Implementazione della protezione da sovraccarico per gateway ed elementi di rete adiacenti sull'appliance ASR serie 5x00

Sommario

Introduzione

Questo documento descrive come implementare le funzionalità di protezione disponibili per i gateway (GW) e gli elementi di rete adiacenti sul Cisco Aggregated Services Router (ASR) serie 5x00 per proteggere le prestazioni complessive della rete.

Controllo della congestione per GW

Controllo congestione è una funzione di autoprotezione generica. Viene utilizzato per proteggere il sistema dai sovraccarichi di utilizzo delle seguenti risorse:

- Utilizzo CPU nell'elaborazione delle schede
- Utilizzo della memoria sulle schede di elaborazione

Quando l'utilizzo supera le soglie predefinite, tutte le nuove chiamate (attivazioni PDP (Packet Data Protocol), attivazioni di sessioni PDN (Packet Data Network)) vengono *eliminate* o *rifiutate*, a seconda della configurazione.

Di seguito è riportato un esempio che illustra come monitorare l'utilizzo complessivo della scheda di elaborazione dati:

```
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
```

> **Nota:** Il limite di progettazione del sistema è pari all'80% dell'utilizzo della CPU, che è definito come il limite consigliato che non deve essere superato per garantire il regolare funzionamento del sistema. Un carico superiore al valore può influire sulle operazioni della piattaforma, ad esempio sulla stabilità e sulla prevedibilità, e deve essere evitato con una pianificazione adeguata della capacità.

> **Nota:** Cisco consiglia di utilizzare l'azione *drop* anziché l'azione *rifiuto*, in quanto le chiamate rifiutate provocano tentativi di riconnessione ripetuti immediati dall'Apparecchiatura Utente (UE). Nel caso di un'azione di rilascio, l'UE attende alcuni secondi prima di effettuare ripetuti tentativi di riconnessione, quindi la frequenza delle chiamate diminuisce.

Protezione da sovraccarico di rete per limitazione messaggi GTP-C in ingresso

Questa funzione protegge i processi Packet GW (P-GW)/Gateway GPRS Support Node (GSN) da sovratensioni nella trasmissione e da errori negli elementi della rete. In un P-GW/Serving GPRS Supporting Node (SGSN), il collo di bottiglia principale è correlato all'elaborazione dei dati utente, come l'utilizzo del gestore delle sessioni e l'utilizzo complessivo della CPU e della memoria DPC.

Nell'entità SGSN/Mobility Management Entity (MME) non è configurato alcun valore per limitare i messaggi GTP-C (GPRS Tunneling Protocol-Control) in entrata quando è attivata la protezione da sovraccarico della rete.

> **Nota:** Per utilizzare la limitazione delle interfacce GTP e di diametro, è necessario installare una chiave di licenza valida.

Questa funzione aiuta a controllare la frequenza dei messaggi in entrata/in uscita su P-GW/GSN, garantendo che P-GW/GSN non sia sopraffatto dai messaggi del piano di controllo GTP. Inoltre, aiuta a garantire che il P-GW/GSN non sopraffagga il peer GTP-C con i messaggi del control plane GTP. Questa funzionalità richiede che i messaggi di controllo GTP (versione 1 (v1) e versione 2 (v2)) abbiano la forma o il controllo sulle interfacce Gn/Gp e S5/S8. Questa funzione copre la protezione dal sovraccarico dei nodi P-GW/GSN e degli altri nodi esterni con cui comunica. La limitazione viene applicata solo ai messaggi di controllo a livello di sessione, pertanto i messaggi di gestione dei percorsi non sono affatto limitati.

L'overload del nodo esterno può verificarsi in uno scenario in cui P-GW/GSN genera richieste di segnalazione a una velocità superiore a quella che gli altri nodi possono gestire. Inoltre, se la velocità in entrata è elevata nel nodo P-GW/GSN, potrebbe inondare il nodo esterno. Per questo motivo, è necessaria la limitazione dei messaggi di controllo in entrata e in uscita. Per la protezione dei nodi esterni da un sovraccarico dovuto alla segnalazione di controllo P-GW/GSN,

viene utilizzato un framework per modellare e controllare i messaggi di controllo in uscita verso le interfacce esterne.

