

Voorlopige statische en dynamische snelheidsbeperking met CPS vDRA

Inhoud

[Inleiding](#)

[Voorwaarden](#)

[Vereisten](#)

[Gebruikte componenten](#)

[Achtergrondinformatie](#)

[Probleem](#)

[Oplossing](#)

[Statische snelheidslimiet bij load balancer](#)

[grenswaarde voor de ingangssnelheid](#)

[grenswaarde van de uitstroomsnelheid](#)

[dynamische snelheidslimiet](#)

Inleiding

In dit document worden de opties voor tarieflijmieten beschreven in DRA, een telecomcomponent die Diameter-berichten routeert en netwerkverkeer beheert.

Voorwaarden

Vereisten

Cisco raadt kennis van de volgende onderwerpen aan:

- Cisco Policy Suite (CPS) Diameter Routing Agent (vDRA)
- Diameter Routing Agent Basisprincipes en specificaties

Gebruikte componenten

De informatie in dit document is gebaseerd op Cisco Policy Suite DRA.

De informatie in dit document is gebaseerd op de apparaten in een specifieke laboratoriumomgeving. Alle apparaten die in dit document worden beschreven, hadden een opgeschoonde (standaard)configuratie. Als uw netwerk live is, moet u zorgen dat u de potentiële impact van elke opdracht begrijpt.

Achtergrondinformatie

DRA is een component in telecommunicatienetwerken, met name in de context van op het Diameter-protocol gebaseerde netwerken. De DRA routeert op efficiënte wijze Diameter-berichten tussen verschillende netwerkelementen, zoals beleidsservers, oplaadsystemen en andere apparaten die geschikt zijn voor Diameter. Tarievenbeperking is een netwerkverkeersleidingstechniek die wordt gebruikt om de hoeveelheid verkeer naar of van een netwerkelement te regelen. Het helpt ervoor te zorgen dat netwerkbronnen niet worden uitgeput, handhaaft de kwaliteit van de service en voorkomt misbruik van het netwerk.

Probleem

Elk onderdeel in het netwerk kan de verkeersbelasting verwerken op basis van de nominale capaciteit, maar in realtime kunnen er scenario's zijn waarbij het gegenereerde verkeer meer is dan wat het systeem aankan. Enkele van hen zijn:

- Gebruikersgedrag - Activiteiten zoals streaming-evenementen of software-updates die in korte tijd grote hoeveelheden gegevens genereren. Doorgaans verzonden van Gateway (Gw) naar DRA.
- Netwerkg congestie - In periodes van hoog netwerkgebruik kan congestie zich ophopen, wat leidt tot gegevens in de wachtrij die vervolgens in bursts worden verzonden wanneer capaciteit beschikbaar komt.
- Netwerkw weerstandsmechanismen - Omleiden van verkeer tijdens uitval of onderhoud, waardoor tijdelijke pieken ontstaan. Dit kan van invloed zijn op de verkeersstroom op gekoppelde sites die geen netwerkproblemen hebben.
- Gedrag van netwerkelementen – In het geval van overbelasting en congestie kunt u geen reacties/time-outs van een of meer netwerkelementen zien die opnieuw verbinden kunnen veroorzaken en bijdragen aan verdere overbelasting van het systeem.
- Gateway flushing - Gateway kan de bestaande sessies doorspoelen als gevolg van beleidswijzigingen, wijzigingen in de topologie of onderhouds- of probleemoplossingsactiviteiten. Tijdens deze scenario's worden sessies gewist en kunt u een uitbarsting van Gx Credit Control Request (CCR)-T-verzoeken ontvangen.

Oplossing

DRA kan de belasting verdelen over meerdere Diameter-servers om een efficiënte afhandeling van verzoeken te garanderen en overbelasting van één server te voorkomen. In het geval van een serverstoring kan de DRA berichten doorsturen naar alternatieve servers, waardoor een hoge beschikbaarheid en betrouwbaarheid van de netwerkservices wordt gewaarborgd.

De snelheidsbeperking op de DRA beschermt niet alleen de DRA, maar ook andere entiteiten door een gecontroleerde stroom van berichten te garanderen. De belangrijkste voordelen van snelheidsbeperking zijn:

- Continuïteit van de service — behoud van de continue beschikbaarheid van de service door ervoor te zorgen dat kritieke netwerkkonderdelen niet overbelast raken en uitval wordt voorkomen.
- Schaalbaarheid - Het netwerk kan verschillende belastingen verwerken zonder dat de

prestaties afnemen.

- Naleving van Service Level Agreements (SLA's) - Ervoor zorgen dat het netwerk aan de SLA's voldoet door consistente prestaties en betrouwbaarheidsniveaus te handhaven.

Statische snelheidslimiet bij load balancer

Dit is een eenvoudige benadering, waarbij een vaste drempel wordt ingesteld op basis van de nominale capaciteit van DRA/Packet Gateway (pGW) en andere netwerkelementen en niet verandert op basis van de netwerkomstandigheden of systeembronnen. Door de snelheid van binnenkomende verzoeken te beperken, hebt u een voorspelbare uitkomst voor de hoeveelheid verkeer die door DRA wordt verwerkt.

