Configure the Congestion Control Mechanism on the ASR 5X00

Document ID: 119151

Contributed by Shashank Varshney, Cisco TAC Engineer.

Jul 23, 2015

Contents

Introduction

This document describes how to configure the congestion control mechanism on the Cisco Aggregated Services Router (ASR) 5x00 Series. The congestion control functionality that is described in this document is primarily applied to the Serving General Packet Radio Service (GPRS) Support Node (SGSN) and Mobility Management Entity (MME) network functions.

Prerequisites

Requirements

There are no specific requirements for this document.

Components Used

This document is not restricted to specific software and hardware versions.

The information in this document was created from the devices in a specific lab environment. All of the devices used in this document started with a cleared (default) configuration. If your network is live, make sure that you understand the potential impact of any command.

Congestion Control Overview

At times, an excessive load can be observed in the network, which can result in a license breach, high CPU utilization, high port utilization, or high memory utilization. This can cause performance degradation on the node that is under heavy load, but these conditions are usually temporary and are quickly resolved. Congestion control is used in order to aid in the identification of such conditions and invoke the policies that address the situation when these heavy load conditions persist continuously, or a large number of these conditions exist.

This section describes the congestion control mechanism in the SGSN and the MME, as per the 3rd Generation Partnership Project (3GPP).

MME/SGSN Congestion Control

The MME provides a Non−Access Startum (NAS) level congestion control mechanism, which is based on Access Point Name (APN) or General NAS−level Mobility Management (MM) control.

The APN−based congestion control mechanisms can handle the Evolved Packet System (EPS) Session Management (ESM) and EPS Mobility Management (EMM) signalling that is associated with the User Equipment (UE) that has a particular APN and UE. The network should support this congestion control function. The MME detects the NAS−level congestion control that is associated with the APN, and it starts and stops the APN−based congestion control in accordance with this criteria:

- Maximum number of active EPS bearers per APN
- Maximum number of EPS bearer activations per APN
- One or more Packet Data Network (PDN) Gateways (PGWs) on an APN is not reachable or indicates congestion to the MME

- Maximum number of MM signaling requests are associated with the devices with the subscription for a particular APN
- Network management settings

Note: The MME should not apply congestion control for high priority access and emergency services. General NAS−level MM control can be used in order to reject NAS−level MM signaling requests under a general congestion condition.

APN−Based Session Management Congestion Control

The APN−based session management congestion control can be activated on the MME due to a congestion situation, by OAM, or by a restart/recovery of a PGW. The MME can reject ESM requests from the UE, which can be included in the PDN Connectivity, Bearer Resource Allocation, or Bearer Resource Modification requests. The MME can also deactivate the current PDN connection during congestion conditions and send a session back−off timer to the UE. When this timer is included, the *reactivation request* should not be activated.

The MME can store the Session Management (SM) back−off timer for a particular UE and APN during congestion and immediately reject any subsequent SM messages from the UE that is targeted to that APN until the timer runs out. This is required for the UEs that do not support the SM back−off timer (for UE releases prior to Release 10). The MME first clears this timer if it wants to send a SM message to the UE for which the timer already runs.

The UE can complete these actions while the timer runs:

- If the APN is provided in the rejected EPS SM request message, or if the SM back−off timer is received in the NAS deactivate EPS bearer context request message, the UE should not initiate any SM procedure for the congested APN.
- If an APN is not provided in the rejected EPS SM request message, then the UE shall not initiate any SM requests without the APN.
- These changes do not stop the back−off timer:
    - Cell
    - Tracking Area (TA)
    - Public Land Mobile Network (PLAMN)
    - Radio Access Technology (RAT)
- The UE is allowed to initiate the SM procedures for high−priority access and emergency services even when the SM back−off timer runs.
- If the UE receives a network−initiated EPS SM request message for the congested APN while the SM back−off timer runs, then the UE stops the SM back−off timer that is associated with this APN and responds to the MME.
- If the UE is configured with permission to override low access priority, and the SM back−off timer runs due to a rejection message that is received in response to a request with low access priority, the upper layers in the UE might request the initiation of SM procedures without low access priority.

- The UE is allowed to initiate the PDN disconnection procedure, but it does not delete the related SM back−off timer.
- The back−off timer does not stop the UE from data transmission or the initiation of the service requests for activation of the user plane bearer towards the congested APN.

APN−Based MM Congestion Control

Similar to the SM procedures, the MME also has a MM back−off timer and can reject the attach procedure. The MME should keep the subscriber data for some time after it rejects the attach procedure so that rejection for subsequent requests for the same subscriber can be completed without interaction with the HSS.

While the back−off timer runs, the UE should not initiate any NAS request for the MM procedure, except for high priority access or emergency services. However, the UE is allowed to perform Tracking Area Updates (TAUs) if it is already in *connected* mode.

The MME should select a back−off timer in such a way that all of the UEs should not have the same value of this timer, and the UEs should initiate deferred requests simultaneously. When the mobility back−off timer is received, the UE behavior is not APN−specific.

General NAS−Level Congestion Control

The general NAS−level congestion control is helpful in general overload conditions. It works similar to the APN−based congestion control and has a similar concept for the back−off timer. When the back−off timer runs, the UE can initiate detach requests, high priority requests, and TAUs (while in *connected* mode).

The back−off timer continues to run even after the UE is detached from the network. The MME should stop the back−off timer if the MME wants to page the UE for which the back−off timer already runs, and the UE should stop the back−off timer after it receives the paging request from the MME and initiate the service request.

The MM back−off timer does not affect the Cell/RAT and PLMN change. The TA change does not stop this timer. This timer is stopped when a new PLMN that is not equivalent to the PLMN is selected.

When the UE receives a handover command, it should proceed with the handover regardless of the back−off timer status.

If the MME rejects the TAU request or the service request with a MM back−off timer, which is larger than the sum of the UE periodic TAU timer plus the Implicit Detach timer, the MME should adjust the mobile reachable timer and/or Implicit Detach timer such that the MME does not implicitly detach the UE while the MM back−off timer runs.

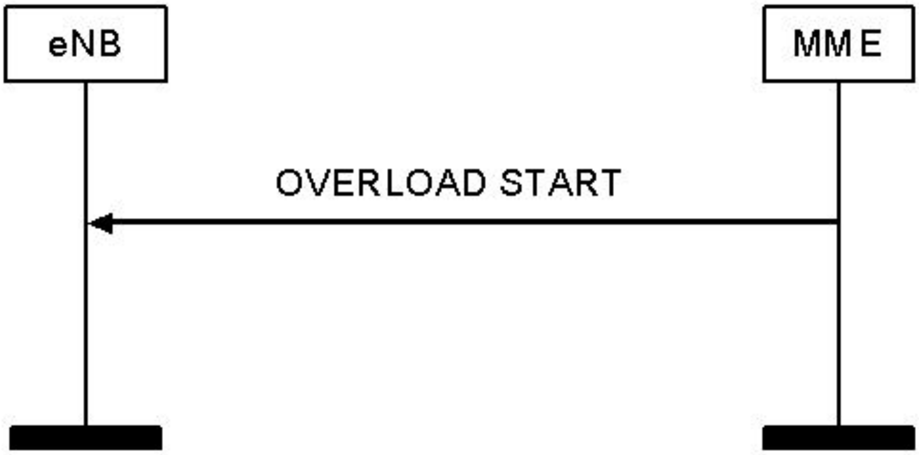
Note: The SGSN congestion control also works in similar manner as that of MME. Refer to 3GPP TS 23.060 for more details about the SGSN congestion control mechanism, and 3GPP TS 23.401 for more details about the MME congestion control mechanism.

