

Architecture du routeur Internet de la gamme Cisco 12000 : Commutation de paquets

Contenu

[Introduction](#)

[Conditions préalables](#)

[Conditions requises](#)

[Composants utilisés](#)

[Conventions](#)

[Informations générales](#)

[Commutation de paquets : Aperçu](#)

[Commutation de paquets : Linecards de l'engine 0 et de l'engine 1](#)

[Commutation de paquets : Linecards d'Engine 2](#)

[Commutation de paquets : Cellules de changement à travers la matrice](#)

[Commutation de paquets : Paquets de transmission](#)

[Résumé d'écoulement de paquet](#)

[Informations connexes](#)

[Introduction](#)

Ce document examine les éléments architecturaux les plus importants du Routeur Internet de la série Cisco 12000 -- paquets de changement. Les paquets de changement sont radicalement différents de la mémoire partagée l'une des ou des architectures Cisco basées sur bus. À l'aide d'une matrice de barre transversale, le Cisco 12000 fournit très un grand nombre de bande passante et d'évolutivité. En outre, les 12000 files d'attente de sortie virtuelle d'utilisations pour éliminer la tête de la ligne bloquant dans la matrice de commutateur.

[Conditions préalables](#)

[Conditions requises](#)

Aucune spécification déterminée n'est requise pour ce document.

[Composants utilisés](#)

Les informations dans ce document sont basées sur le matériel suivant :

- [Routeur Internet de la gamme Cisco 12000](#)

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 des documents, référez-vous aux [Conventions utilisées pour les conseils techniques de Cisco](#).

Informations générales

(La décision de commutation sur un Cisco 12000 est faite par les linecards (LCS). Pour un certain LCS, un circuit intégré spécifique (ASIC) dédié commute réellement les paquets. Cisco Express Forwarding distribué (dCEF) est la seule méthode de commutation disponible.

Remarque : Les engines 0, 1, et 2 ne sont pas les dernières engines développées par Cisco. Il y a également l'engine 3, 4, et les linecards 4+, avec plus à suivre. Les linecards de l'engine 3 sont capables d'exécuter des caractéristiques de périphérie à la ligne débit. Plus l'engine de la couche 3 est élevée, plus les paquets obtiennent commuté dans le matériel. Vous pouvez trouver quelques informations utiles au sujet des différents linecards disponibles pour le routeur de gamme Cisco 12000 et l'engine sur lesquels ils sont basés au [Routeur Internet de la série Cisco 12000 : Forums aux questions](#).

Commutation de paquets : Aperçu

Des paquets sont toujours expédiés par la carte de ligne d'entrée (LC). Le de sortie LC exécute seulement le Qualité de service (QoS) sortant qui est file d'attente-dépendant (par exemple, Détection précoce directe pondérée (WRED) ou Fonction Committed Access Rate (CAR)). La plupart des paquets sont commutés par le LC utilisant Cisco Express Forwarding distribué (dCEF). Seulement les paquets de contrôle (tels que des mises à jour de routage) sont envoyés au processeur de route Gigabit (GRP) pour le traitement. Le chemin de commutation par paquets dépend du type de moteurs de commutation utilisés sur le LC.

C'est ce qui se produit quand un paquet est livré dans :

1. Un paquet entre dans le module d'interface de couche physique (PLIM). Les diverses choses se produisent ici : Un émetteur-récepteur transforme les signaux Optiques en les électriques (la plupart des linecards CSR ont des connecteurs fibre) Le tramage L2 est enlevé (RAISONNABLE, Mode de transfert asynchrone (ATM), Ethernets, High-Level Data Link Control (HDLC)/protocole point-à-point - le PPP) Des cellules atmosphère sont rassemblées Des paquets qui échouent le contrôle de redondance cyclique (CRC) sont jetés
2. Car le paquet est reçu et traité, c'est mémoire directe accédée à dans une petite mémoire de mémoire tampon de Maximum Transmission Unit approximativement 2 x ((MTU)) a appelé mémoire éclatée la « du First In, First Out (FIFO) ». La quantité de cette mémoire dépend du type de LC (de 128 KO à 1 Mo).
3. Une fois que le paquet est complètement dans la mémoire FIFO, un circuit intégré spécifique (ASIC) sur le PLIM entre en contact avec la gestion de mémoire tampon ASIC (BMA) et demande une mémoire tampon pour mettre le paquet dedans. Le BMA est dit quelle taille le paquet est, et alloue une mémoire tampon en conséquence. Si le BMA ne peut pas obtenir une mémoire tampon de la bonne taille, le paquet est lâché et « ignorez » le compteur est incrémenté sur l'interface entrante. Il n'y a aucun mécanisme de repli comme avec quelques