Configuraties voor statische snelheidslimiet zijn afhankelijk van de use case waarvoor het wordt toegepast.

grenswaarde voor de ingangssnelheid

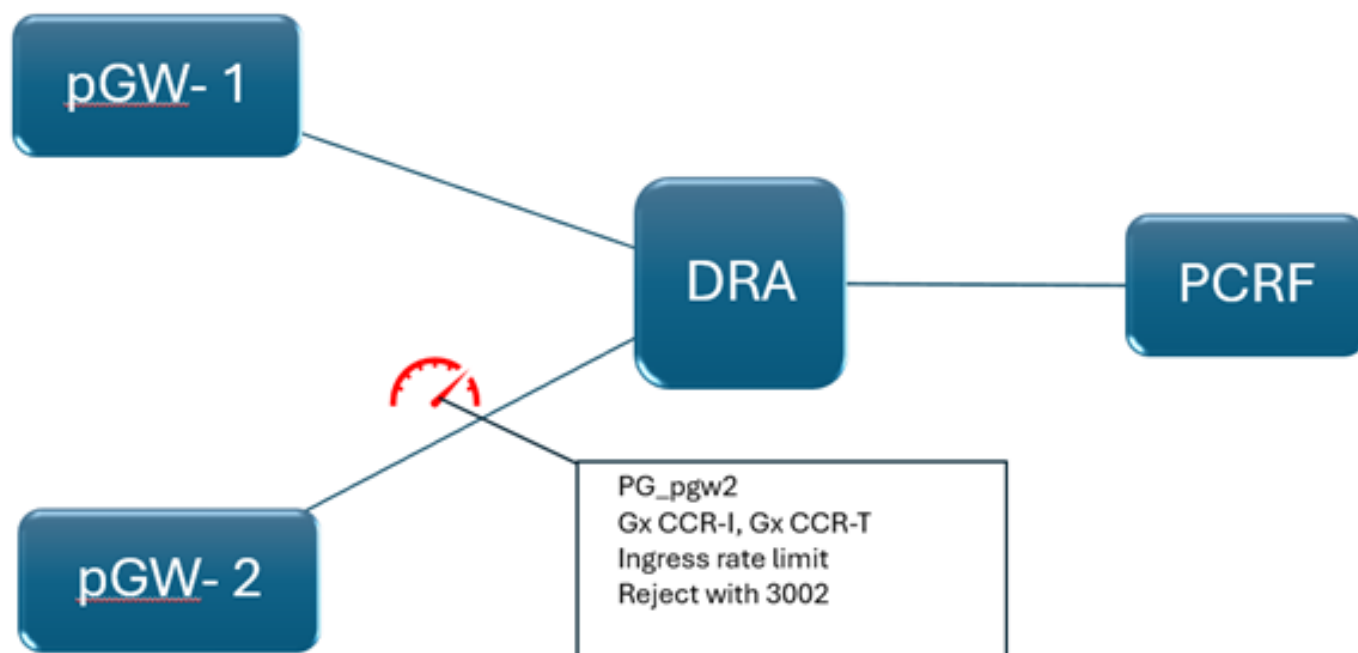
Scenario: uitbarstingen van pGW

Drempelwaarden die specifiek zijn voor pGW's die gevoelig zijn voor deze verkeersexplosies, worden geconfigureerd. De waarde moet worden berekend op basis van de reguliere verkeer / piek verkeer nummers die kunnen worden gezien tijdens deze uitbarstingen.

De drempelnummers kunnen specifiek worden gedefinieerd voor elk berichttype om ervoor te zorgen dat alleen het burst-verkeer wordt ingeperkt, zoals alleen Gx CCR-I- en Gx CCR-T-verzoeken van een GW moeten worden ingeperkt, maar Gx CCR-U-verkeer of Gy-verkeer wordt doorgestuurd zoals ontvangen.

In dit geval kunt u de ingangszijde afknijpen, dat wil zeggen dat DRA het bericht afknijpt zodra het het ontvangt, omdat het doel hier is om te weigeren op basis van het netwerkelement waarvan het het verzoek ontvangt en om te voorkomen dat een groter aantal verzoeken wordt verwerkt dan de DRA aankan.

Het gaspedaal gedrag kan zijn om het bericht met een bepaalde Error-Code en Error-Message af te wijzen of te laten vallen.



Dit gedrag kan worden ingeschakeld in CPS vDRA door aangepaste referentiegegevens (CRD) tabellen 'Peer Rate Limit Profile' en 'Message Rate Limit Profile' te configureren. In deze CRD-tabellen moet u de volgende waarden configureren:

groep van gelijken	Een peer group is een logische groepering van Diameter-knooppunten op basis van hun rijk en host. U moet de peer-groep configureren die moet worden afgekapd.
Volledig gekwalificeerde domeinnaam (FQDN)	FQDN (exacte of regex-match) voor de peers in de peer-groep die u moet beoordelen.
berichtrichting	Richting van het gaspedaal - Ingress of Egress. In dit geval: Ingress.
Profiel van snelheidslimiet	De naam van het profiel voor de berichtsnelheidslimiet die wordt gebruikt om het berichttype te definiëren dat moet worden afgegrensd.
Peer Rate Limit	Percentage aanvragen dat is toegestaan voor deze Peer Group. Dit is inclusief alle berichttypen van die Peer Group.
Afdankgedrag	U kunt ervoor kiezen om het verzoek te laten vallen of af te wijzen met een foutcode.

