Comparaison de la politique et de la forme du trafic pour limiter la bande passante

Contenu

Introduction

Ce document décrit les différences fonctionnelles entre le formatage du trafic et la réglementation du trafic qui limitent toutes deux le débit de sortie.

Conditions préalables

Conditions requises

Aucune spécification déterminée n'est requise pour ce document.

Conventions

Pour plus d'informations sur les conventions utilisées dans ce document, reportez-vous à Conventions relatives aux conseils techniques Cisco.

Components Used

Ce document n'est pas limité à des versions de matériel et de logiciel spécifiques.

The information in this document was created from the devices in a specific lab environment. All of the devices used in this document started with a cleared (default) configuration. Si votre réseau est en ligne, assurez-vous de bien comprendre l’incidence possible des commandes.

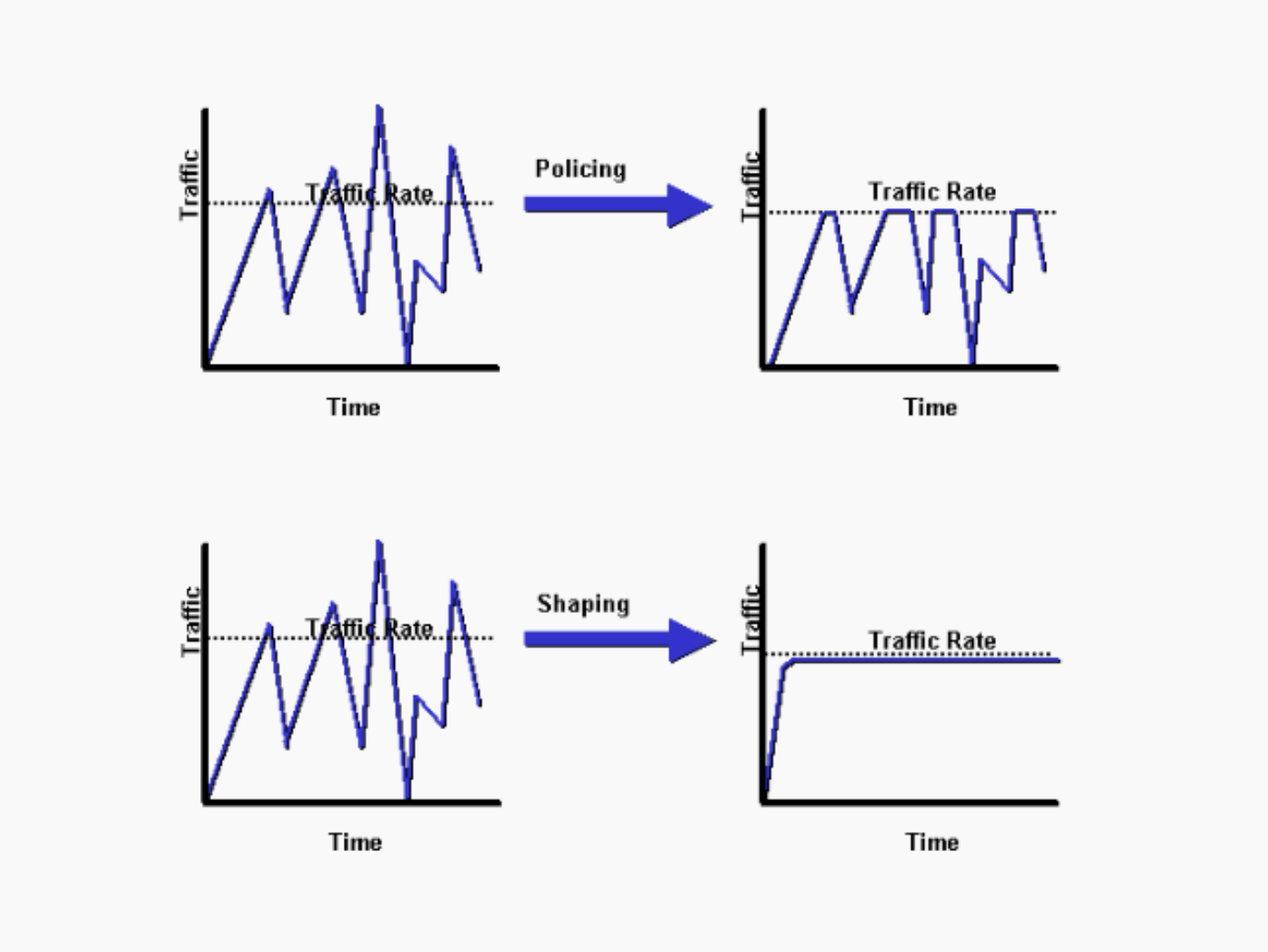
Informations générales

Ce document clarifie les différences fonctionnelles entre le formatage du trafic et la réglementation. Les deux limitent fonctionnellement le débit de sortie du trafic. Les deux mécanismes utilisent un saut de jeton comme indicateur de trafic pour mesurer le débit de paquets. Pour plus d'informations sur les compartiments de jetons, consultez Qu'est-ce qu'un compartiment de jetons ?

Réglementation et formatage

La réglementation de trafic propage des rafales. Quand le débit de trafic atteint le débit maximal configuré, le trafic excessif est extrait (ou marqué une nouvelle fois). Le résultat est un débit en sortie qui apparaît en dents de scie avec des hauts et des bas. Contrairement à la réglementation, le formatage de trafic retient les paquets en excès dans une file d'attente, puis les programme pour une transmission postérieure sur des incréments de temps. Le résultat du formatage de trafic est un débit en sortie en douceur de paquets.

Le schéma suivant illustre les principales différences entre les deux options de trafic.



Réglementation et mise en forme

Le formatage implique l'existence d'une file d'attente et de la mémoire suffisante pour mettre en mémoire tampon des paquets différés, contrairement à la réglementation. Les files d'attente sont un concept sortant ; les paquets qui quittent une interface sont mis en file d'attente et peuvent être formatés. Seule la réglementation peut être appliquée au trafic entrant sur une interface. Assurez-vous de disposer de suffisamment de mémoire lorsque vous activez le formatage. En outre, la mise en forme nécessite une fonction qui planifie la transmission ultérieure de tout paquet retardé.