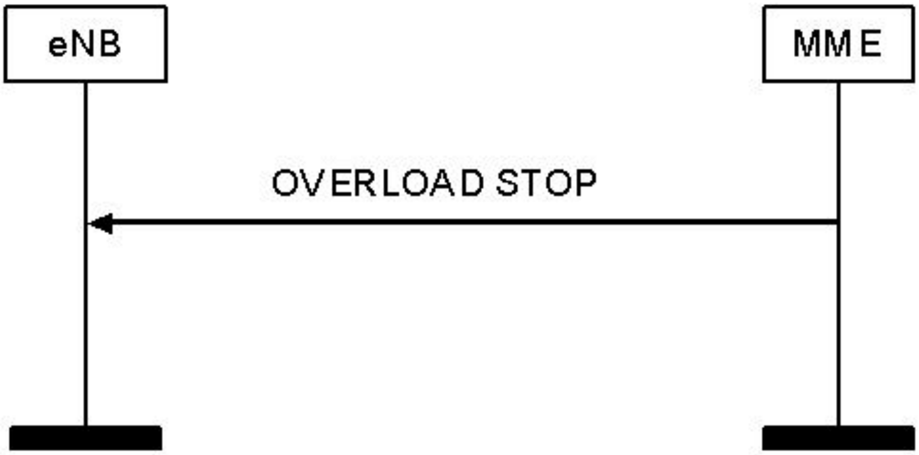
Overload Reduction by MME on S1−MME Interface

The MME can send an *Overload Start* message to the E−NodeB (eNB) in order to reduce the signaling load. This procedure uses non−UE associated signaling. The overload Action Information Element (IE) has an Overload Response IE within the Overload Start message, which contains information about rejection criteria, and the eNB takes action appropriately.

Tip: For more information, refer to the 3GPP Technical Specifications (TS) 36.413.

In order to indicate the end of the overload situation, the MME sends an Overload Stop message to the eNB:





Note: The SGSN also has a similar mechanism for signaling reduction, which is mentioned in 3GPP TS 25.413.

PGW Control of Overload

The PGW can reject a PDN connection during overload scenarios. The PGW can detect an overload condition and start or stop overload control based on criteria such as:

- The maximum number of active bearers per APN
- The maximum rate of bearer activations per APN

The PGW can specify a PGW back−off timer towards the MME for a specific APN, and the MME should reject the PDN connection requests for that APN during this time period. The MME can select another PGW instead of rejection during that time period, unless there is already a current PDN connection to the same APN for that UE.

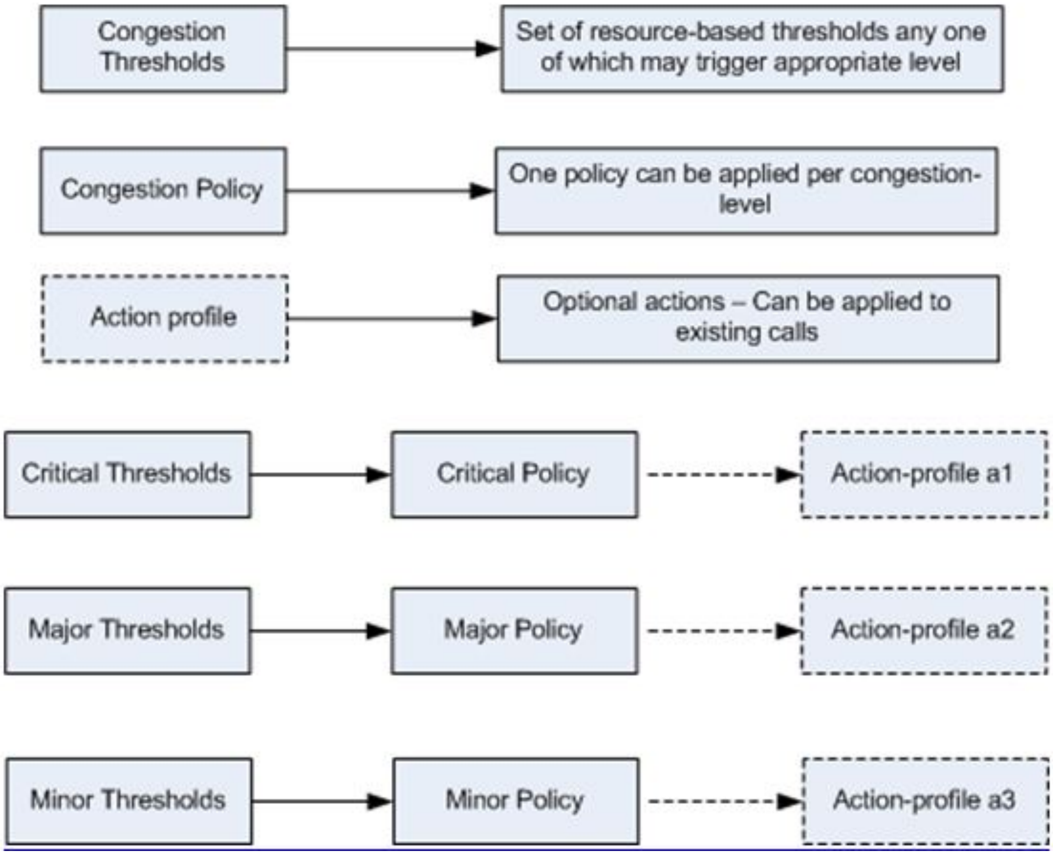
Note: The GGSN congestion control mechanism is similar to that on the PGW, which is mentioned in 3GPP TS 23.060. The PGW congestion control mechanism is mentioned in 3GPP TS 23.401.

Congestion Control Operation on the ASR 5x00

The congestion control operation is based on the configuration of these additional features:

- Call disconnect on overload
- Congestion control condition thresholds
- Service congestion policies

Here is an example:



Call Disconnect on Overload

This functionality allows the system to enable or disable the policy for disconnection of passive calls (chassis−wide) during an overload situation. It also allows you to fine−tune the overload disconnection congestion policy.

