Implementierung von Overload Protection für Gateways und benachbarte Netzwerkelemente der Serie ASR5x00

Inhalt

Einführung

In diesem Dokument wird beschrieben, wie Sie die Schutzfunktionen implementieren, die für Gateways (GWs) und benachbarte Netzwerkelemente auf der Cisco Aggregated Services Router (ASR) 5x00-Serie verfügbar sind, um die Netzwerkleistung insgesamt zu schützen.

Überlastungskontrolle für GWs

Die Überlastungskontrolle ist eine allgemeine Selbstschutzfunktion. Sie wird verwendet, um das System vor dem Anstieg der Auslastung dieser Ressourcen zu schützen:

- CPU-Auslastung auf Verarbeitungskarten
- Speichernutzung auf Verarbeitungskarten

Wenn die Auslastung die vordefinierten Grenzwerte überschreitet, werden alle neuen Aufrufe (PDP-Aktivierungen (Packet Data Protocol), PDN-Sitzungsaktivierungen (Packet Data Network) *verworfen* oder *abgelehnt*, je nach Konfiguration.

Das folgende Beispiel zeigt, wie die allgemeine Nutzung der Data Processing Card (DPC) überwacht wird:

```
congestion-control threshold system-cpu-utilization 85

congestion-control threshold system-memory-utilization 85

congestion-control policy ggsn-service action drop

congestion-control policy sgw-service action drop

congestion-control policy pgw-service action drop
```

> **Hinweis**: Das System-Engineering-Limit beträgt 80% der CPU-Auslastung, was als empfohlener Engineering-Limit definiert ist, der nicht überschritten werden darf, um den regulären Betrieb des Systems zu gewährleisten. Eine über den Wert hinausgehende Last kann sich auf den Betrieb der Plattform auswirken, z. B. auf die Stabilität und Vorhersehbarkeit, und sollte bei angemessener Kapazitätsplanung vermieden werden.

> **Hinweis**: Cisco empfiehlt, die *Drop*-Aktion anstelle der *Ablehnungsaktion* zu verwenden, da die abgelehnten Anrufe sofortige wiederholte Verbindungsversuche der Benutzergeräte (UE) verursachen. Bei einer Drop-Aktion wartet die EU einige Sekunden, bevor sie wiederholt Verbindungsversuche unternimmt, um die Anrufrate zu reduzieren.

Netzwerküberlastungsschutz für eingehende GTP-C-Nachrichtendrosselung

Diese Funktion schützt die Prozesse des Packet GW (P-GW)/Gateway GPRS Supporting Node (GGSN) vor Überlastungen der Übertragung und Ausfällen von Netzwerkelementen. In einem P-GW/Serving GPRS Supporting Node (SGSN) besteht der größte Engpass in der Verarbeitung der Benutzerdaten, z. B. in der Sitzungsverwaltung und in der gesamten CPU- und Speichernutzung des DPCs.

Für die SGSN/Mobility Management Entity (MME) wird ein *No value* konfiguriert, um die eingehenden GPRS Tunneling Protocol-Control (GTP-C)-Meldungen bei Aktivierung des Netzwerküberlastungsschutzes zu drosseln.

> **Hinweis**: Für die Verwendung von GTP und der Schnittstellendrosselung mit Durchmesser muss ein gültiger Lizenzschlüssel installiert werden.

Diese Funktion unterstützt die Steuerung der Geschwindigkeit von ein-/ausgehenden Nachrichten im P-GW/GGSN, wodurch sichergestellt wird, dass der P-GW/GGSN nicht durch die GTP-Nachrichten für den Kontrollplan überlastet wird. Darüber hinaus wird sichergestellt, dass der GTP-GW/GGSN den GTP-C-Peer nicht mit den GTP-Nachrichten auf Kontrollebene überlastet. Für diese Funktion müssen die GTP-Kontrollmeldungen (Version 1 (v1) und Version 2 (v2)) über die Gn/Gp- und S5/S8-Schnittstellen konfiguriert/geregelt werden. Diese Funktion behandelt den Überlastungsschutz der P-GW/GGSN-Knoten und der anderen externen Knoten, mit denen sie kommuniziert. Throttling wird nur für Kontrollnachrichten auf Sitzungsebene durchgeführt, sodass die Übertragungsrate für Pfadmanagementnachrichten nicht begrenzt ist.