Cette fonctionnalité de planification vous permet d'organiser la file d'attente de formatage en différentes files d'attente. Exemples de cette fonctionnalité : Class Based Weighted Fair Queuing (CBWFQ) et faible latence Queuing (LLQ).

Critères de sélection

Le tableau suivant répertorie les différences entre la mise en forme et la réglementation pour vous aider à choisir la solution de trafic appropriée.

	Mise En Forme	Contrôle
Objectif	Mettre en mémoire tampon et en file d'attente les paquets excédentaires sur les débits validés.	Abandonner (ou marquer) les paquets en excès pa rapport aux débits validés. Pas de mise en mémoi tampon.*
Fréquence d'actualisation des jetons	Incrémentée au début d'un intervalle de temps. (Un nombre minimal d'intervalles est requis.)	Continu basé sur la formule : `1 / débit d'informat garanti`
Valeurs des jetons	Configurées en bits par seconde.	Configurées en octets.
Options de configuration	• **Commande shape** dans l'interface de ligne de commande QoS modulaire (MQC) pour mettre en application le formatage basé sur les classes. • **Commande frame-relay traffic-shape** pour mettre en application le formatage de trafic de relais de trame (FRTS). • **t traffic-shape** pour implémenter Generic Traffic Shaping (GTS).	• **Commande police** dans MQC pour mettre en application la réglementation basée sur les cla • **Commande rate-limit** pour mettre en applicatio Committed Access Rate (CAR).
Applicable en entrée	Non	Oui
Applicable en sortie	Oui	Oui
Rafales	Contrôle les rafales et lisse le débit de sortie sur au moins huit intervalles de temps. Utilise un saut percé pour retarder le trafic, ce qui aboutit à un effet de lissage.	Propage les rafales. Pas de lissage.
Avantages	Moins de risques d'extraction des paquets excédentaires puisque ceux-ci sont mis en mémoire tampon. (Met les paquets en mémoire tampon sur toute la longueur de la file d'attente. Des abandons peuvent se produire si le trafic excédentaire est maintenu à des taux élevés.) Évite typiquement les retransmissions en raison des paquets extraits.	Contrôle le débit en sortie par des extractions de paquets. Évite les retards dus à queuing.
Inconvénients	Peut introduire un délai dû à queuing,	Abandonne les paquets excédentaires (une fois

	en particulier les files d'attente longues.	configurés), limite la taille des fenêtres TCP et rédu taux de sortie global des flux de trafic affectés. Des tailles de salves trop agressives peuvent entraîner pertes de paquets excessives et limiter le débit de global, en particulier avec les flux basés sur TCP.
Nouveau marquage facultatif des paquets	Non	Oui (avec fonctionnalité CAR existante).