Congestion Condition Thresholds

Various congestion control thresholds can be defined, which dictate the conditions for which congestion control is to be enabled. It also establishes the limits for the definition of the system state that is congested or cleared. When these thresholds are reached, not only is a Simple Network Management Protocol (SNMP) trap (congestion) generated, but a congestion policy is also invoked.

A threshold tolerance is used in order to dictate the percentage under the configured threshold that must be reached before a condition is considered cleared and an SNMP trap (CongestionClear) is triggered.

Service Congestion Policies

The congestion service policies are configurable for each service, such as Packet Data Serving Node (PDSN), Gateway GPRS Support Node (GGSN), and Serving GPRS Support Node (SGSN). These policies dictate the manner in which the services respond when congestion is detected on the system due to a congestion threshold breach.

Configure

This section describes the configurations that are required in order to enable the congestion control and the basic tuning of congestion control.

Enable Congestion Control

The congestion control is disabled by default on the chassis. Enter the ***congestion−control*** command in *global configuration* mode in order to enable it:

```
[local]host_name(config)# congestion−control
```

Congestion Control Overload Disconnect

The congestion control overload disconnect enables or disables the policy for disconnection of the chassis−wide passive calls during an overload situation. This is disabled by default. It allows disconnection of the passive calls in stages and in iterations from the chassis until congestion control is cleared. The threshold for the *license−utilization* and *max−sessions−per−service−utilization*, along with the threshold value, can be configured.

For example, if the threshold is configured with at value of 90% and a tolerance of 5%, then the system stops the passive call disconnect when the number of calls drops below 85% of the total allowed calls for that service.

Here is the CLI syntax that can be used in order to enable the congestion control overload disconnect, which is always configured in *global configuration* mode:

```
congestion−control overload−disconnect
```

```
congestion−control overload−disconnect [ iterations−per−stage <integer> | percent
 <percentage_value> | threshold { license−utilization <percentage_value> |
 max−sessions−per−service−utilization <percentage_value> | tolerance <number> } ]
```

Here are some notes about this syntax:

- ***Iterations−per−stage***: This parameter defines the number of calls to be disconnected during the defined number of seconds. This value can range between two and eight.
- ***Percent***: This parameter specifies the percentage of calls to be disconnected in stages during an overload situation. This value can range between zero and one hundred, with five as the default value.
- ***Threshold***: This parameter defines the threshold values for the license and the maximum session utilization. It also allows for a definition of the tolerance value.
    - ♦ ***License−utilization***: This specifies the license utilization percentage threshold for overload situations. In case of a trigger, the passive calls are disconnected. This value ranges between one and one hundred, with 80 as the default value.

- ♦ ***Max−sessions−per−service−utilization***: This specifies the percentage of max sessions per service utilization threshold. Once it exceeds the defined value, the system begins to disconnect the passive calls. This value ranges between one and one hundred, with 80 as the default value.

- ♦ ***Tolerance***: This defines the percentage of calls that the system disconnects below the defined values set for *license−utilization* and *max−sessions−per−service−utilization*. This value ranges between one and 25, with ten as the default value. A clear trap message is only sent when the utilization falls below the defined tolerance values.

Congestion Control Policy Configuration

You can configure the congestion control policy on a per−service basis. The policy can cause the system to take actions such as drop, none, redirect, and reject on new sessions when any of the defined congestion control thresholds are exceeded, which activates congestion control.

This configuration allows a more granular definition of the congestion control policy for the MME and SGSN service and allows configuration of different stages of congestion control, such as critical, major, and minor (along with the association of action profiles).

Congestion Control Policy

Here is the congestion control policy configuration CLI syntax (except for MME services):

```
congestion−control policy { asngw−service | asnpc−service | cscf−service | fng−service
 | epdg−service | samog−service | ggsn−service | ha−service | hnbgw−service |
 hsgw−service | ipsg−service | lma−service | lns−service | mipv6ha−service |
 pcc−af−service | pcc−policy−service | pdg−service | pdif−service | pdsn−service |
 pdsnclosedrp−service | pgw−service | phsgw−service | phspc−service | saegw−service
 | sgsn−service | sgw−service | wsg−service } action { drop | none | redirect |
 reject }
```

Here are some notes about this syntax:

- ***Service type***: This parameter defines the service name for which the congestion control policy is being defined. The services that are applicable for this CLI command is specified in the previously mentioned CLI syntax.

- ***Action***: This parameter defines the action to be taken when the congestion control threshold is breached for the specified service. These four types of actions can be configured:

    - ♦ ***Drop***: This action causes the system to drop the new session requests. No rejection/failure response is sent.

    - ♦ ***Reject***: This action causes a rejection of the new session requests. A rejection response is sent. This option is not applicable to the IPSG service.

    - ♦ ***None***: This option is used when you want to configure the system so that no action is taken.

    - ♦ ***Redirect***: This action causes a redirection of the new session requests towards an alternate device. This is applicable only to the CSCF, HSGW, HA, and PDSN services. The IP address of the alternate device should be configured with the ***policy overload redirect*** command.

Policy Overload Redirect

This should be configured if a redirect action is configured for the Call Session Control Function (CSCF), HRPD Serving Gateway (HSGW), Home Agent (HA), or PDSN service.

- The CSCF service has this command configured under the CSCF policy rules configuration.
- The HSGW service, HA service, and PDSN service has this command configured under the respective service configurations.

Congestion Control Policy for MME Service

Prior to Release 14.0, the congestion control policy for the MME service can be defined similarly to the CLI syntax that is mentioned in the previous section, but with some additional options. Here is the CLI syntax:

```
congestion−control policy mme−service action { drop | none | reject | report−overload
 { permit−emergency−sessions | reject−new−sessions | reject−non−emergency−sessions }
 enodeb−percentage <percentage> }
```

In addition to the drop, none, and reject actions, the MME service also has the option to report overload conditions for the eNodeBs. The MME invokes the S1 overload procedure with the *S1AP Overload Start* message in order to report an overload condition to the specified proportion of eNodeBs to which the MME has an S1 interface connection. The MME selects the eNodeBs at random. Two overloaded MMEs in the same pool do not send overload messages to the same eNodeBs. When the MME has recovered and can increase its load, then it sends an *S1AP Overload Stop* message. In addition, these actions can be completed when a report overload action is configured:

- ***Permit−emergency−sessions***: This action allows only emergency sessions on the MME during an overload period.
- ***Reject−new−sessions***: This action causes a rejection of all new sessions inbound towards the MME during an overload situation.
- ***Reject−non−emergency−sessions***: This action causes all non−emergency sessions to be rejected on the MME during an overload period.
- ***Enodeb−percentage***: This action configures the percentage of known eNodeBs that receive the overload report. The percentage can range between one and one hundred.

