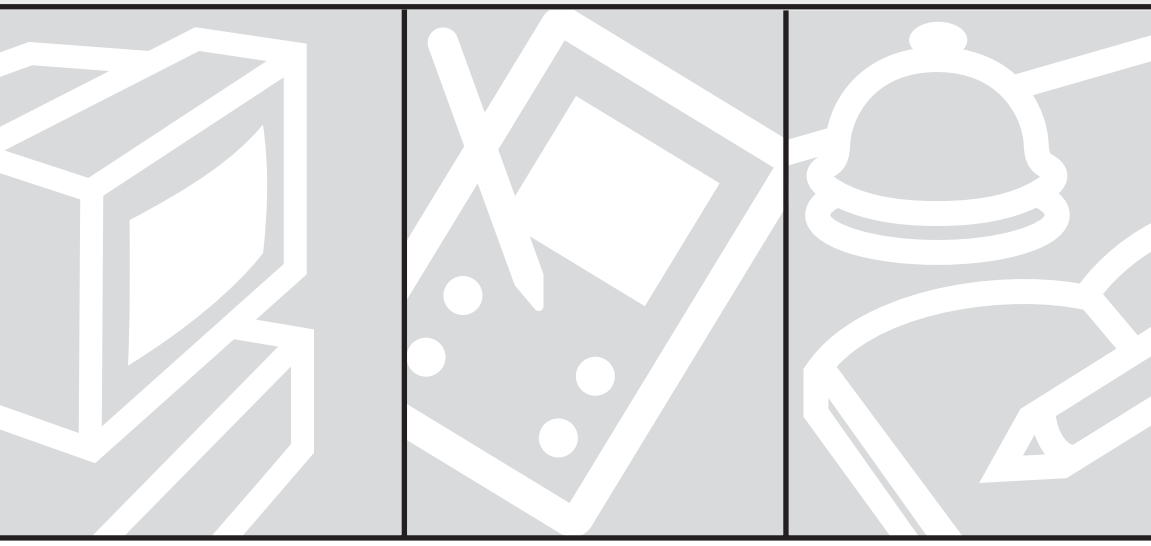


QOS Section 6
Weighted Random Early Detection (WRED)

The previous section addressed queuing, which is a congestion-management QoS mechanism. However, this section focuses on congestion avoidance. Specifically, you do not want a queue to fill to capacity, because all arriving traffic flows would be dropped and enter Transport Control Protocol (TCP) slow start. Some negative consequences arise from having multiple flows simultaneously enter TCP slow start. To prevent such an occurrence, you can configure Weighted Random Early Detection (WRED).

These flash cards review the need for WRED. You also must be familiar with the inner workings of the industry-standard Random Early Detection (RED) approach, in addition to the configuration of WRED. Even though WRED can be configured from interface-configuration mode or as part of a Modular Quality of Service Command-Line Interface (MQC) configuration, these flash cards focus specifically on WRED's MQC implementation. Finally, you need to recall the characteristics and configuration of the Explicit Congestion Notification (ECN) mechanism.

Question 1

What is TCP slow start?

Question 2

Which QoS tool does Cisco use to prevent a queue from filling to capacity?

Question 3

Describe the following WRED parameters: minimum threshold, maximum threshold, and Mark Probability Denominator (MPD).

Question 1 Answer

TCP slow start occurs if a sender does not receive an acknowledgment from its receiver within a certain time. When this occurs, the TCP window size is reduced to 1. The TCP window size then increases exponentially up to one-half of the original congestion window size, after which the window size increases linearly.

Question 2 Answer

Cisco IOS can use WRED to prevent a queue from filling to capacity by discarding traffic more aggressively as the queue begins to fill. Dropping decisions are made based on the traffic's priority markings.

Question 3 Answer

- The minimum threshold specifies the number of packets that must be in a queue before the queue considers discarding packets that have a particular marking.
- The probability of discard increases until the queue depth reaches the maximum threshold. After a queue depth exceeds the maximum threshold, all other packets with a particular marking that attempt to enter the queue are discarded.
- However, the probability of packet discard when the queue depth equals the maximum threshold is 1/(MPD). For example, if the MPD were set to 10, when the queue depth reached the maximum threshold, the probability of discard, for the specified marking, would be 1/10 (that is, a 10 percent chance of discard).

