

GTP-C Load and Overload Control on MME

- Feature Description, on page 1
- How it Works, on page 3
- Configuring GTP-C Load and Overload Control on MME, on page 3
- Monitoring and Troubleshooting the GTP-C Load and Overload Control Feature, on page 6

Feature Description

Overload of packet core network nodes in the network results in service de-gradation. Overload conditions can occur in various network scenarios. Overload issue can be addressed through improved load distribution over the network.

GTP-C load and overload control feature adds MME support for GTP-C load and overload control mechanism on S11 interface. GTP-C load and overload control is a standard driven (3GPP TS 29.807 V12.0.0 and 3GPP TS 29.274 V d30) feature. For standards compliance information see the Standards Compliance section in this feature chapter.



Important This feature is license controlled. The "EPC Support for GTP Overload Control" license is required for successfully configuring and enabling this feature. Please consult your Cisco Account Representative for information about the specific license.

GTP-C Overload Issues and Resultant effects

A GTP-C overload occurs when the number of incoming requests exceeds the maximum request throughput supported by the receiving GTP-C entity. The GTP-C is carried over UDP transport, and it relies on the re-transmissions of unacknowledged requests. When a GTP-C entity experiences overload (or severe overload) the number of unacknowledged GTP-C messages increase exponentially and this leads to a node congestion or even node collapse. An overload or failure of a node further leads to an increase of the load on the other nodes in the network and in some cases into a network issue.

Listed below are some examples of GTP-C signaling based scenarios which lead to GTP-C overload:

- A traffic flood resulting from the failure of a network element, inducing a signaling spike.
- A traffic flood resulting from a large number of users performing TAU/RAU or from frequent transitions between idle and connected mode.

- An exceptional event locally generating a traffic spike for example a large amount of calls (and dedicated bearers) being setup almost simultaneously.
- Frequent RAT re-selection due to scattered Non-3GPP (for example, Wi-Fi) coverage or a massive mobility between a 3GPP and Non-3GPP coverage. This may potentially cause frequent or massive inter-system change activities.

GTP-C overload may result in any of the following service impacts:

- Loss of PDN connectivity (IMS, Internet and so on) and associated services.
- Loss of ability to setup and release radio and core network bearers necessary to support services, for example GBR bearers for VoLTE.
- Loss of ability to report to the PGW/PCRF user's information changes, for example location information for emergency services and lawful intercept, changes in RAT or QoS.
- Billing errors which result in loss of revenue.

Overview

GTP-C load control and overload control are complimentary concepts which can be supported and activated independently on the network. This feature uses the existing EGTPC infrastructure to gather and distribute load and overload control information across session managers. In broad terms GTP-C load control can be described as a preventive action and GTP-C overload control can be described as a corrective action. A GTP-C entity is termed as overloaded when it operates over and above its signaling capacity resulting in a diminished performance (including impacts to handling of incoming and outgoing traffic).

The advantages of enabling GTP-C load control are listed below:

- Load control allows better balancing of the session load; this prevents an GTP-C overload scenario.
- Load control enables a GTP-C entity (for example SGW or PGW) to send its load information to a GTP-C peer (for example a MME or SGSN, ePDG, TWAN) to adaptively balance the session load across entities supporting the same function (for example SGW cluster) according to their effective load. The load information reflects the operating status of the resources of the GTP-C entity.
- Load control does not trigger overload mitigation actions even if the GTP-C entity reports a high load.

The advantages of enabling GTP-C overload control are listed below:

- Overload control prevents a GTP-C entity from becoming or being overloaded to gracefully reduce its incoming signaling load by instructing its GTP-C peers to reduce sending traffic according to its available signaling capacity to successfully process the traffic.
- Overload control aims at shedding the incoming traffic as close to the traffic source as possible when an overload has occurred.

Operational Benefits of GTP-C Load and Overload Control Support on MME:

- Improved load distribution on SGW and PGW this in turn reduces the occurrence of PGW/SGW overload.