In Releases 14.0 and later, the MME service can have three different policies and associated action profiles. Here is the CLI syntax:

```
congestion−control policy { critical mme−service action−profile <action_profile_name> |
 major mme−service  action−profile <action_profile_name> | minor mme−service
 action−profile <action_profile_name> }
```

There are three policy types that can be configured for the MME in Releases 14.0 and later:

- ***Critical***: This defines the critical congestion control threshold for the MME service.
- ***Major***: This defines the major congestion control threshold for the MME service.
- ***Minor***: This defines the minor congestion control threshold for the MME service.

Note: The ***action−profile*** parameter defines the action profile that is associated with the previously mentioned policy type (minor, major, or critical).

MME Congestion Control Policy Action Profile

The MME congestion control policy action profile is configurable under the *lte−policy*. Here is the CLI syntax:

```
configure > lte-policy

congestion-action-profile <profile_name>
```

The sections that follow describe the available actions that can be configured under the congestion action profile.

Drop

This action causes a drop of new session requests when the congestion control threshold is reached. Here is the CLI syntax:

```
drop { addn-brr-requests | addn-pdn-connects | brr-ctxt-mod-requests |
 combined-attaches | handovers | ps-attaches | s1-setups | service-request |
 tau-request } [ lapi ] [ apn-based ]
```

It allows more granular control in regards to the type of requests/call events that should be dropped. Here are the details:

- ***Addn−brr−request***: This drops packets that contain UE−initiated bearer resource requests. This is a licensed keyword.
- ***Addn−pdn−connect***: This drops packets that contain additional PDN context connections. This is a licensed keyword.
- ***Brr−ctxt−mod−requests***: This drops packets that contain bearer context modification requests. This is a licensed keyword.
- ***Combined−attaches***: This drops packets that contain combined attach requests.
- ***Handovers***: This drops packets that contain handover attempts.
- ***Ps−attaches***: This drops packets that contain packet−switched attach requests.
- ***S1−setups***: This drops packets that contain S1 setup attempts. This is a licensed keyword.
- ***Service−requests***: This drops packets that contain all service requests. This is a licensed keyword.
- ***Tau−requests***: This drops packets that contain all of the tracking area update requests.

These two options can also be configured with the previously mentioned call event type (both of these options are license−controlled):

- ***Lapi***: This indicates that requests with Low Access Priority Indication (LAPI) will be dropped for the call events; otherwise, both LAPI and non−LAPI events will be dropped. Here is the CLI syntax:

  ```
  drop <call-event> lapi
  ```

- ***Apn−based***: This indicates that requests for the Access Point Names (APNs) that are configured for the congestion control in the operator policy will be dropped. Here is the CLI syntax:

  ```
  drop <call-event> lapi
  ```

  Note: The ***apn network−identifier*** command in the operator policy is used in order to configure the congestion control for an APN.

Note: If the congestion action profile is configured with both the LAPI and APN−based options, then call events will be dropped only if both conditions are matched.

Exclude Emergency Events

This allows the emergency requests to be processed even when the threshold has been exceeded. Here is the CLI syntax:

```
exclude-emergency-events
```

When this is configured, the congestion action rejects and drops are not applied for these messages in emergency−attached UEs:

- TAU requests
- Service requests
- Handovers
- ADDN−PDN requests

Exclude Voice Events

This allows voice calls to be processed even when the threshold has been exceeded. Here is the CLI syntax:

```
exclude-voice-events
```

None

This specifies that no congestion control action should be taken for inbound requests when the congestion control threshold has been reached. Here is the CLI syntax:

```
none { addn-brr-requests | addn-pdn-connects | combined-attaches | handovers |
 psattaches | s1-setups | service-request | tau-request }
```

Here are the details of the call events that can be configured for this action (*none* is the default action for all of these call events):

- ***Addn−brr−request***: This causes no congestion control action to be completed for packets that contain UE−initiated bearer resource requests.
- ***Addn−pdn−connect***: This causes no congestion control action to be completed for additional Packet Data Network (PDN) context connections.
- ***Brr−ctxt−mod−requests***: This causes no congestion control action to be completed for packets that contain bearer context modification requests.

- ***Combined−attaches***: This causes no congestion control action to be completed for packets that contain combined attach requests.

- ***Handovers***: This causes no congestion control action to be completed for packets that contain handover attempts.

- ***Ps−attaches***: This causes no congestion control action to be completed for packets that contain packet−switched attach requests.

- ***S1−setups***: This causes no congestion control action to be completed for packets that contain S1 setup attempts. This is a licensed keyword.

- ***Service−requests***: This causes no congestion control action to be completed for packets that contain all of the service requests. This is a licensed keyword.

- ***Tau−requests***: This causes no congestion control action to be completed for packets that contain all of the tracking area update requests.

Reject

This causes the inbound requests to be rejected and a *reject message* response to be sent when the congestion control threshold has been reached. Here is the CLI syntax:

```
reject { addn−brr−requests | addn−pdn−connects | brr−ctxt−mod−requests |
 combined−attaches | handovers | ps−attaches | s1−setups time−to−wait
 { 1 | 10 | 2 | 20 | 50 | 60 } | service−request | tau−request }[ lapi ]
 [ apn−based ]
```

Here are the details of the call events that can be configured with the *reject* action:

- ***Addn−brr−request***: This rejects packets that contain UE−initiated bearer resource requests. This is a licensed keyword.

- ***Addn−pdn−connect***: This rejects packets that contain additional PDN context connections. This is a licensed keyword.

- ***Brr−ctxt−mod−requests***: This rejects packets that contain bearer context modification requests. This is a licensed keyword.

- ***Combined−attaches***: This rejects packets that contain combined attach requests.

- ***Handovers***: This rejects packets that contain handover attempts.

- ***Ps−attaches***: This rejects packets that contain packet−switched attach requests.

- ***S1−setups time−to−wait { 1 | 10 | 2 | 20 | 50 | 60 }***: This rejects packets that contain S1 setup attempts after 1, 2, 10, 20, 50, or 60 seconds. This is a licensed keyword.

- ***Service−requests***: This rejects packets that contain all of the service requests. This is a licensed keyword.

- ***Tau−requests***: This rejects packets that contain all of the tracking area update requests.

These two options can also be configured with the previously mentioned call event type (both of these options are license−controlled):

- ***Lapi***: This indicates that requests with LAPI will be rejected for the call events; otherwise, both LAPI and non−LAPI events will be rejected. Here is the CLI syntax:

  ```
  reject <call-event> lapi
  ```
- ***Apn−based***: This indicates that requests for the APNs that are configured for the congestion control in operator policy will be rejected. Here is the CLI syntax:

  ```
  reject <call-event> lapi
  ```

  Note: The ***apn network−identifier*** command in the operator policy is used in order to configure the congestion control for an APN.

Note: If the congestion action profile is configured with both the LAPI and APN−based options, then the call events are rejected only if both conditions are matched.

Report Overload

This enables the MME to report overload conditions to the eNodeBs in order to alleviate congestion scenarios. The MME invokes the S1 overload procedure with the *S1AP Overload Start* message in order to report the overload condition to the specified proportion of eNodeBs to which the MME has an S1−interface connection.