Die Überlastung des externen Knotens kann in einem Szenario auftreten, in dem der P-GW/GGSN Signalisierungsanforderungen mit einer höheren Geschwindigkeit generiert, als die anderen Knoten verarbeiten können. Wenn die eingehende Rate am P-GW/GGSN-Knoten hoch

ist, kann sie auch den externen Knoten überfluten. Aus diesem Grund ist die Drosselung sowohl der ein- als auch der ausgehenden Kontrollnachrichten erforderlich. Zum Schutz der externen Knoten vor Überlastung durch die P-GW/GGSN-Kontrollsignalisierung wird ein Framework verwendet, um die ausgehenden Kontrollnachrichten an die externen Schnittstellen zu gestalten und zu steuern.

Konfigurieren der Empfangs-GTP-C-Nachrichtendrosselung

Geben Sie den folgenden Befehl ein, um die GTP-C-Eingangsdrosselung zu konfigurieren:

```
gtpc overload-protection Ingress
```

Dadurch wird der Überlastungsschutz des GGSN/PGW konfiguriert, indem eingehende GTPv1- und GTPv2-Steuerungsansagen über die Gn/Gp-Schnittstelle (GTPv1) oder die S5/S8-Schnittstelle (GTPv2) mit den anderen Parametern für die Services gedrosselt werden, die im Kontext konfiguriert und auf den GSN und den PGGW angewendet werden.

Wenn Sie den vorherigen Befehl eingeben, wird diese Eingabeaufforderung generiert:

```
[context_name]host_name(config-ctx)# gtpc overload-protection ingress
{msg-rate msg_rate} [delay-tolerance dur] [queue-size size]
[no] gtpc overload-protection Ingress
```

Hier einige Hinweise zu dieser Syntax:

- **Nein**: Dieser Parameter deaktiviert die eingehende GTP-Steuermeldungsdrosselung für die GGSN-/PGW-Dienste in diesem Kontext.
- **msg-rate msg_rate**: Dieser Parameter legt die Anzahl der eingehenden GTP-Nachrichten fest, die pro Sekunde verarbeitet werden können. Bei *msg_rate* handelt es sich um eine Ganzzahl zwischen 100 und 12.000.
- **Verzögerungstoleranz**: Dieser Parameter legt die maximale Anzahl an Sekunden fest, die eine eingehende GTP-Nachricht in die Warteschlange gestellt werden kann, bevor sie verarbeitet wird. Nachdem diese Toleranz überschritten wurde, wird die Nachricht verworfen. Der *Maßstab* ist eine ganze Zahl zwischen 1 und 10.
- **Größe der Warteschlange**: Dieser Parameter legt die maximale Warteschlangengröße für eingehende GTP-C-Nachrichten fest. Wenn die Warteschlange die definierte Größe überschreitet, werden alle neuen eingehenden Nachrichten verworfen. Die *Größe* ist eine ganze Zahl zwischen 100 und 10.000.