Question 4

When configuring WRED, you configure an MPD of 4 for traffic that is marked with a Differentiated Services Code Point (DSCP) value of 46. When the queue depth is at the maximum threshold for DSCP 46 traffic, what is the probability (in percent) that a packet marked with a DSCP value of 46 will be discarded?

Question 5

Identify the QoS marking that WRED references if the following command is issued:

```
Router(config-if)#random-detect
```

Question 6

Define *global synchronization.*

Question 4 Answer

The probability of packet discard when the queue depth equals the maximum threshold is 1/(MPD). Therefore, if the MPD = 4, the probability of discard is 1/4, which equals 0.25 (that is, 25 percent).

Question 5 Answer

The **dscp-based** or **prec-based** option can be used with the **random-detect** command to specify the QoS markings that WRED should reference when making discard decisions. However, if neither of these parameters is used, WRED defaults to referencing IP Precedence values.

Question 6 Answer

Global synchronization occurs when multiple TCP flows simultaneously go into TCP slow start. This can occur if a queue is full, because all newly arriving packets are discarded. Note that global synchronization is sometimes referred to as *TCP synchronization*.

Question 7

Describe the three drop modes of RED: no drop, random drop, and full drop.

Question 8

Besides policy-map-class configuration mode, list two other configuration modes from which WRED can be configured.

Question 9

Identify what is typically the best location in an enterprise network to enable WRED.

Question 7 Answer

The three drop modes of RED are as follows:

- "No drop" occurs if the queue depth is at or below the minimum threshold. In this mode, there is no chance of discard.
- "Random drop" occurs if the queue depth is above the minimum threshold and equal to or below the maximum threshold. The likelihood of discard depends on the queue depth and the MPD. When the queue depth equals the maximum threshold, the probability of discard is 1/MPD. If the queue depth currently equaled the maximum threshold and the MPD = 4, the probability of discard would be 1/4 = 0.25 = 25 percent.
- "Full drop" occurs if the queue depth the maximum threshold. In this mode, there is a 100 percent chance of discard.

Question 8 Answer

In addition to Class-Based WRED (CB-WRED), in which WRED is configured in policy-map-class configuration mode, WRED can be configured from interface-configuration mode or from virtual-circuit-configuration mode.

Question 9 Answer

WRED typically is used on router interfaces that are likely to experience congestion, such as a WAN interface. However, WRED can be enabled within the core of an enterprise network, too. Because WRED is not processor intensive (for example, it does not compare packets to an access list or alter bits in a packet's ToS byte), WRED could be enabled throughout an enterprise. However, it typically serves the greatest benefit at the WAN edge.

Question 10

What policy-map-class configuration-mode command enables WRED and specifies that the WRED profiles should be based on DSCP values, as opposed to IP Precedence values?

Question 11

Which WRED command specifies a minimum threshold of 24, a maximum threshold of 45, and an MPD of 10 for an IP Precedence value of 2?

Question 12

Which command can you use to see the WRED parameters that are associated with a particular interface using the MQC process?

Question 10 Answer

The **random-detect dscp-based** command is used in policy-map-class configuration mode to enable WRED for the class of traffic and to instruct WRED to use drop profiles based on DSCP values. If the **random-detect** command had been issued without the **dscp-based** option, WRED would base its drop profiles instead on IP Precedence values.

Question 11 Answer

The syntax to configure WRED parameters for an IP Precedence value is as follows:

```
Router(config-pmap-c)#random-detect precedence precedence_value
minimum-threshold maximum-threshold mark-probability-denominator
```

Therefore, in this example, the command would be as follows:

```
Router(config-pmap-c)#random-detect precedence 2 24 45 10
```

Question 12 Answer

To view the parameters of a policy-map (including WRED parameters) that are associated with an interface, use the command:

```
show policy-map interface interface-identifier
```

Question 13

Define *Explicit Congestion Notification (ECN).*

Question 14

What are the names of the ECN bits that are located in the 2 right-most bit positions of the ToS byte?

Question 15

What happens to packets that are sent between two ECN-capable routers when the queue depth for those packets is below the minimum threshold?