The MME selects the eNodeBs at random. Two overloaded MMEs in the same pool do not send overload messages to the same eNodeBs. When the MME has recovered and can increase its load, it sends an *S1AP overload Stop* message. Here is the CLI syntax:

```
report-overload { permit-emergency-sessions-and-mobile-terminated-services |
 permit-highpriority-sessions-and-mobile-terminated-services |
 reject-delay-tolerant-access | reject-new-sessions |
 reject-non-emergency-sessions } enodeb-percentage <percent>
```

These are the options that can be configured with this action:

- ***permit−emergency−sessions−and−mobile−terminated−services***: This specifies in the overload message to the eNodeB that only emergency sessions are allowed to access the MME during the overload period.

- ***permit−high−priority−sessions−and−mobile−terminated−services***: This specifies in the overload message to the eNodeB that only high priority sessions and mobile−terminated services are allowed to access the MME during the overload period.

- ***reject−delay−tolerant−access***: This specifies in the overload message to the eNodeB that delay−tolerant access destined for the MME should be rejected during the overload period.

- ***reject−new−sessions***: This specifies in the overload message to the eNodeB that all new connection requests destined for the MME should be rejected during the overload period.

- ***reject−non−emergency−sessions***: This specifies in the overload message to the eNodeB that all non−emergency sessions should be rejected during the overload period.

- ***enobeb−percentage***: This configures the percentage of known eNodeBs that will receive overload report.

Congestion Control Policy for SGSN with Releases 17.0 and Later

In Releases 17.0 and later, the SGSN also required a congestion control policy similar to that of the MME. The SGSN can have three congestion control actions, and each action is associated with an action profile. Here is the CLI syntax:

```
congestion−control policy { critical | major | minor }
 sgsn−service action−profile <action_profile_name>
```

These three *policy types* can be configured for the MME in Releases 14.0 and later:

- ***Critical***: This defines the critical congestion control threshold for the MME service.
- ***Major***: This defines the major congestion control threshold for the MME service.
- ***Minor***: This defines the minor congestion control threshold for the MME service.

Note: The ***action−profile*** parameter defines the action profile that is associated with the *policy type* (minor, major, or critical).

SGSN Congestion Control Policy Action Profile

The SGSN congestion control policy action profile is configured in *sgsn−global* configuration mode. It defines the action to be completed for these types of call/message events when any congestion control threshold has been reached in the SGSN node:

- Active calls
- New calls
- SM messages

Here is the syntax for the configuration of the SGSN congestion control policy action profile:

```
configure > sgsn−global > congestion−control
```

```
congestion−action−profile <action_profile_name>
```

The sections that follow describe the various policies that can be configured under the SGSN congestion action profile.

Active Call Policy

This specifies the drop or reject of any active call messages when congestion occurs during an active call. A drop or reject of active calls can only be defined as LAPI for the message. Here is the CLI syntax:

```
active−call−policy { rau | service−req } { drop | reject } [ low−priority−ind−ue ]
```

Here are some notes about this syntax:

- ***Message Type/call Event***: These message types or call events can be defined for an active call policy:
    - ***RAU***: This defines the Routing Area Update (RAU) message that is received by the SGSN.

    - ***Service−req***: This defines the SR message that is received by the SGSN.

- ***Actions***: This defines the actions to be taken when the SGSN receives the previously mentioned messages during the active calls when the congestion control threshold has been reached.

    - ***Drop***: This instructs the SGSN to drop the defined message when the congestion control threshold has been reached.

    - ***Reject***: This instructs the SGSN to reject the defined message when the congestion control threshold has been reached.

  Note: Drop and reject actions can further be refined for LAPI. The ***low−priority−ind−ue*** keyword is used with a drop/reject action.
- ***low−priority−ind−ue:*** This instructs the SGSN to reject/drop the defined message, only if a message from the UE includes a LAPI, when the congestion control threshold has been reached.

New Call Policy

This specifies the drop or rejection of any new call messages when congestion occurs. The drop or reject actions for new calls (attach request or new inter−SGSN RAU) can be refined to LAPI or APN−based, or both. Here is the CLI syntax:

```
new−call−policy { drop | reject } [ apn−based ] [ low−priority−ind−ue ]
```

Here are some notes about this syntax:

- ***Message Type/call Event***: When a new call policy is defined, it is taken for all of the *attach requests* or *Inter−SGSN RAUs*. For this reason, no message/call event type is required in this CLI command.

- ***Actions***: This defines the actions to be completed when the SGSN receives the previously mentioned messages during the active calls when the congestion control threshold has been reached.

    - ***Drop***: This instructs the SGSN to drop the new call messages when the congestion control threshold has been reached.

    - ***Reject***: This instructs the SGSN to reject the new call messages when the congestion control threshold has been reached.

  Note: The drop and reject actions can further be refined for LAPI and APN−based. The ***low−priority−ind−ue*** and ***apn−based*** keywords are used with the drop/reject actions.
- ***low−priority−ind−ue***: This instructs the SGSN to reject/drop the defined message, only if a message from the UE includes a LAPI, when the congestion control threshold has been reached.

- ***apn−based***: This instructs the SGSN to reject/drop the new call messages based on the APN if the congestion control threshold has been reached. This only occurs if an APN is configured under the operator policy with congestion control.

  Note: If the congestion action profile is configured with both the LAPI and APN−based options, then new call events will be rejected only if both conditions are matched.

SM Messages

This defines the policy for the SM messages, such as *active* or *modification* requests. The response from the SGSN can only be *reject*, and this can be refined to LAPI or APN−based, or both. Here is the CLI syntax:

```
sm−messages reject [ apn−based] [ low−priority−ind−ue ]
```

Here are some notes about this syntax:

- ***Message Type/call Event***: When the SM messages policy is defined, it is applied to all of the *activate* or *modification* requests. For this reason, the message/call event type is required in this CLI command.

- ***Actions***: This defines the actions to be completed when the SGSN receives the previously mentioned message and the congestion control threshold has been reached. The *reject* action instructs the SGSN to reject the SM messages when the congestion control threshold has been reached.

  Note: The reject actions can further be refined for LAPI and APN−based. The ***low−priority−ind−ue*** and ***apn−based*** keywords are used with the drop/reject actions.
- ***low−priority−ind−ue***: This instructs the SGSN to reject the SM message only if the message from the UE includes a LAPI when the congestion control threshold has been reached.

- ***apn−based***: This instructs the SGSN to reject the SM messages based on the APN if the congestion control threshold has been reached. This only occurs if the APN is configured under the operator policy with congestion control.

  Note: If the congestion action profile is configured with both the LAPI and APN−based options, then the new call events are rejected only if both of the conditions are matched.