Sie können diesen Befehl verwenden, um die Einschränkung der eingehenden GTP-Kontrollmeldung für die GGSN-/PGW-Dienste zu aktivieren, die im gleichen Kontext konfiguriert sind. Beispielsweise aktiviert dieser Befehl die eingehenden GTP-Kontrollnachrichten in einem Kontext mit einer Nachrichtenrate von *1.000* pro Sekunde, einer Nachrichtenwarteschlangengröße von *10.000* und einer Verzögerung von *einer* Sekunde:

```
gtpc overload-protection ingress msg-rate 1000 delay-tolerance 1 queue-size 10000
```

Schutz von Netzwerkelementen der Nachbarschaft

Viele benachbarte Netzwerkelemente nutzen ihre eigenen Mechanismen, um sich selbst zu schützen, und ein zusätzlicher Schutz vor Netzwerküberlastungen auf der ASR5x00-Seite ist möglicherweise nicht erforderlich. Der Schutz der Netzwerkelemente des Nachbarn kann erforderlich sein, wenn die Gesamtstabilität des Netzwerks nur dann erreicht werden kann, wenn die Nachrichtendrosselung auf der Ausgangsseite angewendet wird.

Netzwerküberlastungsschutz mit trottling für eine S6a-Schnittstelle

Diese Funktion schützt die S6a- und S13-Schnittstellen in Ausgangs-Richtung. Sie schützt den Home Subscriber Server (HSS), den Diameter Routing Agent (DRA) und das Equipment Identity Register (EIR). Die Funktion verwendet die RLF-Funktion (Rate Limiting Function).

Berücksichtigen Sie die folgenden wichtigen Hinweise, wenn Sie die Endpunktkonfiguration für den Durchmesser anwenden:

- Dem Peer muss eine RLF-Vorlage zugeordnet werden.
- Ein RLF wird nur auf Peer-Basis (einzeln) angehängt.

Konfigurieren der Diameter-Drosselung an einer S6a-Schnittstelle

Die folgende Befehlssyntax wird verwendet, um die Drosselung des Durchmessers auf einer S6a-Schnittstelle zu konfigurieren:

```
[context_na>me]host_name(config-ctx-diameter)#>peer [*] peer_name [*]
[ realm realm_name ] { address ipv4/ipv6_address [ [ port port_number ]
[connect-on-application-access] [ send-dpr-before-disconnect disconnect-cause
disconnect_cause ] [ sctp ] ] + | fqdn fqdn [ [ port port_number ]
[ send-dpr-before-disconnect disconnect-cause disconnect_cause ]
[ rlf-template rlf_template_name ] ] }

no peer peer_name [ realm realm_name ]
```

Hier einige Hinweise zu dieser Syntax:

- **Nein**: Dieser Parameter entfernt die angegebene Peer-Konfiguration.
- **[*] peer_name [*]**: Dieser Parameter gibt den Peernamen als alphanumerische Zeichenfolge an, die ein bis 63 Zeichen lang ist (Satzzeichen sind zulässig).**Hinweis**: Der Serverendpunkt mit Durchmesser kann nun ein Peername mit Platzhalterzeichen (*) sein. Die Client-Peers, die das freigegebene Muster erfüllen, werden als gültige Peers behandelt, und die Verbindung wird akzeptiert. Das wildkartierte Token gibt an, dass der Peername ein Platzhalter ist, und jedes *-Zeichen in der Zeichenfolge, das davor liegt, wird als Platzhalter behandelt.
- **realm_name**: Dieser Parameter gibt den Bereich dieses Peers als alphanumerische Zeichenfolge an, die zwischen einem und 127 Zeichen reicht. Der Name des Bereichs kann ein Firmen- oder Servicename sein.
- **Adresse ipv4/ipv6_address**: Dieser Parameter gibt die Peer-IP-Adresse mit dem Durchmesser

in IPv4-Notation mit Dezimalpunkten oder IPv6-Notation mit Doppelkommazahl und Hexadezimalzeichen an. Diese Adresse muss die IP-Adresse des Geräts sein, mit dem das Chassis kommuniziert.

- **fqdn fqdn**: Dieser Parameter gibt den Durchmesser des Peer Fully Qualified Domain Name (FQDN) als alphanumerische Zeichenfolge an, die zwischen einem und 127 Zeichen reicht.

- **Port-Nummer**: Dieser Parameter gibt die Portnummer für diesen Peer mit Durchmesser an. Bei der Portnummer muss es sich um eine ganze Zahl zwischen 1 und 65.535 handeln.

