

Overview

The Evolved Packet Core (EPC) network is evolving and moving toward Control User Plane Separation (CUPS) based architecture where User Plane and Control Plane are separate nodes for P-GW, S-GW, and TDF products. The User Plane and Control Plane combined together provide functionality of a node for other elements in the EPC network. However, keeping it separate has numerous advantages from the network point of view – support different scaling for Control Plane and User Plane, support more capacity on per session level in User Plane, and so on.

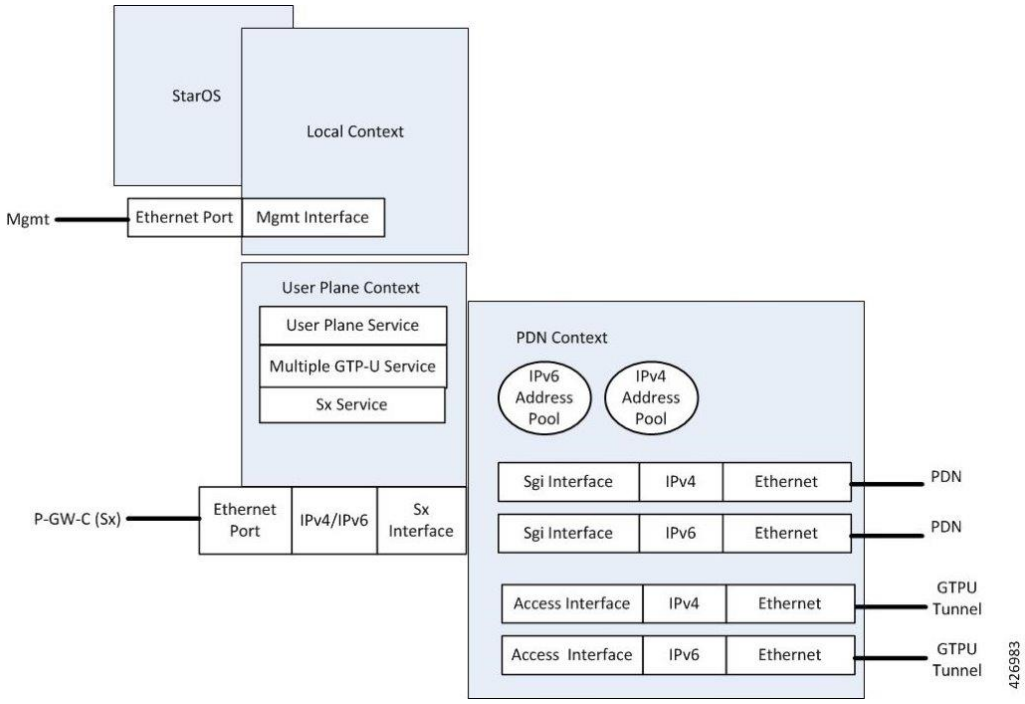
This chapter highlights high-level details, call flows, and configurations related to Control Plane implementation for P-GW, S-GW, and SAEGW products.

- Product Description, on page 1
- Supported Features and Functionality, on page 3
- How It Works, on page 15

Product Description

The SAEGW-U Virtualized Network Function (VNF) can be hosted in Cisco Ultra Services Platform (USP) on COTS hardware or on ASR 5500/DPC2 chassis. The SAEGW-U can be collocated with SAEGW-C in the same data center or can be located remotely in a different data center.

Following is a high-level architecture of User Plane as a service.



Some important points that describe the User Plane as a service:

- User Plane can be programmed from Control Plane.
- Single User Plane service can serve both SGW-U and P-GW-U type sessions.
- Two or more separate User Plane services can be defined for each node type, SGW-U and PGW-U, respectively.
- A group of SAEGW-Us can be explicitly associated with an APN. If no group is associated, a default group is used which includes all the registered User Planes that are registered to SAEGW-C but are not part of any configured SAEGW-U group.
- User Plane service is associated with Sx service for the Control Plane interface, and GTP-U service for receiving GTP-U packets.

Important Currently, each User Plane Service is associated with only single Sx service to interface with Control Plane.

- User Plane service can be associated with four GTP-U services which can be extended to support SaMOG, GGSN, and ePDG.
- Multiple peers of Control Plane services use single User Plane service.
- To associate the IP pool and its configuration, APN configuration is required.

Important Currently, User Plane supports APN and pool configuration. The IP addresses are allocated from the Control Plane and are validated in the User Plane.

Licensing Information

The Cisco UPC CUPS supports Cisco Smart Licensing, a cloud-based licensing model that simplifies the purchase, deployment, and management of Cisco software assets. For more information, refer to the Smart Licensing collection page.

Supported Features and Functionality

3GPP ULI Enhanced Reporting Support

This feature enhancement covers ULI-related gaps in P-GW and GGSN as per 3GPP standards.

S4SGSN reports ULI to the P-GW through S-GW. P-GW determines the changes in the ULI with previously received ULI. If P-GW detects any change and the change request is from the PCRF as an event trigger, then the P-GW reports the ULI to the PCRF.

SGSN reports ULI to the GGSN. GGSN determines the changes in the ULI with previously received ULI. If GGSN detects any change and the change request is from the PCRF as an event trigger, then the GGSN reports the ULI to the PCRF. This feature also supports the detection of the change in RAI received as part of the ULI field at GGSN.

For more information on 3GPP ULI Reporting Support Enhancement, refer the *3GPP ULI Reporting Support Enhanced* section in the *StarOS P-GW Administration Guide*.

AAA Server Group

The AAA Server Group feature is used to create and manage the Diameter/RADIUS server groups within the context or system. The AAA server group facilitates management of group (list) of servers at per subscriber/APN/realm-level for AAA functionality.

Note The AAA Server Group is an existing feature that is supported in non-CUPS architecture. With this release, the feature is qualified in CUPS architecture.

For additional information about CLI configurations related to AAA server group, refer the *AAA Server Group Configuration Mode Commands* chapter in the *Command Line Interface Reference*.

APN Configuration Support

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

Revision Details	Release
First introduced	Pre 21.24

The CLI commands **radius-group**, **cc-home behaviour 0x10 profile 2** and **mediation-device** are qualified and validated in the CUPS architecture to support APN configuration.

radius-group

Under this functionality validation the CUPS architecture supports 800 Radius Server Groups each group configured with RADIUS Authentication and Accounting server.

cc-home { behavior *bits* | profile *index* }

Configures the home subscriber charging characteristics (CC) used by the GGSN when those from the SGSN will not be accepted. The values configured in the CLI are taken into precedence by CUPS SAEGW service and populated appropriately in the GTPP CDR records.

NOTES:

- **behavior** *bits*: Specifies the behavior bit for the home subscriber charging characteristic. bits can be configured to any unique bit from 001H to FFFH (0001 to 1111 1111 1111 bin) where the least-significant bit corresponds to B1 and the most-significant bit corresponds to B12.
- **profile** *index*: Specifies the profile index for the home subscriber charging characteristic. index can be configured to any integer value between 0 and 15. Default: 8
- For more information, refer to the **cc-home** command under *APN Configuration Mode Commands* chapter in the *Command Line Interface Reference A-B* document

mediation-device [context-name *context_name*] [delay-GTP-response] [no-early-PDUs] [no interims] +

This command and all associated sub section CLIs are supported in CUPS. This CLI enables use of **mediation device** and all associated configuration that can be used for a given APN by CUPS SAEGW service.

NOTES:

- **context-name** *context_name*: Configures the mediation VPN context for this APN as an alphanumeric string of 1 through 79 characters that is case sensitive. If not specified, the mediation context is the same as the destination context of the subscriber. Default: The subscribers destination context.
- **delay-GTP-response**: When enabled, delays the CPC response until an Accounting Start response is received from the mediation device. Default: Disabled.
- **no-early-pdus**: Specifies that the system delays PDUs from the MS until a response to the GGSN accounting start request is received from the mediation device. The PDUs are queued, not discarded. Default: Disabled.
- **no-interim**: Disables sending interims to the mediation server. Default: Disabled.
- For more information, refer to the **mediation-device** command under *APN Configuration Mode Commands* chapter in the *Command Line Interface Reference A-B* document

Asynchronous Core Transfer Support for egtpinmgr

Asynchronous core transfer support for egtpinmgr has been added in CUPS to optimize outage time during an egtpinmgr restart.

Previously, when the egtpinmgr restarted, the recovery process began only after a core dump file was created and transferred. However, the time taken to transfer the core file was significant. The outage time during an egtpinmgr restart was equal to the egtpinmgr recovery time plus the core file transfer time.

Support for Asynchronous Core Transfer has been added in CUPS to include the egtpinmgr during the recovery process. Now, recovery begins when the egtpinmgr process crashes without waiting for the kernel to complete a core dump file transfer and release its resources. As a result, the outage time during an egtpinmgr restart is equal to the egtpinmgr recovery time only.

With this enhancement, outage time during an egtpinmgr restart is reduced. The outage time consists only of the time required to recover the egtpinmgr. The time taken to create and transfer the core file no longer contributes to the outage time.



Note The Asynchronous Core Transfer Support for egtpinmgr is an existing feature that is supported in non-CUPS architecture. With this release, the feature is qualified in CUPS architecture.

Charging Data Records to HDD

A Charging Data Record (CDR) is a formatted collection of information about a chargeable event. The GTPP accounting CDRs that are generated are sent to an external node for storage. The CDRs are written to files in formats supported by the external node and stored on the hard disk (HDD). From the HDD, CDR files can be pushed or pulled using FTP or SFTP protocols.



Note It is strongly recommended that you do not use the system directories created by StarOS under */hd-raid/records/* for deployment use cases such as backups. If such directories are used, this could impact the normal functioning of the product.

CDR is an existing feature that is supported in the non-CUPS architecture, and qualified in the CUPS architecture. For additional information, see the *HDD Storage* chapter in the *GTPP Interface Administration and Reference*.

GTP-C Path Failure Enhancements and Improved Debugging Tools

In CUPS architecture, enhancements have been added to optimize GTP-C path failure functionality, and to improve the debug capability of the system for GTP-C path failure problems. These features will help Operators and Engineers to debug different aspects of the system that will help in identifying the root cause of GTP-C path failures in the network. These enhancements affect path failure detection via the S5, S8, S2b, and S2a interfaces.

These enhancements are part of this feature:

- The node can be configured so that it does not detect a path failure if a low restart counter is received due to incorrect or spurious messages. This prevents call loss. The option to disable path failure due to Echo Request/Response and Control Message Request/Response messages is also available so that call loss is prevented in the event of a false path failure detection.
- More granularity has been added to GTP-C path failure statistics so that the root cause of issues in the network can be diagnosed more quickly.

- A path failure history for the last five path failures per peer is available to assist in debugging path failures in the network.
- Seamless path failure handling is implemented so that call loss is avoided during redundancy events.

Note The GTP-C Path Failure Enhancements and Improved Debugging Tools is an existing feature that is supported in non-CUPS architecture. With this release, the feature is qualified in CUPS architecture. For additional information, refer the *GTP-C Path Failure Enhancements and Improved Debugging Tools* section in the *P-GW Administration Guide*.

