

Planification de la sortie QoS sur les commutateurs des gammes Catalyst 6500/6000 exécutant le logiciel système CatOS

Contenu

[Introduction](#)

[Conditions préalables](#)

[Conditions requises](#)

[Composants utilisés](#)

[Conventions](#)

[Informations générales](#)

[Pertes de la file d'attente de sortie](#)

[Tape de la Mise en file d'attente qui sont impliqués dans la planification de sortie sur le Catalyst 6500/6000](#)

[Perte de destination](#)

[Dépistage précoce aléatoire et dépistage précoce aléatoire pesé](#)

[Circulaire pesé](#)

[File d'attente prioritaire stricte](#)

[Capacité de la file d'attente de sortie de différents linecards sur le Catalyst 6000](#)

[capacités de commande de show port](#)

[Comprenez la capacité de mise en file d'attente d'un port](#)

[Créer QoS sur le Catalyst 6500/6000](#)

[Mécanisme de planification de sortie sur le Catalyst 6500/6000](#)

[Configuration, surveillance, et planification de sortie sur le Catalyst 6500/6000](#)

[Configuration par défaut pour QoS sur le Catalyst 6500/6000](#)

[Configuration](#)

[Le Scheduling de sortie moniteur et vérifie la configuration](#)

[Planification de sortie d'utilisation pour réduire le retard et instabilité](#)

[Réduisez le retard](#)

[Réduisez le jitter](#)

[Informations connexes](#)

Introduction

La planification de sortie s'assure que l'important trafic n'est pas abandonné en cas de surabonnement lourd. Ce document discute tous les techniques et algorithmes qui sont impliqués dans la planification de sortie sur les Commutateurs de gamme Cisco Catalyst 6500/6000 qui exécutent le logiciel système OS de Catalyst (CatOS). Ce document fournit également une brève présentation de la capacité de mise en file d'attente de Commutateurs du Catalyst 6500/6000 et comment configurer les différents paramètres de la planification de sortie.

Remarque: Si vous exécutez le logiciel de Cisco IOS® sur votre Catalyst 6500/6000, référez-vous à la [planification de sortie de QoS sur des Commutateurs de gamme Catalyst 6500/6000 exécutant le](#) pour en savoir plus de [logiciel système de Cisco IOS](#).

Conditions préalables

Conditions requises

Aucune spécification déterminée n'est requise pour ce document.

Composants utilisés

Les exemples dans ce document ont été créés d'un Catalyst 6000 avec une engine 1A de superviseur et une carte de fonctionnalité de stratégie (PFC). Mais les exemples sont également valides pour un Supervisor Engine 2 avec un PFC2 ou pour une engine 720 de superviseur avec un PFC3.

Les informations contenues dans ce document ont été créées à partir des périphériques d'un environnement de laboratoire spécifique. Tous les périphériques utilisés dans ce document ont démarré avec une configuration effacée (par défaut). Si votre réseau est opérationnel, assurez-vous que vous comprenez l'effet potentiel de toute commande.

Conventions

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

Informations générales

Pertes de la file d'attente de sortie

Les pertes en sortie sont provoquées par une interface congestionnée. Une cause classique de ceci pourrait être le trafic d'une liaison à large bande passante qui est commuté à un lien ou à un trafic de bande passante inférieure des plusieurs liens d'arrivée qui sont commutés à un lien sortant simple.

Par exemple, si un grand volume de trafic en rafales entre sur une interface Gigabit et est commuté vers une interface 100Mbps, cela pourrait entraîner l'incréméntation des suppressions en sortie sur l'interface 100Mbps. C'est parce que la file d'attente de sortie sur cette interface est accablée par le trafic excédentaire dû à la non-concordance de vitesse entre les bandes passantes entrantes et sortantes. Le débit de trafic sur l'interface sortante ne peut pas recevoir tous les paquets qui devraient être envoyés.

La solution ultime pour résoudre le problème est d'augmenter la vitesse de la ligne. Cependant, il y a des façons d'empêcher, de diminuer ou de contrôler les pertes en sortie quand vous ne voulez pas augmenter la vitesse de la ligne. Vous pouvez empêcher les pertes en sortie seulement si elles sont une conséquence de brèves rafales de données. Si les pertes en sortie sont provoquées par un flux constant à haut débit, vous ne pouvez pas empêcher les pertes. Cependant, vous pouvez les contrôler.

[Tape de la Mise en file d'attente qui sont impliqués dans la planification de sortie sur le Catalyst 6500/6000](#)

[Perte de destination](#)

La perte de destination est un mécanisme de base de manière d'éviter d'encombrement. La perte de destination traite tout le trafic également et ne le différencie pas entre les classes de service (CoSs) quand les files d'attente commencent à remplir au cours des périodes d'encombrement. Quand la file d'attente de sortie est pleine et la perte de destination est en vigueur, des paquets sont lâchés jusqu'à ce que l'encombrement soit éliminé et la file d'attente n'est plus pleine. La perte de destination est le type le plus fondamental de manière d'éviter d'encombrement et ne prend pas en considération n'importe quel paramètre de QoS.

Le Catalyst 6000 a mis en application une version avancée de la manière d'éviter d'encombrement de perte de destination qui relâche tous les paquets avec certain cos quand un certain pourcentage de taux de remplissage de la mémoire tampon est atteint. Avec le Weighted Tail Drop, vous pouvez définir un ensemble de seuils et associer le cos avec chaque seuil. Dans l'exemple dans cette section, il y a quatre seuils possibles. Les définitions de chaque seuil sont :

- Le seuil 1 est atteint quand 50 pour cent de la mémoire tampon sont remplis. Le cos 0 et 1 est assigné à ce seuil.
- Le seuil 2 est atteint quand 60 pour cent de la mémoire tampon sont remplis. Le cos 2 et 3 est assigné à ce seuil.
- Le seuil 3 est atteint quand 80 pour cent de la mémoire tampon sont remplis. Le cos 4 et 5 est assigné à ce seuil.
- Le seuil 4 est atteint quand 100 pour cent de la mémoire tampon sont remplis. Le cos 6 et 7 est assigné à ce seuil.