- **Zugriff auf eine Verbindung über eine Anwendung**: Dieser Parameter aktiviert den Peer beim erstmaligen Zugriff auf die Anwendung.

- **send-dpr-before disconnect**: Dieser Parameter sendet die Disconnect-Peer-Request (DPR).

- **Trennungsursache**: Mit diesem Parameter wird der DPR mit dem angegebenen Trennungsgrund zum angegebenen Peer beendet. Bei der Trennungsursache muss es sich um eine Ganzzahl zwischen 0 und zwei handeln, die den folgenden Ursachen entspricht:

  0 â Â

  1 â Â

  2 â Â DO_NOT_WANT_TO_TALK_TO_YOU

- **rlf-template rlf_template_name**: Dieser Parameter gibt die RLF-Vorlage an, die diesem Durchmesser Peer zugeordnet werden soll. Bei *rlf_template_name* muss es sich um eine alphanumerische Zeichenfolge mit einer Länge von 1 bis 127 Zeichen handeln.

  Hinweis: Für die Konfiguration einer RLF-Vorlage ist eine RLF-Lizenz erforderlich.

Netzwerküberlastungsschutz mit Durchmesserdrosselung an einer Gx/Gy-Schnittstelle

Diese Funktion schützt die Gx- und Gy-Schnittstellen in Ausgangs-Richtung. Sie schützt die Policy and Charging Rules Function (PCRF) und das Online Charging System (OCS) und verwendet RLF.

Berücksichtigen Sie die folgenden wichtigen Hinweise, wenn Sie die Endpunktkonfiguration für den Durchmesser anwenden:

- Dem Peer muss eine RLF-Vorlage zugeordnet werden.

- Ein RLF wird nur auf Peer-Basis (einzeln) angehängt.

Dieser Befehl wird zur Konfiguration des Netzwerküberlastungsschutzes verwendet:

```
[context_name]host_name(config-ctx-diameter)# rlf-template rlf_template_name
```

Hinweis: Für die Konfiguration einer RLF-Vorlage ist eine RLF-Lizenz erforderlich.

Konfigurieren der Durchmesserdrosselung auf einer Gx/Gy-Schnittstelle

Sie können die Verwendung des RLF für Durchmesser-Schnittstellen in Betracht ziehen. Hier ein Beispiel für eine Konfiguration:

```
rlf-template rlf1

msg-rate 1000 burst-size 100

threshold upper 80 lower 60

delay-tolerance 4

#exit


diameter endpoint Gy

use-proxy

origin host Gy address 10.55.22.3

rlf-template rlf1

peer peer1 realm foo.com address 10.55.22.1 port 3867 rlf-template rlf2

peer peer2 realm fo.com address 10.55.22.1 port 3870

#exit
```

Hier einige Hinweise zu dieser Konfiguration:

- Der Peer mit dem Namen *Peer1* ist an *RFL2* gebunden, und die übrigen Peers unter dem Endpunkt sind an *RLF1* gebunden.
- Die RLF-Vorlage auf Peer-Ebene hat Vorrang vor der Vorlage auf Endpunktebene.
- Die Anzahl der ausgesendeten Nachrichten beträgt maximal 1.000 pro Sekunde (msg-rate). Diese Überlegungen gelten auch für

  Alle einhundert Millisekunden werden nur hundert Nachrichten (Burst-Size) gesendet (um die 1.000 Nachrichten pro Sekunde zu erreichen).

  Wenn die Anzahl der Nachrichten in der RLF-Warteschlange 80 % der Nachrichtenrate überschreitet (80 % von 1.000 = 800), wechselt die RLF in den Status *OVER_THRESHOLD*.

  Wenn die Anzahl der Nachrichten in der RLF-Warteschlange die Nachrichtenrate (1.000) überschreitet, wechselt die RLF in den Status *OVER_LIMIT*.