Configurazione della limitazione dei messaggi GTP-C in ingresso

Immettere questo comando per configurare la limitazione dei messaggi GTP-C in entrata:

```
gtpc overload-protection Ingress
```

In questo modo viene configurata la protezione dall'overload del GSN/PGW limitando i messaggi di controllo GTPv1 e GTPv2 in entrata sull'interfaccia Gn/Gp (GTPv1) o S5/S8 (GTPv2) con gli altri parametri per i servizi configurati in un contesto e applicati ai GSN e PGW.

Quando si immette il comando precedente, viene generato questo prompt:

```
[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
```

Ecco alcune note su questa sintassi:

- **no:** Questo parametro disabilita la limitazione dei messaggi di controllo in entrata GTP per i servizi GSN/PGW in questo contesto.
- **msg-rate velocità_messaggio**: Questo parametro definisce il numero di messaggi GTP in entrata che è possibile elaborare al secondo. *msg_rate* è un numero intero compreso tra 100 e 12.000.
- **ritardo-tolleranza durante**: Questo parametro definisce il numero massimo di secondi durante i quali un messaggio GTP in entrata può essere inserito in coda prima di essere elaborato. Quando questa tolleranza viene superata, il messaggio viene eliminato. La *durata* è un numero intero compreso tra uno e dieci.
- **dimensioni coda**: Questo parametro definisce le dimensioni massime della coda per i messaggi GTP-C in ingresso. Se la coda supera le dimensioni definite, tutti i nuovi messaggi in ingresso vengono eliminati. La *dimensione* è un numero intero compreso tra 100 e 10.000.

Èpossibile utilizzare questo comando per abilitare la limitazione GTP dei messaggi di controllo in entrata per i servizi GSN/PGW configurati nello stesso contesto. Ad esempio, questo comando abilita i messaggi di controllo GTP in entrata in un contesto con una velocità di *1.000* messaggi al secondo, una dimensione della coda di messaggi di *10.000* e un ritardo di *un* secondo:

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

Protezione degli elementi di rete adiacenti

Molti elementi della rete adiacente utilizzano i propri meccanismi per proteggersi e potrebbe non essere necessaria una protezione aggiuntiva da sovraccarico di rete sul lato ASR5x00. La protezione degli elementi della rete adiacente potrebbe essere necessaria nei casi in cui la

stabilità complessiva della rete può essere raggiunta solo quando viene applicata la limitazione dei messaggi sul lato uscita.

Protezione da sovraccarico di rete con limitazione del diametro su un'interfaccia S6a

Questa funzione protegge le interfacce S6a e S13 nella direzione di uscita. Protegge il Home Subscriber Server (HSS), il Diameter Routing Agent (DRA) e il Registro di Identità dell'Apparecchiatura (EIR). La funzione utilizza la funzione di limitazione della velocità (RLF).

Quando applicate la configurazione dell'estremità del diametro, tenete presenti le seguenti note importanti:

- Ènecessario associare un modello RLF al peer.
- L'associazione di un fattore di miglioramento viene eseguita solo per peer (singolarmente).

Configurazione della limitazione del diametro su un'interfaccia S6a

Di seguito è riportata la sintassi del comando utilizzata per configurare la limitazione del diametro su un'interfaccia S6a:

```
[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 ]
```

Ecco alcune note su questa sintassi:

- **no:** Questo parametro rimuove la configurazione peer specificata.
- **[*] nome_peer [*]**: Questo parametro specifica il nome del peer come stringa alfanumerica compresa tra uno e 63 caratteri (sono consentiti i caratteri di punteggiatura).**Nota:** L'endpoint del server del diametro può ora essere un nome peer con caratteri jolly (il carattere * è un carattere jolly valido). I peer client che soddisfano il modello con caratteri jolly vengono trattati come peer validi e la connessione viene accettata. Il token con caratteri jolly indica che il nome del peer è con caratteri jolly e qualsiasi carattere * nella stringa che precede viene considerato come carattere jolly.
- **nome_area di autenticazione**: Questo parametro specifica l'area di autenticazione del peer come stringa alfanumerica compresa tra uno e 127 caratteri. Il nome dell'area di autenticazione può essere un nome di società o di servizio.
- **indirizzo ipv4/ipv6_address**: Questo parametro specifica il diametro dell'indirizzo IP peer in notazione IPv4 decimale con punti o IPv6 con valori esadecimali separati da due punti. Questo indirizzo deve essere l'indirizzo IP del dispositivo con cui comunica lo chassis.

- **fqdn fqdn**: Questo parametro specifica il nome di dominio completo (FQDN, Fully Qualified Domain Name) del peer del diametro come stringa alfanumerica compresa tra uno e 127 caratteri.

- **port numero_porta**: Questo parametro specifica il numero di porta per questo peer di diametro. Il numero di porta deve essere un numero intero compreso tra 1 e 65.535.

- **connessione all'applicazione**: Questo parametro attiva il peer all'accesso iniziale all'applicazione.

- **send-dpr-before-disconnect**: Questo parametro invia la richiesta Disconnect-Peer-Request (DPR).

- **causa-disconnessione**: Questo parametro termina il DPR al peer specificato, con il motivo di disconnessione specificato. La causa di disconnessione deve essere un numero intero compreso tra zero e due, che corrisponde alle seguenti cause:

  0 ÂÂ RIAVVIO

  1 ÂÂ OCCUPATO

  2 ÂÂ DO_NOT_WANT_TO_TALK_TO_YOU

- **rlf-template nome_rlf**: Questo parametro specifica il modello RLF da associare a questo peer di diametro. *rlf_template_name* deve essere una stringa alfanumerica di lunghezza compresa tra uno e 127 caratteri.

  Nota: Per configurare un modello RLF è necessaria una licenza RLF.

Protezione da sovraccarico di rete con limitazione del diametro su interfaccia Gx/Gy

Questa funzione protegge le interfacce Gx e Gy in uscita. Protegge la funzione PCRF (Policy and Charging Rules Function) e il sistema di caricamento online (OCS) e utilizza RLF.

Quando applicate la configurazione dell'estremità del diametro, tenete presenti le seguenti note importanti:

- Ènecessario associare un modello RLF al peer.

- L'associazione di un fattore di miglioramento viene eseguita solo per peer (singolarmente).

Questo comando è usato per configurare la protezione dal sovraccarico della rete:

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

Nota: Per configurare un modello RLF è necessaria una licenza RLF

Configurazione della limitazione del diametro su un'interfaccia Gx/Gy

Si potrebbe prendere in considerazione l'uso di RLF per le interfacce di diametro. Di seguito è riportato un esempio di configurazione:

```
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
```

Ecco alcune note su questa configurazione:

- Il peer denominato *peer1* è associato a *RFL2*, mentre gli altri peer sotto l'endpoint sono associati a *RLF1*.
- Il modello RLF di livello peer ha la precedenza sul modello di livello endpoint.
- Il numero di messaggi viene inviato a una velocità massima di 1.000 al secondo (velocità msg). Queste considerazioni si applicano anche:

  Ogni cento millisecondi (per raggiungere i 1000 messaggi al secondo) vengono inviati solo cento messaggi (dimensione burst).

  Se il numero di messaggi nella coda RLF supera l'80% della velocità (80% di 1.000 = 800), RLF passa allo stato *OVER_THRESHOLD*.

  Se il numero di messaggi nella coda RLF supera la velocità dei messaggi (1.000), RLF passa allo stato *OVER_LIMIT*.

  Se il numero di messaggi nella coda RLF diminuisce al di sotto del 60% della velocità dei messaggi (60% di 1.000 = 600), RLF torna allo stato *READY*.

  Il numero massimo di messaggi che possono essere accodati è uguale alla velocità dei messaggi moltiplicata per la tolleranza di ritardo (1.000 x 4 = 4.000).