Dans le diagramme dans la [figure 1](#), tous les paquets avec cos de 0 ou de 1 sont lâchés si la mémoire tampon est de 50 pour cent remplis. Tous les paquets avec cos de 0, de 1, de 2, ou de 3 sont lâchés si les mémoires tampons sont de 60 pour cent remplis. Des paquets avec cos de 6 ou de 7 sont lâchés quand les mémoires tampons sont complètement remplis.

Figure 1

Remarque: Dès que le taux de remplissage de la mémoire tampon chutera au-dessous d'un certain seuil, des paquets avec le cos associé ne sont plus lâchés.

[Dépistage précoce aléatoire et dépistage précoce aléatoire pesé](#)

Le Détection précoce directe pondérée (WRED) est un mécanisme de manière d'éviter d'encombrement qui relâche aléatoirement des paquets avec une certaine Priorité IP quand les mémoires tampons atteignent un seuil remplissant défini. WRED est une combinaison de ces deux caractéristiques :

- Perte de destination
- Détection précoce aléatoire (RED)

Le ROUGE n'est pas priorité-averti ou Cos-averti. Le ROUGE utilise un des seuils simples quand la valeur seuil pour la mémoire tampon rempli. Des débuts de ROUGE pour relâcher aléatoirement des paquets (mais non tous les paquets, comme dans la perte de destination) jusqu'au seuil (maximum) maximum est atteints. Après que le seuil maximum soit atteint, tous les

paquets sont lâchés. La probabilité qu'un paquet est des augmentations lâchées linéairement avec l'augmentation de taux de remplissage de la mémoire tampon au-dessus du seuil. Le diagramme dans la [figure 2](#) affiche la probabilité de perte de paquets :

Figure 2 – Probabilité de rejet de paquet

Remarque: La probabilité de marque dans ce diagramme est réglable en ROUGE, ainsi il signifie que la pente de la probabilité de perte linéaire est réglable.

Le ROUGE et les WRED sont les mécanismes très utiles de manière d'éviter d'encombrement pour le trafic basé sur TCP. Pour d'autres types de trafic, le ROUGE n'est pas très efficace. C'est parce que le ROUGE tire profit du mécanisme de fenêtrage que le TCP l'utilise pour gérer l'encombrement. Le ROUGE évite l'encombrement typique qui se produit sur un routeur quand les plusieurs sessions TCP passent par le même port de routeur. Le mécanisme s'appelle la synchronisation de réseau global. Le diagramme dans la [figure 3](#) affiche comment le ROUGE exerce un effet doux sur le chargement :

Figure 3 – ROUGE pour la manière d'éviter d'encombrement

Pour plus d'informations sur la façon dont le ROUGE peut réduire l'encombrement et lisser le trafic par le routeur, référez-vous [le comment le routeur interagit avec la](#) section de [TCP de l'aperçu de manière d'éviter d'encombrement de](#) document.

WRED est semblable au ROUGE du fait chacun des deux définissent quelques seuils et, quand ces seuils minimum sont atteints, paquets (minimum) minimum sont aléatoirement relâchés. WRED définit également certains seuils maximum et, quand ces seuils maximum sont atteints, tous les paquets sont lâchés. WRED est également Cos-averti, ainsi il signifie qu'un ou plusieurs valeurs CoS sont ajoutées à chaque seuil minimum/paire seuil maximum. Quand le minute-seuil est dépassé, des paquets sont aléatoirement lâchés avec le cos qui est assigné. Considérez cet exemple avec deux seuils dans la file d'attente :

- Le cos 0 et 1 est assigné au seuil minimum 1 et au seuil maximum 1. le seuil minimum 1 est placé à 50 pour cent de taux de remplissage de la mémoire tampon, et le seuil maximum 1 est placé à 80 pour cent.
- Le cos 2 et 3 est assigné au seuil minimum 2 et au seuil maximum le seuil 2 de 2. minutes est placé à 70 pour cent de taux de remplissage de la mémoire tampon, et le seuil maximum 2 est placé à 100 pour cent.

Dès que la mémoire tampon dépassera le seuil minimum 1 (50 pour cent), les paquets avec le début du cos 0 et 1 à abandonner aléatoirement. Plus de paquets sont lâchés pendant que l'utilisation de la mémoire tampon se développe. Si le seuil minimum 2 (70 pour cent) est atteint, des paquets avec le cos 2 et 3 commencent à être aléatoirement lâchés.

Remarque: À ce stade, la probabilité de perte pour des paquets avec le cos 0 et 1 est beaucoup de supérieur à la probabilité de perte pour des paquets avec le cos 2 ou le cos 3.

Toutes les fois que le seuil maximum 2 est atteint, les paquets avec le cos 0 et 1 sont tous lâchés, alors que des paquets avec le cos 2 et 3 continuent à être aléatoirement lâchés. En conclusion, quand 100 pour cent sont atteints (le seuil maximum 2), tous les paquets avec le cos 2 et 3 sont abandonnés.

Les diagrammes dans la [figure 4](#) et la [figure 5](#) montrent un exemple de ces seuils :

Figure 4 – WRED avec deux ensembles de seuils de minute et de seuils maximum (deux services) Figure 5 – WRED avec deux ensembles de services, mais égal minimum 0 de les deux

seuils

L'implémentation tôt de CatOS de WRED a seulement placé le seuil maximum, alors que le seuil minimum était dur codé à 0 pour cent. La partie inférieure du diagramme dans la [figure 5](#) met en valeur le comportement en résultant.