GTPP Suppress-CDR No Zero Volume

This feature allows suppression of CDRs with zero byte data count, so that the OCG node is not overloaded with a flood of CDRs. The CDRs can be categorized as follows:

- Final-cdrs: These CDRs are generated at the end of a context.
- Internal-trigger-cdrs: These CDRs are generated due to internal triggers such as volume limit, time limit, tariff change, or user-generated interims through the CLI commands.
- External-trigger-cdrs: These CDRs are generated due to external triggers such as QoS Change, RAT change and so on. All triggers which are not considered as final-cdrs or internal-trigger-cdrs are considered as external-trigger-cdrs.

The customers can select the CDRs they want to suppress.

The CLI command mentioned below helps suppress CDRs on different CDR triggers supported in CUPS:

- **[default | no] gtpp suppress-cdrs zero-volume { external-trigger-cdr | final-cdr | internal-trigger-cdr }**

Location Based DNS and PCSCF IP Address Selection

Location-based DNS and P-CSCF Selection provides an option to the operator to manage the DNS server address and P-CSCF IP address according to location information.

P-GW gathers the DNS server address and P-CSCF IP address information by Tracking Area Identifier (TAI), which is achieved through the TAC-based Virtual APN (VAPN) selection.

When UE sends the PCO request in session creation, P-GW selects the Virtual APN (VAPN) with the received location information. The selected VAPN (with DNS server address and P-CSCF IP address configured in it) with PCO IE is sent in the Create session response.

Following are the CLI commands for enabling the Location-based DNS and PCSCF IP address selection:

Command	Description
Tracking-area-code-range from <*start value*> **to** <*end value*>	Provides the tracking area code range, starting from 0 through 65536. The end value is always greater than the start value.

Command	Description
P-cscf priority <*priority*> **ip/ipv6** <*IPv4/Ipv6 address*>	Specifies the priority for P-CSCF address for the APN. Address_ priority is an integer 1–3. One is the maximum priority. IPv4_address is in IPv4 dotted-decimal notation. IPv6_address is in IPv6 colon-separated-hexadecimal notation.
Show apn name <*APN Name*>	To show PCSCF IP address at APN
dns primary <*IPv4 address*> **Dns secondary** <*IPv4 address*> **ipv6 dns primary** <*IPv6 address*> **ipv6 dns secondary** <*IPv6 address*>	Primary: Configures the primary DNS server for the APN. Secondary: Configures the secondary DNS server for the APN. Only one secondary DNS server is configurable. Address: Configures the IP address of the DNS server expressed in IPv4 dotted-decimal notation. Default: primary = 0.0.0.0, secondary = 0.0.0.0 dns_address: Specifies the IP address of the DNS server to remove, expressed in IPv4 dotted-decimal notation.
Show apn name <*APN Name*>	To show DNS IP address at APN

MPRA Support

P-GW supports negotiation of Multiple-Presence Reporting Area feature in Feature-List-ID 2 over Gx interface with PCRF. The CNO-ULI feature works only when the P-GW and/or the PCRF doesn't support Multiple-PRA and both P-GW and PCRF support CNO-ULI.

For Multiple-PRA feature support during the lifetime of the IP-CAN session, P-GW handles the change of UE Presence in Reporting Areas request from PCRF in PRA-Install AVP including the Presence-Reporting-Area-Information AVPs. Each AVP contains the Presence Reporting Area Identifier within the Presence-Reporting-Area-Identifier AVP

For more information on Presence Reporting Area (PRA) and Multiple PRA, refer the *Presence Reporting Area* chapter in the *StarOS P-GW Administration Guide*.

No udp-checksum Support

This feature supports **no udp-checksum** CLI command for CUPS under GTPU service where **udp-checksum** is disabled in the outer GTPU header for the downlink subscriber packet. When downlink packet arrives from internet, the GTPU header is added on top of the packet and is sent to the access side. The "checksum" value is zero in the outer UDP layer of this packet enabling optimization and therefore, improving the performance throughput.

Use the following configuration to enable the feature.

```
configure
  context context_name
    gtpu-service gtpu_service_name
    [ no ] udp-checksum
    end
```

Show Commands and Outputs

This section provides information about the show CLI commands available in support of the feature.

Use the following command to determine if **GTPU UDP Checksum** is enabled or disabled.

- **show gtpu-service all**—Displays all GTPU services.
- **show gtpu-service name** *service_name*—Displays information for the specific GTPU service name.

QUIC IETF Implementation

In the current framework, Deep Packet Inspection (DPI) is done for every packet in a flow when it reaches the plugin. The DPI is done by analyzing the packets and extracting deterministic patterns. The DPI is done in-order to detect the application and to classify its subtype. Plugin excludes the flow after the DPI. The flow is offloaded after the detection. As part of QUIC IETF, the initial QUIC handshake packets (Client/Server Hello) are encrypted over the network. Hence, there are no deterministic patterns available for detection of the application. Support is added in p2p plugin to decrypt and obtain the SNI (Server Name Indication) for detection.

Configuring QUIC IETF

Use the following configuration to enable or disable the QUIC IETF decryption.

```
configure
  active-charging service acs_service_name
    p2p-detection debug-param protocol-param p2p_quic_ietf_decrypt 1
    end
```



Note By default, the CLI is disabled and there's minimal impact on the performance due to TLS decryption.

Optimization for egtpinmgr Recovery

Previously, when the egtpinmgr task restarted, it took a significant amount of time for it to recover. As a result, the outage time when the SAEGW were unable to accept any new calls during egtpinmgr recovery was high.

The software has been enhanced to optimize the recovery outage window in the event of an egtpinmgr task restart; this has been achieved by optimizing the internal algorithms of egtpinmgr recovery and the data structures required. In addition, recovery time now is dependent only on the number of unique IMSIs and not on the number of sessions for an IMSI.



Note The Optimization for egtpinmgr Recovery is an existing feature that is supported in non-CUPS architecture. With this release, the feature is qualified in CUPS architecture.

Quota Hold Time Support

Quota-Hold-Time (QHT) is an inactivity time duration, after which the Gateway(Diameter client) returns the Charging-Bucket with its usage and reaches a clean-state.

The QHT value is provided by the OCS per Category - Multiple-Services-Credit-Control (MSCC), or the gateway provides an option to configure the default value of the QHT - for enabling the default QHT value for the MSCCs for which the OCS has not provided any QHT AVP.

The QHT timer runs per MSCC bucket. If the QHT timer expires without a packet during run-time, then the usage is reported with the Reporting-Reason: QHT as per 3GPP specification.

The QHT value received in the CP from the OCS, is sent in the "Quota Holding Time" IE defined in the CUPS specification 3GPP TS 29.244. Also along with provisioning the Quota-Holding-Time IE to the UP, the Reporting-Triggers will be sent with the bit corresponding to Quota-Holding-Time SET, so that on QHT expiry the reporting takes place.

The UP on receiving the Quota-Holding-Time IE along with the QHT Reporting-Triggers enabled, starts the timer per URR to monitor the inactivity period. Once the inactivity period exceeds the QHT time, the Usage-Reporting is initiated from the UP for the Trigger : Quota-Holding-Time.

The CP on receiving the QHT event from UP, triggers the QHT reporting to the OCS after updating the usage in the MSCC bucket.

Configuring Quota Hold Time

Use the following configuration to enable Quota Hold Time in CUPS:

```
configure
  require active-charging
  active-charging service service_name
    credit-control group group_name
      quota-hold-time timer_value
      end
```

NOTES:

- **quota-hold-time**: configures the inactivity duration after which the charging bucket reports its usage and have a clean state.

Limitation

The QHT (inactivity-timer) usually is a larger value compared to the flow-idle timer. If the flow-idle timer is larger than QHT, then there is a possibility for the flows present even after the QHT expiry, and is processed by VPP as per the NoQuota Pending-Traffic-Treatment configuration.

S-GW Paging Enhancement

S-GW Paging includes the following scenarios:

Scenario 1: S-GW sends a Downlink Data Notification (DDN) message to the MME/S4-SGSN nodes. MME/S4-SGSN responds to the S-GW with a DDN Ack message. While waiting for the DDN Ack message from the MME/S4-SGSN, if the S-GW receives a high priority downlink data, it does not resend a DDN to the MME/S4-SGSN.

Scenario 2: If a DDN is sent to an MME/S4-SGSN and TAU/RAU MBR is received from another MME/S4-SGSN, S-GW doesn't send DDN.

Scenario 3: DDN is sent to an MME/S4-SGSN and DDN Ack with Cause #110 is received. DDN Ack with cause 110 is treated as DDN failure and standard DDN failure action procedure is initiated.

To handle these scenarios, the following two enhancements are added to the DDN functionality in CUPS architecture:

- High Priority DDN at S-GW
- MBR-DDN Collision Handling

These enhancements support the following:

- Higher priority DDN on S-GW and SAEGW, which helps MME/S4-SGSN to prioritize paging.
- Enhanced paging KPI and VoLTE services.
- DDN message and mobility procedure so that DDN isn't lost.
- MBR guard timer, which is started when DDN Ack with temporary HO is received. A CLI command **ddn temp-ho-rejection mbr-guard-timer** has been introduced to enable the guard timer to wait for MBR once the DDN Ack with cause #110 (Temporary Handover In Progress) is received.
- TAU/RAU with control node change triggered DDNs.

In addition, to be compliant with 3GPP standards, support has been enhanced for Downlink Data Notification message and Mobility procedures. As a result, DDN message and downlink data which triggers DDN is not lost. This helps improve paging KPI and VoLTE success rates in scenarios where DDN is initiated because of SIP invite data.



Note For information on Downlink Data Notification (DDN) messages with support for DDN Delay and DDN Throttling, refer the *SAEGW Idle Buffering with DDN Delay and DDN Throttling* chapter in this guide.

For more information on how S-GW Paging Enhancement feature works, configuration, monitoring and troubleshooting, refer the *S-GW Paging Enhancements* chapter in the StarOS *S-GW Administration Guide*.

Session Recovery in User Plane

Support is added to recover the Session Manager process in the event of any crash. The recovered Session Manager has all the existing subscriber session on the recently crashed Session Manager process.

Uplink and Downlink data flow is processed on the newly recovered Session Manager process for all recovered subscriber sessions.

SRVCC PS to CS Handover Indication and the QoS Class Index IMS Media Configuration Support

This feature notifies the PCRF about the cause for PCC rule deactivation on Voice bearer deletion. This notification helps the PCRF to take further action appropriately.

This feature ensures the compliance for SRVCC. This feature also supports the PS-to-CS handover indication after release of the voice bearers.

SRVCC service for LTE lets a single radio User Equipment (UE) accessing IMS-anchored voice call services to switch from LTE network to Circuit Switched domain. The UE switches the network while it can transmit or receive on only one of the access networks then. The SRVCC service removes the need for a UE to have multiple Radio Access Technology (RAT) capabilities.

After handing over the PS sessions to the target, the source MME removes the Voice Bearers (VB). The MME removes the VB by deactivating the voice bearers. The MME bars the VB towards S-GW/P-GW and sets the VB flag of Bearer Flags IE in the Delete Bearer Command message (TS 29.274 v9.5.0).

If the IP-CAN bearer termination happens due to PS to CS handover. The PCEF reports the related PCC rules for this IP-CAN bearer by including the Rule-Failure-Code AVP set to the value: PS_TO_CS_HANDOVER (TS 29.212 v10.2.0 and TS 23.203 v10.3.0).