autres Plateformes. Tandis que ceci va en fonction, le PLIM pourrait recevoir un autre paquet dans la mémoire de rafale FIFO, qui est pourquoi c'est 2xMTU dans la taille.

4. S'il y a une mémoire tampon libre disponible dans la file d'attente droite, le paquet est enregistré par le BMA dans la liste de file d'attente libre de la taille appropriée. Cette mémoire tampon est placée sur la file d'attente de paquets non traités, qui est examinée par le Salsa ASIC ou la CPU R5K. La CPU R5K détermine la destination du paquet en consultant sa table locale de dCEF dans la mémoire vive dynamique (mémoire vive dynamique), et puis déplace la mémoire tampon à partir de la file d'attente de paquets non traités à une file d'attente de ToFabric correspondant à l'emplacement de destination. Si la destination n'est pas dans la table CEF, le paquet est lâché. Si le paquet est un paquet de contrôle (par exemple, conduisant des mises à jour), il est mis à la file d'attente du GRP et sera traité en file d'attente par le GRP. Il y a 17 tofab queue (unicast 16, plus 1 Multidiffusion). Il y a un tofab queue par linecard (ceci inclut le RP). Ces files d'attente sont connues en tant que « files d'attente de sortie virtuelle », et sont importantes de sorte que le blocage de tête-de-ligne ne se produise pas.
5. Le ToFab BMA coupe le paquet en parties 44-byte, qui sont la charge utile pour ce qui sera par la suite connu en tant que « cellules de Cisco ». Ces cellules sont données une en-tête 8-byte et l'en-tête de la mémoire tampon 4-byte par le frFab BMA (taille de données totale jusqu'ici = 56 octets), et puis mises en file d'attente dans le tofab queue approprié (quel point, le #Qelem contre- dans le groupe la mémoire tampon est provenu descend par une, et le compteur de tofab queue monte par un). Le « décideur » dépend du type de moteurs de commutation : Sur des cartes de l'engine 2+, une offre spéciale ASIC est utilisée pour améliorer la manière que les paquets sont commutés. Les paquets normaux (IP/Tag, aucune options, somme de contrôle) sont traités directement par la commutation de paquets ASIC (PSA), puis sautent la combinaison crue queue/CPU/Salsa et sont mis en file d'attente directement sur le tofab queue. Seulement les 64 premiers octets du paquet sont traversés la commutation de paquets ASIC. Si le paquet ne peut pas être commuté par la PSA, le paquet est mis en file d'attente au RawQ à manipuler par la CPU du LC comme expliqué précédemment. En ce moment, la décision de commutation a été prise et le paquet a été mis en file d'attente sur la file d'attente de sortie appropriée de ToFab.
6. Les accès directs à la mémoire du toFab BMA (accès direct à la mémoire) les cellules du paquet dans de petits tampons FIFO dans l'interface de matrice ASIC (la FIA). Il y a 17 tampons FIFO (un par tofab queue). Quand la FIA obtient une cellule du toFab BMA, il ajoute un CRC 8-byte (taille totale de cellules - 64 octets ; la charge utile de 44 octets, en-tête de cellule de 8 octets, 4 octets mettent en mémoire tampon l'en-tête). La FIA a l'interface de ligne série (SLI) ASIC qui exécutent alors le codage 8B/10B sur la cellule (comme le Fiber Distributed Data Interface (FDDI) 4B/5B), et dispose au transmettre au-dessus de la matrice. Ceci peut sembler comme beaucoup de temps système (44 octets de données obtient transformé en 80 octets à travers la matrice !), mais ce n'est pas une question puisque la capacité de matrice provisioned en conséquence.
7. Maintenant qu'une FIA est prête à transmettre, la FIA demande l'accès à la matrice du actuellement - le programmeur de carte à puce et l'horloge (CSC). Le CSC travaille sur un algorithme plutôt complexe d'équité. L'idée est qu'on ne permet à aucun LC pour monopoliser la bande passante sortante de n'importe quelle autre carte. Notez que même si un LC veut transmettre des données sur un de ses propres ports, il doit encore passer par la matrice. C'est important parce que si ceci ne se produisait pas, un port sur un LC pourrait monopoliser toute la bande passante pour un port donné sur ce même LC. Il ferait également la conception de commutation plus compliquée. La FIA envoie des cellules à travers la

