WRED Functionality

WRED drops packets, based on the average queue length exceeding a specific threshold value, to indicate congestion. ECN is an extension to WRED in that ECN marks packets instead of dropping them when the average queue length exceeds a specific threshold value. When configured with the WRED -- Explicit Congestion Notification feature, routers and end hosts would use this marking as a signal that the network is congested and slow down sending packets.

As stated in RFC 3168, *The Addition of Explicit Congestion Notification (ECN) to IP,*implementing ECN requires an ECN-specific field that has two bits--the ECN-capable Transport (ECT) bit and the CE (Congestion Experienced) bit--in the IP header. The ECT bit and the CE bit can be used to make four ECN field combinations of 00 to 11. The first number is the ECT bit and the second number is the CE bit. The table below lists each of the ECT and CE bit combination settings in the ECN field and what the combinations indicate.

Table 1: ECN Bit Setting

ECT Bit	CE Bit	Combination Indicates
0	0	Not ECN-capable
0	1	Endpoints of the transport protocol are ECN-capable
1	0	Endpoints of the transport protocol are ECN-capable
1	1	Congestion experienced

The ECN field combination 00 indicates that a packet is not using ECN.

The ECN field combinations 01 and 10--called ECT(1) and ECT(0), respectively--are set by the data sender to indicate that the endpoints of the transport protocol are ECN-capable. Routers treat these two field combinations identically. Data senders can use either one or both of these two combinations. For more information about these two field combinations, and the implications of using one over the other, refer to RFC 3168, *The Addition of Explicit Congestion Notification (ECN) to IP.*

The ECN field combination 11 indicates congestion to the endpoints. Packets arriving a full queue of a router will be dropped.

How Packets Are Treated When ECN Is Enabled

If the number of packets in the queue is below the minimum threshold, packets are transmitted. This happens whether or not ECN is enabled, and this treatment is identical to the treatment a packet receives when WRED only is being used on the network.

If the number of packets in the queue is between the minimum threshold and the maximum threshold, one of the following three scenarios can occur:

- If the ECN field on the packet indicates that the endpoints are ECN-capable (that is, the ECT bit is set to 1 and the CE bit is set to 0, or the ECT bit is set to 0 and the CE bit is set to 1)--and the WRED algorithm determines that the packet should have been dropped based on the drop probability--the ECT and CE bits for the packet are changed to 1, and the packet is transmitted. This happens because ECN is enabled and the packet gets marked instead of dropped.
- If the ECN field on the packet indicates that neither endpoint is ECN-capable (that is, the ECT bit is set to 0 and the CE bit is set to 0), the packet may be dropped based on the WRED drop probability. This is the identical treatment that a packet receives when WRED is enabled without ECN configured on the router.
- If the ECN field on the packet indicates that the network is experiencing congestion (that is, both the ECT bit and the CE bit are set to 1), the packet is transmitted. No further marking is required.

If the number of packets in the queue is above the maximum threshold, packets are dropped based on the drop probability. This is the identical treatment a packet receives when WRED is enabled without ECN configured on the router.

Proxy Queue Drain Rates

When the proxy queue reaches a threshold that indicates congestion, Explicit Congestion Notification (ECN) marking is performed so that the receiver of the packet echoes the congestion indication to the sender. The sender must respond as though the congestion had been indicated by packet drops. The proxy queue drain rate is configured to ensure that during congestion at egress ports only a certain amount of packets are drained. For example, on a 10 Gigabot port, you can configure a drain rate of 9900 Mbps ensuring that not all packets are allowed to drain.

Recommended ECN Thresholds and Proxy Queue Drain Rates

The following table describes the recommended proxy-ques drain rates and maximum and minimum Explicit Congestion Notification (ECN) threshold values.

Parameter	10 Gigabit Port	40 Gigabit Port
ECN min-threshold	64000 bytes	4000 bytes
ECN max-threshold	128000 bytes	256000 bytes
Proxy-queue drain rate	9900 Mbps	39900 Mbps

How to Configure WRED Explicit Congestion notification

Configuring WRED-Explicit Congestion Notification

To configure WRED-ECN you specify interface thresholds, enable ECN, and specify the proxy queue drain rate.

Before you begin

Before you configure Weighted Random Early Detection (WRED) Explicit Congestion Notification (ECN) on the device, you must configure a Quality of Service (QoS) group. In addition, the following restrictions apply:

- Explicit Congestion Notification (ECN) parameters are configurable only at system level.
- Weighted Random Early Detection (WRED) cannot be configured alone on a Quality of Service (QoS) group. ECN is enabled by default.
- You must configure WRED thresholds for 10 G interfaces and 40 G interfaces even when no interfaces are present.

Procedure

	Command or Action	Purpose
Step 1	switch# **configure terminal**	Enters global configuration mode.
Step 2	switch(config)# **hardware random-detect min-thresh 10g** *10g-min-threshold* **40g** *40g-min-threshold* **max-thresh 10g** *10g-max-threshold* **40g** *40g-max-threshold* **ecn qos-group** *qos-group-number*	Configures the minimum and maximum thresholds for 10 Gigabit and 40 Gigabit interfaces and enables ECN on particular QoS group. The thresholds for both 10 G and 40 G interfaces can range from 1 to 67108863 bytes. The QoS group number specifies which QoS group is being configured, the range is from 0 (class default) to 5.
Step 3	switch(config)# **hardware pq-drain 10g** *10g-drain-rate* **40g** *40g-drain-rate*	Configures the proxy queue drain rate for 10 Gigabit and 40 Gigabit ports. When congestion occurs at egress ports the drain rate values specify the maximum amount of packets that can be drained. The drain rate for 10 G interfaces can range from 1 to 20000 Mbps. The drain rate for 40 G interfaces can range from 1 to 80000 Mbps.

Example for WRED Explicit Congestion Notification

Configuration Example for WRED Explicit Congestion Notification

The following example shows how to configure Weighted Random Early Detection (WRED) Explicit Congestion Notification (ECN):

```
switch# configuration terminal
switch(config)# hardware random-detect min-thresh 10g 64000 40g 4000 max-thresh 10g 128000
 40g 256000 ecn qos-group 2
switch(config)# hardware pq-drain 10g 9900 40g 39900
switch(config)# exit
switch(config)# copy running-config startup-config
```