Support for new AVP PS-to-CS-Session-Continuity (added in 3GPP Release 11) inside Charging Rule Install indicates the bearer support for PS to CS continuity.

QCI IMS-Media Configuration Support

Specifies the QoS Class Index (QCI) value to mark the IMS media bearers for preferential treatment during session recovery and ICSR switchover.

Mode

Exec > Global Configuration > Context Configuration > APN Configuration

configure > context <*context_name* > **apn** <*apn_name*>

Entering the above command sequence results in the following prompt:

[context_name]host_name(config-apn)#

Syntax

qci *value_bytes* **ims-media**

no qci *value_bytes* **ims-media**



Note

- *no*: Disables this IMS QCI feature.
- *ims_media*: Marks bearers classified as IMS media for preferential treatment during session recovery and ICSR switchover.
- *value_bytes*: Specifies the QCI value an integer from 1 through 254.

Usage Guidelines

Use this command to specify the QCI value to be used to mark bearers classified as IMS media for preferential treatment during session recovery and ICSR switchover.

The following prerequisites apply to the implementation of this feature:

- A dedicated APN must be reserved for VoLTE traffic.
- A call connected to this APN will not be classified as Active VoLTE unless there is a dedicated bearer matching the VoLTE-configured QCI.

- Preferential treatment would be given to only those calls which are active VoLTE.
- A GGSN call connected to this APN will not be classified as Active VoLTE unless there is network initiated bearer matching the VoLTE-configured QCI.
- VoLTE marking is preserved across a Gn-Gp handoff.

When this feature is enabled via a CLI command, the actions are taken:

- During bearer creation
  - New bearer QCI is matched against APN configuration.
  - If the QCI matches an APN configuration, the bearer is marked for preferential treatment.
  - Flow_entries are modified with this information (if this is first VoLTE bearer).
  - Egtpu_session is updated with the VoLTE tag during a rx_setup request.
  - An indication message informs ECS about the VoLTE tagging.
- During bearer deletion
  - Flow_entry is updated with VoLTE information if this is the last VoLTE bearer.
  - ECS is informed of the deletion via an indication message.

The following command enables preferential treatment for IMS bearers with a QCI of 9:

qci 9 ims-media

Support for ip hide-service-address CLI Command

The **ip hide-service-address** CLI command is supported in CUPS.

When enabled, this CLI renders the IP address of the GGSN unreachable from mobile stations (MSs) using this APN. This command is configured on a per-APN basis.

Use the following configuration to enable or disable the feature.

```
configure
  context context_name
    apn apn_name
      [ default | no ] ip hide-service-address
      end
```

- **default**: Does not allow the mobile station to reach the GGSN IP address using this APN.
- **no**: Allows the mobile station to reach the GGSN IP address using this APN.
- Use this command to prevent subscribers from using traceroute to discover the network addresses that are in the public domain and configured on services.

Support for regardless-of-other-triggers CLI Command

This feature supports **regardless-of-other-triggers** option in CLI for CUPS. **regardless-of-other-triggers** option enables eG-CDR or P-GW-CDR generation at the fixed time interval irrespective of any other eG-CDR or P-GW-CDR triggers that may occur in between. Therefore, when you enable this option although other CDR triggers occur, the Time Limit CDR gets triggered dynamically at every *interval* in seconds, that is, the Time Threshold calculation is based on the sum of the last threshold time and the interval. This option supports session recovery and ICSR.

Use the following configuration to enable the feature.

```
configure
  active-charging service service_name
    rulebase rulebase_name
      egcdr threshold interval interval regardless-of-other-triggers
      end
```

The following steps are carried out when a new call comes in:

- When you enable, **regardless-of-other-triggers** even if any other usage report triggers in between, timer will not be reset, and the session usage report for the time threshold occurs for every interval time configured.

Show Commands and Outputs

This section provides information about the show CLI commands available in support of the feature.

- **show active-charging rulebase name** *name*
- **show active-charging rulebase all**

The output of these CLI commands includes the following fields to support this feature.

- Interval Threshold: <seconds> (secs) Regardless of Other Triggers

TFT Suppression for Default Bearer

Feature Description

TFT Suppression for default bearer is supported in the UPC CUPS architecture. Following CLI commands are added in support of this feature.

- **policy-control update-default-bearer**
- **no tft-notify-ue-def-bearer**

The preceding CLI commands are used to bind all the predefined rules received from PCRF without QoS and ARP or with the same QoS and ARP as that of the default bearer, to the default bearer.



Important This CLI is applicable to all the rulebases in the chassis configuration. If the rulebase is changed to some other rulebase in the interim period or anytime later, this CLI will continue to apply to the current new rulebase too.

Configuring TFT Suppression

Configuring TFT Suppression in Default Bearer for Predefined Rules

Use the following commands to configure TFT Suppression for default bearers.

```
configure
   require active-charging
   require active-charging service_name
      [ default | no ] policy-control update-default-bearer
      end
```

Caution Upon executing this CLI command "**no policy-control update-default-bearer**", system crash is likely to occur if the TFT information is not added to the charging-action.

Configuring TFT Suppression in Default Bearer

Use the following commands to configure TFT Suppression for default bearers.

```
configure
   require active-charging
   require active-charging service_name
      rulebase rulebase_name
         [ default | no ] tft-notify-ue-def-bearer
      end
```

Note

- **default:** Configures this command with its default setting.

  Disables only binding those rules having QoS of default bearer to the default bearer and specifies to not ignore other rules. Rules having respective QoS gets attached to the relevant bearers. Also, TFT updates towards UE (access side) is not suppressed.

- **no:** Enables binding rules having QoS of default bearer to the default bearer and specifies to ignore other rules.

  In case no QoS is specified the rule gets attached to default bearer. Also, TFT updates towards UE (access side) is suppressed for default bearer. So only one default-bearer is ever be created.

Zero-byte EDR Suppression

The Zero-byte Event Data Record (EDR) Suppression, a CLI-controlled feature, enables or disables creation of EDRs when there is no data for the flow. A zero-byte EDR is typically possible when two successive EDRs are generated for a flow. The CLI command suppresses the second such EDR for the flow.

Use the following configuration to enable or disable the suppression of zero-byte EDRs.

```
configure
   active-charging service service_name
      rulebase rulebase_name
```

```
[ default | no ] edr suppress-zero-byte-records
end
```

NOTES:

- **default**: Configures this command with its default setting.

  Default: Disabled; same as **no edr suppress-zero-byte-records**

- **no**: Disables the suppression of zero-byte EDRs.
- **edr suppress-zero-byte-records**: Suppresses zero-byte EDRs.
- The "Total zero-byte EDRs suppressed" field in the output of the following CLI command can be used to verify if the zero-byte EDRs are suppressed: **show user-plane-service statistics rulebase name** *rulebase_name*.

How It Works

This section describes the Call Flows for User Plane service.

Call Flows

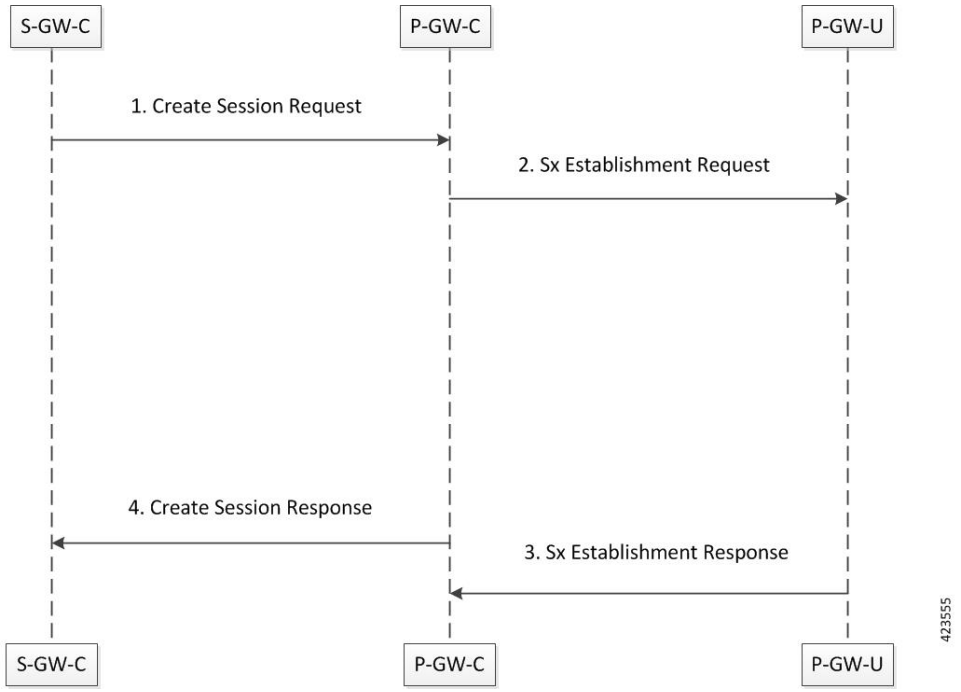
This section describes the User Plane Call Flows in the CUPS architecture.

P-GW Data Session

This section describes the P-GW initial attach procedure.

Initial Attach Procedure (Pure P)

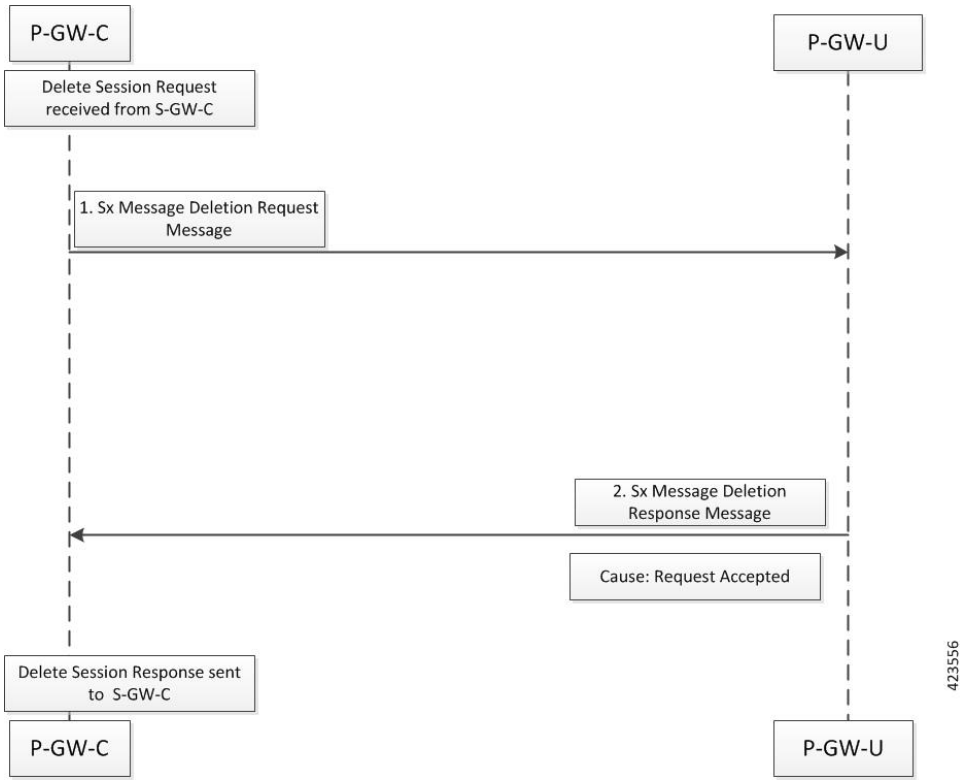
Following call flow illustrates, at a high-level, the initial attach procedure for a Pure-P PDN.



