

Load/Overload and UP Data Throttling Support on Sx

Table 1: Feature History

Feature Name	Release Information	Description
Support for load and overload control in multi-Sx scenario	2025.02.0	This feature extends the load and overload control functionality to multi-Sx scenario. This allows other UPs to share the load of incoming new calls if an existing UP fails due to any error. **Default Setting:** Always enabled.

- Feature Description, on page 1
- How It Works, on page 2
- Configuring Load and Overload Support, on page 3
- Monitoring and Troubleshooting, on page 8

Feature Description

The Load/Overload support is implemented in the UPC CUPS architecture. This support is handled between the Contol Plane (CP) and User Plane (UP).

Load control enables UP to send its load information to CP to adaptively balance the PFCP session load across the UP functions according to their effective load, whereas Overload control enables throttling of new session requests towards a particular UP.

Load and overload control in multi-Sx scenario

CUPS also supports load and overload control in multi-Sx scenario. The multi-Sx scenario indicates a mesh architecture where multiple CPs are connected with multiple UPs .

It enables single UP to establish multiple N4 or Sx interfaces with any number of control plane NFs such as SMF, SAEGW-C, PGW-C, SGW-C, and cnSGWc.

How It Works

User Plane Selection

When Load/Overload support is enabled, UP selection is implemented as given below, with a UP group:

- If none of the UP is in overload condition, Load Control Information (LCI) is used for UP selection. In this case, the least loaded UP will be selected.
- If all UPs are in the overload state, UP selection is based on the Overload Control Information (OCI). In this case the least overloaded UP is selected.
  - After a particular UP is selected, the reduction metric is still applied to this UP for throttling.
  - If throttling needs to be dropped, UP selection request is rejected for that PDN connection.
- In some scenarios where some of the UPs are in Overload condition and some of the UPs are not in Overload condition, the selection is done based on the OCI value.
- If the LCI or the OCI value are the same for a peer node, the session count information is used for UP selection.

Node-level Load/Overload Support

The CP informs the UP about the Load/Overload support enabled. Based on this information UP decides to send the Load/Overload information towards CP peer or not.

Load/Overload support at CP is configured as part of the Sx-Service node configuration. This information is sent to the UP during Sx Association Response or Sx Association Update request, if the information has changed through dynamix configuration.



Note If CP does not support Load/Overload feature through supported CLI, it ignores the reported Load/Overloadby UP. In that case, UP selection continues with the session count information.

Sx Establishment Request Throttling at CP in Overload State

Once the UP is in Overload situation, CP starts throttling the Sx Establishment Request message towards UP. This avoids new calls (Low priority/non-emergency) towards the overloaded UP.

Throttling happens based on reported OCI values – Overload Reduction metric value. The value is calculated in percentage. It randomly drops the required percentage of Sx Establishment Request towards that UP Peer. This results in call drop at CP with "sx-no-resource" disconnect reason. Also, respective statistics are incremented for the same.



Note The eMPS (high priority) subscribers' Session/ Emergency Subscribers' session is not throttled.

Sx Establishment Request Throttling at UP in Self-Protection

Once the UP is in Self-Protection state, it starts rejecting all the new sessions (non-eMPS session only), Sx Establishment Request, and Sx Modification Request for the existing sessions (non-eMPS session only).

Session Termination Trigger from UP in Self-Protection

Being in Self-Protection mode, if there are no improvement in the Load condition at UP, it starts triggering Session Termination Request towards CP in a staggered manner through Sx Report Request message indicating that UP is in Self-Protection. Based on this, CP starts initiating Sx Termination Request for those sessions.

Self Protection Termination Request (SPTER): This bit is set from UP towards CP for initiating Self-Protection based termination. The CP releases the call with disconnect reason as "graceful-term-up-self-protectn".

Note When Actual Load value is greater than the Session-Termination-Start-Threshold value, Session termination is triggered towards CP.

Limitation

This feature has the following known limitations:

- The maximum number of Load/Overload profiles supported on UP is 8.
- If the Load/Overload profiles are not configured in all UPs in the UP group, it can lead to uneven distribution of sessions. It is recommended that all UPs must be configured with the Load/Overload support in a single UP group.
- After session recovery, SessMgr instance gets to relearn Load/Overload values from SxDemux. The SxDemux communicates these values only when there is a change in Load/Overload values.
- Toggling the Load Control configurations on the fly (enable to disable or disable to enable) is not supported.
- All UPs in a UP group should either be enabled or disabled. Having one of the UP enabled and another as disabled in a UP group is not supported as it can result in improper Load/Overload values on CP.
- It's recommended to disable Load Control for all IMS UPs.

