

Enhanced Congestion Control and Overload Control

- Feature Description, on page 1
- Configuring Enhanced Congestion Control, on page 2
- Monitoring and Troubleshooting, on page 8

Feature Description

The MME provides an enhanced congestion control and overload control feature set.

This feature builds on the base congestion control functionality provided on the MME. Refer to the *Congestion Control* and *Overload Control* sections in the *MME Overview* chapter for more information about the basic functionality.

To use this feature, you need a valid license key (MME Resiliency) installed. Contact your Cisco Account or Support representative for information on how to obtain a license.

Enhanced Congestion Control and Overload Control

To allow greater control during overload conditions, the MME supports the configuration of three separate levels (critical, major, minor) of congestion thresholds for the following system resources:

- System CPU usage
- System service CPU usage (Demux-Card CPU usage)
- System Memory usage
- License usage
- Maximum Session per service

The MME can, in turn, be configured to take specific actions when any of these thresholds are crossed, such as:

- Drop or reject the following S1-AP/NAS messages: S1 Setup, Handover events, TAU request, Service request, PS-Attach request, Combined-attach request, Additional PDN request, or UE initiated bearer resource allocation.
- Allow voice or emergency calls/events.
- Initiate S1AP overload start to a percentage of eNodeBs with options to signal any of the following in the Overload Response IE:

- reject non-emergency sessions
- reject new sessions
- permit emergency sessions
- permit high-priority sessions and mobile-terminated services
- reject delay-tolerant access.

Relationships to Other Features

This license-enabled feature builds on the base congestion control functionality provided on the MME.

Refer to the *Congestion Control* and *Overload Control* sections in the *MME Overview* chapter for more information about the basic functionality.

Additional information is also provided in the *Congestion Control* chapter in the *System Administration Guide*.

Limitations

The base congestion control functionality also can monitor congestion of the following resources:

- Port-specific RX and TX utilization
- Port RX and TX utilization
- Message queue utilization
- Message queue wait time

The license-enabled Enhanced Congestion Control funtionality on the MME does not support the monitoring of these resources using three different threshold levels (critical, major and minor). Only a single threshold level (critical) can be monitored for these resources.

Configuring Enhanced Congestion Control

This feature requires that a valid license key be installed. Contact your Cisco Account or Support representative for information on how to obtain a license.

Configuring Enhanced Congestion Control

This section includes configuration procedures for the following:

- Configuring Thresholds and Tolerances
- License Utilization Thresholds
- Maximum Session Per Service Thresholds
- Service Control CPU Thresholds
- System CPU Thresholds
- System Memory Thresholds
- Configuring a Congestion Action Profile
- Associating a Congestion Action Profile with Congestion Control Policies
- Configuring Overload Control
- Configuring Congestion SNMP Traps

Configuring Thresholds and Tolerances

Congestion threshold values must be defined to establish when a congestion condition is reached. Congestion threshold tolerances must also be configured to establish when a congestion condition is cleared. Individual thresholds values and tolerances can be defined for *critical* , *major* and *minor* thresholds.

The default tolerance window for critical thresholds is 10. The default for major and minor thresholds is 0.

If the tolerance is configured greater than threshold, then the tolerance will be treated as zero.

When configuring thresholds and tolerances for critical, major and minor congestion levels, the threshold levels and tolerances should never overlap. Consider the following example configuration, where the following threshold levels do not overlap:

- Critical congestion will trigger at 80 and will clear at 70
- Major congestion will trigger at 70 and will clear at 60
- Minor congestion will trigger at 60 and will clear at 50.

```
configure
   congestion-control threshold tolerance critical 10
  congestion-control threshold max-sessions-per-service-utilization major
 70
   congestion-control threshold tolerance major 10
  congestion-control threshold max-sessions-per-service-utilization minor
 60
   congestion-control threshold tolerance minor 10
   congestion-control threshold max-sessions-per-service-utilization
critical 80
   end
```

For information about all of the congestion control commands available, refer to the *Global Configuration Mode Commands* chapter of the *ASR 5500 Command Line Interface Reference* .

License Utilization Thresholds

The license-utilization threshold is calculated based on the configured license values for the chassis.

In this example configuration, the minor threshold will be triggered at 4000 calls, major threshold will be triggered at 6000 calls, and critical threshold will be triggered at 8000 calls.

```
congestion-control threshold license-utilization critical 80
congestion-control threshold license-utilization major 60
congestion-control threshold license-utilization minor 40
```

Maximum Session Per Service Thresholds

This threshold is configured across all MME services.

```
config
   congestion-control threshold max-sessions-per-service-utilization critical 80
```

When there are multiple MME services configured with different max-subscribers parameters, chassis congestion will be calculated using the minimum of max-subscribers configured in each of the different MME services.