- P-GW receives a Create Session Request message including an APN, on the S5/S8 interface.
- P-GW-C initiates an Sx establishment request on Sxb interface towards selected P-GW-U with PRDs, FARs information to establish the data path. PGW-C does not support TEID (Tunnel Identifier) allocation; it is allocated by PGW-U.
- Once the resources are allocated (TEID and so on), P-GW-U sends an Sx establish response message towards P-GW-C.
- P-GW responds to the S-GW with a Create Session Response message including the assigned address, TEID, and additional information.
- The S5/S8 data plane tunnel is established and the PGW-U can forward and receive packets to and from the PDN.

Initial Detach Procedure (Pure P)

Following call flow illustrates, at a high-level, the initial detach procedure for a Pure-P PDN.



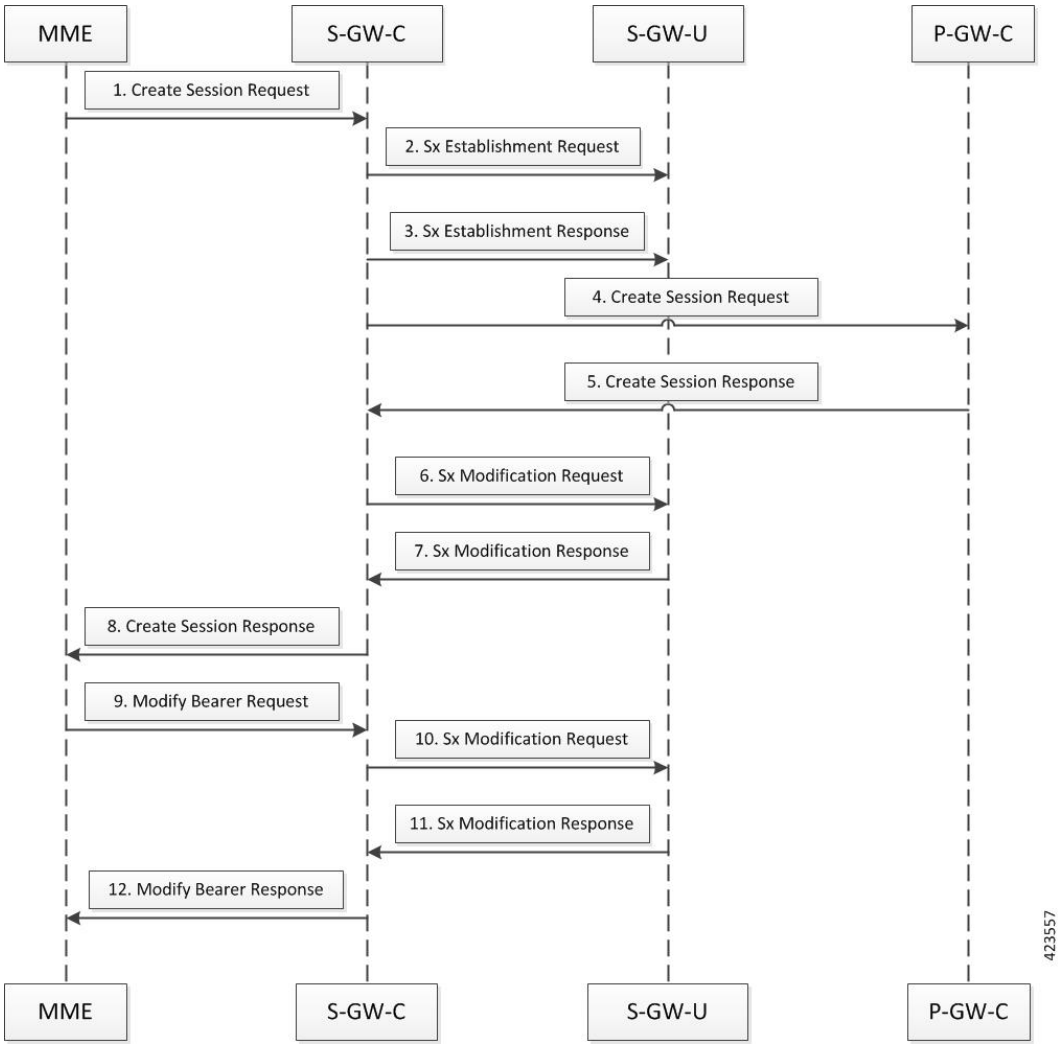
- P-GW receives a Create Session Request message including an APN, on the S5/S8 interface.
- P-GW-C initiates an Sx establishment request on Sxb interface towards selected P-GW-U with PRDs, FARs information to establish the data path. PGW-C does not support TEID (Tunnel Identifier) allocation; it is allocated by PGW-U.
- Once the resources are allocated (TEID and so on), P-GW-U sends an Sx establish response message towards P-GW-C.
- P-GW responds to the S-GW with a Create Session Response message including the assigned address, TEID, and additional information.
- The S5/S8 data plane tunnel is established and the PGW-U can forward and receive packets to and from the PDN.

S-GW Data Session

This section describes the S-GW initial attach procedure.

Initial Attach Procedure (Pure S)

Following call flow illustrates, at a high-level, the initial attach procedure for a Pure-S PDN.

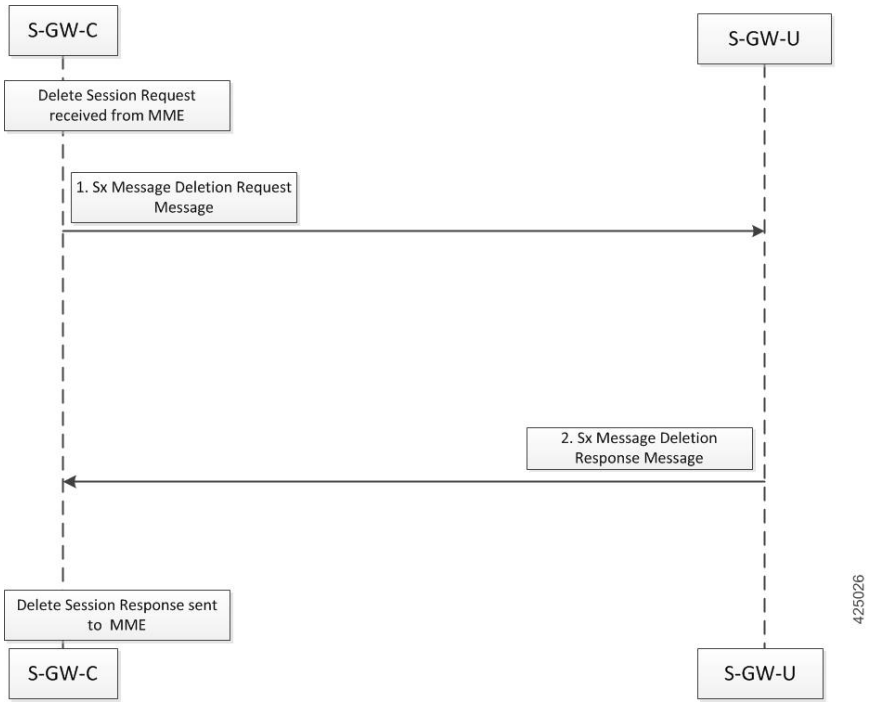


- On S11 interface, S-GW-C receives a Create Session Request message including an Access Point Name (APN) from MME.
- S-GW-C initiated the Sx establishment request on the Sxa interface towards the selected S-GW-U with PDRs, FARs information to establish data path. Here, the S-GW-C does not support the TEID (Tunnel Identifier) allocation. It is allocated by the S-GW-U.
- After allocation of resources such as egree TEID and so on, S-GW-U sends the Sx establishment response message towards the S-GW-C.
- The S-GW-C initiates the Create Session Request towards the selected P-GW-C.
- P-GW-C responds with the Create Session Response with the IP address and default bearer related information.
- SGW-C initiates Sx Modification Request message towards SGW-U to update FAR (Forwarding action) information for the existing session.
- SGW-U provides Sx Modification Response with success after updating information.
- SGW-C sends Create Session Response with all necessary information for Default Bearer towards MME.

- MME initiates Modify Bearer Request message towards SGW-C, once it received eNodeB's F-TEID information.
- SGW-C initiates Sx Modification Request towards SGW-U for updating FAR information on eNodeB's F-TEID.
- After successfully updating, Sx Modification Response is sent to SGW-C.
- SGW-C inturn will send Modification Response message towards MME to complete attach procedure.
- SGW-U has established S1U side data tunnel towards eNodeB and S5/S8 side data tunnel towards PGW-U. Now, SGW-U can forward and receive packets to/from PGW as well eNodeB.
- The S5/S8 data plane tunnel is established and the PGW-U can forward and receive packets to/from the PDN.

Initial Detach Procedure (Pure S)

Following call flow illustrates, at a high-level, the initial detach procedure for a Pure-S PDN.



- Once Delete Session Request is received from MME, SGW initiate Sx Delete Request message towards SGW-U.
- SGW-U clears all allocated User-Plane resources and responds back with Cause Success to SGW-C.
- SGW-C responds back to MME with Delete Session Response message.

Support for Addition, Deletion and Updation of Dedicated Bearers for S-GW

Feature Description

Addition, Deletion and Updation of Dedicated Bearers for Pure-S calls is supported in the CUPS architecture.

The following functionality is added in support of this feature:

- SAEGW-CP supports Create Bearer Request for dedicated bearer for Pure-S Call.
- SAEGW -CP supports multiple bearer contexts in single Create Bearer Request.
- SAEGW-CP supports multiple Create Bearer request in parallel for different PDN; these PDN can be Pure-S PDN or Collapsed and Pure-S combinations.
- SAEGW-UP creates uplink and downlink bearer stream at VPP for Pure-S call per bearer. Number of streams per direction depends on the GTP-U Service IP address.
- SAEGW-CP supports Release Access Bearer Request (RAB) with dedicated bearer, all FAR corresponding to all bearer is modified.
- SAEGW-CP supports Modify Bearer Request (Idle mode, Connected mode) with dedicated bearer.
- SAEGW-CP supports Create Bearer Response Failure handling from MME.
- SAEGW-CP and SAEGW-UP supports DSCP marking for default and dedicated bearer with VPP.
- SAEGW-CP and SAEGW-UP supports Delete Bearer Request for dedicated bearer. SAEGW-UP removes bearer stream and TEP entries belonging to those bearers.
- SAEGW-CP supports Pure-S Dedicated Bearer Creation when call is in IDLE state.
- SAEGW-CP supports Pure-S Dedicated Bearer S-GW Relocation (both X2 and S1-based).
- SAEGW-CP supports Pure-S Dedicated Bearer Update success scenarios.
- SAEGW-CP supports Piggybacking of Create Bearer Request for dedicated bearer for Pure-S call along with Create Session Response.
- SAEGW-CP supports Piggybacking of Create Bearer Response for dedicated bearer for Pure-S call with Modify Bearer Request.
- SAEGW-CP supports Pure-S Dedicated Bearer Creation if P-GW receives bearer creation as part of CCA-I, where P-GW does not send Piggyback request, which results in Create Session Response followed by Create Bearer Request.
- SAEGW-CP supports Session Recovery and ICSR with Pure-S dedicated bearer.
- SAEGW-CP supports Create Bearer Request and Delete Bearer Request (default bearer) collision.
- SAEGW-CP supports Create Bearer Request and Delete Session Request collision.
- SAEGW-CP supports Create Bearer Response and Delete Bearer Request (default bearer) collision.
- SAEGW-CP supports Create Bearer Response and Delete Session Request collision.
- SAEGW-CP supports End Marker with Pure-S default and dedicated bearer.
- SAEGW-UP supports Session Recovery with Pure-S default and dedicated bearer.