Configuring Load and Overload Support

The Load and Overload support is configured using the following configurations:

- User Plane Load Control Profile Configuration
- User Plane Overload Control Profile Configuration
- Association of Load Profile to a User Plane Service
- Sx Protocol Configuration on Control Plane

User Plane Load Control Profile Configuration

Use the following commands to configure load control profile.

```
configure
   userplane-load-control-profile profile_name
   end
```

Configuring User Plane Load Control Profile Parameters

Use the following configuration to configure UP Load profile parameters:

```
configure
   userplane-load-control-profile profile_name
      system-weightage system-cpu-utilization utilization_value
system-memory-utilization utilization_value license-session-utilization
utilization_value
      sessmgr-weightage sessmgr-cpu-utilization utilization_value
system-memory-utilization utilization_value
     inclusion-frequency advertisement-interval interval_value change-factor
 changefactor_value
      end
```

NOTES:

- **inclusion-frequency**: Configures parameters to decide inclusion frequency of load control information IE.
- **advertisement-interval**: Advertisement interval is the periodic interval after which the load value is advertised. Configures the advertisement interval for load control. The default value is 300. Set the value as 0 (zero) to include LCI IE in every applicable message.
- **change-factor**: Change-factor is the delta increase or decrease in the load value based on which the load advertisement occurs. Configures the change factor for load control. The default value is 5.
- **sessmgr-weightage**: Configures sessmgr weightage for various load control parameters. Total weightage of all the parameters should be 100. The default ratio is 65% weightage to sessmgr-cpu-utilization and 35% weightage to sessmgr-memory-utilization.
- **sessmgr-cpu-utilization**: Configures session manager CPU utilization weightage in percentage. Default weightage in load factor is 35%.
- **sessmgr-memory-utilization**: Configures session manager memory utilization weightage in percentage. Default weightage in load factor is 65%.
- **system-weightage**: Configures system weightage for various load control parameters. Total weightage of all the parameters should be 100. The default values are 40% weightage to system-cpu-utilization, 30% weightage to system-memory-utilization and 30% weightage to license-session-utilization.
- **system-cpu-utilization**: Configures system CPU utilization weightage in percentage. Default weightage in load factor is 40%.

Note The value displayed in the **show cpu table** CLI command is based on the average value of 5 minutes, 10 minutes, and 15 minutes. Use the result of the average value for system CPU utilization to verify the utilization manually.

- **system-memory-utilization**: Configures system memory utilization weightage in percentage. Default weightage in load factor is 30%.

Note The value displayed in the **show cpu table** CLI command is based on the average value of 5 minutes, 10 minutes, and 15 minutes. Use the result of the average value for system memory utilization to verify the utilization manually.

- **license-session-utilization**: Configures license session utilization weightage for User Plane service in percentage. Default weightage in load factor is 30%. The license utilization percentage is equal to the utilization percentage of the current UP sessions out of the maximum UP sessions.

User Plane Overload Control Profile Configuration

Use the following commands to configure overload control profile.

```
configure
  userplane-overload-control-profile profile_name
  end
```

Configuring User Plane overload Control Profile Parameters

Use the following configuration to configure UP overload profile parameters:

```
configure
  userplane-overload-control-profile profile_name
    overload-threshold system lower-limit limit_value upper-limit limit_value
 sessmgr lower-limit limit_value upper-limit limit_value vpp-cpu lower-limit
 limit_value upper-limit limit_value
     system-weightage system-cpu-utilization utilization_value
system-memory-utilization utilization_value license-session-utilization
utilization_value
     sessmgr-weightage sessmgr-cpu-utilization utilization_value
system-memory-utilization utilization_value
    inclusion-frequency advertisement-interval interval_value change-factor
 changefactor_value
     tolerance tolerance_value
     validity-period validity_period
     end
```

NOTES:

- **inclusion-frequency**: Configures parameters to decide inclusion frequency of overload control information IE.