However, congestion actions will be applied to each individual service based on its corresponding max-session-per-service parameters.

For example:

```
configure
   context ingress
      mme-service mmesvc1
         bind s1-mme ipv4-address 10.10.10.2 max-subscribers 10000
         exit
      exit
      mme-service mmesvc2
         bind s1-mme ipv4-address 10.10.10.3 max-subscribers 1000
         exit
      exit
      mme-service mmesvc3
         bind s1-mme ipv4-address 192.80.80.3 max-subscribers 20000
         end
```

In the above example, chassis level critical congestion will get triggered when the number of subscribers in mmesvc2 is at 800. Corresponding SNMP traps will be generated. However, congestion policies will not be applied for mmesvc1 and mmesvc3. When the number of subscribers in mmesvc1 and mmesvc3 reaches 8000 and 16000 respectively, then congestion policies will be applied for mmesvc1 and mmesvc3.

Chassis congestion will be cleared only when the congestion is cleared in all MME services.

Similarly, when minor, major and critical threshold are configured for max-session-per-service for many MME services, the maximum value of the threshold will be considered for chassis level congestion.

For example, if mmesvc1 reaches the major threshold, mmesvc2 reaches the critical threshold and mmesvc3 reaches the minor threshold, then chassis congestion state will be critical.

Service Control CPU Thresholds

This threshold is calculated from the system's demux CPU. The threshold is calculated based on a five minute average CPU usage.

The highest CPU usage value of two CPU cores of the demux CPU is considered. For example, if CPU core 0 has a five minute CPU usage of 40 and CPU core 1 has a five minute CPU usage of 80, then CPU core 1 will be considered for threshold calculation.

The following example configuration shows threshold levels of 80, 60, and 40 usage:

```
congestion-control threshold service-control-cpu-utilization critical 80
congestion-control threshold service-control-cpu-utilization major 60
congestion-control threshold service-control-cpu-utilization minor 40
```

System CPU Thresholds

This threshold is calculated using the five minute CPU usage average of all CPUs (except standby CPU and SMC CPU).

The highest CPU usage value of two CPU core of all CPU will be considered.

The following example configuration shows threshold levels of 80, 60, and 40 usage:

```
congestion-control threshold system-cpu-utilization critical 80
congestion-control threshold system-cpu-utilization major 60
congestion-control threshold system-cpu-utilization minor 40
```

System Memory Thresholds

This threshold is calculated using the five minute memory usage average of all CPUs (except standby CPU and SMC CPU).

The following example configuration shows threshold levels of 80, 60, and 40 usage:

```
congestion-control threshold system-memory-utilization critical 80
congestion-control threshold system-memory-utilization major 60
congestion-control threshold system-memory-utilization minor 40
```

Configuring a Congestion Action Profile

Congestion Action Profiles define a set of actions which can be executed after the corresponding threshold is crossed.

Use the following example configuration which creates a congestion action profile named *critical_action_profile* and defines several actions for this profile:

```
configure
   lte-policy
      congestion-action-profile critical_action_profile
         reject s1-setups time-to-wait 60
         drop handovers
         reject combined-attaches
         report-overload permit-emergency-sessions enodeb-percentage 50
         end
```

See the *Congestion Action Profile Configuration Commands* chapter in the *Command Line Reference* for details about all the congestion action profile commands available.

Refer to *Configuring Overload Control* in this chapter for more information about the **report-overload** keyword and associated functionality.

Associating a Congestion Action Profile with Congestion Control Policies

Each congestion control policy (critical, major, minor) must be associated with a congestion control profile.

The following example configuration to associate the congestion action profile named *critical_action_profile* with the **critical** congestion control policy:

```
configure
   congestion-control policy critical mme-service action-profile critical_action_profile
```

Separate congestion action profiles can be associated with major and minor congestion control policies, for example:

```
   congestion-control policy major mme-service action-profile major_action_profile
   congestion-control policy minor mme-service action-profile minor_action_profile
```

Configuring Overload Control

When an overload condition is detected on an MME, the system can be configured to report the condition to a specified percentage of eNodeBs and take the configured action on incoming sessions.