Remarque: La probabilité de perte pour un paquet est toujours non nulle parce que cette probabilité est toujours au-dessus du seuil minimum. Ce comportement a été corrigé dans la version de logiciel 6.2 et plus tard.

[Circulaire pesé](#)

Circulaire pesé (WRR) est un autre mécanisme pour la planification de sortie sur le Catalyst 6000. Travaux WRR entre deux files d'attente ou plus. Les files d'attente pour WRR sont vidées d'une permutation circulaire, et vous pouvez configurer le poids pour chaque file d'attente. Par défaut, les ports ont deux files d'attente WRR sur le Catalyst 6000. Le par défaut est :

- Pour servir le WRR prioritaire alignez 70 pour cent de temps
- Pour servir le WRR non prioritaire alignez 30 pour cent de temps

Le diagramme dans la [figure 6](#) affiche un WRR qui a trois files d'attente qui sont servies d'une mode WRR. La file d'attente prioritaire (paquets rouges) envoie plus de paquets que les deux autres files d'attente :

Figure 6 – Planification de sortie : WRR

Remarque: La plupart des 6500 linecards implémentent WRR par bande passante. Cette implémentation de WRR par bande passante signifie que chaque fois que le programmeur permet à une file d'attente pour transmettre des paquets, on laisse un certain nombre d'octets pour être transmis. Ce nombre d'octets peut représenter plus d'un paquet. Par exemple, si vous envoyez 5120 octets à un tour, vous pouvez envoyer trois paquets 1518-byte, pour un total de 4554 octets. Les octets supérieurs sont perdus ($5120 - 4554 = 566$ octets). Par conséquent, avec un certain poids extrême (comme 1 pour cent pour la file d'attente 1 et 99 pour cent pour la file d'attente 2), le poids configuré précis ne peuvent être atteints. Ce manque d'atteindre pour exiger le poids est souvent la caisse pour de plus grands paquets.

Quelques linecards de la nouvelle génération, comme le 6548-RJ-45, surmontent cette limite par l'implémentation de circulaire pesé par déficit (DWRR). DWRR transmet des files d'attente mais ne meurt pas de faim la file d'attente à basse priorité. DWRR maintient la file d'attente à basse priorité qui est sous la transmission et compense dans le prochain Round.

[File d'attente prioritaire stricte](#)

Un autre type de file d'attente dans le Catalyst 6000, une file d'attente prioritaire stricte, est toujours vidé d'abord. Dès qu'il y aura un paquet dans la file d'attente prioritaire stricte, le paquet est envoyé.

Les files d'attente WRR ou WRED sont vérifiées seulement après que la file d'attente prioritaire stricte est vidée. Après que chaque paquet soit transmis de la file d'attente WRR ou de la file d'attente WRED, la file d'attente prioritaire stricte est vérifiée et vidée, s'il y a lieu.

Remarque: Tous les linecards avec un type de Mise en file d'attente semblable 1p2q1t, 1p3q8t, et 1p7q8t à l'utilisation DWRR. L'autre utilisation WRR standard de linecards.

Capacité de la file d'attente de sortie de différents linecards sur le Catalyst 6000

capacités de commande de show port

Si vous n'êtes pas sûr au sujet de la capacité de mise en file d'attente d'un port, vous pouvez émettre la commande de **show port capabilities**. C'est la sortie de la commande sur un linecard WS-X6408-GBIC :

```
Model                WS-X6408-GBIC
Port                 4/1
Type                 No GBIC
Speed                1000
Duplex                full
Trunk encap type     802.1Q,ISL
Trunk mode            on,off,desirable,auto,nonegotiate
Channe                yes
Broadcast suppression percentage(0-100)
Flow control          receive-(off,on,desired),send-(off,on,desired)
Security              yes
MembershIP            static,dynamic
Fast start            yes
QOS scheduling        rx-(1q4t),tx-(2q2t)
CoS rewrite           yes
ToS rewrite           DSCP
UDLD                  yes
SPAN                  source,destination
COPS port group      none
```

Ce port a un type de Mise en file d'attente sorti qui s'appelle le 2q2t.

Comprenez la capacité de mise en file d'attente d'un port

Il y a plusieurs types de files d'attente qui sont disponibles sur des Commutateurs du Catalyst 6500/6000. Les tables dans cette section peuvent devenir inachevées pendant que de nouveaux linecards sont libérés. Les nouveaux linecards peuvent introduire de nouvelles combinaisons de Mise en file d'attente. Pour une description en cours de toute la Mise en file d'attente qui est disponible pour des modules de commutation du Catalyst 6500/6000, référez-vous à la section de *configuration de QoS* pour votre version de CatOS de [documentation du logiciel de gamme Catalyst 6500](#).

Remarque: Le module de supports de communication de Cisco (CMM) ne prend en charge pas toutes les caractéristiques de QoS. Vérifiez les notes de mise à jour pour votre version logicielle spécifique afin de déterminer les caractéristiques qui sont prises en charge.

Cette table explique la notation de l'architecture de QoS de port :

$\frac{T}{R}$ x2 ide	Notation de file d'attente	Non des files d'attente	File d'attente prioritaire	Non des files d'attente WRR	No. et type de seuil pour des files d'attente WRR
Tx	2q2t	2		2	perte de destination 2

					configurable
Tx	1p2q2t	3	1	2	2 WRED configurables
Tx	1p3q1t	4	1	3	1 WRED configurable
Tx	1p2q1t	3	1	2	1 WRED configurable
Rx	1q4t	1		1	perte de destination 4 configurable
Rx	1p1q4t	2	1	1	perte de destination 4 configurable
Rx	1p1q0t	2	1	1	Non configurable
Rx	1p1q8t	2	1	1	8 WRED configurables

¹ Tx = transmettent.