resultaatcode	Resultaatcodewaarde voor het geval u de berichten afwijst. Niet van toepassing in het geval dat de berichten worden verwijderd.
Fouttekenreeks	De fouttekenreeks die wordt gebruikt in het antwoordbericht van het verzoek dat wordt afgewezen. Niet van toepassing in het geval dat het bericht wordt verwijderd.
applicatie-id	Toepassings-ID van het bericht dat moet worden geblokkeerd.
opdrachtcode	Opdrachtcode van het bericht dat moet worden afgekapd.
Type bericht/aanvraag	Aanvraag-ID en aanvraagtype van de verzoeken die moeten worden afgegrensd.
berichtsnelheidslimiet	TelePresence Server (TPS) van het verzoek van dat berichttype dat door DRA wordt verwerkt. Verzoeken buiten deze TPS worden beperkt. Deze waarde is per peer in de Peer Group.

Message Rate Limit Profile



Filter by All Visible Columns

CCR_I_T_Limit					
Rate Limit Profile Name *	Application Identifier *	Command Code *	Message/Request Type *	Message Rate Limit *	Actions
CCR_I_T_Limit	16777238	272	1	200	
CCR_I_T_Limit	16777238	272	3	200	

Showing 2 out of 2

Show 50 rows 1 out of 1

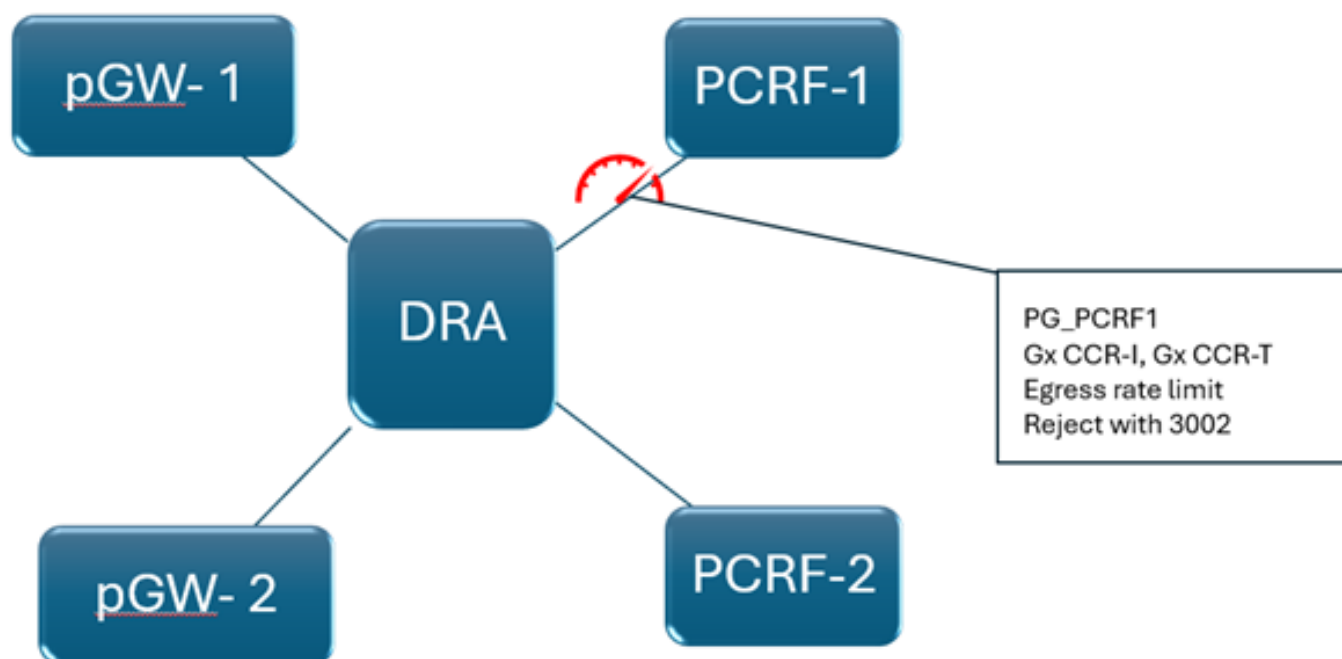
grenswaarde van de uitstroomsnelheid

Scenario: bescherming van netwerkelement dat verzoek ontvangt

Neem een voorbeeld van een Gx-transactie waarbij het verzoek wordt ontvangen van pGW en doorgestuurd naar de functie Beleid en oplaadregels (PCRF). Als er beperkingen zijn aan de hoeveelheid gegevens die PCRF kan verwerken, kunt u, hoewel DRA het binnenkomende verkeer kan verwerken, DRA gebruiken om het bericht bij DRA te verstikken in plaats van het verzoek door te sturen naar PCRF en het te overbelasten.

Hier moet je aan de uitgang-kant knippen, dat wil zeggen dat DRA het bericht afknijpt vlak voordat het doorstuurt naar de PCRF, op basis van de PCRF Peer Group die is geïdentificeerd op basis van de DRA-routeringslogica.

Het gaspedaal gedrag kan zijn om het bericht met een bepaalde Error-Code en Error-Message af te wijzen of te laten vallen.