matrice de commutateur à leur LC sortant (spécifié par des données dans l'en-tête de cellule de Cisco mise là en le moteur de commutation). L'algorithme d'équité est également conçu pour apparier optimal ; si la carte 1 veut transmettre pour carder 2, et la carte 3 veut transmettre pour carder 4 en même temps, ceci se produit en parallèle. C'est la grande différence entre une matrice de commutateur et une architecture de bus. Pensez à lui comme analogue à un commutateur ethernet contre un hub ; sur un commutateur, si le port A veut envoyer pour mettre en communication B, et le C de port veut parler pour mettre en communication D, ces deux écoulements se produisent indépendamment de l'un l'autre. Sur un hub, il y a les questions bidirectionnelles-alternées telles que des collisions et l'interruption et relance des algorithmes.

8. Les cellules de Cisco qui sortent de la matrice passent par le SLI traitant pour retirer le codage 8B/10B. Si là des erreurs ici, elles apparaîtraient dans la sortie de commande FIA de show controller en tant que « parité de cellules ». Voyez [comment lire la sortie de l'ordre FIA de show controller](#) pour information les informations complémentaires.
9. Ces cellules de Cisco sont DMA'd dans FIFOs sur le FIAS de frFab, et puis dans une mémoire tampon sur le frFab BMA. Le frFab BMA est celui qui fait réellement le réassemblage des cellules dans un paquet. Comment le frFab BMA connaît-il quelle mémoire tampon pour mettre les cellules dedans avant qu'elle les rassemble ? C'est une autre décision prise en le moteur de commutation entrant de linecard ; puisque toutes les files d'attente sur la case entière sont la même taille et dans la même commande, le moteur de commutation fait juste mettre le Tx LC le paquet dans la même file d'attente de nombre de laquelle il a présenté le routeur. Les files d'attente du frFab BMA SDRAM peuvent être visualisées avec la commande de file d'attente de frfab de show controller sur le LC. Voyez [comment lire la sortie du frfab de show controller | commandes de tofab queue sur un Routeur Internet de la série Cisco 12000](#) pour des détails. C'est fondamentalement la même idée que la sortie du toFab BMA. Des paquets entrés et sont placés en paquets qui sont retirés de la file d'attente de leurs files d'attente libre respectives. Ces paquets sont placés dans la file d'attente de-matrice, mise en file d'attente sur la file d'attente d'interface (il y a une file d'attente par port physique) ou le rawQ pour le traitement de sortie. Pas beaucoup se produit dans le rawQ : réplification de Multidiffusion de par-port, Mécanisme MDRR (Modified Deficit Round Robin) - la même idée que la mise en file d'attente pondérée distribuée (DWFQ), et CAR de sortie. Si la file d'attente de transmission est pleine, le paquet est lâché et le compteur de suppressions de sortie est incrémenté.
10. Le frFab BMA attend jusqu'à ce que la partie TX du PLIM soit prête à envoyer un paquet. Le frFab BMA fait la réécriture réelle de MAC (basée, souvenez-vous, sur les informations contenues dans l'en-tête de cellule de Cisco), et des accès directs à la mémoire le paquet plus d'à une petite (de nouveau, 2xMTU) mémoire tampon dans les circuits PLIM. Le PLIM fait l'atmosphère SAR et SONET s'encapsule, le cas échéant, et transmet le paquet.
11. Le trafic atmosphère est rassemblé (par le SAR), segmenté (par le tofab BMA), rassemblé (par le fromfab BMA) et de nouveau segmenté (par le fromfab SAR). Ceci se produit très rapidement.