² Rx = reçoivent.

Ce tableau présente tous les modules et la file d'attente saisit le côté de Rx et de Tx de l'interface ou du port :

Module	Files d'attente de Rx	Files d'attente de Tx
WS-X6K-S2-PFC2	1p1q4t	1p2q2t
WS-X6K-SUP1A-2GE	1p1q4t	1p2q2t
WS-X6K-SUP1-2GE	1q4t	2q2t
WS-X6501-10GEX4	1p1q8t	1p2q1t
WS-X6502-10GE	1p1q8t	1p2q1t
WS-X6516-GBIC	1p1q4t	1p2q2t
WS-X6516-GE-TX	1p1q4t	1p2q2t
WS-X6416-GBIC	1p1q4t	1p2q2t
WS-X6416-GE-MT	1p1q4t	1p2q2t
WS-X6316-GE-TX	1p1q4t	1p2q2t
WS-X6408A-GBIC	1p1q4t	1p2q2t
WS-X6408-GBIC	1q4t	2q2t
WS-X6524-100FX-MM	1p1q0t	1p3q1t
WS-X6324-100FX-SM	1q4t	2q2t
WS-X6324-100FX-MM	1q4t	2q2t
WS-X6224-100FX-MT	1q4t	2q2t
WS-X6548-RJ-21	1p1q0t	1p3q1t
WS-X6548-RJ-45	1p1q0t	1p3q1t

WS-X6348-RJ-21	1q4t	2q2t
WS-X6348-RJ21V	1q4t	2q2t
WS-X6348-RJ-45	1q4t	2q2t
WS-X6348-RJ-45V	1q4t	2q2t
WS-X6148-RJ-45V	1q4t	2q2t
WS-X6148-RJ21V	1q4t	2q2t
WS-X6248-RJ-45	1q4t	2q2t
WS-X6248A-TEL	1q4t	2q2t
WS-X6248-TEL	1q4t	2q2t
WS-X6024-10FL-MT	1q4t	2q2t

[Créez QoS sur le Catalyst 6500/6000](#)

Trois champs sur le Catalyst 6500/6000 sont utilisés pour faire QoS :

- La Priorité IP — Les trois premiers bits du Type de service (ToS) mettent en place dans l'en-tête IP
- Le point de code de Différenciation de services (DSCP) — les six premiers bits du tos mettent en place dans l'en-tête IP
- Le cos — Les trois bits utilisés au niveau de la couche 2 (L2) Ces trois bits sont l'un ou l'autre de partie de l'en-tête de Liaison inter-commutateurs (ISL) ou sont à l'intérieur de la balise du 802.1Q d'IEEE (dot1q). Il n'y a aucun cos à l'intérieur d'un paquet Ethernet non-marqué.

[Mécanisme de planification de sortie sur le Catalyst 6500/6000](#)

Quand une trame est envoyée du bus de données à transmettre, le cos du paquet est le seul paramètre qui est considéré. Le paquet passe alors par un programmeur, qui choisit la file d'attente dans laquelle le paquet est mis. , Souvenez-vous par conséquent que la planification de sortie et tous les mécanismes que ce document discute sont seulement Cos-avertis.

Le Catalyst 6500/6000 avec une carte de commutation multicouche (MSFC) emploie un DSCP interne afin de classer le paquet. Le Catalyst 6500/6000 que cela est configuré avec QoS a activé assigne une valeur DSCP quand la décision d'expédition est prise au niveau PFC. Ce DSCP est assigné à n'importe quel paquet, qui inclut les paquets non-IP, et est tracé au cos afin d'activer la planification de sortie. Vous pouvez configurer le mappage du DSCP aux valeurs CoS sur le Catalyst 6500/6000. Si vous laissez la valeur par défaut, vous pouvez dériver le cos du DSCP. La formule est :

DSCP_value / 8

En outre, la valeur DSCP est tracée dans le cos du paquet sortant, si le paquet est un paquet IP qui est ISL ou dot1q (VLAN non-indigène) étiquetés. La valeur DSCP est également écrite à l'intérieur du champ de tos de l'en-tête IP.

Le diagramme dans la [figure 7](#) affiche une file d'attente 1p2q2t. Les files d'attente WRR sont vidées avec l'utilisation du programmeur WRR. Il y a également un arbitre qui vérifie entre chaque paquet des files d'attente WRR afin de déterminer s'il y a quelque chose dans la file d'attente prioritaire stricte.

Figure 7

1. Le champ de tos est réécrit dans l'en-tête IP et le cos 802.1p/ISL met en place.
2. La file d'attente et le seuil de établissement du programme sont sélectionnés sur la base du cos, par une carte configurable.
3. Chaque file d'attente a la taille et les seuils configurables, et quelques files d'attente ont WRED.
4. Utilisations retirantes de la file d'attente WRR entre deux files d'attente.
5. L'encapsulation sortante peut être dot1q, ISL, ou aucun.

[Configuration, surveillance, et planification de sortie sur le Catalyst 6500/6000](#)

[Configuration par défaut pour QoS sur le Catalyst 6500/6000](#)

Cette section fournit la sortie témoin de la configuration QoS par défaut sur un Catalyst 6500/6000, en plus des informations sur ce qui moyen de ces valeurs et la façon dont vous pouvez accorder les valeurs.