- The MME pro-actively advertises its overload information so that the peer nodes SGW/PGW can reduce the traffic.
- The MME can reduce the traffic towards the peer SGW/PGW if they are overloaded.

Relationships to Other Features

This feature utilizes the existing EGTPC infrastructure to gather and distribute load and overload control information across session managers.

How it Works

This section describes the detailed working mechanism of this feature.

Limitations

- APN level load and overload control is not supported

Standards Compliance

The GTP-C load and overload control feature complies with the following standards:

- 3GPP TS 29.807,version 12.0.0
- 3GPP TS 29.274, version 13.3.0 and beyond

Configuring GTP-C Load and Overload Control on MME

The following configuration procedures are used to configure the GTP-C load and overload control feature.

Configuring GTP-C Load or Overload Control Profile

The **associate** command associates or disassociates supportive services and policies with an MME service. New keywords **gtpc-load-control-profile** and **gtpc-overload-control-profile** are introduced in the **associate** command to configure the GTP-C load control profile and GTP-C overload control profile.

```
configure
  context context_name
    mme-service service_name
     associate { { egtp-service egtp_svc_name | egtp-sv-service egtp_sv_svc_name
 | foreign-plmn-guti-mgmt-db db_name |  gtpc-load-control-profile profile_name
 |  gtpc-overload-control-profile profile_name | henbgw-mgmt-db db_name |
hss-peer-service hss_svc_name | ipne-service ipne_svc_name | location-service
 location_svc_name | lte-emergency-profile profile_name |
network-global-mme-id-mgmt-db | s102-service  s102_svc_name [ context
context_name ] | sbc-service sbc_svc_name | sctp-param-template template_name |
```

```
sgs-service sgs_svc_name | sgtpc-service sgtpc_svc_name } [ context ctx_name ]
| subscriber-map map_name | tai-mgmt-db database_name }
      no associate { egtp-service | egtp-sv-service |
foreign-plmn-guti-mgmt-db |  gtpc-load-control-profile |
gtpc-overload-control-profile| henbgw-mgmt-db | hss-peer-service  |
ipne-service | location-service | lte-emergency-profile |
network-global-mme-id-mgmt-db  | s102-service | sctp-param-template |
sgs-service | sgtpc-service | subscriber-map | tai-mgmt-db }
```

Notes:

- The keyword **gtpc-load-control-profile** is used to configure GTP-C Load Control Profile for this MME service.
- The keyword **gtpc-overload-control-profile** is used to configure GTP-C Overload Control Profile for this MME service.
- The *profile_name* is a string of size 1 up to 64.
- These CLI commands are not enabled by default.

Configuring Usage of GTP-C Load Information in SGW/PGW Selection

The gw-selection command configures the parameters controlling the gateway selection process. A new keyword **gtp-weight** is introduced as a part of this feature which is the weight value calculated from the Load Control Information received from the GTP peers.

```
configure
  call-control-profile profile_name
    [remove] gw-selection {co-location [weight [prefer{ sgw | pgw }]]|
gtp-weight | pgw weight | sgw weight | topology [ weight [ prefer { sgw
| pgw } ] ] }
    exit
```

Notes:

- The option enables the MME selection of SGW and PGW based on the advertised load control information.
- This configuration can be applied selectively to subscribers.
- This CLI is not enabled by default.

