

Rate Limits for Interfaces Overview

To configure rate limiting for interfaces, you first create a rate-limit profile, which is a set of bandwidth attributes and associated actions. Your router supports two types of rate-limit profiles—one-rate and two-rate—for IP, IPv6, LT2P, and MPLS Layer 2 transport traffic. You next create a policy list with a rule that has rate limit as the action and associate a rate-limit profile with this rule.

You configure rate limit profiles from Global Configuration Mode.



NOTE: Commands that you issue in Rate Limit Profile Configuration mode do not take effect until you exit from that mode.

When packets enter an interface that has a rate-limit profile applied, the router performs the following:

- Counts the number of bytes (packets) over time
 - Categorizes each packet as committed, conformed, or exceeded
 - Assigns a transmit, drop, or mark action
-



NOTE: Mark actions and mask values are supported only on IP, IPv6, and MPLS rate-limit profiles. They are not supported on hierarchical rate limits, but are replaced by color-mark profiles.

An additional function of rate limiting is to apply a color code to packets assigned to each category: green for committed, yellow for conformed, and red for exceeded. The system uses the color code internally to indicate drop preference when an outbound interface is congested.

Rate limiters are implemented using a dual token bucket scheme: a token bucket for conformed (yellow) packets and a token bucket for committed (green) packets. One token is synonymous with one byte. The capacity of the buckets is the maximum number of tokens that can be placed in each bucket.

You configure the bucket capacity with the peak burst parameter or the committed burst parameter. The burst parameters are in bytes (not bytes per second), which is the number of tokens in a full bucket. When a packet passes through a rate limiter, its size is compared to the contents of both buckets, the packet is categorized, and the rate-limiter action is taken on the packet.

Peak rate and committed rate determine the fill rate of their respective buckets. If you set the committed rate to 128,000 bps, tokens are added to the committed (green) bucket at a rate of 128,000 bps (16 K bytes per second), regardless of the traffic. If no traffic passes through the rate limiter, the bucket continues to fill until it reaches the committed burst setting.

Traffic passes through the rate limiter causing a draining of tokens. The drain rate is dependent on how large the packets are and how much time elapses between

packets. At any given instant the level of tokens in each bucket is a function of the fill rate, size of packets, and elapsed time between packets.

When packets are received on an interface with a rate limiter applied, the level of tokens in each bucket dynamically changes in both of the following ways:

- Tokens are added every 100-ms sample period
- Tokens are removed based on the size and rate of incoming packets