QoS est désactivé par défaut quand vous émettez cette commande :

```
set qos disable
```

Les commandes dans cette liste affichent à l'affectation implicite pour le chaque le cos dans un port 2q2t. La file d'attente 1 a le cos 0 et 1 assigné à son premier seuil et a le cos 2 et 3 assignés à son deuxième seuil. La file d'attente 2 a le cos 4 et 5 assignés à son premier seuil et a le cos 6 et 7 assignés à son deuxième seuil :

```
set qos map 2q2t tx 1 1 cos 0
```

```
set qos map 2q2t tx 1 1 cos 1
```

```
set qos map 2q2t tx 1 2 cos 2
```

```
set qos map 2q2t tx 1 2 cos 3
```

```
set qos map 2q2t tx 2 1 cos 4
```

```
set qos map 2q2t tx 2 1 cos 5
```

```
set qos map 2q2t tx 2 2 cos 6
```

```
set qos map 2q2t tx 2 2 cos 7
```

Ces commandes affichent le seuil d'avertissement par défaut sur un port 2q2t pour chaque file d'attente :

```
set qos drop-threshold 2q2t tx queue 1 80 100
```

```
set qos drop-threshold 2q2t tx queue 2 80 100
```

Vous pouvez assigner le poids par défaut à chacune des files d'attente WRR. Émettez cette commande afin d'assigner les poids par défaut pour la file d'attente 1 et la file d'attente 2 :

Remarque: La file d'attente à basse priorité est servie 5/260 pour cent du temps, et la file d'attente prioritaire est servie 255/260 pour cent du temps.

```
set qos wrr 2q2t 5 255
```

Toute la disponibilité de la mémoire tampon est séparée parmi les deux files d'attente. La file d'attente à basse priorité est correctement assignée à 80 pour cent des mémoires tampons qui sont disponibles parce que c'est la file d'attente qui est le plus susceptible pour avoir des paquets mis en mémoire tampon et se reposants pendant quelque temps. Émettez cette commande afin de définir la Disponibilité :

```
set qos txq-ratio 2q2t 80 20
```

Vous pouvez visualiser les configurations semblables pour le port 1p2q2t dans cette configuration :

```
set qos map 1p2q2t tx 1 1 cos 0
```

```
set qos map 1p2q2t tx 1 1 cos 1
```

```
set qos map 1p2q2t tx 1 2 cos 2
```

```
set qos map 1p2q2t tx 1 2 cos 3
```

```
set qos map 1p2q2t tx 2 1 cos 4
```

```
set qos map 1p2q2t tx 3 1 cos 5
```

```
set qos map 1p2q2t tx 2 1 cos 6
```

```
set qos map 1p2q2t tx 2 2 cos 7
```

```
set qos wrr 1p2q2t 5 255
```

```
set qos txq-ratio 1p2q2t 70 15 15
```

```
set qos wred 1p2q2t tx queue 1 80 100
```

```
set qos wred 1p2q2t tx queue 2 80 100
```

Remarque: Par défaut, le cos 5 (le trafic vocal) est assigné à la file d'attente prioritaire stricte.

Configuration

La première étape de configuration est d'activer QoS. Souvenez-vous que QoS est désactivé par défaut. Quand QoS est désactivé, la cartographie de cos est inutile. Il y a une file d'attente simple qui est servie de FIFO, et tous les paquets obtiennent relâché là.

```
bratan> (enable) set qos enable
```

```
QoS is enabled
```

```
bratan> (enable) show qos status
```

```
QoS is enabled on this switch
```

La valeur CoS doit être assignée à la file d'attente ou le seuil pour toute la file d'attente tape. Le mappage qui est défini pour un type 2q2t de port n'est appliqué à aucun port 1p2q2t. En outre, le mappage qui est fait pour 2q2t est appliqué à tous les ports qui ont un mécanisme de mise en file

d'attente 2q2t. Émettez la commande suivante :

```
set qos map queue_type tx Q_number threshold_number cos value
```

Remarque: Des files d'attente sont toujours numérotées pour commencer par la plus basse possible file d'attente prioritaire et pour finir avec la file d'attente prioritaire stricte qui est disponible. Voici un exemple :

- La file d'attente 1 est la file d'attente non prioritaire WRR
- La file d'attente 2 est la file d'attente prioritaire WRR
- La file d'attente 3 est la file d'attente prioritaire stricte

Vous devez répéter cette exécution pour tous les types de files d'attente. Autrement, vous gardez l'affectation par défaut de cos. Voici un exemple pour 1p2q2t :

Configuration

```
set qos map 1p2q2t tx 1 1 cos 0
!--- This is the low-priority WRR queue threshold 1, CoS 0 and 1. set qos map 1p2q2t tx 1 1 cos
1 and 1

set qos map 1p2q2t tx 1 2 cos 2
!--- This is the low-priority WRR queue threshold 2, CoS 2 and 3. set qos map 1p2q2t tx 1 2 cos
3 and 3

set qos map 1p2q2t tx 2 1 cos 4
!--- This is the high-priority WRR queue threshold 1, CoS 4. set qos map 1p2q2t tx 3 1 cos 5
!--- This is the strict priority queue, CoS 5. set qos map 1p2q2t tx 2 1 cos 6
!--- This is the high-priority WRR queue threshold 2, CoS 6. set qos map 1p2q2t tx 2 2 cos 7 and
7
```

Sortie de console

```
tamer (enable) set qos map 1p2q2t tx 1 1 cos 0
```

```
QoS tx priority queue and threshold mapped to cos successfully
```

Vous devez configurer le poids WRR pour les deux files d'attente WRR. Émettez la commande suivante :

```
set qos wrr Q_type weight_1 weight_2
```

Weight_1 associe pour aligner 1, qui devrait être la file d'attente non prioritaire WRR. *Weight_1* doit toujours être inférieur à *weight_2*. Le poids peut prendre n'importe quelle valeur entre 1 et 255. Vous pouvez assigner le pourcentage avec ces formules :

- File d'attente 1 :

```
set qos wrr Q_type weight_1 weight_2
```
- File d'attente 2 :

```
set qos wrr Q_type weight_1 weight_2
```

Vous devez également définir le poids pour les divers types de files d'attente. Le poids n'a pas besoin d'être identique. Par exemple, pour 2q2t, où la file d'attente 1 est servie 30 pour cent du temps et la file d'attente 2 est servie 70 pour cent du temps, vous pouvez émettre cette commande afin de définir le poids :

```
set qos wrr 2q2t 30 70
```

!--- This ensures that the high-priority WRR queue is served 70 percent of the time !--- and that the low-priority WRR queue is served 30 percent of the time.