Congestion Control Threshold

The congestion control threshold defines the threshold values for the various parameters that can invoke congestion control when the threshold is exceeded. Here is the CLI syntax:

```
congestion-control threshold { license-utilization percent |
 max-sessions-per-service-utilization <percent> |  message-queue-utilization <percent>
 | message-queue-wait-time <time> | port-rx-utilization <percent> | port-specific
 { <slot/port> | all } [ tx-utilization <percent> ] [ rx-utilization <percent> ]
 port-specific-rx-utilization critical | port-specific-tx-utilization critical |
 port-tx-utilization <percent> | service-control-cpu-utilization

<percent> | system-cpu-utilization <percent> | system-memory-utilization <percent>
 | tolerance <percent> }
```

Here are the different parameters that can be configured with threshold values and can trigger congestion control when the threshold has been reached:

- ***License−utilization***: This parameter defines the percent utilization of the licensed capacity, as measured in ten−second intervals. This value is formatted as a percentage and can range between zero and one hundred (the default value is one hundred).

- ***max−sessions−per−service−utilization***: This parameter defines the percent utilization of the maximum sessions allowed per service, as measured in real−time. This threshold is based on the maximum number of sessions, or the PDP context that is configured for a particular service. This value ranges between zero and one hundred, with a default value of 80.

- ***message−queue−utilization***: This parameter defines the percent utilization of the DEMUX manager software task message queue, as measured in ten−second intervals. This queue has the capability to store 10,000 messages. This value ranges between zero and one hundred, with a default value of 80.

- ***message−queue−wait−time***: This parameter defines the maximum time (in seconds) that a message can remain in the queue, as measured by the packet time stamps. This value ranges between one and

30 seconds, with a default value of five seconds.

- ***port−rx−utilization***: This parameter defines the average percent utilization of the port resources for all of the ports, by received data, as measured in five−minute intervals. This value ranges between zero and one hundred, with a default value of 80. This threshold parameter can be disabled with the ***no*** command.
- ***port−specific***: This parameter defines the port−specific thresholds. When any individual port−specific threshold is reached, the congestion control is applied system−wide. This is disabled by default for each particular port number or for all of the ports for which the ***all*** keyword can be used. This parameter has two sub−options that can be defined:
    - ***rx−utilization***: The default value for this option is 80%. It measures the average percent utilization of port resources for the specific port, by received data, as measured in five−minute intervals. The values range between zero and one hundred.
    - ***tx−utilization***: The default value for this option is 80%. It measures the average percent utilization of port resources for the specific port, by transmitted data, as measured in five−minute intervals. The value ranges between one and one hundred.
- ***port−tx−utilization***: This parameter defines the average percent utilization of the port resources for all of the ports, by transmitted data, as measured in five−minute intervals. This value ranges between zero and one hundred, with a default value of 80. This threshold parameter can be disabled via the ***no*** version of this command.
- ***service−control−cpu−utilization***: This parameter defines the average percent utilization of CPUs on which a DEMUX manager software task instance runs, as measured in ten−second intervals. This value ranges between zero and one hundred, with a default value of 80.
- ***system−cpu−utilization***: This parameter defines the average percent utilization for all PSC/PSC2 CPUs that are available to the system, as measured in ten−second intervals. This value ranges between zero and one hundred, with a default value of 80. This can be disabled with ***no congestion−control threshold system−cpu−utilization*** CLI command.
- ***system−memory−utilization***: This parameter defines the average percent utilization for all of the CPU memory that is available to the system, as measured in ten−second intervals. This value ranges between zero and one hundred, with a default value of 80.
- ***Tolerance***: This parameter defines the percentage under a configured threshold that dictates the point at which the condition is cleared. This value ranges between zero and one hundred, with a default value of ten. For example, if the threshold is configured with a value of 90 and the congestion control is triggered, then the trigger is cleared at 80 if the default value of ten for the tolerance is defined.

Congestion Control Threshold Values for MME and SGSN

This section defines the configuration of the threshold for the MME and the SGSN when three different triggers, along with congestion control profiles, are defined.

This information is applicable to MME Releases 14.0 and later, and SGSN Releases 17.0 and later. These are the three different levels of triggers that are available for the MME and SGSN, which are further associated with the congestion control policies that correspond:

- ***Critical***: This trigger level defines the critical threshold values for different parameters. The value of this trigger level should be the largest among all three levels of thresholds. The critical thresholds include pre−configured default values.

- ***Major***: This trigger level defines the major threshold values for different triggers. The values of this trigger level should be greater than the minor threshold and less than the critical. The default value is zero.

- ***Minor***: This trigger level defines the minor threshold values for different triggers. The values of this trigger should be least among all three thresholds. The default value is zero.

The three threshold values can be defined for all of the parameters/triggers that are mentioned in the previous section. Here is the CLI syntax that is used in order to define the thresholds for the different parameters:

```
congestion−control threshold license−utilization { critical <percent> | major
 <percent>t | minor <percent> }
```

```
congestion−control threshold max−sessions−per−service−utilization { critical
 <percent> | major <percent> | minor <percent> }
```

```
congestion−control threshold message−queue−utilization { critical <percent> |
 major <percent> | minor <percent> }
```

```
congestion−control threshold message−queue−wait−time { critical <time> |
 major <time> | minor <time> }
```

```
congestion−control threshold port−rx−utilization { critical <percent> | major
 <percent> | minor <percent> }
```

```
congestion−control threshold port−specific { <slot/port> [ tx−utilization {
 critical <percent> | major <percent> | minor <percent> ] [ rx−utilization {
 critical <percent> | major <percent> | minor <percent> } | all { critical
 <percent> | major <percent> | minor <percent> } }
```

```
congestion−control threshold port−tx−utilization { critical <percent> | major
 <percent> | minor <percent> }
```

```
congestion−control threshold service−control−cpu−utilization { critical
 <percent> | major <percent> | minor <percent >}
```

```
congestion−control threshold system−cpu−utilization { critical <percent> |
 major <percent> | minor <percent> }
```

```
congestion−control threshold system−memory−utilization { critical <percent> |
 major <percent> | minor <percent> }
```

```
congestion−control threshold tolerance { critical <percent> | major
 <percent> | minor <percent> }
```

Note: The critical threshold values for the different parameters (except the ***license−utilization***) use default values that are the same as those that are described in the previous section. The ***license−utilization*** parameter has a default value for the critical profile as *80%*.

Verify

Use the information that is described in this section in order to verify your congestion control configuration.

Congestion Control Configuration Verification

Enter the ***show congestion−control configuration | more*** CLI command in order to verify the configuration of the congestion control. The sections that follow provide example command outputs for the various stages of congestion control.