Question 13 Answer

ECN indicates the presence of congestion through signaling rather than dropping. By using the 2 right-most bits in the ToS byte, an ECN-capable device can indicate whether congestion is being experienced. (Reference RFC 3168 for more information.)

Question 14 Answer

The 7th bit in the ToS byte is the ECN-Capable Transport (ECT) bit. The 8th bit in the ToS byte is the Congestion Experienced (CE) bit. Note that the bit combinations of 10 or 01 are viewed identically (that is, as an ECN-capable device that is not currently experiencing congestion).

Question 15 Answer

If the queue depth is below the WRED minimum threshold, the packets are sent normally, just as with WRED.

Question 16

What is being indicated when a packet is received with both of its ECN bits set to 1?

Question 17

What happens to packets that are sent between two ECN-capable routers when the queue depth is above the packets' maximum threshold?

Question 18

What happens when an ECN-marked packet is received by a router that is not ECN capable?

Question 16 Answer

The ECT bit that is being set to 1 indicates that the device that is sending the packet is capable of using ECN, and the CE bit that is being set to 1 indicates that congestion is currently being experienced. When a packet is received with both of its ECN bits (that is, the ECT and CE bits) set to 1, the sender is ECN capable and is currently experiencing congestion.

Question 17 Answer

If the queue depth is above the configured maximum threshold, the packets are dropped, just as with WRED.

Question 18 Answer

If a packet is received by a router that is not ECN capable, the destination router does not examine the 7th or 8th bits (that is, the ECT and CE bits). Therefore, the destination router treats the packet with normal WRED behavior, assuming that the destination router is configured for WRED.

Question 19

Under what configuration mode (or modes) can you configure ECN?

Question 20

Which command enables ECN?

Question 19 Answer

Even though you can configure WRED under interface-configuration mode or in virtual-circuit-configuration mode, you can configure ECN only under policy-map-class configuration mode.

Question 20 Answer

The command to enable ECN is as follows:

```
Router(config-pmap-c)#random-detect ecn
```

In this example, the **wrr-queue** command is assigning the weights 1, 2, 3, and 4 to the switch's four queues. The first queue, with a weight of 1, for example, only gets one-third the bandwidth that is given to the third queue, which has a weight of 3. The **wrr cos-map 4 5** command is instructing frames that are marked with a CoS of 5 to enter the fourth queue.

To verify how a Catalyst 2950 is mapping CoS values to DSCP values (or vice versa), use the following command:

```
Switch#show mls qos maps [cos-dscp | dscp-cos]
```

You can use the following command to view the weight that is assigned to each queue:

```
Switch#show wrr-queue bandwidth
```

Another useful WRR command, which shows how CoS values are being mapped to switch queues shows is as follows:

```
Switch#show wrr-queue cos-map
```

Finally, you can see the QoS configuration for an interface (for example, trust state and the interface's default CoS value) with the following command:

```
Switch#show mls qos interface [interface-identifier]
```

Weighted Random Early Detection (WRED)

Whereas queuing provides congestion management, mechanisms such as WRED provide congestion avoidance. Specifically, WRED can prevent an output queue from ever filling to capacity, which would result in packet loss for all incoming packets. This section examines the need for and the configuration of WRED on Cisco routers.

How TCP Handles Drops

Recall from your early studies of networking technology how Transport Control Protocol (TCP) windowing functions. A sender sends a single segment, and if the sender receives a successful acknowledgment from the receiver, it then sends two segments (that is, a "windows size" of 2). If those two segments were acknowledged successfully, the sender sends four segments, and so on, increasing the window size exponentially.

However, if one of the segments is dropped, the TCP flow goes into TCP slow start, where the window size is reduced to 1. The TCP flow then exponentially increases its window size until the window size reaches half of the window size when congestion originally occurred. At that point, the TCP flow's window size increases linearly.

TCP slow start is relevant to QoS, because when an interface's output queue is full, all newly arriving packets are discarded (that is, "tail dropped"), and all of those TCP flows simultaneously go into TCP slow start.