Sortie de console

```
tamer (enable) set qos wrr 2q2t 30 70
```

QoS wrr ratio is set successfully

Vous devez également définir le rapport de la file d'attente de transmission, qui se rapporte à la manière dont les mémoires tampons sont séparé parmi les différentes files d'attente. Émettez la commande suivante :

```
set qos txq-ratio port_type queue1_val queue2_val ... queueN_val
```

Remarque: Si vous avez trois files d'attente (1p2q2t), vous devez placer la file d'attente prioritaire WRR et la file d'attente prioritaire stricte au même niveau pour des raisons matérielles.

Configuration

```
set qos txq-ratio 1p2q2t 70 15 15
```

!--- This gives 70 percent of the buffer of all 1p2q2t ports to the low-priority WRR !--- queue and gives 15 percent to each of the other two queues. set qos txq-ratio 2q2t 80 20 !--- This gives 80 percent of the buffer to the low-priority queue, !--- and gives 20 percent of the buffer to the high-priority queue.

Sortie de console

```
tamer (enable) set qos txq-ratio 1p2q2t 70 15 20
```

Queue ratio values must be in range of 1-99 and add up to 100

Example: set qos txq-ratio 2q2t 20 80

```
tamer (enable) set qos txq-ratio 1p2q2t 70 30 30
```

Queue ratio values must be in range of 1-99 and add up to 100

Example: set qos txq-ratio 2q2t 20 80

```
tamer (enable) set qos txq-ratio 1p2q2t 80 10 10
```

QoS txq-ratio is set successfully

Pendant que cette sortie de console illustre, la somme des valeurs de file d'attente doit être 100. Laissez la plus grande partie des mémoires tampons pour le WRR non prioritaire pour s'aligner parce que cette file d'attente a besoin de la plupart de mise en mémoire tampon. Les autres files d'attente sont servies avec la haute priorité.

La dernière étape est de configurer le seuil d'avertissement pour la file d'attente WRED ou pour la file de perte de destination. Émettez les commandes suivantes :

```
set qos wred port_type [tx] queue q_num thr1 thr2 ... thrn
```

```
set qos drop-threshold port_type tx queue q_num thr1 ... thr2
```

Configuration

```
set qos drop-threshold 2q2t tx queue 1 50 80
```

!--- For low-priority queues in the 2q2t port, the first threshold is defined at 50 !--- percent

and the second threshold is defined at 80 percent of buffer filling. `set qos drop-threshold 2q2t tx queue 2 40 80`
!--- For high-priority queues in the 2q2t port, the first threshold is defined at 40 !---
percent and the second threshold is defined at 80 percent of buffer filling. `set qos wred 1p2q2t tx queue 1 50 90`
!--- The commands for the 1p2q2t port are identical. `set qos wred 1p2q2t tx queue 2 40 80`

Sortie de console

```
tamer (enable) set qos drop-threshold 2q2t tx queue 1 50 80
```

Transmit drop thresholds for queue 1 set at 50% 80%

```
tamer (enable) set qos drop-threshold 2q2t tx queue 2 40 80
```

Transmit drop thresholds for queue 2 set at 40% 80%

```
tamer (enable) set qos wred 1p2q2t tx queue 1 50 90
```

WRED thresholds for queue 1 set to 50 and 90 on all WRED-capable 1p2q2t ports

```
tamer (enable) set qos wred 1p2q2t tx queue 2 40 80
```

WRED thresholds for queue 2 set to 40 and 80 on all WRED-capable 1p2q2t ports

La file d'attente 2 du tx 1p2q2t wred par set qos 40 80 travaux de commande en même temps que le cos pour le mappage de seuil. Par exemple, quand vous émettez les commandes dans la liste ci-dessous, vous vous assurez que — sur le port 1p2q2t dans la direction de transmission — des paquets avec le cos 0,1, 2, et 3 sont introduits la première file d'attente (la file d'attente de bas WRR). Quand les mémoires tampons dans cette file d'attente sont de 50 pour cent remplis, WRED commence à relâcher des paquets avec le cos 0 et les paquets 1. avec le cos 2 et 3 est relâché seulement quand les mémoires tampons sur la file d'attente sont de 90 pour cent remplis.

```
set qos map 1p2q2t tx 1 1 cos 0
```

```
set qos map 1p2q2t tx 1 1 cos 1
```

```
set qos map 1p2q2t tx 1 2 cos 2
```

```
set qos map 1p2q2t tx 1 2 cos 3
```

```
set qos wred 1p2q2t tx queue 1 50 90
```

Le Scheduling de sortie moniteur et vérifie la configuration

Une commande simple de utiliser afin de vérifier la configuration d'exécution en cours pour la planification de sortie d'un port est *modèle de délai d'exécution de show qos info/port*. La commande affiche ces informations :

- Le type de Mise en file d'attente sur le port
- Le mappage du cos aux différents files d'attente et seuils
- Partager de mémoire tampon
- Le poids WRR