* Bien que la réglementation n'applique pas de tampon, un queuing s'applique aux paquets **conformes** qui peuvent avoir besoin d'être mis en file d'attente pendant qu'ils attendent d'être sérialisés à l'interface physique.

Fréquence d'actualisation des jetons

Une différence clé entre la mise en forme et la réglementation est la vitesse à laquelle les jetons sont réapprovisionnés. La mise en forme et la réglementation utilisent toutes deux la métaphore du compartiment de jetons. Un saut à jetons lui-même n'a aucune stratégie de rejet ni de priorité.

Avec la fonctionnalité token bucket :

- Les jetons sont placés dans le saut à un certain débit.
- Chaque jeton est une autorisation pour la source d'envoyer un certain nombre de bits dans le réseau.
- Pour envoyer un paquet, le régulateur de trafic doit pouvoir retirer du saut un certain nombre de jetons égal dans la représentation à la taille de paquet.
- Si le nombre de jetons dans le saut est insuffisant pour envoyer un paquet, le paquet attend que le saut ait assez de jetons (dans le cas d'un modélisateur), ou le paquet est ignoré ou démarqué (dans le cas d'un régulateur).
- Le saut lui-même a une capacité spécifique. Si le compartiment atteint sa capacité maximale, les nouveaux jetons qui arrivent sont rejetés et ne sont pas disponibles pour les paquets futurs. Ainsi, à tout moment, la plus grande rafale qu'une source peut envoyer dans le réseau est approximativement proportionnelle à la taille du saut. Un saut de jeton permet la rafale mais la limite.

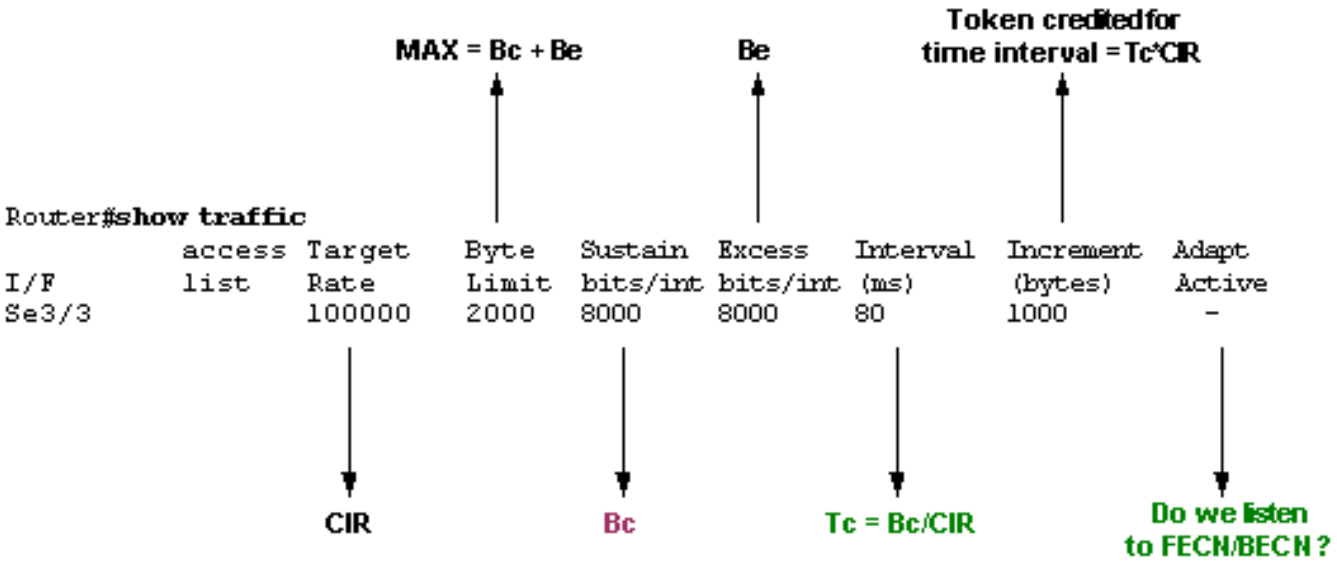
La mise en forme incrémente le compartiment de jeton à des intervalles temporisés qui utilisent une valeur de bits par seconde (bits/s). Un modélisateur utilise la formule suivante :

```
Tc = Bc/CIR (in seconds)
```