  Wenn die Anzahl der Nachrichten in der RLF-Warteschlange unter 60 % der Nachrichtenrate sinkt (60 % von 1.000 = 600), wechselt die RLF zurück in den *READY*-Status.

Die maximale Anzahl von Nachrichten, die in die Warteschlange gestellt werden können, entspricht der Nachrichtenrate multipliziert mit der Verzögerungstoleranz (1.000 x 4 = 4.000).

Wenn die Anwendung mehr als 4.000 Nachrichten an den RLF sendet, werden die ersten 4.000 in die Warteschlange gestellt, und der Rest wird verworfen.

Die verworfenen Nachrichten werden von der Anwendung in angemessener Zeit erneut versucht bzw. an das RLF gesendet.

Für die Anzahl der Wiederholungen ist der Antrag verantwortlich.

- Die Vorlage kann mit dem Parameter *no rlf-template* vom Endpunkt *getrennt* werden. Beispielsweise würde die Bindung von *RLF1* von *peer2* aufgehoben.

- Verwenden Sie den Parameter *no rlf-template rlf1* im *Endpunktkonfigurationsmodus*, da die CLI versucht, die RLF-Vorlage *RLF1* zu löschen. Dieser CLI-Befehl ist Teil der globalen Konfiguration und nicht der Endpunktkonfiguration.

- Die Vorlage kann über einen der folgenden Befehle an die einzelnen Peers gebunden werden:

```
no peer peer2 realm foo.com

peer peer2 realm foo.com address 10.55.22.1 port 3867
```

- Der RLF kann nur für Durchmesser-Endpunkte verwendet werden, in denen Diamproxy verwendet wird.

- Die konfigurierte Nachrichtenrate wird pro Diamant implementiert. Wenn beispielsweise die Nachrichtenrate 1.000 beträgt und 12 Diamanten aktiv sind (voll bestücktes Chassis = 12 aktive Packet Services Card (PSC) + 1 Demux + 1 Standby-PSC), beträgt die effektive Übertragungsrate pro Sekunde (TPS) 12.000. Sie können einen dieser Befehle eingeben, um die RLF-Kontextstatistiken anzuzeigen:

```
show rlf-context-statistics diamproxy

show rlf-context-statistics diamproxy verbose
```

Schutz vor Netzwerküberlastung durch Seitendrosselung mit RLF

Die Seitendrosselungsfunktion begrenzt die Anzahl der Paging-Nachrichten, die aus dem SGSN gesendet werden. Er bietet dem Operator Flexibilität und Kontrolle, der nun die Anzahl der Paging-Nachrichten, die vom SGSN gesendet werden, entsprechend der Netzwerkbedingungen reduzieren kann. An einigen Standorten ist die Anzahl der Paging-Nachrichten, die vom SGSN initiiert werden, aufgrund schlechter Funkbedingungen sehr hoch. Eine höhere Anzahl an Paging-Nachrichten führt zu einer erhöhten Bandbreitennutzung im Netzwerk. Diese Funktion bietet eine konfigurierbare Ratenbeschränkung, bei der die Paging-Nachricht auf folgenden Ebenen gedrosselt wird:

- Weltweite Ebene für 2G- und 3G-Zugriff

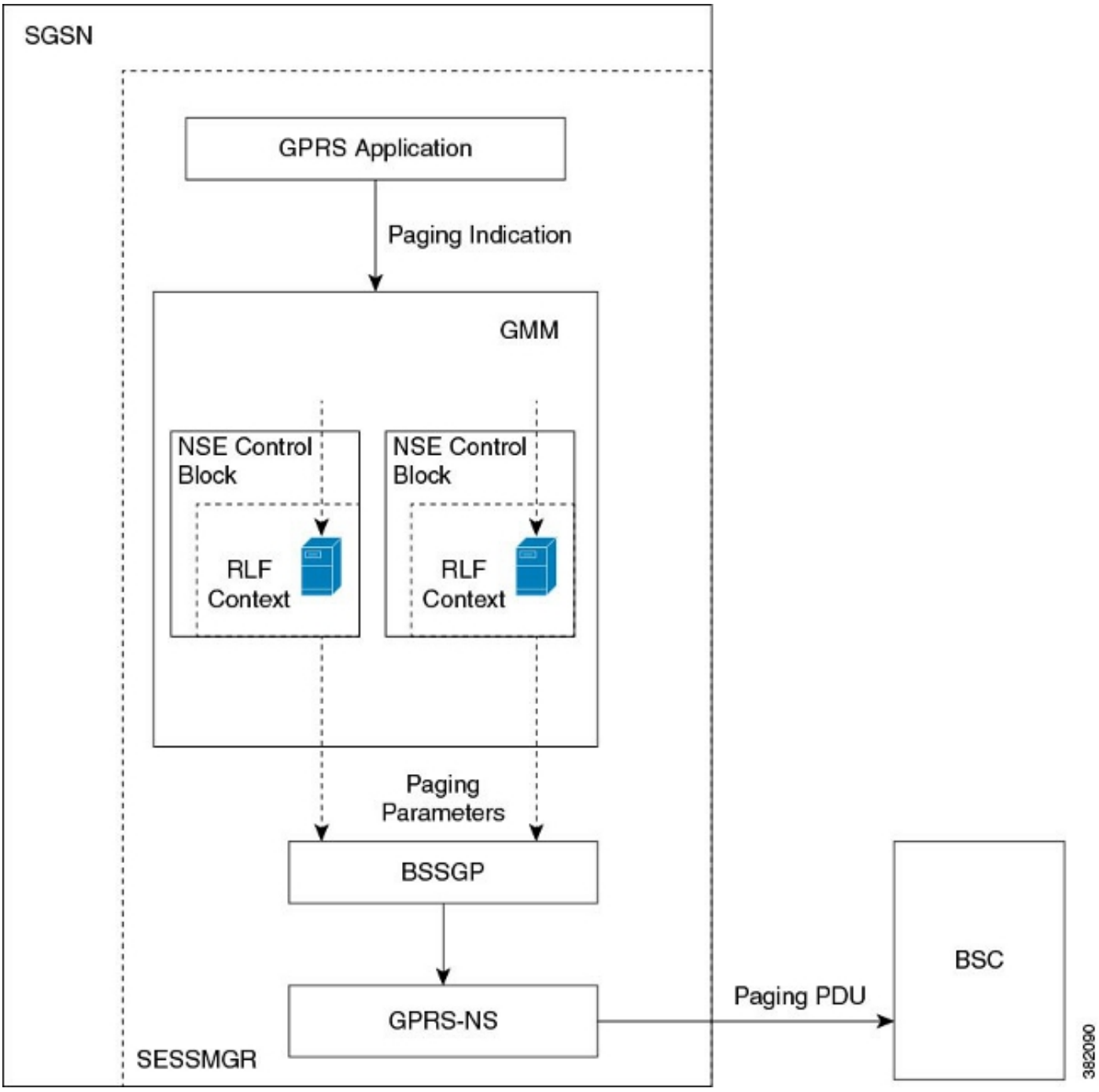
- Die Ebene der Netzwerkserviceeinheit (NSE) nur für 2G-Zugriff

- Die Radio Network Controller (RNC)-Ebene dient nur für 3G-Zugriff.

