

Rate Limiting Function (RLF)

This chapter contains the following topics:

- Revision History, on page 1
- Feature Description, on page 1

Revision History

Note Revision history details are not provided for features introduced before release 21.24.

Revision Details	Release
First introduced	Pre 21.24

Feature Description

The RLF feature implements a generic framework that can be used by multiple interfaces and products for rate-limiting/throttling outgoing messages like Diameter messages on Gx, Gy interface towards PCRF.

Important The working of RLF feature, including the CLI commands, in the CUPS architecture is similar to how it works in the non-CUPS environment.

When applications send messages to peers at a high rate (for example, when a large number of sessions goes down at the same time), accounting stop messages for all the sessions are generated at the same time) the peer may not be able to handle the messages at such high rates. To overcome this situation, the Rate Limiting Function (RLF) framework is developed so that the application sends messages at an optimal rate such that peer is capable of receiving all the messages and does not enter an overload condition.

This feature can be enabled using the **rlf-template** CLI command in the Global Configuration mode. The users can define the rate limiting configurations within this template. For more information on the command, see the *Command Line Interface Reference*.

Important RLF template cannot be deleted if it is bound to any application (peers/endpoints).

When RLF feature is enabled, all the messages from the application are pushed to the RLF module for throttling and rate control, and depending on the message-rate configured the RLF module sends the messages to the peer. Once the rate or a threshold value is reached, the RLF module notifies the application to slow down or stop sending messages. RLF module also notifies the application when it is capable of accepting more messages to be sent to the peer. RLF module typically uses a Token Bucket Algorithm to achieve rate limiting.

Currently in the deployment of the Diameter applications (Gx, Gy, and so on), many operators make use of **max-outstanding** *number* CLI command as a means of achieving some rate-limiting on the outgoing control traffic. With RLF in place, this is no longer required since RLF takes care of rate-limiting in all cases. If both RLF and **max-outstanding** is used, there might be undesirable results.

Important If RLF is being used with a **diameter endpoint**, then set the **max-outstanding** value of the peer to be 255.

To use the template, Diameter or any other applications must be associated with the template. The RLF provides only the framework to perform the rate limiting at the configured Transactions Per Second (TPS). The applications (like Diameter) should perform the configuration specific to each application.