Dans cette équation, Bc représente la rafale validée et CIR représente le taux d'informations obligatoires. (Référez-vous à Configuration du formatage de trafic de relais de trame pour plus d'informations.) La valeur de Tc définit le délai pendant lequel vous envoyez les bits Bc afin de maintenir le débit moyen de CIR en secondes.

La plage de Tc se situe entre 10 ms et 125 ms. Avec le formatage de trafic distribué (DTS) sur la gamme Cisco 7500, le Tc minimum est de 4 ms. Le routeur calcule en interne cette valeur selon les valeurs CIR et Bc. Si Bc/CIR est inférieur à 125 ms, il utilise le Tc calculé à partir de cette équation. Si Bc/CIR est supérieur ou égal à 125 ms, il utilise une valeur Tc interne si Cisco IOS®

détermine que le flux de trafic peut être plus stable avec un intervalle plus petit. Utilisez la commande **show traffic-shape** pour déterminer si votre routeur utilise une valeur interne pour Tc ou la valeur que vous avez configurée à la ligne de commande. L'exemple de sortie suivant de la commande **show traffic-shape** est expliqué dans les commandes show pour le formatage du trafic Frame Relay.



show traffic output

Quand la rafale en excès (Be) est configurée à une valeur différente de 0, le modélisateur permet de stocker des jetons dans le saut, jusqu'à Bc + Be. La plus grande valeur que le saut à jetons peut jamais atteindre est Bc + Be, et les jetons de dépassement sont extraits. La seule façon d'avoir plus de Bc jetons dans le saut est de ne pas utiliser tous les Bc jetons pendant un ou plusieurs Tc. Puisque le saut à jetons est réapprovisionné tous les Tc avec Bc jetons, vous pouvez accumuler les jetons inutilisés pour une utilisation ultérieure jusqu'à Bc + Be.

En revanche, la réglementation basée sur les classes et leslimiting ajoute des jetons en continu au compartiment. Spécifiquement, le débit d'arrivée des jetons est calculé comme suit :

```
(time between packets<which is equal to t-t1>* policer rate)/8 bits per byte
```

En d'autres termes, si l'arrivée précédente du paquet était à t1 et l'heure actuelle est t, le saut est mis à jour avec la valeur t-t1 des octets selon le débit d'arrivée des jetons. Notez qu'un régulateur de trafic utilise des valeurs de rafale spécifiées en octets, et la formule précédente convertit les bits en octets.

Voici un exemple qui utilise un débit de données garanti (CIR) de 8 000 bits/s et une rafale normale de 1 000 octets :

```
Router(config)# policy-map police-setting
Router(config-pmap)# class access-match
Router(config-pmap-c)# police 8000 1000 conform-action transmit exceed-action drop
```

Les sauts à jetons sont pleins à 1 000 octets. Si un paquet de 450 octets arrive, le paquet est conforme parce qu'assez d'octets sont disponibles dans le saut à jetons. L'action de conformité (transmission) est effectuée par le paquet et 450 octets sont supprimés du compartiment à jetons (et laissent 550 octets). Si le paquet suivant arrive .25 secondes plus tard, 250 octets sont ajoutés au compartiment de jeton selon la formule suivante :

```
(0.25 * 8000)/8
```
Le calcul laisse 700 octets dans le saut à jetons. Si le paquet suivant est de 800 octets, le paquet dépasse et la mesure exceed (drop) est prise. Aucun octet n'est pris du saut à jetons.