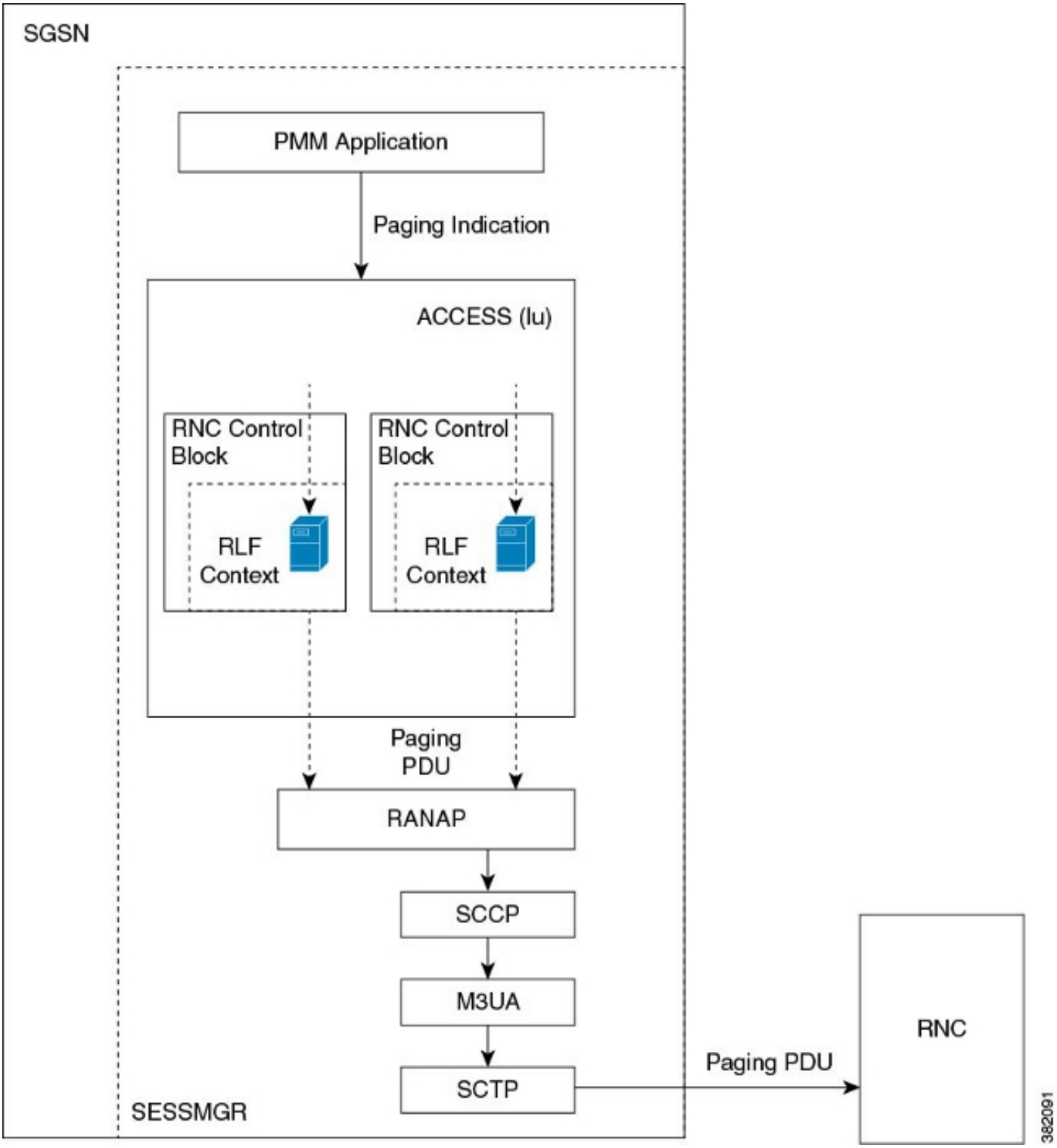
Diese Funktion verbessert die Bandbreitennutzung an der Funkschnittstelle.

Hinweis: Für die Konfiguration einer RLF-Vorlage ist eine RLF-Lizenz erforderlich.