Dans cet exemple, les valeurs sont à 20 pour cent WRR pour la file d'attente 1 et 80 pour cent WRR pour la file d'attente 2 :

```
tamer (enable) show qos info runtime 1/1
```

Run time setting of QoS:

```

QoS is enabled
Policy Source of port 1/1: Local
Tx port type of port 1/1 : lp2q2t
Rx port type of port 1/1 : lp1q4t
Interface type: port-based
ACL attached:
The qos trust type is set to untrusted
Default CoS = 0
Queue and Threshold Mapping for lp2q2t (tx):
Queue   Threshold   CoS
-----  -
1        1             0 1
1        2             2 3
2        1             4 6
2        2             7
3        1             5
Queue and Threshold Mapping for lp1q4t (rx):
All packets are mapped to a single queue
Rx drop thresholds:
Rx drop thresholds are disabled
Tx drop thresholds:
Tx drop-thresholds feature is not supported for this port type
Tx WRED thresholds:
Queue #           Thresholds - percentage (* abs values)
-----  -
1             80% (249088 bytes) 100% (311168 bytes)
2             80% (52480 bytes) 100% (61440 bytes)
Queue Sizes:
Queue #           Sizes - percentage (* abs values)
-----  -
1             70% (311296 bytes)
2             15% (65536 bytes)
3             15% (65536 bytes)
WRR Configuration of ports with speed 1000Mbps:
Queue #           Ratios (* abs values)
-----  -
1             20 (5120 bytes)
2             80 (20480 bytes)
(*) Runtime information may differ from user configured setting
due to hardware granularity.
tamer (enable)

```

Dans l'exemple suivant, notez que les poids WRR ne sont pas la valeur par défaut de 1. Les poids ont été placés aux valeurs de 20 pour la file d'attente 1 et 80 pour la file d'attente 2. Cet exemple utilise un générateur du trafic pour envoyer 2 gigaoctets du trafic à un Catalyst 6000. Ces 2 gigaoctets du trafic devraient quitter par le port 1/1. Puisque le port 1/1 est oversubscribed, beaucoup de paquets sont lâchés (GBP 1). La commande de **show mac** prouve qu'il y a beaucoup de suppression de sortie :

```
tamer (enable) show mac 1/1
```

```

Port           Rcv-Unicast           Rcv-Multicast           Rcv-Broadcast
-----  -
1/1           0                     1239                    0

Port           Xmit-Unicast           Xmit-Multicast           Xmit-Broadcast
-----  -
1/1           73193601              421                     0

Port           Rcv-Octet             Xmit-Octet
-----  -
1/1           761993                100650803690

```

MAC	Dely-Exced	MTU-Exced	In-Discard	Out-Discard
1/1	0	-	0	120065264

Last-Time-Cleared

Fri Jan 12 2001, 17:37:43

Considérez les paquets qui sont lâchés. C'est comment la conformation du trafic suggérée est séparée :

- 1 gigaoctet du trafic avec la Priorité IP 0
- Mi-bande 250 du trafic avec la Priorité IP 4
- Mi-bande 250 du trafic avec la Priorité IP 5
- Mi-bande 250 du trafic avec la Priorité IP 6
- Mi-bande 250 du trafic avec la Priorité IP 7

Selon le cos traçant, ce trafic est envoyé :

- 1 gigaoctet du trafic pour aligner 1 seuil 1
- 0 mis-bande du trafic pour aligner 1 seuil 2
- Mi-bande 500 du trafic pour aligner 2 le seuil 1
- Mi-bande 250 du trafic pour aligner 2 le seuil 2
- Mi-bande 250 du trafic pour aligner 3 (file d'attente prioritaire stricte)

Le commutateur doit faire confiance au trafic reçu de sorte que la Priorité IP entrante soit préservée dans le commutateur et soit utilisée pour tracer à la valeur CoS pour la planification de sortie.

Remarque: La priorité d'IP par défaut à la cartographie de cos est cos d'égaux de Priorité IP.

Émettez la commande **stat 1/1 de show qos** afin de voir les paquets qui ont été lâchés et le pourcentage approximatif :

- En ce moment, aucun paquet n'est lâché dans la file d'attente 3 (cos 5).
- 91.85 pour cent des paquets abandonnés sont cos 0 paquets dans la file d'attente 1.
- 8 pour cent des paquets abandonnés sont le cos 4 et 6 dans la file d'attente 2, le seuil 1.
- 0.15 pour cent des paquets abandonnés est le cos 7 dans la file d'attente 2, le seuil 2.

Cette sortie illustre l'utilisation de la commande :

```
tamer (enable) show qos stat 1/1
```

```
Tx port type of port 1/1 : lp2q2t
```

```
Q3T1 statistics are covered by Q2T2.
```

```
Q #      Threshold #:Packets dropped
---      -
1        1:110249298 pkts, 2:0 pkts
2        1:9752805 pkts, 2:297134 pkts
3        1:0 pkts
```

```
Rx port type of port 1/1 : lp1q4t
```

```
Rx drop threshold counters are disabled for untrusted ports
```

```
Q #      Threshold #:Packets dropped
---      -
1        1:0 pkts, 2:0 pkts, 3:0 pkts, 4:0 pkts
2        1:0 pkts
```

Si vous changez le poids WRR de nouveau à la valeur par défaut après que les compteurs aient été effacés, seulement 1 pour cent des paquets abandonnés se produit dans la file d'attente 2 au

lieu des 8 pour cent qui sont apparus précédemment :