Configuring MME Manager and IMSI Manager CPU Utilization to Calculate Overload Factor

This new command allows the user to configure the inclusion of CPU utilization of Session Manager, Demux Manager, IMSI Manager and MME Manager under GTP-C overload control profile for overload factor calculation.

```
configure
   gtpc-overload-control-profileprofile_name
     cpu-utilization {sessmgr-card | demuxmgr-card | imsimgr | mmemgr}
     no cpu-utilization
```

```
default cpu-utilization
exit
```

Notes:

- The **default** behavior for the above CLI is to include the average CPU utilization of Sessmgr cards and Demuxmgr card in the overload factor calculation.
- The **no** keyword disables the configuration of CPU utilization of Sessmgr/Demuxmgr/IMSImgr/MMEmgr under GTP-C overload control profile for overload factor calculation.
- The keyword **sessmgr-card** configures the inclusion of average cpu-utilization of SessMgr cards for overload factor calculation.
- The keyword **demuxmgr-card** configures the inclusion of average cpu-utilization of Demux Manager card for overload factor calculation.
- The keyword **imsimgr** configures the inclusion of cpu-utilization of IMSIMgr(s) proclet for overload factor calculation.
- The keyword **mmemgr** configures the inclusion of cpu-utilization of MMEMgr(s) proclet for overload factor calculation.

Sample Configuration

```
configure
  call-control-profile profile_name
    gw-selection topology weight prefer pgw
    gw-selection gtp-weight
exit
operator-policy namepolicy_name
  associate call-control-profile profile_name
exit
lte-policy
  subscriber-mapmap_name
    precedence xmatch-criteria all operator-policy-name policy_name
  exit
exit
context ingress
  mme-serviceservice_name
    bind s1-mme ipv4-address x.x.x.x
    s1-mme sctp portxx
    mme-id group-idxxxxx mme-codex
    plmn-id mcc xxx mnc xxx
    associate egtp-service egtp_mme context ingress
    associate mme-hss-service mme_hss contexthss
    associate subscriber-mapmap_name
    associate gtpc-load-control-profileprofile_name
    associate gtpc-overload-control-profile profile_name
  exit
exit
end
```

Verifying the Configuration

The configuration of this feature can be verified using the following show commands.

Execute the **show configuration** command to verify the configuration, the part of the output related Call Control Profile displays the following parameters based on the configuration:

- **gw-selection: gtp-weight**

The parameter **gtp-weight** identifies GTP-C load based SGW or PGW selection.

Execute the **show configuration** command to verify the configuration, the part of the output related to MME Service Configuration displays the following parameters based on the configuration:

- **associate gtpc-load-control-profile profile_name**
- **associate gtpc-overload-control-profile profile_name**

Execute the **show configuration** command to verify the configuration, the part of the output related to GTP-C overload control profile Configuration displays the following parameters based on the configuration:

- **cpu-utilization: sessmgr-card demuxmgr-card imsimgr mmemgr**

The **cpu-utilization** is used to configure the inclusion of CPU utilization of Session Manager, Demux Manager, IMSI Manager and MME Manager under GTP-C overload control profile for overload factor calculation.

Monitoring and Troubleshooting the GTP-C Load and Overload Control Feature

This section provides information regarding show commands and bulk statistics for this feature.

Troubleshooting

Listed below are the troubleshooting steps for any issues encountered during configuration or functioning of the GTP-C Load and Overload control feature:



Important All configuration parameters newly added will be recovered during Session Manager re-starts. The LCI/OCI information of SGW/PGW peer nodes are recovered during Session Manager restarts.

Step 1:

If the CLI commands required to enable the feature are not visible, ensure that the GTP-C Load and Overload feature control license is present and enabled.

Step 2:

If MME selection of SGW and PGW based on load control information is not working, verify the following:

- Ensure that the load control profile is associated with the MME service, for more information see the section on “Configuring GTP-C Load or Overload Control Profile” in this feature chapter.