Das folgende Beispiel zeigt den Paging-Prozess mit 2G-Zugriff und Ratenbegrenzung:



Das folgende Beispiel zeigt den Paging-Prozess mit 3G-Zugriff und Ratenbegrenzung:



Seitendrosselung mit RLF konfigurieren

Die in diesem Abschnitt beschriebenen Befehle werden zum Konfigurieren der Seitendrosselungsfunktion verwendet. Diese CLI-Befehle werden verwendet, um die RLF-Vorlage für die Seitendrosselung auf globaler Ebene, der NSE-Ebene und der RNC-Ebene auf dem SGSN zuzuordnen/zu entfernen.

Ordnen Sie den RNC-Namen der RNC-Kennung zu.

Der Befehl **interface** dient zum Konfigurieren der Zuordnung zwischen RNC Identifier (ID) und RNC-Name. Sie können die *paging-rlf-template* entweder über den RNC-Namen oder die RNC-ID

konfigurieren. Die Syntax lautet wie folgt:

```
config
sgsn-global
interface-management
[ no ] interface {gb
peer-nsei | iu peer-rnc} {name <value> | id <value>}
exit
```

> **Hinweis**: Die *no*-Form des Befehls entfernt die Zuordnung und andere Konfiguration, die der RNC-*Paging-rlf-template*-Konfiguration zugeordnet ist, aus dem SGSN und setzt das Verhalten auf den Standardwert für diesen RNC zurück.

Hier ein Beispiel für eine Konfiguration:

```
[local]asr5000# configure
[local]asr5000(config)# sgsn-global
[local]asr5000(config-sgsn-global)# interface-management
[local]asr5000(config-sgsn-interface-mgmt)# interface
iu peer-rnc id 250 name bng_rnc1
[local]asr5000(config-sgsn-interface-mgmt)# end
[local]asr5000#
```

Zuordnen einer Paging-RLF-Vorlage

Mit diesem Befehl kann der SGSN eine RLF-Vorlage entweder auf globaler Ebene zuordnen, wodurch die Paging-Meldungen, die sowohl auf 2G- (NSE-Ebene) als auch auf 3G-Zugriff (RNC-Ebene) initiiert werden, eingeschränkt werden, oder auf Ebene einzelner Einheiten, die entweder auf der RNC-Ebene für 3G-Zugriff oder auf der NSE-Ebene für 2G-Zugriff erfolgt. Die Syntax lautet wie folgt:

```
config
sgsn-global
interface-management
[no] paging-rlf-template {template-name <template-name>} {gb
peer-nsei | iu peer-rnc} {name <value> | id <value>}
exit
```

> **Hinweis**: Wenn einem bestimmten NSE/RNC keine RLF-Vorlage zugeordnet ist, wird die Auslagerungslast auf Basis der globalen RLF-Vorlage begrenzt, die (falls vorhanden) zugeordnet ist. Wenn keine globale RLF-Vorlage zugeordnet ist, wird keine Ratenbegrenzung auf die Auslagerungslast angewendet.

Hier ein Beispiel für eine Konfiguration:

```
[local]asr5000(config)# sgsn-global
[local]asr5000(config-sgsn-global)# interface-management
[local]asr5000(config-sgsn-interface-mgmt)# paging-rlf-template
template-name rlf1
[local]asr5000(config-sgsn-interface-mgmt)# end
[local]asr5000#
[local]asr5000# configure
[local]asr5000(config)# sgsn-global
[local]asr5000(config-sgsn-global)# interface-management
```

```
[local]asr5000(config-sgsn-interface-mgmt)# paging-rlf-template
template-name rlf2 gb peer-nsei id 1
[local]asr5000(config-sgsn-interface-mgmt)# end
[local]asr5000#
[local]asr5000# configure
[local]asr5000(config)# sgsn-global
[local]asr5000(config-sgsn-global)# interface-management
[local]asr5000(config-sgsn-interface-mgmt)# paging-rlf-template
template-name rlf2 iu peer-rnc name bng_rnc1
[local]asr5000(config-sgsn-interface-mgmt)# end
[local]asr5000#
```