Congestion Control Before Activation

```
[local]st40−sim# show congestion−control configuration | more
Congestion−control: disabled
...................
```

Congestion Control After Activation

```
[local]st40−sim# configure
[local]st40−sim(config)# congestion−control
[local]st40−sim(config)# end
[local]st40−sim# show congestion−control configuration | more
Congestion−control: enabled
............
```

Congestion Control After Overload Disconnect Activation

```
[local]st40−sim# configure
[local]st40−sim(config)# congestion−control overload−disconnect
[local]st40−sim(config)# end
[local]st40−sim# show congestion−control configuration | more
Congestion−control: enabled
............

Overload−disconnect: enabled

Overload−disconnect threshold parameters
  license utilization:                  80%
  max−session−per−service utilization:  80%
  tolerance:                            10%
  session disconnect percent:           5%
  iterations−per−stage:                  8

 ............
```

Congestion Control After Activation of Policies Other Than SGSN and MME

The configuration of the ***congestion−control policy <service−name> action <action>*** parameter changes the value of the ***congestion control policy*** section, as per the configuration. Here is one example configuration of an ***action drop*** for the ***ggsn−service***:

```
[local]st40−sim(config)# congestion−control policy ggsn−service action drop
[local]st40−sim(config)# end
[local]st40−sim# show congestion−control configuration | more
Congestion−control: enabled
............

Congestion−control Policy
  pdsn−service: none
  hsgw−service: none
  ha−service:   none
  ggsn−service: drop
  closedrp−service: none
.............
```

Congestion Control Threshold for Major and Minor Profiles

This section describes the congestion control threshold configuration verification for the major and minor profiles. The critical profile already has some default values, which can be changed as required, but the major and minor thresholds are required to be configured. These three profiles can later be used along with a congestion control policy.

```
[local]st40-sim# configure
[local]st40-sim(config)# congestion-control threshold license-utilization major 70
[local]st40-sim(config)# congestion-control threshold license-utilization minor 60
[local]st40-sim(config)# congestion-control threshold
 max-sessions-per-service-utilization major 70
[local]st40-sim(config)# congestion-control threshold
 max-sessions-per-service-utilization minor 60
[local]st40-sim(config)# congestion-control threshold mes
message-queue-utilization     message-queue-wait-time
[local]st40-sim(config)# congestion-control threshold
 message-queue-utilization major 70
[local]st40-sim(config)# congestion-control threshold
 message-queue-utilization minor 60
[local]st40-sim(config)# congestion-control threshold message-queue-wait-time major 4
[local]st40-sim(config)# congestion-control threshold message-queue-wait-time minor 3
[local]st40-sim(config)# congestion-control threshold port-rx-utilization major 70
[local]st40-sim(config)# congestion-control threshold port-rx-utilization minor 60
[local]st40-sim(config)# congestion-control threshold port-tx-utilization major 70
[local]st40-sim(config)# congestion-control threshold port-tx-utilization minor 60
[local]st40-sim(config)# congestion-control threshold
 service-control-cpu-utilization major 70
[local]st40-sim(config)# congestion-control threshold
 service-control-cpu-utilization minor 60
[local]st40-sim(config)# congestion-control threshold syst
system-cpu-utilization        system-memory-utilization
[local]st40-sim(config)# congestion-control threshold system-cpu-utilization major 70
[local]st40-sim(config)# congestion-control threshold system-cpu-utilization minor 60
[local]st40-sim(config)# congestion-control threshold
 system-memory-utilization major 70
[local]st40-sim(config)# congestion-control threshold
 system-memory-utilization minor 60
[local]st40-sim(config)# congestion-control threshold tolerance major 5
[local]st40-sim(config)# congestion-control threshold tolerance minor 2
[local]st40-sim(config)# end
[local]st40-sim# show congestion-control configuration | more
Congestion-control: enabled

Congestion-control Critical threshold parameters
  system cpu utilization:                80%
  service control cpu utilization:       80%
  system memory utilization:             80%
  message queue utilization:             80%
  message queue wait time:               5 seconds
  port rx utilization:                   80%
  port tx utilization:                   80%
  license utilization:                   100%
  max-session-per-service utilization:   80%
  tolerance limit:                       10%

Congestion-control Major threshold parameters
  system cpu utilization:                70%
  service control cpu utilization:       70%
  system memory utilization:             70%
  message queue utilization:             70%
  message queue wait time:               4 seconds
  port rx utilization:                   70%
  port tx utilization:                   70%
  license utilization:                   70%
  max-session-per-service utilization:   70%
  tolerance limit:                       5%

Congestion-control Minor threshold parameters
  system cpu utilization:                60%
  service control cpu utilization:       60%
  system memory utilization:             60%
  message queue utilization:             60%
  message queue wait time:               3 seconds
  port rx utilization:                   60%
  port tx utilization:                   60%
  license utilization:                   60%
  max-session-per-service utilization:   60%
  tolerance limit:                       2%

Overload-disconnect: enabled

Overload-disconnect threshold parameters
  license utilization:                   80%
  max-session-per-service utilization:   80%
  tolerance:                             10%
  session disconnect percent:            5%
  iterations-per-stage:                   8
............
```

Congestion Control Policy Activation for SGSN

Use this information in order to verify the congestion control policy activation for the SGSN:

```
[local]st40-sim# configure
[local]st40-sim(config)# sgsn-global
[local]st40-sim(config-sgsn-global)# congestion-control
[local]st40-sim(config-congestion-ctrl)# end
[local]st40-sim# configure
[local]st40-sim(config)# congestion-control
[local]st40-sim(config)# end
[local]st40-sim# configure
[local]st40-sim(config)# sgsn-global
[local]st40-sim(config-sgsn-global)# congestion-control
[local]st40-sim(config-congestion-ctrl)# congestion-action-profile sgsn_critical
[local]st40-sim(config-cong-act-prof-sgsn_critical)# active-call-policy rau reject
[local]st40-sim(config-cong-act-prof-sgsn_critical)# active-call-policy
 service-req reject
[local]st40-sim(config-cong-act-prof-sgsn_critical)# new-call-policy reject
[local]st40-sim(config-cong-act-prof-sgsn_critical)# sm-messages reject
[local]st40-sim(config-cong-act-prof-sgsn_critical)# exit
[local]st40-sim(config-congestion-ctrl)# congestion-action-profile sgsn_major
[local]st40-sim(config-cong-act-prof-sgsn_major)# active-call-policy rau drop
[local]st40-sim(config-cong-act-prof-sgsn_major)# active-call-policy
 service-req drop
[local]st40-sim(config-cong-act-prof-sgsn_major)# new-call-policy drop
[local]st40-sim(config-cong-act-prof-sgsn_major)# sm-messages reject
 low-priority-ind-ue
[local]st40-sim(config-cong-act-prof-sgsn_major)# exit
[local]st40-sim(config-congestion-ctrl)# congestion-action-profile sgsn_minor
[local]st40-sim(config-cong-act-prof-sgsn_minor)# exit
[local]st40-sim(config-congestion-ctrl)# exit
[local]st40-sim(config-sgsn-global)# exit
[local]st40-sim(config)# congestion-control policy critical sgsn-service
 action-profile sgsn_critical
[local]st40-sim(config)# congestion-control policy major sgsn-service
 action-profile sgsn_major
[local]st40-sim(config)# congestion-control policy minor sgsn-service
 action-profile sgsn_minor
[local]st40-sim(config)#end

[local]st40-sim# show congestion-control configuration | more
Congestion-control: enabled
............
```

```
  pdsn-service: none
  hsgw-service: none
  ha-service:   none
  ggsn-service: drop
  closedrp-service: none
  lns-service: none
  cscf-service: reject
  pdif-service: none
  wsg-service: none
  pdg-service: none
  epdg-service: none
  fng-service: none
  sgsn-service:
    Critical Action-profile : sgsn_critical
    Major    Action-profile : sgsn_major
    Minor    Action-profile : sgsn_minor
...........
```