C'est le cycle de vie d'un paquet, du commencement jusqu'à la fin. Si vous voulez savoir ce que se sent un GSR comme finalement, lit ce document entier 500,000 fois !

Le chemin de commutation par paquets sur le GSR dépend du type d'engine d'expédition sur le LC. Maintenant nous passerons par toutes les étapes pour l'engine 0, l'engine 1, et les deux LCS.

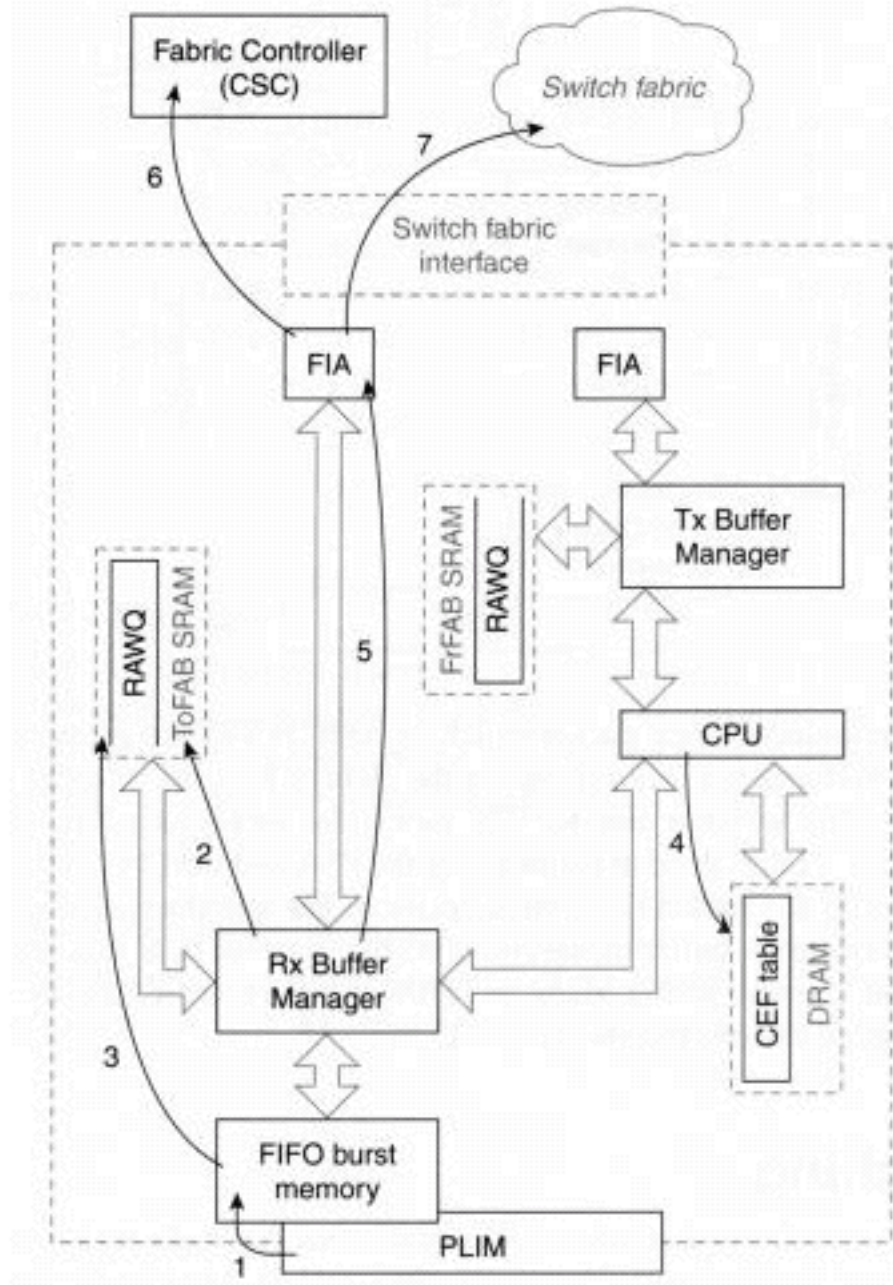
[Commutation de paquets : Linecards de l'engine 0 et de l'engine](#)