Dit gedrag kan worden ingeschakeld in CPS vDRA door CRD-tabellen 'Peer Rate Limit Profile' en 'Message Rate Limit Profile' te configureren. In deze CRD-tabellen moet u de volgende waarden configureren:

groep van gelijken	Een peer group is een logische groepering van Diameter-knooppunten op basis van hun rijk en host. U moet de peer-groep configureren die moet worden afgekap.
Peer FQDN	FQDN (exacte of regex-match) voor de peers in de peer-groep die u moet beoordelen.
berichtrichting	Richting van het gaspedaal - Ingress of Egress. In dit geval - Egress.
Profiel van snelheidslimiet	De naam van het profiel voor de berichtsnelheidslimiet die wordt gebruikt om het berichttype te definiëren dat moet worden afgegrensd.
Peer Rate Limit	Percentage aanvragen dat voor deze Peer Group is toegestaan. Dit is inclusief alle berichttypen van die Peer Group.
Afdankgedrag	U kunt ervoor kiezen om het verzoek te laten vallen of af te wijzen met een foutcode.

resultaatcode	Resultaatcodewaarde voor het geval u de berichten afwijst. Niet van toepassing in het geval dat de berichten worden verwijderd.
Fouttekenreeks	De fouttekenreeks die wordt gebruikt in het antwoordbericht van het verzoek dat wordt afgewezen. Niet van toepassing in het geval dat het bericht wordt verwijderd.
applicatie-id	Toepassings-ID van het bericht dat moet worden geblokkeerd.
opdrachtcode	Opdrachtcode van het bericht dat moet worden afgekapt.
Type bericht/aanvraag	Aanvraag-ID en aanvraagtype van de verzoeken die moeten worden afgegrensd.
berichtsnelheidslimiet	TPS van het verzoek van dat berichttype dat door DRA wordt verwerkt. Verzoeken buiten deze TPS worden beperkt. Deze waarde is per peer in de Peer Group.

Peer Rate Limit Profile 

Filter by

All Visible Columns

CCR_I_T  

Peer Group *	Peer FQDN *	Message Direction *	Rate Limit Profile	Peer Rate Limit	Discard Behavior *	Result Code	Error String	Actions
PG_PCRF1	match=peer- 	Egress	CCR_I_T_Limit	1000	Send Error Answer	3002	DRA rate limit breached	 

Showing 1 out of 1

Show

50

 rows  

1

 out of 1  

Message Rate Limit Profile



Filter by All Visible Columns

CCR_I_T_Limit					
Rate Limit Profile Name *	Application Identifier *	Command Code *	Message/Request Type *	Message Rate Limit *	Actions
CCR_I_T_Limit	16777238	272	1	200	
CCR_I_T_Limit	16777238	272	3	200	

Showing 2 out of 2

Show 50 rows 1 out of 1

Scenario: traagheid in het netwerk resulterend in verkeerscongestie, waardoor DRA volledig / gedeeltelijk mislukt

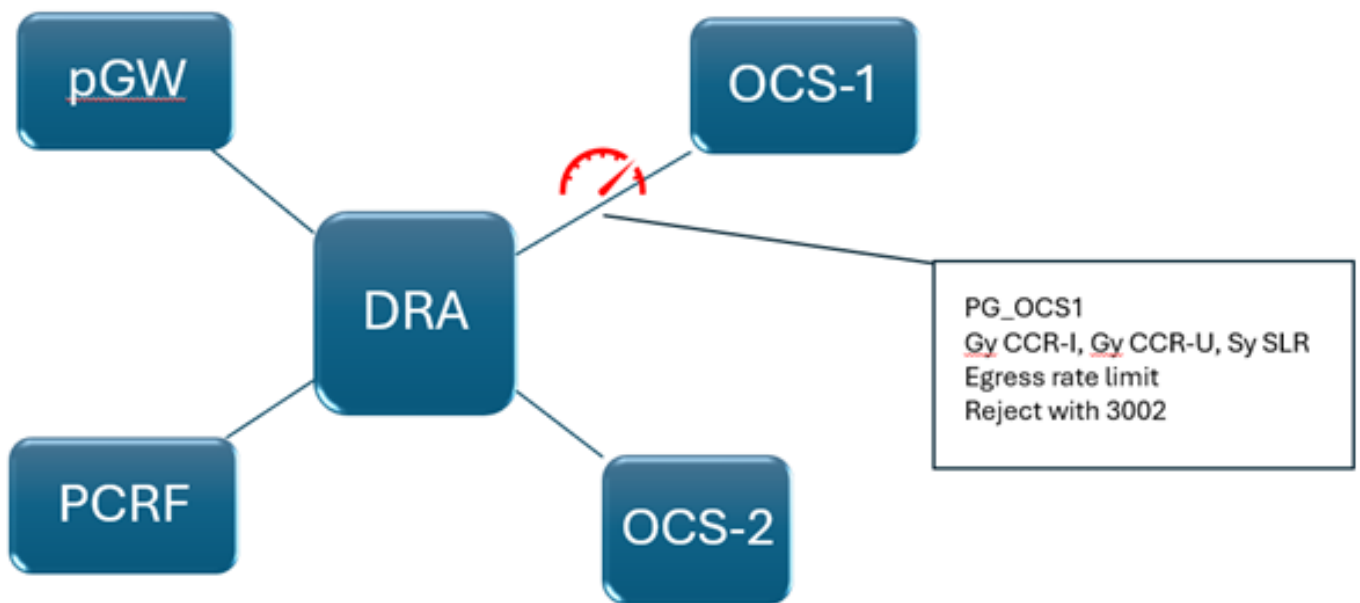
Neem een voorbeeld van een Gy-transactie die wordt uitgewisseld tussen pGW en Online Charging System (OCS). In het geval van traagheid in het netwerk op het DRA-OCS-kanaal (vanwege hoog verkeer van pGW of vanwege een ander netwerkprobleem), wordt het verzoek getimed vanwege SLA-schending. Deze time-outs hebben niet alleen invloed op DRA, maar op het totale netwerk.