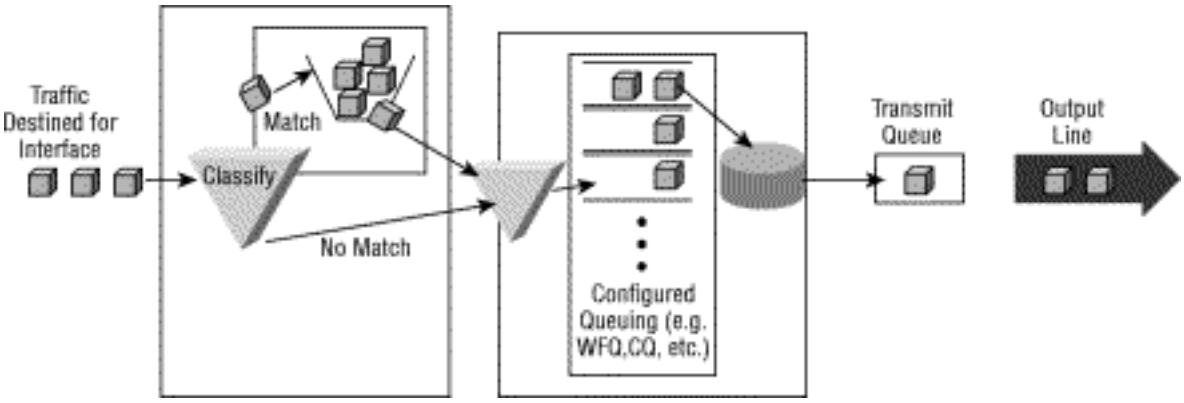
Modélisation du trafic

Cisco IOS prend en charge les méthodes suivantes de formatage du trafic :

- Generic Traffic Shaping
- Formatage de trafic de relais de trame
- Formatage basé sur les classes et Formatage basé sur les classes distribuées

Toutes les méthodes de formatage de trafic sont semblables dans la mise en place, bien que leur interface de ligne de commande (CLI) diffère un peu, et elles emploient différents types de files d'attente pour contenir et formater le trafic qui est reporté. Cisco recommande le formatage basé sur les classes et le formatage distribué, qui sont configurés avec l'interface de ligne de commande modulaire QoS.

Le schéma suivant illustre comment une stratégie QoS organise le trafic en classes et met en file d'attente les paquets qui dépassent les débits de mise en forme configurés.



Contrôle du trafic

Cisco IOS prend en charge les méthodes suivantes de réglementation du trafic :

- Committed Access Rate
- Réglementation basée sur les classes

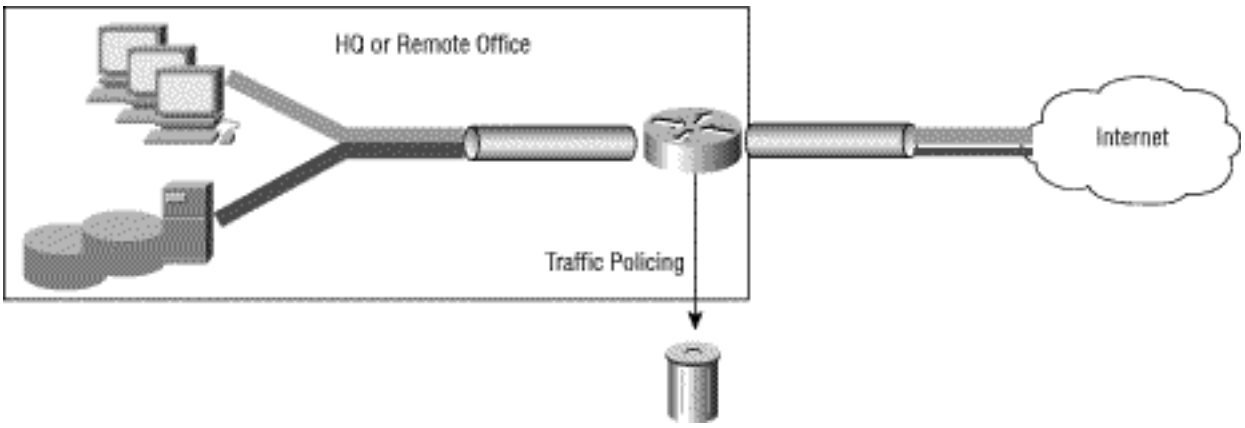
Les deux mécanismes présentent d'importantes différences fonctionnelles, comme expliqué dans la section Comparaison de la réglementation par classe et du débit d'accès garanti. Cisco recommande la réglementation basée sur les classes et d'autres fonctionnalités de l'interface de ligne de commande QoS modulaire lorsque des politiques QoS sont appliquées.