Note that the process of multiple TCP flows simultaneously entering TCP slow start is called *global synchronization* or *TCP synchronization*. When TCP synchronization occurs, the link's bandwidth is underutilized, resulting in wasted bandwidth.

RED Basics

The purpose of Random Early Detection (RED) is to prevent TCP synchronization by randomly discarding packets as an interface's output queue begins to fill. How aggressively RED discards packets depends on the current queue depth.

The following three parameters influence when a newly arriving packet is discarded:

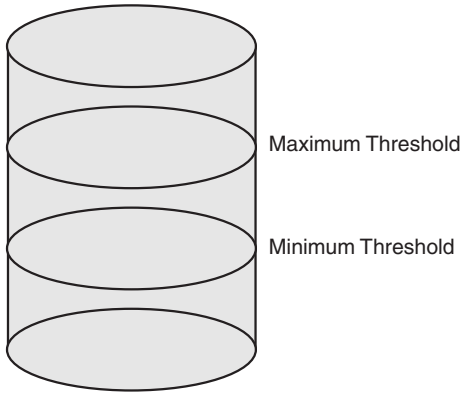
- Minimum threshold
- Maximum threshold
- Mark Probability Denominator (MPD)

The *minimum threshold* specifies the number of packets in a queue before the queue considers discarding packets. The probability of discard increases until the queue depth reaches the *maximum threshold*. After a queue depth exceeds the maximum threshold, all other packets that attempt to enter the queue are discarded.

However, the probability of packet discard when the queue depth equals the maximum threshold is 1/(MPD). For example, if the mark probability denominator were set to 10, when the queue depth reached the maximum threshold, the probability of discard would be 1/10 (that is, a 10 percent chance of discard).

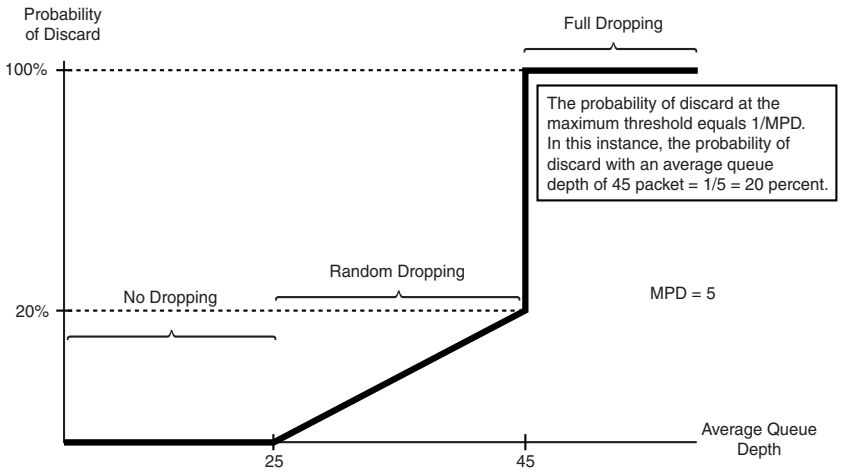
Random Early Detection (RED)

As an output queue fills beyond the minimum threshold, RED begins to discard packets. Those packets are discarded more aggressively as the queue depth increases. When the queue depth exceeds the maximum threshold, all packets are discarded.



The minimum threshold, maximum threshold, and MPD comprise the RED profile. The following figure shows the three distinct ranges in a RED profile: no drop, random drop, and full drop.

RED Drop Ranges



RED is most useful on router interfaces where congestion is likely. For example, a WAN interface might be a good candidate for RED.

CB-WRED

Cisco does not support RED, but fortunately it supports something better: Weighted Random Early Detection (WRED). Unlike RED, WRED has a profile for each priority marking. For example, a packet with an IP Precedence value of 0 might have a minimum threshold of 20 packets, whereas a packet with an IP Precedence of 1 might have a minimum threshold of 25 packets. In this example, packets with an IP Precedence of 0 would start to be discarded before packets with an IP Precedence of 1.