DRA-bronnen worden geblokkeerd wanneer u probeert het verzoek via het trage netwerk naar OCS te verzenden, waardoor de bronnen uitgeput raken. Dit leidt ertoe dat meerdere aanvragen worden afgewezen door DRA, hoewel de nominale capaciteit van DRA niet wordt overschreden.

Dit heeft ook invloed op verkeer dat zich niet op het DRA-OCS-kanaal bevindt. Deze afwijzingen / time-outs en drops leiden tot herverbinding op meerdere netwerkelementen.

In dit geval moet u aan de uitgangskant het bericht afknijpen - DRA verstoort het bericht vlak voordat het wordt doorgestuurd naar het OCS, op basis van de OCS Peer Group (die capaciteitsbeperkingen of netwerkproblemen heeft).

Het gaspedaal gedrag kan zijn om het bericht met een bepaalde Error-Code en Error-Message af te wijzen of te laten vallen.



Dit gedrag kan worden ingeschakeld in CPS vDRA door CRD-tabellen 'Peer Rate Limit Profile' en 'Message Rate Limit Profile' te configureren. In deze CRD-tabellen moet u de volgende waarden configureren:

groep van gelijken	Een peer group is een logische groepering van Diameter-knooppunten op basis van hun rijk en host. U moet de peer-groep configureren die moet worden afgekapt.
Peer FQDN	FQDN (exacte of regex-match) voor de peers in de peer-groep die u moet beoordelen.
berichtrichting	Richting van het gaspedaal - Ingress of Egress. In dit geval - Egress.
Profiel van snelheidslimiet	De naam van het profiel voor de berichtssnelheidslimiet die wordt gebruikt om het berichttype te definiëren dat moet worden afgegrensd.
Peer Rate Limit	Percentage aanvragen dat is toegestaan voor deze Peer Group. Dit is inclusief alle berichttypen van die Peer Group.
Afdankgedrag	U kunt ervoor kiezen om het verzoek te laten vallen of af te wijzen met een foutcode.

Message Rate Limit Profile



Filter by

All Visible Columns



Gy_Sy_Limit



Rate Limit Profile Name *	Application Identifier *	Command Code *	Message/Request Type *	Message Rate Limit *	Actions
Gy_Sy_Limit	16777302	8388635	1	300	
Gy_Sy_Limit	4	272	1	500	