1

Les sections ci-dessous sont basées sur l'ouvrage *à l'intérieur du Cisco IOS architecture logicielle*, Cisco appuient sur.

La figure 1 ci-dessous montre les différentes étapes pendant la commutation par paquets pour une engine 0 ou l'engine 1 LC.

Figure 1 : Chemin de commutation de l'engine 0 et de l'engine 1



Le chemin de commutation pour l'engine 0 et l'engine 1 LC est essentiellement identique, bien que l'engine 1 LC ait un moteur de commutation et un gestionnaire de tampons améliorés pour des performances accrues. Le chemin de commutation est comme suit :

- **Étape 1** - Le processeur d'interface (PLIM) détecte un paquet sur les supports réseau et commence le copiant dans une mémoire FIFO appelée la **mémoire de rafale** sur le LC. La quantité de mémoire de rafale que chaque interface a dépend du type de LC ; les LCS typiques ont 128 KO à 1 Mo de mémoire de rafale.

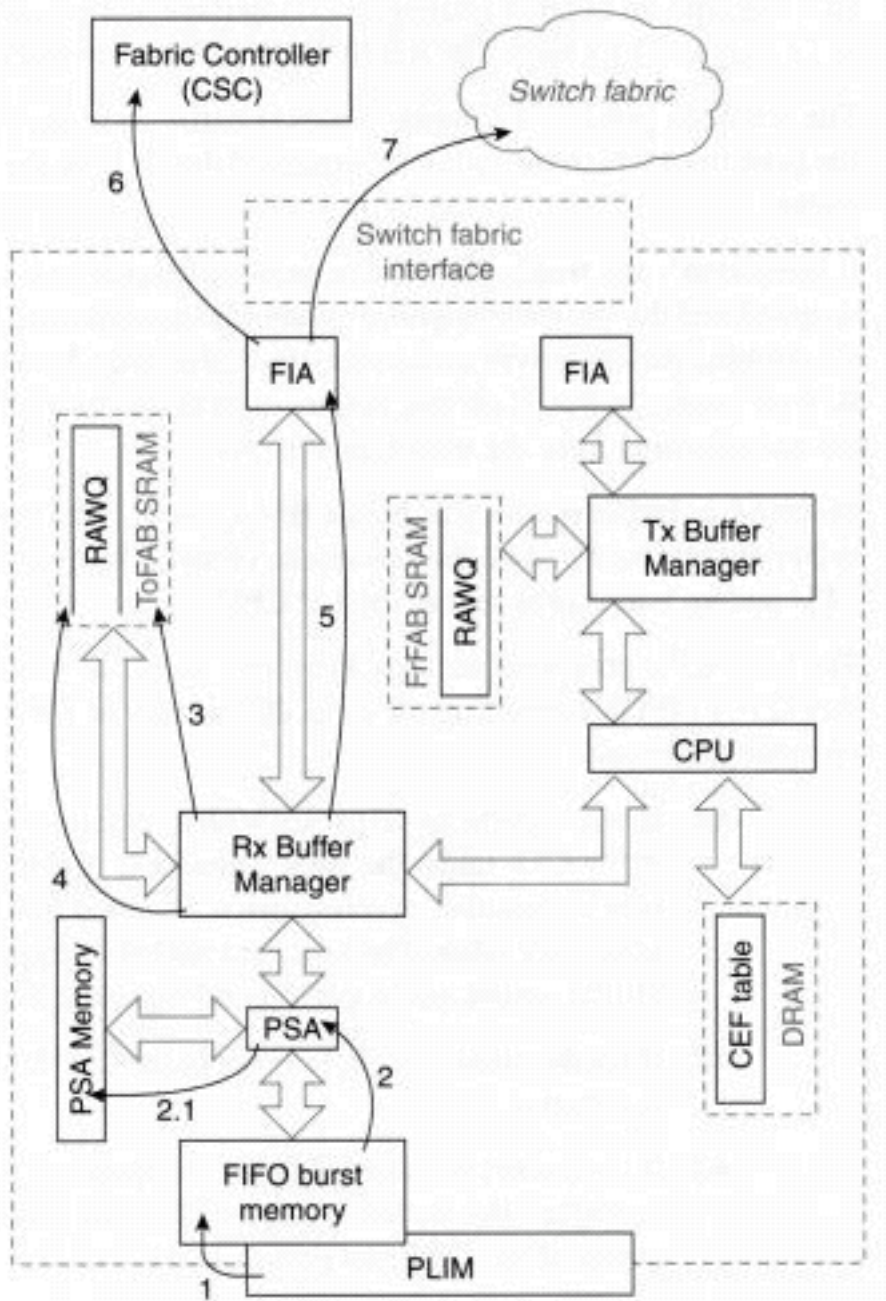
- **Étape 2** - Le processeur d'interface demande un tampon de paquets de la réception BMA ; le groupe auquel la mémoire tampon est demandée dépend de la longueur du paquet. S'il n'y a pas aucune mémoire tampon libre, l'interface est abandonnée et l'interface « ignorent » le compteur est incrémentée. Par exemple, si un paquet 64-byte arrive dans une interface, les essais BMA pour allouer un tampon de paquets de 80 octets. Si mémoire tampon libre n'existe pas dans le groupe de 80 octets, des mémoires tampons ne sont pas allouées du prochain groupe disponible.