Remarque: La valeur par défaut est 5 pour la file d'attente 1 et 255 pour la file d'attente 2.

```
tamer (enable) show qos stat 1/1
```

```
TX port type of port 1/1 : lp2q2t
Q3T1 statistics are covered by Q2T2
Q #           Threshold #:Packets dropped
-----
1             1:2733942 pkts, 2:0 pkts
2             1:28890 pkts, 2:6503 pkts
3             1:0 pkts
Rx port type of port 1/1 : lp1q4t
Rx drop threshold counters are disabled for untrusted ports
Q #           Threshold #:Packets dropped
-----
1             1:0 pkts, 2:0 pkts, 3:0 pkts, 4:0 pkts
2             1:0 pkts
```

Planification de sortie d'utilisation pour réduire le retard et instabilité

L'exemple dans le [Scheduling de sortie moniteur de](#) section [et vérifient la configuration](#) explique l'avantage de l'implémentation de planification de sortie, qui évite une baisse de VoIP ou de trafic crucial en cas du surabonnement du port de sortie. Le surabonnement se produit rarement dans un réseau normal, en particulier sur une liaison Gigabit. Habituellement, le surabonnement se produit seulement pendant des temps du trafic maximal ou pendant les rafales de trafic au cours très d'une courte période.

Même sans n'importe quel surabonnement, la planification de sortie peut être d'un grand secours dans un réseau où QoS est de bout en bout mis en application. Aides de planification de sortie pour réduire le retard et instabilité. Cette section fournit des exemples de la façon dont la planification de sortie peut aider à réduire le retard et instabilité.

Réduisez le retard

Le retard d'un paquet est augmenté avant que « perdu » dans la mémoire tampon de chaque commutateur pendant l'attente pour la transmission. Par exemple, un petit paquet vocal avec cos de 5 est envoyé hors d'un port pendant une grande sauvegarde ou transfert de fichiers. Si vous n'avez aucun QoS pour le port de sortie, et si vous supposez que le petit paquet vocal est aligné après 10 grands paquets 1500-byte, vous pouvez facilement calculer le temps de vitesse de gigabit qui est nécessaire pour transmettre les 10 grands paquets :

```
tamer (enable) show qos stat 1/1
```

```
TX port type of port 1/1 : lp2q2t
Q3T1 statistics are covered by Q2T2
Q #           Threshold #:Packets dropped
-----
1             1:2733942 pkts, 2:0 pkts
2             1:28890 pkts, 2:6503 pkts
3             1:0 pkts
Rx port type of port 1/1 : lp1q4t
Rx drop threshold counters are disabled for untrusted ports
Q #           Threshold #:Packets dropped
-----
```



```
1          1:0 pkts, 2:0 pkts, 3:0 pkts, 4:0 pkts
2          1:0 pkts
```

Si ce paquet doit croiser huit ou neuf Commutateurs tandis qu'il traverse le réseau, un retard d'environ 1 ms peut résulter. Cette quantité compte seulement des retards dans la file d'attente de sortie du commutateur qui est croisé dans le réseau.

Remarque: Si vous devez faire la queue les mêmes 10 grands paquets sur un 10-Mbps reliant (par exemple, avec un téléphone IP et avec un PC connecté), le retard qui est introduit est :

```
tamer (enable) show qos stat 1/1
```

```
TX port type of port 1/1 : lp2q2t
Q3T1 statistics are covered by Q2T2
Q #          Threshold #:Packets dropped
---          -
1          1:2733942 pkts, 2:0 pkts
2          1:28890 pkts, 2:6503 pkts
3          1:0 pkts
Rx port type of port 1/1 : lp1q4t
Rx drop threshold counters are disabled for untrusted ports
Q #          Threshold #:Packets dropped
---          -
1          1:0 pkts, 2:0 pkts, 3:0 pkts, 4:0 pkts
2          1:0 pkts
```

L'implémentation de planification de sortie s'assure que des paquets vocaux avec cos de 5 sont mis dans la file d'attente prioritaire stricte. Ce les placements s'assure que ces paquets sont envoyés avant tous les paquets avec cos de moins de 5, qui réduit les retards.

Réduisez le jitter

Un autre important avantage d'implémentation de planification de sortie est qu'il réduit le jitter. Le jitter est la variation du retard qui est observé pour des paquets dans le même écoulement. Le diagramme dans la [figure 8](#) affiche un exemple de scénario de la façon dont la planification de sortie peut réduire le jitter :

Figure 8

Dans ce scénario, il y a deux flots qu'un port à sortie unique doit envoyer :

- Un flux voix qui est entrant sur un port Ethernet 10-Mbps
- Un flot de FTP qui est entrant sur une liaison ascendante Ethernet 1-Gbps

Les deux flots laissent le commutateur par le même port de sortie. Cet exemple affiche ce qui peut se produire sans utilisation de la planification de sortie. Tous les grands paquets de données peuvent être intercalés entre deux paquets vocaux, qui crée le jitter dans la réception du paquet vocal du même flot. Il y a un plus grand retard entre la réception du paquet n et du paquet $n+1$ car le commutateur transmet le grand paquet de données. Cependant, le retard entre $n+1$ et $n+2$ est négligeable. Ceci a comme conséquence le jitter dans le flot du trafic vocal. Vous pouvez facilement éviter ce problème avec l'utilisation d'une file d'attente prioritaire stricte. Assurez-vous que la valeur CoS des paquets vocaux est tracée à la file d'attente prioritaire stricte.

Informations connexes

- [Planification de la sortie QoS sur les commutateurs des gammes Catalyst 6500/6000 exécutant le logiciel système Cisco IOS](#)

- [Présentation de Qos \(Qualité de service\) sur les commutateurs de la gamme Catalyst 6000](#)
- [Pages de support pour les produits LAN](#)
- [Page de support sur la commutation LAN](#)
- [Support et documentation techniques - Cisco Systems](#)