- SAEGW-UP supports movement of IP transport from IPv4 to IPv6, or IPv6 to IPv4, during IDLE->Active and Handover procedure on S1U interface. Transport selected on S1U at the time of Attach is supported. For example, eNode handover from IPv4 eNodeB to IPv6 eNodeB will work.
- SAEGW-CP supports CBRsp with Cause Partially Accepted and Context Not Found.
- SAEGW-CP supports Downlink Data Notification for Pure-S Call, so when UE moves to IDLE state for Pure-S call, FAR action is set as BUFFER.
- SAEGW-CP supports Update Bearer Response with cause PARTIALLY_ACCEPTED and context not Found.
- SAEGW-CP supports the Error and Failure handling from other peer nodes including User Plane node.

Limitations

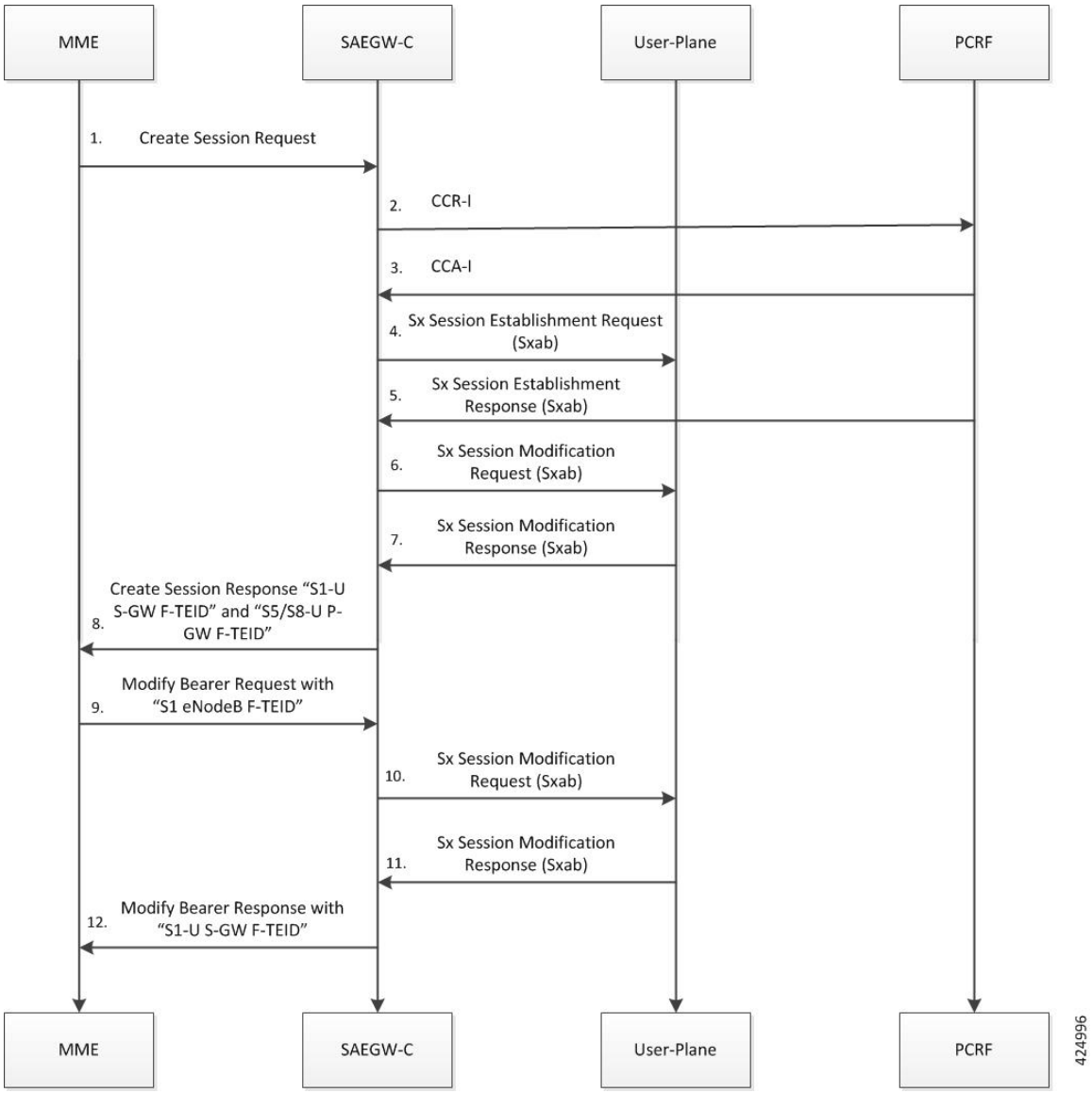
For Pure-S calls, Idle Session timeout is not supported.

Support for Collapse Call

Following call flow illustrates, at a high-level, the detach procedure for UE initiated Collapsed PDN.

Initial Attach Procedure (Collapsed PDN)

The following call flow illustrates, at a high-level, the initial attach procedure for Collapsed PDN.



1. For CUPS SAEGW collapsed call, SAEGW-C does the following:
   - After Gx interaction, performs Gx communication (CCR-I and CCA-I).
   - Performs User Plane selection based on **user-plane-profile** configured with IP Pool (APN associated with IP Pool).
   - Establishes GTP-U session (required for RA/RS, in case of IPv6/IPv4v6 PDN).
   - Performs Sxab interaction with selected User Plane.
2. Sx Establishment Request contains the following information:
   - Create PDR/FAR information for S-GW Uplink and Downlink data path (Sxa Type PDR).
   - Create PDR/FAR/URR information for Uplink and Downlink data path (Sxb Type PDR): For dynamic/pre-defined/static rules.

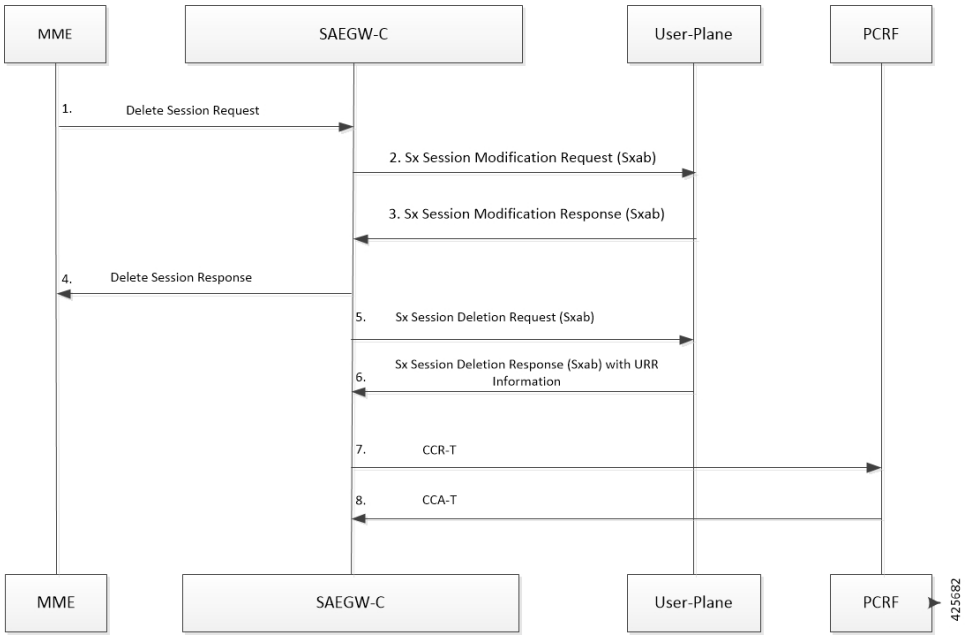
- Create PDR/FAR for RA/RS (Sxb Type PDR): Required for IPv6/IPv4v6 PDN Type.
- Additionally, Control Plane requests User Plane to allocate F-TEID for:
  - S-GW Ingress "S1-U S-GW F-TEID",
  - S-GW Egress "S5/S8-U S-GW F-TEID", and
  - P-GW Ingress PDR "S5/S8-U P-GW F-TEID"

3. User Plane provides following information as part of Sx Session Establishment Response:
   - Created PDR: S-GW Ingress PDR "S1-U S-GW F-TEID",
   - Created PDR: S-GW Egress PDR "S5/S8-U S-GW F-TEID", and
   - Created PDR: P-GW Ingress PDR "S5/S8-U P-GW F-TEID"
4. On receipt of successful Sx Session Establishment Response, the Control Plane triggers Sx Modification Request with the following information:
   - To update P-GW (Sxb) "Uplink PDR" with "Outer Header Removal" based on IP address information in "S5/S8-U S-GW F-TEID"
   - To update P-GW (Sxb) "Downlink FAR" with "Outer Header Creation" as "S5/S8-U S-GW F-TEID"
   - To update S-GW (Sxa) "Uplink FAR" with "Outer Header Creation" as "S5/S8-U P-GW F-TEID"
   - To update S-GW (Sxa) "Downlink PDR" with "Outer Header Removal" based on IP address information in "S5/S8-U P-GW F-TEID"
5. On receipt of Sx Session Modification Response, the SAEGW-C sends Create Session Response toward MME with "S1-U S-GW F-TEID" and "S5/S8-U P-GW F-TEID".
6. On receipt of Modify Bearer Request (MBR), the SAEGW-C does the following:
   - Trigger Sx Session Modification Request:
     - To update Downlink FAR with "Outer Header Creation" as "S1 eNodeB F-TEID".
     - To update Uplink PDR with "Outer Header Removal" based on IP address information in "S1 eNodeB F-TEID".
7. On receipt of Sx Session Modification Response, the SAEGW-SGW-C sends MBR with "S1-U S-GW F-TEID".