Utilisez la commande **police** pour spécifier qu'une classe de trafic doit avoir un débit maximal qui lui est imposé, et si ce débit est dépassé, une action immédiate doit être prise. En d'autres termes, avec la commande **police, ce n'est pas une option de mettre en mémoire tampon le paquet pour l'envoyer plus tard, comme c'est le cas pour la commande shape.**

En outre, avec la réglementation, le saut à jetons détermine si un paquet dépasse le débit appliqué ou est conforme. Dans les deux cas, la réglementation implémente une action

configurable, qui inclut la priorité IP ou le DSCP (Differentiated Services Code Point).

Le schéma suivant illustre une application courante de la réglementation du trafic au niveau d'un point de congestion, où les fonctionnalités QoS s'appliquent généralement.



Contrôles de bande passante minimum et maximum

Les deux commandes **shape et police limitent le débit en sortie à une valeur maximale en Kbits/s.** Essentiellement, aucun mécanisme ne fournit une garantie de bande passante minimale au cours des périodes d'encombrement. Utilisez la commande **bandwidth ou priority pour fournir de telles garanties.**

Une stratégie hiérarchique utilise deux stratégies de service : une stratégie parent pour appliquer un mécanisme QoS à un agrégat de trafic et une stratégie enfant pour appliquer un mécanisme QoS à un flux ou à un sous-ensemble de l'agrégat. Les interfaces logiques, telles que les sous-interfaces et les interfaces de tunnel, nécessitent une stratégie hiérarchique avec le traficlimiting au niveau parent et mise en file d'attente aux niveaux inférieurs. Le trafic-limiting réduit le débit de sortie et (vraisemblablement) crée un encombrement, comme le montre queuing paquets en excès.

La configuration suivante est sous-optimale et est montrée pour illustrer la différence entre la commande **police** et la commande **shape** quand limiting un trafic agrégé, dans ce cas class-default, à un débit maximal. Dans cette configuration, la commande **police** envoie des paquets à partir des classes enfants en fonction de la taille du paquet et du nombre d'octets qui restent dans les compartiments de jeton conformes et supérieurs. (Référez-vous à Réglementation du trafic.) Il en résulte que les débits donnés aux classes Voix sur IP (VoIP) et Protocole Internet (IP) ne peuvent pas être garantis puisque la fonctionnalité de **police** supplante les garanties apportées par la fonctionnalité de **priorité**.

Cependant, si la commande shape est utilisée, le résultat est un système de mise en file d'attente hiérarchique, et toutes les garanties sont faites. En d'autres termes, quand la charge offerte dépasse le débit de format, les classes VoIP et IP ont leur débit garanti, et le trafic class-default (au niveau enfant) n'encourt aucune extraction.

> **Attention** : Cette configuration n'est pas recommandée et est montrée pour illustrer la différence entre la commande **police** et la commande **shape** quand elle limite un agrégat de trafic.

```
class-map match-all IP
   match ip precedence 3
```

```
class-map match-all VoIP
   match ip precedence 5

policy-map child
  class VoIP
     priority 128
   class IP
     priority 1000

 policy-map parent
   class class-default
     police 3300000 103000 103000 conform-action transmit exceed-action drop
     service-policy child
```

Pour que la configuration précédente ait un sens, la réglementation doit être remplacée par une mise en forme. Exemple :

```
policy-map parent
    class class-default
     shape average 3300000 103000 0
      service-policy child
```

> **Note**: Pour en savoir plus sur les stratégies parent et enfant, reportez-vous à Politique de service QoS enfant pour la classe de priorité .

Informations connexes

- Assistance technologique pour la qualité de service (QoS)
- Support et documentation techniques - Cisco Systems

À propos de cette traduction
Cisco a traduit ce document en traduction automatisée vérifiée par une personne dans le cadre d'un service mondial permettant à nos utilisateurs d'obtenir le contenu d'assistance dans leur propre langue.

Il convient cependant de noter que même la meilleure traduction automatisée ne sera pas aussi précise que celle fournie par un traducteur professionnel.