Se l'applicazione invia più di 4.000 messaggi al file RLF, i primi 4.000 vengono accodati e gli altri vengono eliminati.

I messaggi eliminati vengono ritentati o inviati nuovamente dall'applicazione all'RLF in un periodo di tempo appropriato.

Il numero di tentativi è responsabilità dell'applicazione.

- Èpossibile annullare l'associazione del modello dall'endpoint con il parametro *no rlf-template*. Ad esempio, *RLF1* verrebbe dissociato da *peer2*.
- Non utilizzare il parametro *no rlf-template rlf1* nella modalità di *configurazione dell'endpoint*, in quanto la CLI tenta di eliminare il modello RLF *RLF1*. Questo comando CLI fa parte della configurazione globale, non della configurazione dell'endpoint.
- Il modello può essere associato ai singoli peer tramite uno dei seguenti comandi:

```
no peer peer2 realm foo.com

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

- L'RLF può essere utilizzato solo per gli endpoint con diametro in cui viene utilizzato il proxy a rombo.
- La velocità dei messaggi configurata è implementata per Diamproxy. Ad esempio, se la velocità dei messaggi è 1.000 e sono attivi 12 diamanti (chassis completamente popolato = 12 PSC (Packet Services Card) attivi + 1 Demux + 1 PSC in standby), il valore TPS (Transmissions Per Second) effettivo è 12.000. È possibile immettere uno di questi comandi per visualizzare le statistiche del contesto RLF:

```
show rlf-context-statistics diamproxy

show rlf-context-statistics diamproxy verbose
```

Protezione dall'overload della rete tramite limitazione delle pagine con RLF

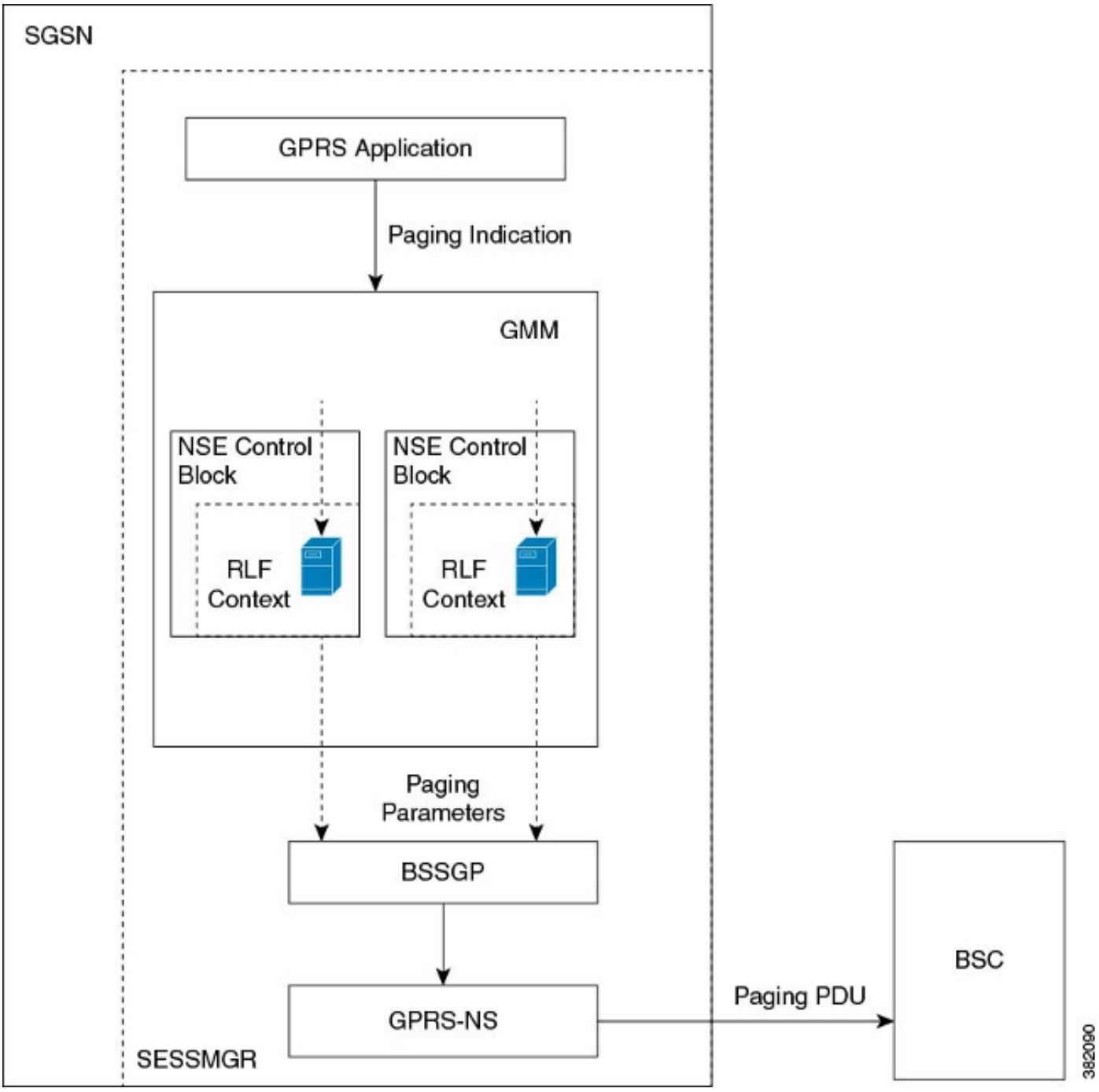
La funzionalità di limitazione delle pagine limita il numero di messaggi di paging inviati dall'SGSN. Offre flessibilità e controllo all'operatore, che ora può ridurre il numero di messaggi di paging inviati dall'SGSN in base alle condizioni della rete. In alcune località, la quantità di messaggi di paging avviati dal SGSN è molto elevata a causa di condizioni radio non valide. Un numero maggiore di messaggi di paging determina l'utilizzo della larghezza di banda nella rete. Questa funzione fornisce un limite di velocità configurabile, in cui il messaggio di paging viene limitato ai seguenti livelli:

- Livello globale per l'accesso 2G e 3G
- Livello NSE (Network Service Entity) solo per accesso 2G
- Livello di Radio Network Controller (RNC) solo per accesso 3G

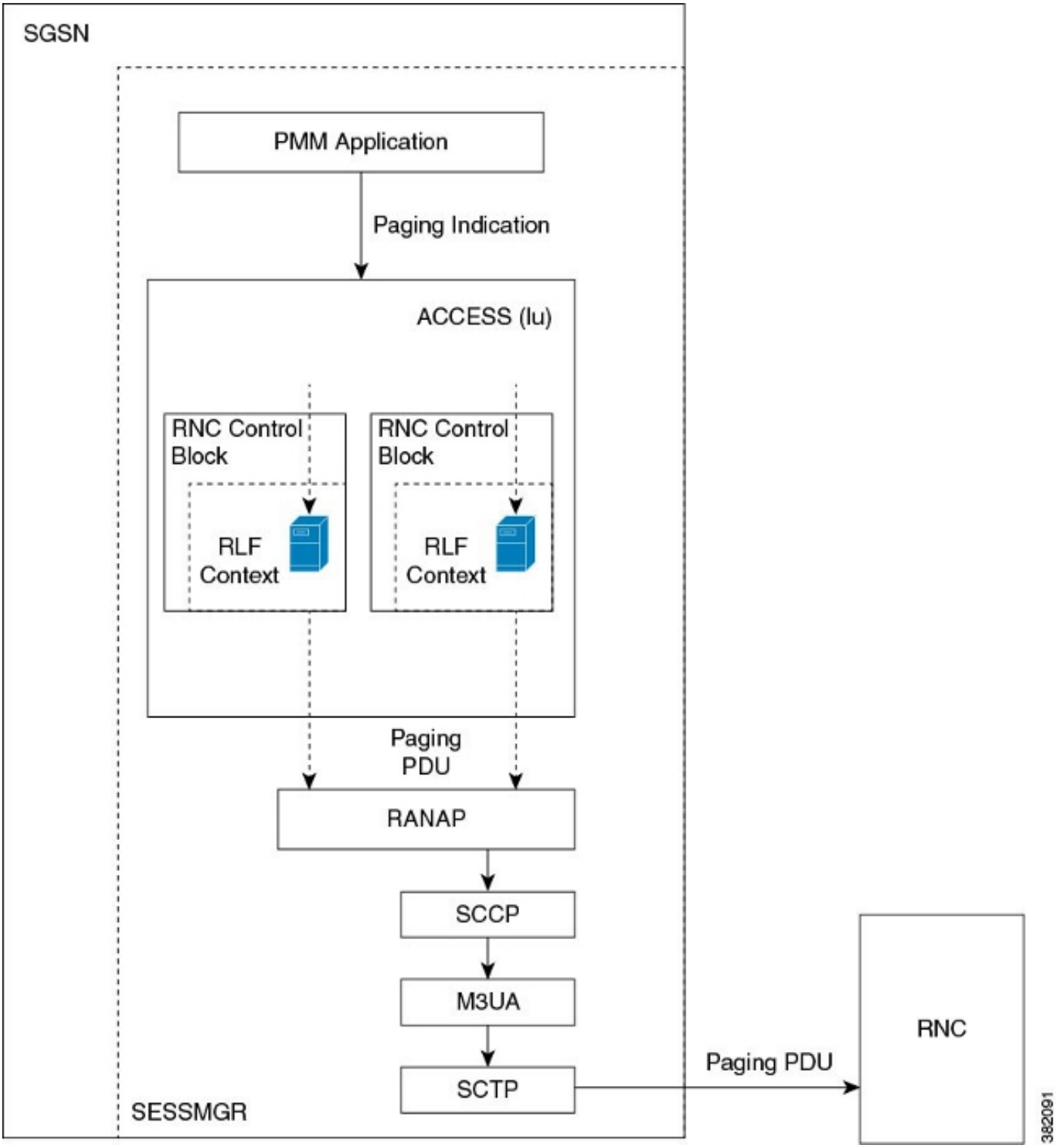
Questa funzione migliora il consumo della larghezza di banda sull'interfaccia radio.

Nota: Per configurare un modello RLF è necessaria una licenza RLF.

Di seguito è riportato un esempio del processo di paging con accesso 2G e limitazione delle velocità:



Di seguito è riportato un esempio del processo di paging con accesso 3G e limitazione delle velocità:



Configurare la limitazione delle pagine con RLF

I comandi descritti in questa sezione vengono usati per configurare la funzione di limitazione delle pagine. Questi comandi CLI vengono usati per associare/rimuovere il modello RLF per la limitazione delle pagine a livello globale, il livello NSE e il livello RNC su SGSN.

Mappare il nome RNC all'identificatore RNC

Il comando **interface** viene usato per configurare il mapping tra l'identificatore RNC (ID) e il nome RNC. È possibile configurare il *paging-rlf-template* in base al nome RNC o all'ID RNC. Di seguito è riportata la sintassi utilizzata:

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

Nota: La forma *no* del comando rimuove la mappatura e l'altra configurazione associata alla configurazione *paging-rlf-template* RNC da SGSN e ripristina il comportamento predefinito per tale RNC.

Di seguito è riportato un esempio di configurazione:

```
[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#
```

Associare un modello RLF di paging

Questo comando consente al SGSN di associare un modello RLF a livello globale, limitando i messaggi di paging che vengono avviati sia nell'accesso 2G (livello NSE) che 3G (livello RNC), o a livello per entità, che è a livello RNC per l'accesso 3G o a livello NSE per l'accesso 2G. Di seguito è riportata la sintassi utilizzata:

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

Nota: Se a un NSE/RNC specifico non è associato alcun modello RLF, il carico di paging è limitato in base al modello RLF globale associato (se presente). Se non è associato alcun modello RLF globale, non viene applicata alcuna limitazione di velocità al carico di paging.

Di seguito è riportato un esempio di configurazione:

```
[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#
```