Showing 2 out of 2

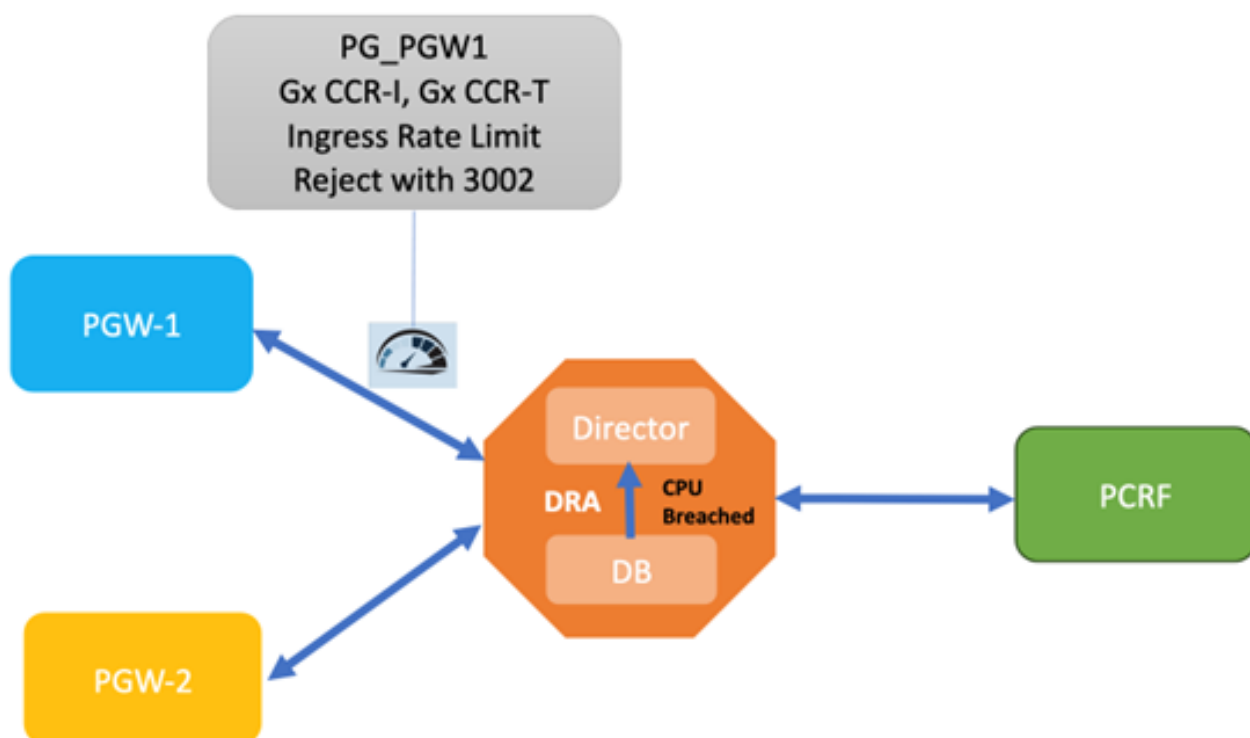
Show 50 rows 1 out of 1

dynamische snelheidslimiet

Wanneer CCR-I of CCR-T barsten optreden, kan er een overbelasting van de database (DB), die kan leiden tot het systeem te destabiliseren. Om dit te verhelpen ondersteunt DRA dynamische snelheidsbeperking (alleen voor Gx CCR-I en Gx CCR-T) op basis van de beschikbare DB-capaciteit.

DRA bewaakt het CPU-gebruik van de DB en wanneer de drempel wordt overschreden, worden de binnenkomende verzoeken afgetopt. De CPU-drempels voor throttling en het binnenkomende verkeer dat moet worden afgeknepen, zijn configureerbaar.

Verschillende CPU-drempels met bijbehorende gasklepperpercentages kunnen worden geconfigureerd. DRA past het throttling-niveau aan op basis van het huidige CPU-gebruik van de DB. Wanneer het CPU-gebruik stabiel wordt, stopt het afknijpen geleidelijk.



Dit gedrag kan worden ingeschakeld in CPS vDRA door CRD-tabellen 'Peer Rate Limit Profile', 'Message Rate Limit Profile', 'Dynamic Peer Rate Limit Profile' en 'Dynamic Throttling DB CPU Profile' te configureren. In deze CRD-tabellen moet u de volgende waarden configureren:

groep van gelijken	Een peer group is een logische groepering van Diameter-knooppunten op basis van hun rijk en host. In dit voorbeeld configureert u de peergroep van de pGW.
Peer FQDN	FQDN (exacte of regex-match) voor de peers in de peer-groep die u moet beoordelen.
Type bericht/aanvraag	Aanvraag-ID en aanvraagtype van de verzoeken die moeten worden afgegrensd. In dit voorbeeld Gx CCR-I, Gx CCR-T.
berichtrichting	Richting van het gaspedaal - Ingress of Egress. In dit geval: Ingress.
Profiel van snelheidslimiet	De naam van het profiel voor de berichtssnelheidslimiet die wordt gebruikt om het berichttype te definiëren dat moet worden afgegrensd.

Peer Rate Limit	Percentage aanvragen dat is toegestaan voor deze Peer Group. Dit is inclusief alle berichttypen van die Peer Group.
Afdankgedrag	U kunt ervoor kiezen om het verzoek te laten vallen of af te wijzen met een foutcode.
resultaatcode	Resultaatcodewaarde voor het geval u de berichten afwijst. Niet van toepassing in het geval dat de berichten worden verwijderd.
Fouttekenreeks	De fouttekenreeks die wordt gebruikt in het antwoordbericht van het verzoek dat wordt afgewezen. Niet van toepassing in het geval dat het bericht wordt verwijderd.
applicatie-id	Toepassings-ID van het bericht dat moet worden geblokkeerd.
opdrachtcode	Opdrachtcode van het bericht dat moet worden afgekapt.
berichtsnelheidslimiet	TPS van het verzoek van dat berichttype dat door DRA wordt verwerkt. Verzoeken buiten deze TPS worden beperkt. Deze waarde is per peer in de Peer Group.
Dynamic Throttling DB CPU-profiel	Dit verwijst naar de naam van het CPU-profiel, die wordt gebruikt om het gaspedaalpercentage voor verschillende CPU-bereiken te definiëren.
Gebruiksdrempel voor DB CPU	U kunt de juiste waarde van CPU-niveaus kiezen die zijn geconfigureerd als leklimieten volgens het verkeerspatroon in uw implementatie.
gaskleppercentage	Het percentage van de snelheidslimiet dat wordt toegepast wanneer het corresponderende CPU-niveau wordt overschreden.