- **advertisement-interval**: Advertisement interval is the periodic interval after which the overload value is advertised. Configures the advertisement interval for overload control. The default value is 300. Set the value as 0 (zero) to include LCI IE in every applicable message.
- **change-factor**: Change-factor is the delta increase or decrease in the overload value based on which the overload advertisement occurs. Configures the change factor for overload control. The default value is 5.
- **tolerance**: Configures the Overload tolerance limits.
- **validity-period**: Configures validity of overload control information. Default value is 600.
- **overload-threshold**: Configures Overload thresholds limits for system, sessmgr and vpp-cpu.
- **system**: Configures overload system threshold after which node moves to self-protection mode.
- **vpp-cpu**: Configures the overload vpp-cpu threshold after which node moves to self-protection mode.
- **sessmgr**: Configures the overload threshold for session manager after which node moves to self-protection mode.
- **upper-limit** *limit_value*: Configures the various upper limit values. Following are the various upper limit values:
  - System Threshold Upper Limit : Configures overload system threshold after which node moves to self-protection mode. Default limit value is 80%.
  - Sessmgr Threshold Upper Limit : Configures overload SessMgr threshold after which node moves to self-protection mode. Default limit value is 60%.
  - vpp-cpu Threshold Upper Limit : Configures overload vpp-cpu threshold L2 after which node moves to self-protection mode. Default limit value is 60%.
- **lower-limit** *limit_value*: Configures the various lower limit values. Following are the various lower limit values:
  - System Threshold Lower Limit: Configures overload system threshold after which node moves to self-protection mode. Default limit value is 60%.
  - Sessmgr Threshold Lower Limit: Configures overload SessMgr threshold after which node moves to self-protection mode. Default limit value is 50%.
  - vpp-cpu Threshold Lower Limit : Configures overload vpp-cpu threshold L1 after which node moves to self-protection mode. Default limit value is 50%.
- **sessmgr-weightage**: Configures sessmgr weightage for various overload control parameters. Total weightage of all the parameters should be 100. The default ratio is 65% weightage to sessmgr-cpu-utilization and 35% weightage to sessmgr-memory-utilization.
- **sessmgr-cpu-utilization**: Configures session manager CPU utilization weightage in percentage. Default weightage in overload factor is 35%.
- **sessmgr-memory-utilization**: Configures session manager memory utilization weightage in percentage. Default weightage in overload factor is 65%.
- **system-weightage**: Configures system weightage for various overload control parameters. Total weightage of all the parameters should be 100. The default values are 40% weightage to system-cpu-utilization, 30% weightage to system-memory-utilization and 30% weightage to license-session-utilization.

- **system-cpu-utilization**: Configures system CPU utilization weightage in percentage. Default weightage in overload factor is 40%.

Note The value displayed in the **show cpu table** CLI command is based on the average value of 5 minutes, 10 minutes, and 15 minutes. Use the result of the average value for system CPU utilization to verify the utilization manually.

- **system-memory-utilization**: Configures system memory utilization weightage in percentage. Default weightage in overload factor is 30%.

Note The value displayed in the **show cpu table** CLI command is based on the average value of 5 minutes, 10 minutes, and 15 minutes. Use the result of the average value for system memory utilization to verify the utilization manually.

- **license-session-utilization**: Configures license session utilization weightage for User Plane service in percentage. Default weightage in overload factor is 30%. The license utilization percentage is equal to the utilization percentage of the current UP sessions out of the maximum UP sessions.

Associating a Load Control Profile with a User Plane Service

Use the following commands to associate the Overload Control profile to a use plane service.

```
configure
  context context_name
    user-plane-service service_name
    [ no ] associate userplane-load-control-profile profile_name
```

NOTES:

- **associate**: This command associates the user plane overload control profile with a user plane service.

Sx Protocol Configuration on Control Plane

The CP Function Features IE indicates the features supported by CP. Only features having an impact on the (system-wide) UP function behaviour are signalled in this IE.

The following features are supported by CP:

- LOAD (Load Control)
- OVRL (Overload Control)

Use the following configuration to configure the supported features on CP through the Sx Protocol:

```
configure
  context context_name
    sx-service service_name
      sx-protocol supported-features { load-control | overload-control
```

```
}
      no sx-protocol supported-features [ load-control | overload-control
 ]
       end
```

NOTES:

- **supported-features**: Configures supported features for Sx interface by CP. Default value is Disabled.
- **load-control**: Enables or disables Load control feature support on CP function.
- **overload-control**: Enables or disables the Overload control feature on CP function.