Although WRED can be configured from interface-configuration mode or from virtual-circuit-configuration mode, these Quick Reference Sheets focus on an MQC-based WRED configuration. To enable WRED and to specify the marking that WRED pays attention to (that is, IP Precedence or DSCP), issue the following policy-map-class configuration-mode command:

```
Router(config-pmap-c)#random-detect [dscp-based | prec-based]
```

If neither **dscp-based** nor **prec-based** is specified, WRED defaults to **prec-based**. After WRED is configured, the IOS assigns default minimum threshold, maximum threshold, and MPD values. However, you can alter those default parameters with the following commands:

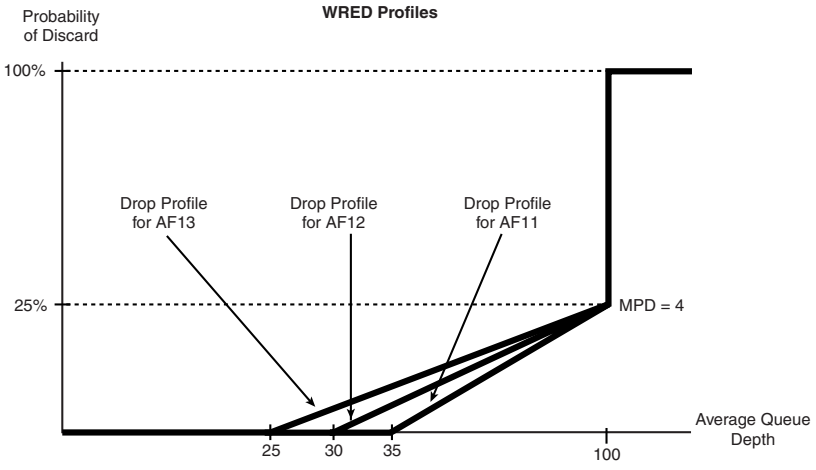
```
Router(config-pmap-c)#random-detect precedence precedence_value minimum-
  threshold maximum-threshold mark-probability-denominator
```

(Used for prec-based WRED)

```
Router(config-pmap-c)#random-detect dscp dscp_value minimum-threshold maximum-
  threshold mark-probability-denominator
```

(Used for dscp-based WRED)

To reinforce this syntax, consider the following example, where the goal is to configure WRED for the **WREDTEST** class-map. After the class-map's queue depth reaches 25 packets, a DSCP value of AF13 might be discarded. Packets that are marked with a DSCP value of AF12 should not be discarded until the queue depth reaches 30 packets, and finally, packets that are marked with a DSCP value of AF11 should have no chance of discard until the queue depth reaches 35 packets. If the queue depth exceeds 100 packets, there should be a 100 percent chance of discard for these three DSCP values. However, when the queue depth is exactly 100 packets, the chance of discard for these various packet types should be 25 percent. Also, CB-WRED requires that CB-WFQ be configured for the interface. So, as an additional requirement, you make 25 percent of the interface's bandwidth available to the **WREDTEST** class of traffic.



```
Router(config-pmap)#class WREDTEST
Router(config-pmap-c)#bandwidth percent 25
Router(config-pmap-c)#random-detect dscp-based
Router(config-pmap-c)#random-detect dscp af13 25 100 4
Router(config-pmap-c)#random-detect dscp af12 30 100 4
Router(config-pmap-c)#random-detect dscp af11 35 100 4
```

Examine the solution, and notice that the MPD is 4. This value was chosen to meet the requirement of a 25 percent chance of discard when the queue depth equals the maximum threshold (that is, 1/4 = .25). Also, notice that a DSCP value of AF13 is dropped before a DSCP value of AF12, which is dropped before a DSCP value of AF11. This approach is consistent with the definition of the per-hop behaviors (PHBs), because the last digit in the Assured Forwarding (AF) DSCP name indicates its drop preference. For example, a value of AF13 would drop before a value of AF12.

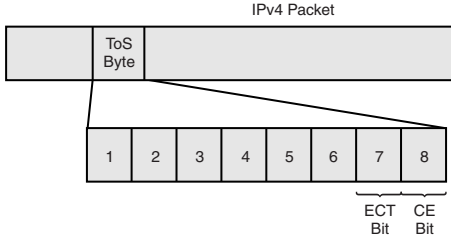
To view the minimum threshold, maximum threshold, and MPD settings for the various IP Precedence or DSCP values, you can issue the **show policy-map interface** ***interface-identifier*** command.