To create a congestion control policy with overload reporting, apply the following example configuration:

```
configure
   lte-policy
      congestion-action-profile  <profile_name>
      congestion-action-profile  <profile_name>
      end
```

```
configure
   congestion-control policy critical mme-service action report-overload reject-new-sessions
 enodeb-percentage <percentage>
   end
```

Notes:

- The following overload actions are also available (in addition to **reject-new-sessions**):
  - **permit-emergency-sessions-and-mobile-terminated-services**
  - **permit-high-priority-sessions-and-mobile-terminated-services**
  - **reject-delay-tolerant-access**
  - **reject-non-emergency-sessions**

See the *Congestion Action Profile Configuration Mode Commands* chapter in the *Command Line Reference* for details about all the congestion action profile commands available.

Configuring Enhanced Congestion SNMP Traps

When an enhanced congestion condition is detected, an SNMP trap (notification) is automatically generated by the system.

To disable (suppress) this trap:

```
configure
   snmp trap suppress EnhancedCongestion
   end
```

To re-enable generation of the Enhanced Congestion trap:

```
configure
   snmp trap enable EnhancedCongestion target <target-name>
   end
```

Verifying the Congestion Control Configuration

Use the following Exec mode command to display the configuration of the congestion control functionality.

```
show congestion-control configuration
```

The following output is a concise listing of all threshold and policy configurations showing multi-level Critical, Major and Minor threshold parameters and congestion control policies:

```
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:                              10 seconds
    port rx utilization:                                    80
    port tx utilization:                                    80
    license utilization:                                    100
    max-session-per-service utilization:      100
    tolerence limit:                                              10
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:                              10 seconds
    port rx utilization:                                    80
    port tx utilization:                                    80
    license utilization:                                    100
```

```
    max-session-per-service utilization:      100
    tolerence limit:                                    10
Congestion-control Major threshold parameters
    system cpu utilization:                               0
    service control cpu utilization:              0
    system memory utilization:                          0
    message queue utilization:                          0
    message queue wait time:                              0 seconds
    port rx utilization:                                    0
    port tx utilization:                                    0
    license utilization:                                    0
    max-session-per-service utilization:      0
    tolerence limit:                                        0
Congestion-control Minor threshold parameters
    system cpu utilization:                               0
    service control cpu utilization:              0
    system memory utilization:                          0
    message queue utilization:                          0
    message queue wait time:                              0 seconds
    port rx utilization:                                    0
    port tx utilization:                                    0
    license utilization:                                    0
    max-session-per-service utilization:      0
    tolerence limit:                                        0
Overload-disconnect: disabled
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
    mme-service:
        Critical Action-profile : ap3
        Major    Action-profile : ap2
        Minor    Action-profile : ap1
```

Verifying Congestion Action Profiles

To verify the configuration of a congestion action profile, use the following Exec mode command:

show lte-policy congestion-action-profile { name *<profile_name>* **| summary }**

Monitoring and Troubleshooting

This section provides information on how to monitor congestion control.

Congestion Control Show Command(s) and/or Outputs

This section provides information regarding show commands and/or their outputs in support of enhanced congestion control.

show congestion-control statistics mme

The following command shows an overview of all congestion control statistics for the MME.

show congestion-control statistics mme [full | critical | major | minor]

The following output is a concise listing of congestion control statistics. In this example output, only the **Critical** information is shown.

```
Critical Congestion Policy Action
    Congestion Policy Applied                       :    0 times
    PS attaches
                    Rejected            :      0 times
                    Dropped              :      0 times
    PS attaches
                    Rejected            :      0 times
                    Dropped              :      0 times
    Combined attaches
                    Rejected            :      0 times
                    Dropped              :      0 times
    S1-Setup
                    Rejected            :      0 times
                    Dropped              :      0 times
    Handover
                    Rejected            :      0 times
                    Dropped              :      0 times
    Addn-pdn-connect
                    Rejected            :      0 times
                    Dropped              :      0 times
    Addn-brr-connect
                    Rejected            :      0 times
                    Dropped              :      0 times
    Service-Request
                    Rejected            :      0 times
                    Dropped              :      0 times
    TAU-Request
                    Rejected            :      0 times
                    Dropped              :      0 times
    S1AP Overload Start Sent                         :      2 times
    S1AP Overload Stop  Sent                         :      2 times
    Excluded Emergency Events                        :      0 times
    Excluded Voice Events                                :      0 times
```

show congestion-control statistics mme

The following command shows SNMP event statistics for the EnhancedCongestion trap and EnhancedCongestionClear trap .

```
show snmp trap statistics verbose | grep EnhancedCongestion
```