Congestion Control Policy Activation for MME

Use this information in order to verify the congestion control policy activation for the MME:

```
[local]st40-sim# configure
[local]st40-sim(config)# lte-policy
[local]st40-sim(lte-policy)# congestion-action-profile mme_critical
Are you sure? [Yes|No]: yes
[local]st40-sim(congestion-action-profile)# drop addn-brr-requests
[local]st40-sim(congestion-action-profile)# drop s1-setups
[local]st40-sim(congestion-action-profile)# exit
[local]st40-sim(lte-policy)# congestion-action-profile mme_major
Are you sure? [Yes|No]: yes
[local]st40-sim(congestion-action-profile)# reject addn-brr-requests
[local]st40-sim(congestion-action-profile)# reject s1-setups time-to-wait 20
[local]st40-sim(congestion-action-profile)# exit
[local]st40-sim(lte-policy)# congestion-action-profile mme_minor
Are you sure? [Yes|No]: yes
[local]st40-sim(congestion-action-profile)# none addn-brr-requests
[local]st40-sim(congestion-action-profile)# none s1-setups
[local]st40-sim(congestion-action-profile)# exit
[local]st40-sim(lte-policy)# exit
[local]st40-sim(config)# congestion-control policy critical mme-service
 action-profile mme_critical
[local]st40-sim(config)# congestion-control policy major mme-service
 action-profile mme_major
[local]st40-sim(config)# congestion-control policy minor mme-service
 action-profile mme_minor
[local]st40-sim(config)# end

[local]st40-sim# show congestion-control configuration | more
Congestion-control: enabled
............

  pdsn-service: none
  hsgw-service: none
  ha-service:   none
  ggsn-service: drop
  closedrp-service: none
  lns-service: none
  cscf-service: reject
  pdif-service: none
  wsg-service: none
  pdg-service: none
  epdg-service: none
  fng-service: none
  sgsn-service:
    Critical Action−profile : sgsn_critical
    Major    Action−profile : sgsn_major
    Minor    Action−profile : sgsn_minor
  mme−service:
    Critical Action−profile : mme_critical
    Major    Action−profile : mme_major
    Minor    Action−profile : mme_minor
...........
```

Congestion Control Statistics

These commands are used in order to view the statistics and statuses that are related to congestion control:

```
show congestion−control { configuration | statistics { <manager> [ all | instance
 <task_instance> ] } [ | { grep <grep_options> | more } ]

show congestion−control statistics mme { critical | full | major | minor } [ | {
 grep <grep_options> | more } ]
```

The ***<manager>*** option can have these values:

- ***A11mgr***: This is the PDSN service.
- ***asngwmgr***: This is the Access Service Network Gateway (ASN−GW) service.
- ***asnpcmgr***: This is the ASN Paging Control (PC−LR) service.
- ***bindmux***: This is the Bindmux Manager that is used by the PCC service.
- ***egtpinmgr***: This is the Enhanced GPRS Tunneling Protocol (EGTP) ingress DEMUX manager.
- ***gtpcmgr***: This is the GGSN service.
- ***hamgr***: This is for the HA services.
- ***hnbmgr***: This is the Home Node B (HNB) Manager that is used by the HNB−GW service.
- ***imsimgr***: This is the IMSI manager, which is used for the SGSN.
- ***ipsecmgr***: This is the IP Security (IPSec) manager.
- ***ipsgmgr***: This is for the IP Service Gateway (IPSG) managers.
- ***l2tpmgr***: This is for the Layer 2 (L2) Tunneling Protocol (L2TP) managers.

Congestion Control Trigger for SGSN by OAM Intervention

The ***sgsn trigger−congestion level { critical | major | minor }*** command is used in order to manually trigger congestion control in the SGSN. The ***sgsn clear−congestion*** command is used in order to clear the congestion that is initiated by the ***sgsn trigger−congestion*** command.

Here is an example output:

```
[local]st40−sim# sgsn trigger−congestion level critical
[local]st40−sim# show congestion−control statistics imsimgr all full | more
 Current congestion status:                              Cleared
 Current congestion Type  :                              None
 Congestion applied:                                     0 times
 Critical Congestion Control Resource Limits
  system cpu use exceeded:                               No
  service cpu use exceeded:                              No
  system memory use exceeded:                            No
  port rx use exceeded:                                  No
  port tx use exceeded:                                  No
  port specific rx use exceeded:                         No
  port specific tx use exceeded:                         No
  max sess use exceeded:                                 No
  license use exceeded:                                  No
  msg queue size use exceeded:                           No
  msg queue wait time exceeded:                          No
  license threshold exceeded:                            No
  max sess threshold exceeded:                           No
  Sessions disconnected due to overload disconnect:      0

 Major Congestion Control Resource Limits
  system cpu use exceeded:                               No
  service cpu use exceeded:                              No
  system memory use exceeded:                            No
  port rx use exceeded:                                  No
  port tx use exceeded:                                  No
  port specific rx use exceeded:                         No
  port specific tx use exceeded:                         No
  max sess use exceeded:                                 No
  license use exceeded:                                  No
  msg queue size use exceeded:                           No
  msg queue wait time exceeded:                          No

 Minor Congestion Control Resource Limits
  system cpu use exceeded:                               No
  service cpu use exceeded:                              No
  system memory use exceeded:                            No
  port rx use exceeded:                                  No
  port tx use exceeded:                                  No
  port specific rx use exceeded:                         No
  port specific tx use exceeded:                         No
  max sess use exceeded:                                 No
  license use exceeded:                                  No
  msg queue size use exceeded:                           No
  msg queue wait time exceeded:                          No
SGSN Congestion Control:
   MM Congestion Level:                                  Critical
   Congestion Resource:                                  None
   SM Congestion Level:                                  Critical
   O&M Congestion Level:                                 Critical
```

Troubleshoot

There is currently no specific troubleshooting information available for this configuration.

Related Information

- ***3GPP TS 23.401***
- ***3GPP TS 23.060***
- ***3GPP TS 25.413***
- ***3GPP TS 36.413***

- ***Command Line Interface Reference, StarOS Release 17***
- ***Technical Support & Documentation − Cisco Systems***

Updated: Jul 23, 2015 Document ID: 119151