ECN Configuration

WRED discards packets, and that is one way for the router to indicate congestion. However, routers can now indicate a congestion condition by signaling, using an approach called Explicit Congestion Notification (ECN).

ECN uses the 2 last bits in the ToS byte to indicate whether a device is ECN capable, and if so, whether congestion is being experienced.

Explicit Congestion Notification (ECN)



Bit Combination	Meaning
00	Router is not ECN capable.
01	Router is ECN capable but is not experiencing congestion.
10	Router is ECN capable but is not experiencing congestion.
11	Router is ECN capable and is currently experiencing congestion.

Cisco routers can use ECN as an extension to WRED and mark packets that exceed a specified value, instead of dropping the packets. If the queue depth is at or below the WRED minimum threshold, the packets are sent normally, just as with WRED. Also, if the queue depth is above the WRED maximum threshold, all packets are dropped, just as with WRED.

But if the queue depth is currently in the range from the minimum threshold through the maximum threshold, one of the following things can happen:

- If both endpoints are ECN capable, the ECT and CE bits are set to a 1 and sent to the destination, indicating that the transmission rate should be reduced.
- If neither endpoints supports ECN, the normal WRED behavior occurs.
- A packet with its ECN and CE bits marked can reach a destination router that already has a full queue. In such an instance, the notification is dropped.

Use the following command to enable ECN:

```
Router(config-pmap-c)#random-detect ecn
```

Note that although WRED also can be configured in interface-configuration mode, ECN must be configured through MQC. Because ECN is configured by the three-step MQC process, the same verification and troubleshooting commands apply. Specifically, you can use the **show policy-map** and **show policy-map interface** ***interface-identifier*** commands to verify the ECN configuration.

Traffic Conditioners

QoS mechanisms can not only provide for the allocation of a minimum amount of bandwidth for specific traffic but also limit the amount of bandwidth made available to that traffic. This section discusses how policing and shaping mechanisms limit traffic rates.

Policing Versus Shaping

Instead of allocating bandwidth for applications, in some instances, you might want to restrict the amount of bandwidth that is available for specific traffic. For example, you might want to set a "speed limit" for users on the network who are downloading music files from the Internet.

QoS mechanisms that limit bandwidth are called *traffic conditioners*. The two categories of traffic conditioners are policing and shaping. Although both of these approaches limit bandwidth, they have different characteristics, as follows:

- **Policing**—Policing typically limits bandwidth by discarding traffic that exceeds a specified rate. However, policing also can remark traffic that exceeds the specified rate and attempt to send the traffic anyway. Because policing's drop behavior causes TCP retransmits, it is recommended for use on higher-speed interfaces. Also, note that policing can be applied inbound or outbound on an interface.
- **Shaping**—Shaping limits excess traffic, not by dropping it but by buffering it. This buffering of excess traffic can lead to delay. Because of this delay, shaping is recommended for slower-speed interfaces. Unlike policing, shaping cannot remark traffic. As a final contrast, shaping can be applied only in the outbound direction on an interface.

The question becomes this: How do you send traffic out of an interface at a rate that is less than the physical clock rate of the interface? It is impossible for an interface to send at a rate that is slower than the line rate. However, you can send at an "average" rate that is less than the clock rate by using policing or shaping tools that do not transmit all the time. Specifically, these tools send a certain number of bits or bytes at line rate, and then they stop sending until a specific timing interval (for example, 1/8 of a second) is reached. When the timing interval is reached, the interface again sends a specific amount of traffic at line rate, it stops, and then it waits for the next timing interval to occur. This process continually repeats, allowing an interface to send an average bandwidth that can be below the physical speed of the interface. This average bandwidth is called the Committed Information Rate (CIR). The number of bits (the unit of measure that is used with shaping tools) or bytes (the unit of measure that is used with policing tools) that is sent during a timing interval is called the Committed Burst (Bc). The timing interval is written as Tc.

For example, consider that you have a physical line rate of 128 kbps, but the CIR is only 64 kbps. Also consider that there are eight timing intervals in a second (that is, Tc = 1/8 of a second = 125 ms), and during each of those timing intervals, 8000 bits (that is, the committed burst parameter)