- Ensure that the overload control profile is associated with the MME service, for more information see the section on “Configuring GTP-C Load or Overload Control Profile” in this feature chapter.
- Verify if the load and overload feature is associated and feature is enabled, execute the show commands **show mme-service all** and **show egtp-service name** *service_name*.
- Verify if the MME has started learning the LCI/OCI information from peer nodes using following show commands:

```
show egtpc peers
show egtpc peers sessmgrs
show egtpc peers address x.x.x.x
show egtpc statistics sgw-address x.x.x.x
show egtpc statistics remote-peer-address x.x.x.x
show subscribers summary mme-service service-name sgw-address x.x.x.x
show subscribers summary mme-service service-name pgw-address x.x.x.x
show egtpc statistics egtp-service service-name
```

- Ensure the configuration to consider LCI/OCI information in node selection is enabled. For more information see the section “Configuration to use GTPC load information in SGW/PGW selection” in this feature chapter. Execute the show command **show call-control-profile full all** to verify the same.
- For the GTP dynamic weight (that is, LCI/OCI) to work in case of DNS based node selection, following DNS weight based configuration should be present in the call control profile.
  - In case of SGW selection:

    ```
    configure
      call-control-profile profile_name
        gw-selection sgw weight
        end
    ```

    Please refer to the section " DNS SGW selection with Load Control Information" in this feature chapter.
  - In case of PGW selection:

    ```
    configure
      call-control-profile profile_name
        gw-selection pgw weight
        end
    ```

    Please refer to the section " DNS PGW selection with Load Control Information" in this feature chapter.
  - During Topology based selection:

    ```
    configure
      call-control-profile profile_name
        gw-selection topology [ weight [ prefer { pgw|sgw } ] ]
        end
    ```

    Please refer to the section "Topology" in this feature chapter.
  - When Co-location is enabled:

    ```
    configure
      call-control-profile profile_name
    ```

```
gw-selection co-location [ weight [ prefer { pgw|sgw } ] ]
end
```

Please refer to the section "Co-location" in this feature chapter.

Step 3:

If the MME Selection of SGW using DNS does not result in expected session distribution on the SGW.

- Ensure that Steps 1 and 2 are working as explained earlier in this section.
- Collect the data for following show commands at regular intervals of time to observe if the distribution is happening or not:

```
show egtpc peers
show egtpc peers sessmgrs
show egtpc peers address x.x.x.x
show egtpc statistics sgw-address x.x.x.x
show egtpc statistics remote-peer-address x.x.x.x
show subscribers summary mme-service service-name sgw-address x.x.x.x
```

Step 4:

If the MME Selection of PGW using DNS does not result in expected session distribution on the SGW.

- Ensure that Steps 1 and 2 are working as explained earlier in this section.
- Collect the data for following show commands at regular intervals of time to observe if the distribution is happening or not:

```
show egtpc peers
show egtpc peers sessmgrs
show egtpc peers address x.x.x.x
show egtpc statistics remote-peer-address x.x.x.x
show subscribers summary mme-service service-name pgw-address x.x.x.x
show egtpc statistics egtp-service service-name
```

Step 5:

If MME Selection of SGW and PGW using topology is not resulting in expected session distribution on SGW/PGW.

- Ensure that Steps 1 and 2 are working as explained earlier in this section.
- Collect the data for following show commands at regular intervals of time to observe if the distribution is happening or not:

```
show egtpc peers sessmgrs
show egtpc peers address x.x.x.x
show egtpc statistics sgw-address x.x.x.x
show egtpc statistics remote-peer-address x.x.x.x
show subscribers summary mme-service service-name sgw-address x.x.x.x
show subscribers summary mme-service service-name pgw-address x.x.x.x
show egtpc statistics egtp-service service-name
```

Step 6:

If the MME is not reporting overload control information, follow the steps described below:

- Ensure that Step 1 is working as explained earlier in this section.
- Ensure that the overload control profile is associated with the MME Service. For more information see the section “Configuring GTP-C Load or Overload Control Profile” in this feature chapter.
- Execute the following show commands to verify if the feature is enabled:
  - **show mme-service all**
  - **show egtp-service name egtp_mme**
- Verify the parameters configured through the commands inclusion-frequency, message- prioritization, overload-control-handling, overload-control-publishing, self-protection-behavior, tolerance, throttling-behavior, validity-period and weightage under the GTPC Overload Profile Configuration Mode. Execute the following show command to verify the same:

  show gtpc-overload-control-profile *profile_name*