Peer Rate Limit Profile



Filter by All Visible Columns

DRA rate								
Peer Group *	Peer FQDN *	Message Direction *	Rate Limit Profile	Peer Rate Limit	Discard Behavior *	Result Code	Error String	Actions
PG_pGW_1	match=peer-*	Ingress	CCR_I_T	1000	Send Error Answer	3002	DRA rate limit breached	

Showing 1 out of 1

Message Rate Limit Profile



Filter by All Visible Columns

CCR_I_T					
Rate Limit Profile Name *	Application Identifier *	Command Code *	Message/Request Type *	Message Rate Limit *	Actions
CCR_I_T	16777238	272	1	1000	
CCR_I_T	16777238	272	3	1000	

Showing 2 out of 2

Show 50 rows 1 out of 1

Dynamic Peer Rate Limit Profile



Filter by All Visible Columns

DynRateLimit			
Peer Group *	Peer FQDN *	Dynamic Throttling DB CPU Profile	Actions
PG_pGW_1	*	DynRateLimit	

Showing 1 out of 1

Show 50 rows 1 out of 1

Dynamic Throttling DB CPU Profile



Filter by

All Visible Columns



dyn



CPU Profile Name *	DB CPU Utilization Threshold *	Throttle Percentage	Actions
DynRateLimit	50	20	
DynRateLimit	55	30	
DynRateLimit	60	40	
DynRateLimit	65	50	

Showing 4 out of 4

Show 50 rows



1



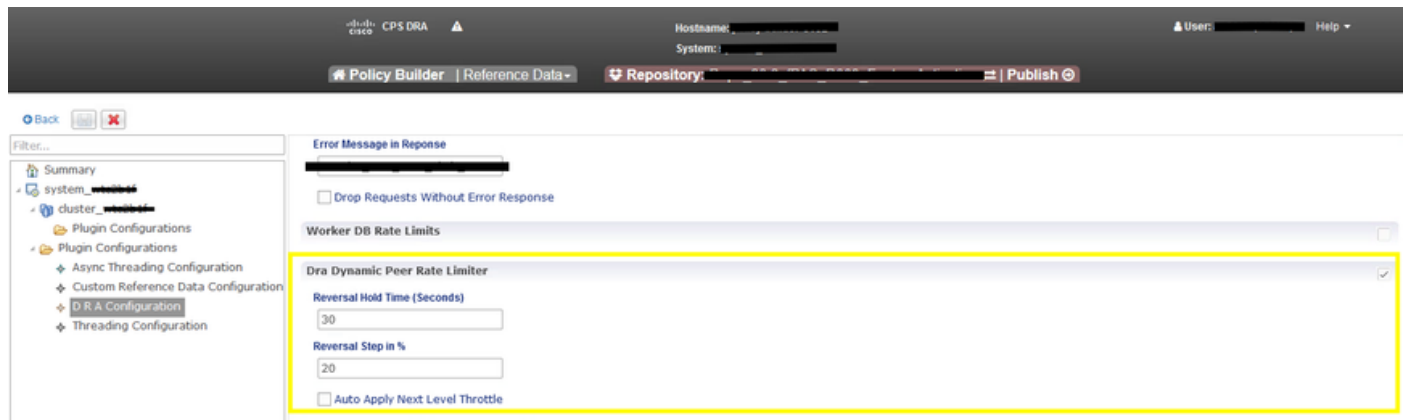
out of 1



Bovendien moet dit gedrag worden ingeschakeld door het selectievakje in Policy Builder onder 'DRA Configuration Plugin' te kiezen onder de sectie 'DRA Dynamic Peer Rate Limiter'.

Reversal Hold Time - Tijdsperiode waarin het CPU-gebruik wordt gecontroleerd voordat omkering wordt toegepast.

Omkeerstap in % - Het percentage van het omkeren van de gasklep.



Scenario: dynamische snelheidsbeperking op basis van CPU-gebruik

Overweeg deze configuratie bij DRA:

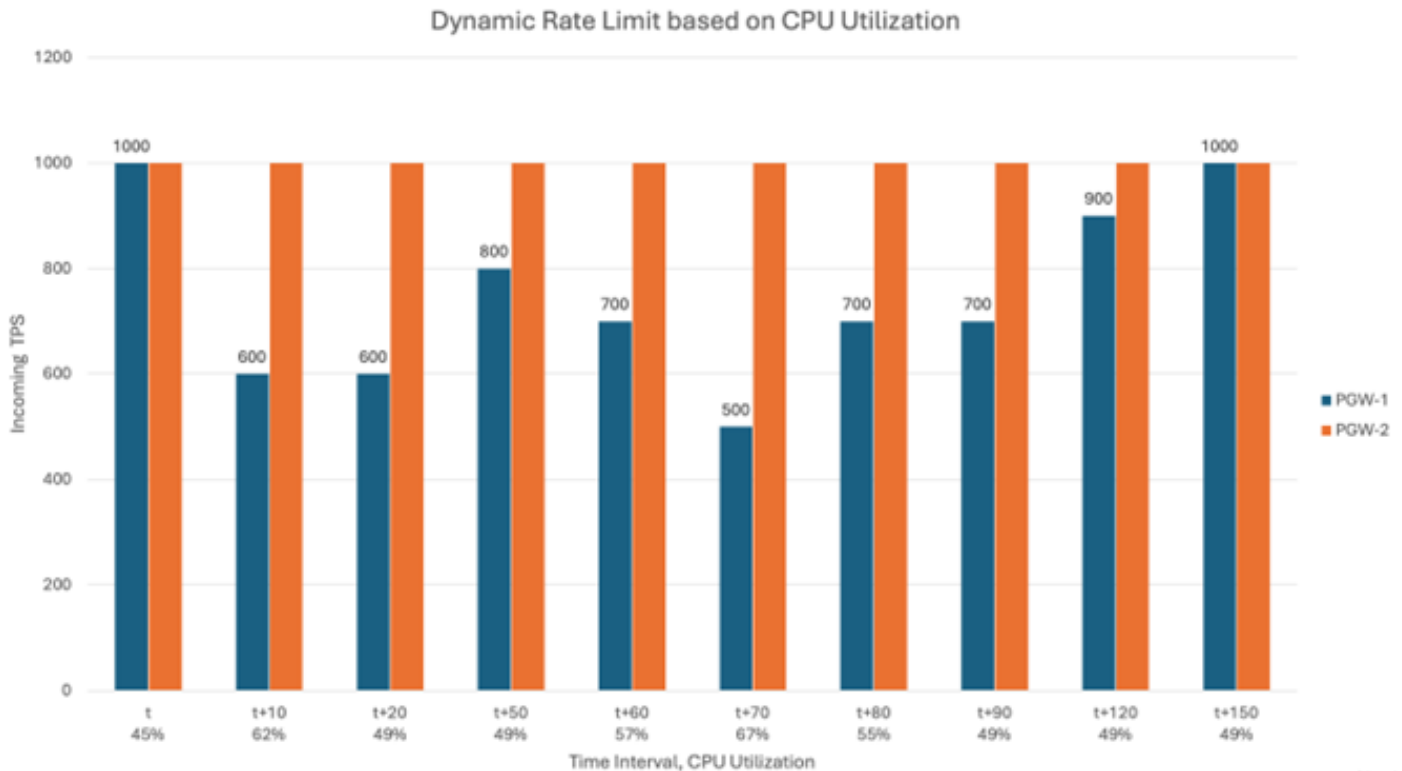
Statische berichtssnelheidslimiet: 1000 (het is dus de waarde van inkomende TPS)