Initial Detach Procedure (Collapsed Call)

Detach Procedure (Collapsed): UE Initiated

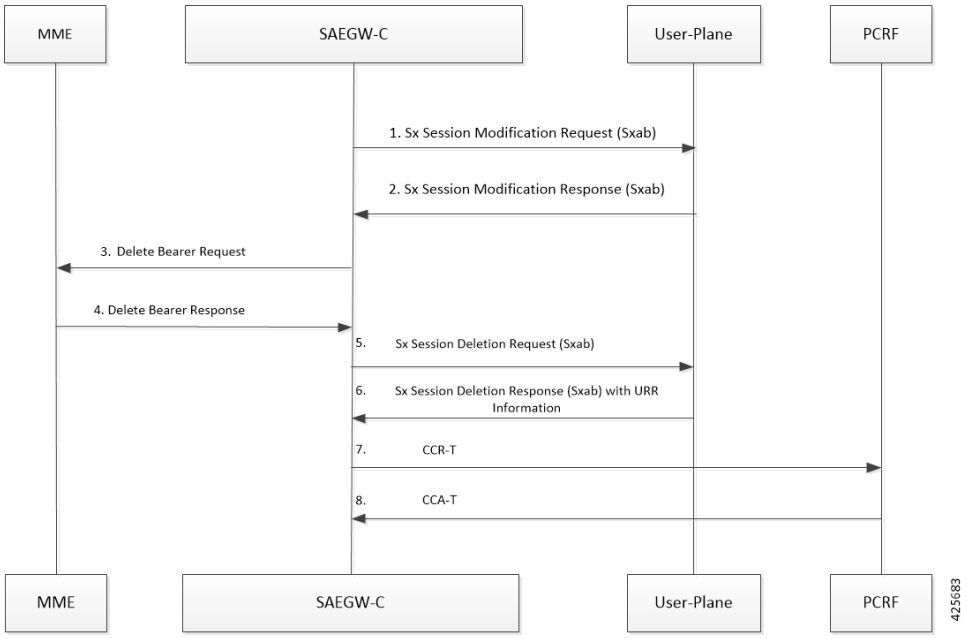
Following call flow illustrates, at a high-level, the detach procedure for UE initiated Collapsed PDN.



1. On receipt of Delete Session Request, the SAEGW-C performs Sxab interaction to update FAR with Apply Action as "DROP" for both Uplink and Downlink data path.
2. On receipt of Sx Session modification Response, SAEGW-C sends Delete Session Response towards MME.
3. For CUPS SAEGW Collapsed call, the SAEGW-C does the following:
   - Removes GTP-U session (required for RA/RS in case of IPv6/IPv4v6 PDN).
   - Performs Sxab interaction with the selected User Plane.
4. On receipt of Sx Session Deletion Response, the SAEGW-C does the following:
   - Performs Gx communication (CCR-T and CCA-T).
   - Generates CDR (Gz) based on URR information received.

Detach Procedure (Collapsed): Network Initiated

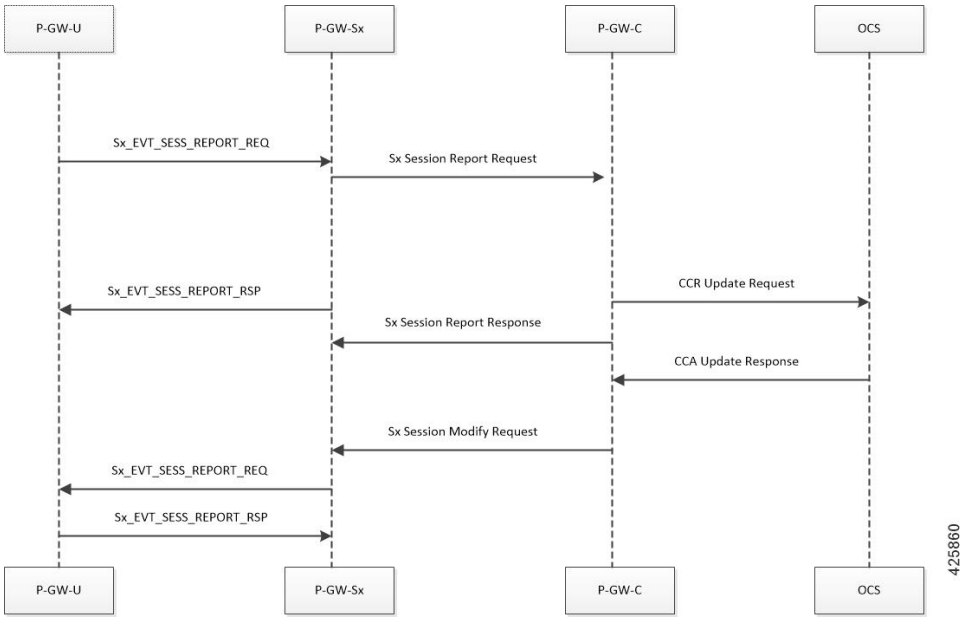
Following call flow illustrates, at a high-level, the detach procedure for network initiated Collapsed PDN.



1. On receipt of Delete Bearer Request (RAR initiated or by the **clear sub all** CLI), SAEGW-C performs Sxab interaction to update FAR with Apply Action as "DROP" for both Uplink and Downlink data path.
2. On receipt of Sx Session Modification Response, SAEGW-C sends Delete Bearer Request toward MME.
3. For CUPS SAEGW Collapsed call, the SAEGW-C does the following:
   - Removes GTP-U session (required for RA/RS in case of IPv6/IPv4v6 PDN).
   - Performs Sxab interaction with the selected User Plane.
4. On receipt of Sx Session Deletion Response, the SAEGW-C does the following:
   - Performs Gx communication (CCR-T and CCA-T).
   - Generates CDR (Gz) based on URR information received.

P-GW Session Reporting with Gy Interface

This section describes P-GW session reporting with Gy interface.



URR Support in Session Establishment Request

- User plane module supports the storage of a list of URRs received as part of session establishment request.
- Each PDR is associated with one or more URRs.
- A particular URR is linked to another URR.
- Each URR contains the measurement method (time or volume), and reporting triggers that indicates the event on which the user plane has to send usage report.
- The URR have both volume-quota and volume-threshold present for the Gy-URRs.

Session Delete Response

This message sent from the User Plane is in response to a session deletion request from control plane. This results in the termination of the Sx session at User plane. Usage Report is included as part of Sx Delete Session Response.

Session Report Request and Response Message

Request Message

- On encountering a time or volume threshold limit, user plane generates an Sx Session Report Request message and sends the same to control plane.
- This message contains the Usage report, which indicates the reason for generating the message, specified by Usage Report Trigger.
- In addition to this, the Usage report contains the time or volume measurement.
- If any other URRs are linked to the URR for which the session report request is being generated, then a session report request is generated for those linked URRs as well. For this release, Gy-URRs are not linked with any of the URRs.

Response Message

This message from the Control plane indicates a successful delivery of the Session Report Request message with a cause code. Currently no specific failure handling is done on receiving a failure cause.

Server-Unreachable Support for Gy

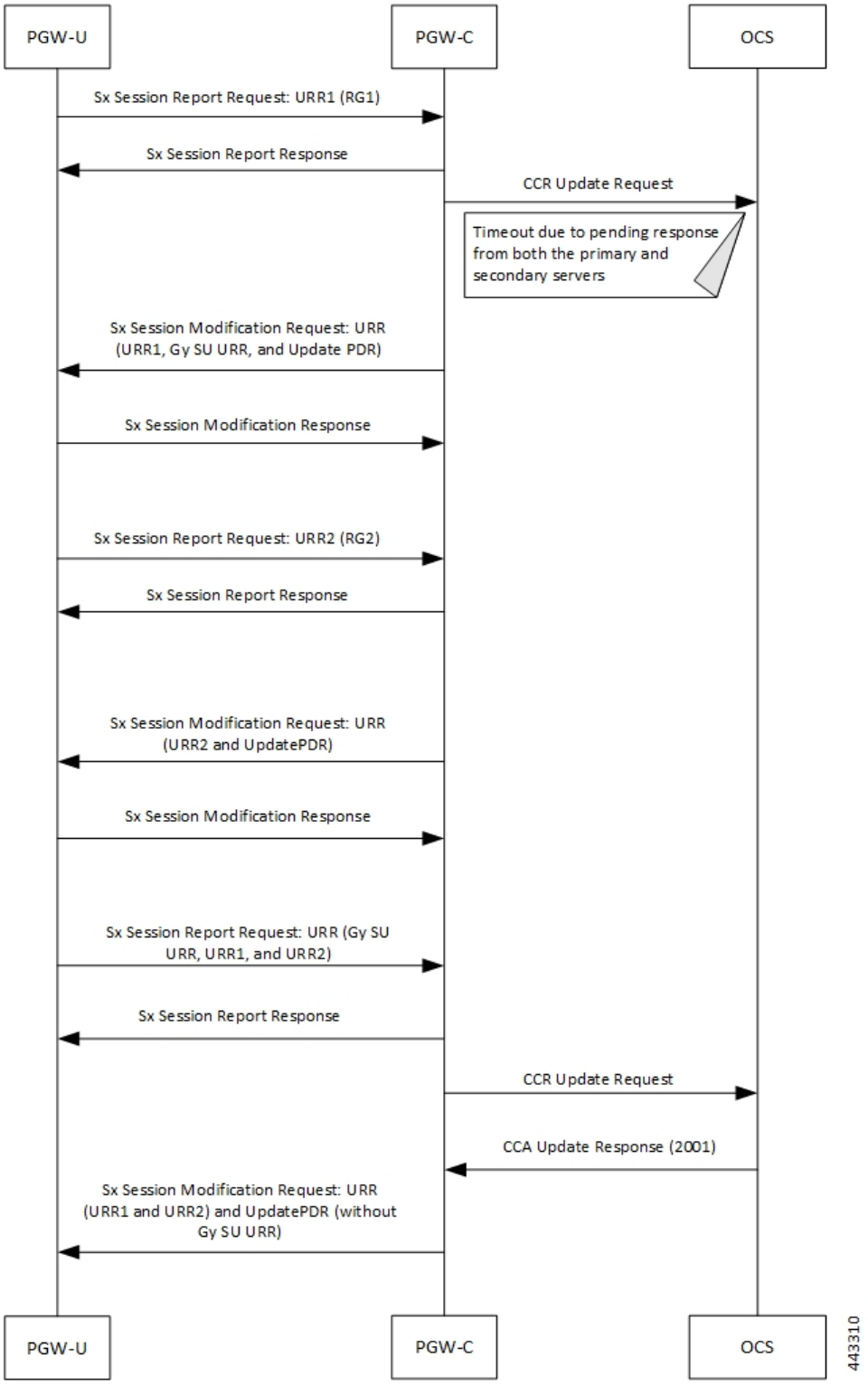
The Server-Unreachable (SU) mechanism is configured on the Control-Plane (CP), for the Gy interface in order to resolve issues that are encountered on the Online Charging System (OCS) or with the connectivity between Policy and Charging Enforcement Function (PCEF) and OCS. The SU configuration provides the options to continue the session even after a failure by providing the option to use configurable interim quota (volume and/or time) and configurable server retries before a session is converted to offline or terminated.

A new Usage Reporting Rule (URR) bucket is created, which contains the SU quota when a Gy session goes into the SU state. The ID for the new URR is generated dynamically when the SU URR is allocated.

In a CUPS User Plane (UP) node, the existing Vector Packet Processor (VPP) streams are modified with a new LC record, which contains the updated SU URR bucket along with the existing set of charging buckets.

When the VPP streams are in the SU state, two quota rows are available, GyURR and SU URR. When GyURR is in the SU state with linked-usage-reporting trigger set, the quota row for the SU URR is linked to the VPP streams.

This section describes the SU call flow in CUPS.