Monitoring and Troubleshooting

Show Commands Input and/or Outputs

This section provides information regarding show commands and their outputs in support of the feature.

show userplane-load-control-profile name *name*

The following fields are displayed in support of this feature:

- User Plane Load Control Profiles
- User Plane Load Control Profile Name
- System Weightage and Thresholds:
  - CPU Utilization Weightage
  - Memory Utilization Weightage
  - License Session Utilization Weightage
  - System Threshold Lower Limit
  - System Threshold Upper Limit
- Sessmgr Weightage and Thresholds:
  - CPU Utilization Weightage
  - Memory Utilization Weightage
  - Sessmgr Threshold Lower Limit
  - Sessmgr Threshold Upper Limit
- VPP Weightage and Thresholds:
  - VPP Utilization Weightage
  - vpp-cpu Threshold Lower Limit
  - vpp-cpu Threshold Upper Limit

- Inclusion Frequency:
  - Change Factor
  - Advertisement Interval

show userplane-overload-control-profile name *name*

The following fields are displayed in support of this feature:

- User Plane Overload Control Profiles
- User Plane Overload Control Profile Name
- System Weightage and Thresholds:
  - CPU Utilization Weightage
  - Memory Utilization Weightage
  - License Session Utilization Weightage
  - System Threshold Lower Limit
  - System Threshold Upper Limit
- Sessmgr Weightage and Thresholds:
  - CPU Utilization Weightage
  - Memory Utilization Weightage
  - Sessmgr Threshold Lower Limit
  - Sessmgr Threshold Upper Limit
- VPP Weightage and Thresholds:
  - VPP Utilization Weightage
  - vpp-cpu Threshold Lower Limit
  - vpp-cpu Threshold Upper Limit
- Inclusion Frequency
  - Change Factor
  - Advertisement Interval
- Validity Period

show user-plane-service statistics all

The following fields are displayed in support of this feature:

- Overload Stats
  - Current State : Normal

- Number of time self-protection condition reached in user plane : 0
- No of Session Eshtablishment Req rejected during self-protection mode : 0
- No of Session Modif Req rejected during self-protection mode : 0
- No of eMPS Session Eshtablishment Req allowed during self-protection mode : 0
- No of eMPS Session Modif Req allowed during self-protection mode : 0
- Overload reduction metric : 0
- Current Overload factor system : 0
- Current Overload factor sessmgr : 0
- Current Overload factor vpp cpu : 0

- Overload Data Stats:
  - Total Packets dropped due to overload : 0
  - Total Bytes dropped due to overload : 0
  - Total Packets dropped in self-protection mode : 0
  - Total Bytes dropped in self-protection mode : 0

- Load Stats:
  - Load metric : 0
  - Current Load factor system : 0
  - Current Load factor sessmgr : 0
  - Current Load factor vpp cpu : 0

- eMPS PDNs Total
  - Active
  - Setup
  - Released
  - Rejected

show sx service statistics all

The following fields are displayed in support of this feature:

- Throttled

Bulk Statistics

Following bulkstats are available in support of Load and Overload Support on Sx feature.

Table 2: Supported Bulk Stats

Bulkstats	Description
num-self-protection-reached	Total number of time self-protection condition reached in UP.
num-session-estab-rejected-on-self-protection	Total number of Session Establishment Request rejected during self-protection mode.
num-session-modif-rejected-on-self-protection	Total number of Session Modification Request rejected during self-protection mode.
num-emps-session-estab-allowed-on-self-protection	Total number of eMPS Session Establishment Request allowed during self-protection mode.
num-emps-session-modif-allowed-on-self-protection	Total number of eMPS Session Modification Request allowed during self-protection mode.
overload-reduction-metric	Overload reduction metric is calculated based on the configured Lower and Upper limit of Overload Condition.
overload-factor-system	Overload factor system is calculated based on the System CPU, Memory, VPP CPU, and other information polled from Resource Manager (RM).
overload-factor-session	UP starts rejecting new sessions and data throttling during self-protection mode.
overload-factor-vpp-cpu	Total average VPP CPU per core during overload.
load-metric	Total number of current Load metric.
load-factor-system	Total number of current sytem load factor.
load-factor-session	Total number of current session load factor.
load-factor-vpp-cpu	Total number of current VPP CPU load factor.
num-packets-dropped-on-overload	Total number of packets dropped during the overload.
num-bytes-dropped-on-overload	Total number of bytes dropped during the self protection mode.
num-packets-dropped-on-self-protection	Total number of packets dropped during the self protection mode.
num-bytes-dropped-on-self-protection	Total number of bytes dropped during the self protection mode.

SNMP Traps

The following SNMP Traps are added in support of this feature:

- UPlaneSelfOverload: When system enters into Self-Protection mode.
- UPlaneSelfOverloadClear: When system is out of Self-Protection mode.