Omkeerbare Hold Time: 30s

Omkeerstap in %: 20 %

Wanneer DB CPU Utilization de drempel overschrijdt, verwijst het naar de 'Dynamic Throttling DB CPU Profile'-configuratie en verstoort het de inkomende TPS dienovereenkomstig door de directeur op de hoogte te stellen. Omdat het wordt verstikt op basis van steeds veranderende

CPU-gebruikswaarden, kunt u zeggen dat de dynamische snelheid het verkeer beperkt.



- In eerste instantie is het CPU-gebruik van de DB onder de limiet, dus er vindt geen afknipping plaats. Ook PGW-2 heeft niet de configuratie van Dynamic Rate Limiting en dus geen afknippen gebeurt daar, ongeacht CPU-gebruik.
- Wanneer het CPU-gebruik van de DB 62% is, wordt het verkeer met 40% beperkt en is de effectieve snelheidslimiet 600 (inkomende TPS is 1000, DRA staat slechts 600 toe).
- Als het CPU-gebruik tussen de 60-65% blijft, blijft een beperking van 40% van toepassing op de geconfigureerde snelheidslimiet van 1000 en is de effectieve snelheidslimiet 600 (inkomende TPS is 1000, DRA staat slechts 600 toe).
- Het CPU-gebruik neemt af tot 49%, het omkeren van de throttling begint bij pGW-1.
- Als het CPU-gebruik 30 seconden lang op 49% of minder blijft, wordt de throttling verminderd met 20% tot 20%. Nu is de effectieve snelheidslimiet 800 (inkomende TPS is 1000, DRA staat slechts 800 toe.) Terwijl omkering, volgens configuratie, gebeurt het in de stappen van 20%.
- Wanneer het CPU-gebruik van de DB toeneemt tot 57%, wordt het verkeer met 30% beperkt en is de effectieve snelheidslimiet 700 (inkomende TPS is 1000, DRA staat slechts 700 toe).
- Wanneer het CPU-gebruik van de DB toeneemt tot 67%, wordt het verkeer met 50% beperkt en is de effectieve snelheidslimiet 500 (inkomende TPS is 1000, DRA staat slechts 500 toe).
- Wanneer het CPU-gebruik van de DB afneemt tot 55%, wordt het verkeer met 30% gereduceerd en is de effectieve snelheidslimiet 700 (inkomende TPS is 1000, DRA staat slechts 700 toe).
- Als de CPU daalt tot 49% of minder voor de volgende 30 seconden interval, dan is de throttling verminderd met 20% tot 10% en effectieve snelheidslimiet is 900 (inkomende TPS is 1000, DRA staat slechts 900.)
- Als de CPU verder op 49% of minder blijft gedurende de volgende 30 seconden interval, dan wordt de throttling verminderd met 20% tot 0 en er is geen snelheidslimiet toegepast als de

omkering is voltooid (inkomende TPS is 1000, DRA staat 1000 toe).

Over deze vertaling

Cisco heeft dit document vertaald via een combinatie van machine- en menselijke technologie om onze gebruikers wereldwijd ondersteuningscontent te bieden in hun eigen taal. Houd er rekening mee dat zelfs de beste machinevertaling niet net zo nauwkeurig is als die van een professionele vertaler. Cisco Systems, Inc. is niet aansprakelijk voor de nauwkeurigheid van deze vertalingen en raadt aan altijd het oorspronkelijke Engelstalige document (link) te raadplegen.