PGW-U
PGW-C
OCS
Sx Session Report Request: URR1 (RG1)
Sx Session Report Response
CCR Update Request
Timeout due to pending response from both the primary and secondary servers
Sx Session Modification Request: URR (URR1, Gy SU URR, and Update PDR)
Sx Session Modification Response
Sx Session Report Request: URR2 (RG2)
Sx Session Report Response
Sx Session Modification Request: URR (URR2 and UpdatePDR)
Sx Session Modification Response
Sx Session Report Request: URR (Gy SU URR, URR1, and URR2)
Sx Session Report Response
CCR Update Request
CCA Update Response (2001)
Sx Session Modification Request: URR (URR1 and URR2) and UpdatePDR (without Gy SU URR)
PGW-U
PGW-C
OCS
443310

Step	Description
1	PGW-U sends a Sx Session Report Request message with URR1 (RG-1) to PGW-C due to triggers like Time and Volume or Quota and Threshold.
2	PGW-C acknowledges the request and sends a Sx Session Report Response message to P
3	PGW-C sends CCR Update (CCR-U) Request message to both the primary and secondar
4	When the CCR-U Request messages fail at both primary and secondary OCS, the Gy ses enters an SU state. The SU URR is created for the Gy session, and it's linked to the relevan PGW-C sends an Sx Session Modification Request message to PGW-U with UpdatePDR includes Gy SU URR in the PDR URR list.
5	PGW-U starts updating the usage in both the Gy buckets (URR1 and Gy SU URR1) and an Sx Session Modification Response message to PGW-C.
6	PGW-U sends another Sx Session Report Request message with URR2 (RG-2) to PGW- the Gy URR bucket exhausts the quota .
7	PGW-C acknowledges the request sends a Sx Session Report Response message to PGW
8	PGW-C sends an Sx Session Modification Request message to PGW-U with UpdatePDR Update RR2. The UpdatePDR has a modified URR list, which contains both URR2 and URR.
9	PGW-U sends an Sx Session Modification Response message to PGW-C.
10	PGW-U sends a Sx Session Report Request message with URR1 (RG-1) and URR2 (RG PGW-C after the Gy SU URR quota is exhausted.
11	PGW-C acknowledges the request and sends a Sx Session Report Response.
12	PGW-C sends CCR Update (CCR-U) Request message to OCS after an SU retry.
13	OCS sends a CCA Update Response message with the Result-Code as 2001.
14	PGW-C sends an Sx Modification Request message with URR1 (RG-1), URR2 (RG-2), UpdatePDR (without Gy SU URR) to PGW-U.

New Behavior in CUPS

The new SU mechanism in CUPS, are as follows:

- In a non CUPS architecture, where a single node (P-GW) processes the Gy session state and the data-traffic, an SU URR is created without any messaging delay. However, in the CUPS mode, the CP forms an additional node, which maintains information about the session state and handles any URR requests from the User Plane (UP). Only the CP can associate a Gy session with an SU URR. This messaging between UP and CP causes a delay and the data packets are treated according to the Pending-Traffic-Treatment configuration to complete the communication.
- In a non CUPS architecture, the SU state timer is processed in a different manner compared to the Time-Quota timer. After an SU quota is exhausted, the retry attempt to OCS occurs and a new next-interim-time-quota is started. However, in the CUPS mode, when the SU Time Quota is used and it is reported to CP for the Quota Exhaust, and if the session goes into Server-Unreachable state again, the time elapsed from the last Usage-Report is accounted in the usage.

- It's not recommended to use the **servers-unreachable after-timer-expiry** *timeout_period* CLI command in CUPS. Instead, use the **servers-unreachable after-interim-time** *timeout_period* **server-retries** *retry_count* to achieve similar behavior but with a single retry (set *retry_count* to 1).

Limit-Reached Postprocessing

Limit-reached-post-processing is a non-3GPP, proprietary behavior supported in both CUPS and non-CUPS architecture. This feature allows redirection or restriction operation implemented when the quota is exhausted for the charging-bucket, however, the OCS server is unable to grant the FUI-Redirect or FUI-Restrict. When using this feature, the operator can combine all the rule-matching criteria that are available—for example, to enable IMSI-based matching criteria, and so on—to selectively apply different handling for different subscribers/traffic. Use the following CLI commands to enable the feature.

```
configure
  active-charging service service_name
    rulebase rulebase_name
      post-processing policy always
      end
```

Also, **rule-application post-processing** CLI command must be configured as limit-reached under ACS Ruledef configuration mode.

PTT no-quota Limited Pass

This feature allows the subscriber to use the network while waiting for the response from OCS. The Limited-Pass configuration allows to specify the Volume which the subscriber can consume while waiting for the quota-response from OCS. The usage is accounted in the respective charging bucket and are adjusted against the next-quota allocation.

Use the following CLI commands to enable the feature:

```
configure
  active-charging service service_name
    credit-control
      pending-traffic-treatment noquota limited-pass volume volume
      end
```

Limited Pass Volume is used only for **noquota** case (Rating Group (RG) seeking quota for the first time) and not for **quota-exhausted**. Limited Pass Volume isn't used for subsequent credit requests.

The traffic is allowed to pass until the Limited-Pass Volume gets exhausted. The usage is counted in the respected charging-bucket and adjusted against the "Quota" granted. If the "Quota" allocation is less than the actual usage, immediate reporting towards OCS with the usage-report occurs requesting for more quota allocation. The subsequent incoming packets are handled as per the "quota-exhausted" PTT configuration.

If the Limited Pass Volume is NOT exhausted before the OCS responds with denial of quota, traffic is blocked after the OCS response. The gateway reports usage on Limited-Pass Volume even in for CCR-U (FINAL) (in non-CUPS) or CCR-T (for CUPS) until the OCS responds.

If the Limited Pass Volume is exhausted before the OCS responds, then the subsequent incoming packets for the session are dropped until quota is granted from OCS.

The default pending-traffic-treatment for **noquota** is Drop. The **default pending-traffic-treatment noquota** command removes any Limited Pass Volume size configured.

PTT Quota-Exhausted Limited Pass

Pending-Traffic-Treatment (PTT) Quota-Exhausted Limited-Pass in CUPS architecture is an alternative to the Buffering option. The Buffering option has practical limitations in the high-speed network. Buffering requires packet buffering for large number of packets at the gateway, causing the risk to run out of memory and affecting the bandwidth speed. The PTT Quota-Exhausted Limited Pass allows the traffic to pass through until it reaches the configured limit on the Quota-Exhaust scenarios.

The PTT allows the traffic until the Limited-Pass volume exhausts. The PTT counts and adjusts the usage in the respected charging-bucket against the granted "Quota". If the "Quota" allocation is less than the actual usage, there's immediate reporting towards OCS with the usage-report and asking for more quota allocation.

If the Limited-Pass Volume doesn't exhaust before the OCS responds with denial of quota, there's traffic blockage after the OCS response. Gateway reports the usage in CCR-U (FINAL).

If the Limited-Pass Volume exhausts before the OCS responds, then further incoming packets for the session are dropped until quota is granted from OCS.

The default behavior of pending-traffic-treatment for quota-exhausted is Drop. The default pending-traffic-treatment quota-exhausted CLI command removes any configured Limited-Pass Volume size.

Use the following CLI command to enable the feature:

```
configure
  active-charging service service_name
    credit-control
     pending-traffic-treatment quota-exhausted limited-pass volume volume

       end
```



Note The above CLI command is applicable only in CUPS architecture.

NOTES:

- **limited-pass**: Enables limited access to subscriber when OCS is unreachable.
- **volume** *volume*: Enables limited volume access to subscriber when OCS is unreachable. *volume* specifies the Default Quota size (in bytes) and must be an integer from 1 through 4294967295

Quota-Validity-Time Handling

When the Quota-Validity-Time is received for an MSCC bucket, the same is sent to the User Plane. Since there is no specific IE that can be used directly, the QVT value is filled in the Time-Quota IE, and URR is sent to the User Plane. The lesser QVT or Time-Quota is set in the Time-Quota IE. And, the Usage-Reporting from the User Plane for the Time-Quota trigger, the interpretation is made and the CCR-Update for the Validity-Timeout is generated.

Supported Functionality and Limitations

Basic call flow with Volume-Quota mechanism is supported with the following limitations on P-GW session reporting for Gy interface:

- Only CCR/CCA-I , CCR/CCA-U and CCR/CCA-T, RAR/RAA messages are supported.
- Dynamic Rules with Online Enabled is supported; both at Session-Setup and Mid-Session.

- Predefined Rules (dynamic-only) is supported; both at Session-Setup and Mid-Session. No restriction on configuring the "preemptively request".
- Static-rules with Online Charging are supported.
- Ignore-service-id is supported.
- Volume-Quota/Volume-Threshold mechanisms are supported.
- Event-Triggers (through which the Query URR occurs), and sending of usage information to the OCS is supported.



Important RAT-change functionality is not validated for this release.

- The "updateURR" procedure, through the Sx-Session-Modification procedure where the OCS grants a fresh Quota, is supported.
- Bearer-Level Gy and Subscriber-Level Gy is supported.
- Pending-Traffic-Treatment (PTT) Drop/Pass is supported with following limitations:
  - The scenarios supported for now are no-quota and quota-exhausted.
  - The trigger/re-authorization scenarios are not supported.
  - The PTT action (Forward/Drop) is considered after the quota-get is exhausted.
- Failure scenarios are qualified, which includes:
  - Failure-Handling Terminate, Continue and Retry, and Terminate: With CC-Group/FHT
  - Handling for the Error-Result-Codes (both at MSCC and Command level) is supported.
- Wall-Clock time-quota mechanism is supported.
- Other Time Quota Mechanisms (Discrete Time Period and Continuous-Time-Period) are not supported.
- Final-Unit-Indication Terminate mechanism is supported.
- FUI-Restrict is not supported.
- Mid-Session Rule Installation/Removal/Modification is supported.
- RAR mechanism is supported.
- Server-Unreachable (SU) mechanism is now supported with minor change in behavior compared to non-CUPS P-GW.
  - When an URR needs quota at UP, the usage-report is generated to CP and until the CP responds with the linked SU_URR, the packets matching this URR are treated with Pending-Traffic-Treatment configuration.
  - When the SU Time Quota is used and it's reported to CP for the Quota Exhaust, and if the session goes into Server-Unreachable state again, the time elapsed from the last Usage-Report is accounted in the usage.
- Pending-Traffic-Treatment Buffer mechanism is not supported.

- The “send-ccri on traffic-start” is supported.
- Quota-Hold-Time is supported.
- Quota-Consumption-Time mechanism is not supported.
- Quota-Validity-Time is supported.
- Triggering Gz records from Gy, when any event in Gy occurs, is supported; Gy-Gz sync is not supported.
- Triggering Rf records from Gy, when any event in Gy occurs, is not supported.
- Configuring different "rating-group" value other than the "content-id" is supported.
  - The RG 0 is not supported.
- Trigger to PCRF for the Out-of-Credit, Reallocation-of-Credit events are not qualified.



Important Event-trigger Out-of-Credit towards PCRF is validated with a limitation of having only one time Grant-Quota (Keeping Total Volume and Granted Volume at same value).

- The delayed response from OCS for the CCR-I is supported.
- Service-Specific-Units are not supported.
- Tariff-Time change is supported as per 3GPP specification.
- Quota-Retry Timer is supported.
- The **diameter mscc-final-unit-action terminate session** CLI command under Credit Control Configuration mode is supported.
- FUI-Redirect is supported with following limitations:
  - Redirection for HTTPs is not supported.
  - The FUI-Redirect with Filter-IDs/Filter-Rules are not supported.
  - The WSP Protocol is not supported.
  - In accordance with 3GPP specification, the Redirected-Traffic also gets redirected if it hits the rule that is in FUI-Redirect. There is no provision to allow the redirected-traffic to pass through.
    - In accordance with 3GPP specification, the CUPS architecture adheres to **no diameter fui-redirected-flow allow** CLI command behavior.
  - The **redirect-require-user-agent** CLI command is not supported; the redirection continues to work even if the user-agent is not present.
  - Appending the original URL is not supported.
  - The **diameter redirect-validity-timer immediate** CLI command is supported. However, **diameter redirect-validity-timer traffic-start** CLI command is not supported.
  - Token based mechanism, to come out of Redirection, is not supported. To end the redirection in CUPS, OCS sends Redirect Validity-Time or RAR.

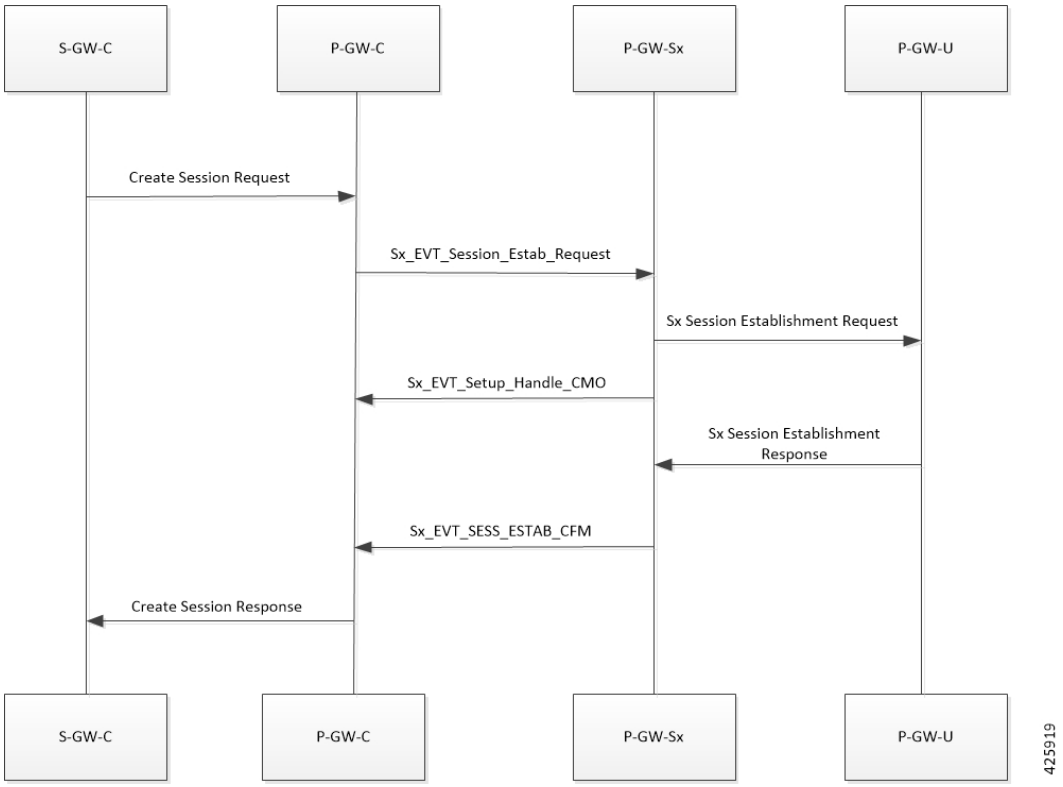
- FUI-Redirection is supported only for the URL, similar to the behavior in non-CUPS architecture.

- Rulebase change from PCRF/OCS is supported.

P-GW Session Reporting with Gz Interface

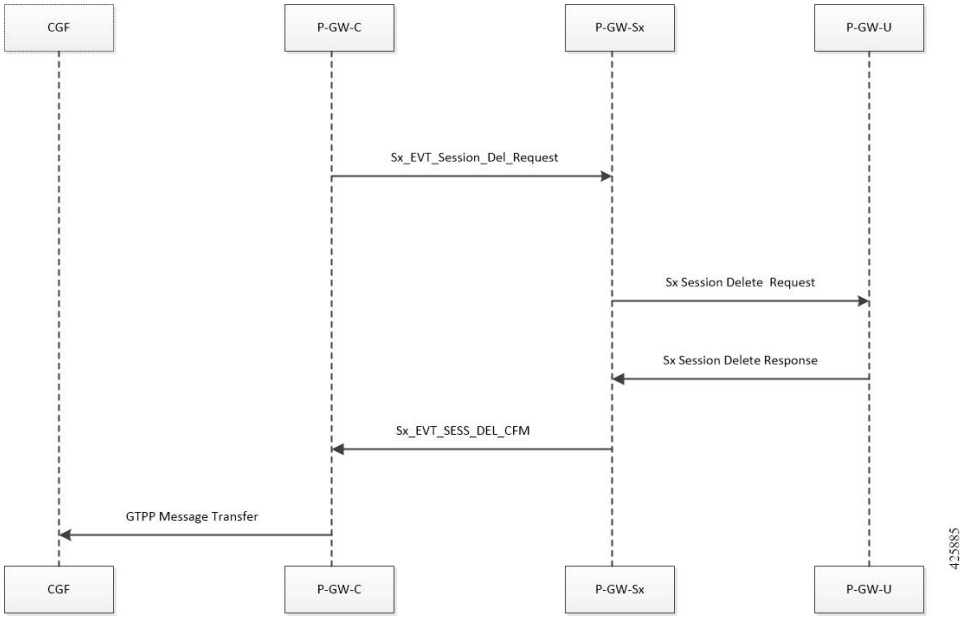
This section describes P-GW session reporting with Gz interface.

URR Support in Session Establishment Request



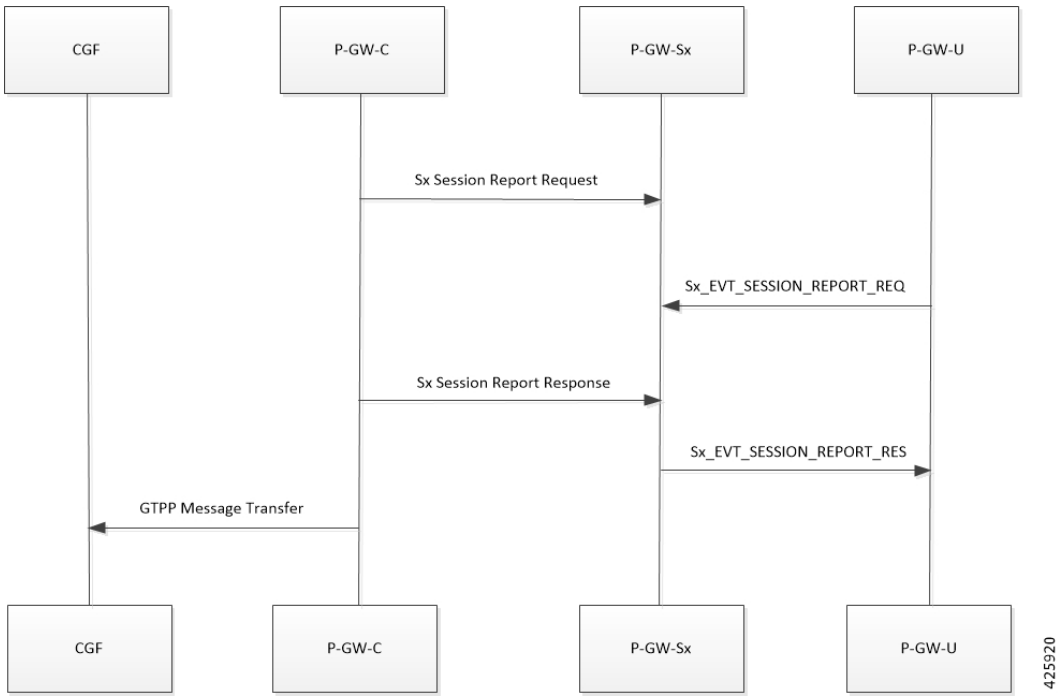
- User plane module supports the storage of a list of URRs received as part of session establishment request.
- Each PDR is associated with one or more URRs.
- A particular URR is linked to another URR.
- Each URR contains the measurement method (time or volume), and reporting triggers that indicates the event on which the user plane has to send usage report.

Session Delete Response



This message sent from the User Plane is in response to a session deletion request from control plane. This results in the termination of the Sx session at User plane. Usage Report is included as part of SX Delete Session Response.

Session Report Request and Response Message



Request Message

- On encountering a time or volume threshold limit, user plane generates an Sx Session Report Request message and sends the same to control plane.
- This message contains the Usage report, which indicates the reason for generating the message, specified by Usage Report Trigger.
- In addition to this, the Usage report contains the time or volume measurement.
- If any other URRs are linked to the URR for which the session report request is being generated, then a session report request is generated for those linked URRs as well.

Response Message

This message from the Control plane indicates a successful delivery of the Session Report Request message with a cause code. Currently no specific failure handling is done on receiving a failure cause.

Bit Rate Mapping Support

P-GW converts the bit rate value that it receives from PCRF from bps to kbps. This conversion may lead to truncation of fractional value to nearest integer (floor) value and lead to loss of information. 3GPP suggested that if the conversion from bps to kbps leads to a fractional value, then it should be rounded up to the nearest integer value (ceil) value and sent to the access side.

Note Design changes are done to ensure rounded down (floor) value from bps to kbps is sent on the PFCP interface.

Standards Compliance

The bit rate mapping feature complies with 3GPP TS 29.274 release 12.

Configuring the Bit Rate Mapping Feature

To configure the rounded up (ceil) value for bit rate from bps to kbps in APN-AMBR, GBR, and MBR on P-GW, perform the following steps:

```
configure
  context context_name
    pgw-service service_name
      [ no ] egtp bitrates-rounded-down-kbps
      end
```

To configure the rounded down (floor) value for bit rate from bps to kbps in APN-AMBR, GBR, and MBR on P-GW, perform the following steps:

```
configure
  context context_name
    pgw-service service_name
      egtp bitrates-rounded-down-kbps
      end
```

New Behavior in CUPS

By default, the rounded up value of bit rate in kbps for APN-AMBR, MBR, and GBR will be sent on the Sx and GTP interfaces. To enable the rounding down behavior, CLI must be configured.

Standards Compliance

The User Plane in CUPS complies with the following standards:

- 3GPP specification 23.214 release 14.0: Universal Mobile Telecommunications System (UMTS); LTE; Architecture enhancements for control and user plane separation of EPC nodes
- 3GPP specification 29.244 release 14.0: LTE; Interface between the Control Plane and the User Plane of EPC Nodes
- 3GPP specification 23.401 release 14.0: 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access