- **Étape 3** - Quand une mémoire tampon libre est allouée par le BMA, le paquet est copié dans la mémoire tampon et est mis en file d'attente sur la **file d'attente de paquets non traités** (RawQ) pour traiter par la CPU. Une interruption est envoyée à la CPU LC.
- **Étape 4** - La CPU Du LC traite chaque paquet dans le RawQ car elle est reçue (le RawQ est un FIFO), consultant la table de Cisco Express Forwarding distribuée par gens du pays dans la mémoire vive dynamique pour prendre une décision de commutation.**4.1** Si c'est un paquet IP d'unicast avec une adresse de destination valide dans la table CEF, l'en-tête de paquet est réécrite avec les nouvelles informations d'encapsulation obtenues de la table de contiguïté CEF. Le paquet commuté est mis en file d'attente sur la file d'attente de sortie virtuelle correspondant à l'emplacement de destination.**4.2** Si l'adresse de destination n'est pas dans la table CEF, le paquet est lâché.**4.3** Si le paquet est un paquet de contrôle (une mise à jour de routage, par exemple), le paquet est mis sur la file d'attente de sortie virtuelle du GRP et traité en file d'attente par le GRP.
- **Étape 5** - La réception BMA fragmente le paquet dans les cellules 64-bytes, et remet ces derniers hors fonction à la FIA pour la transmission au LC sortant.

À la fin de l'étape 5, le paquet qui est arrivé dans une engine 0/1 LC a été commuté et est prêt à être transporté à travers la matrice de commutateur comme cellules. Passez à l'étape 6 dans la [commutation de paquets de](#) section : [Cellules de changement à travers la matrice](#).

[Commutation de paquets : Linecards d'Engine 2](#)

[La figure 2](#) ci-dessous montre le chemin de commutation par paquets quand les paquets arrivent dans un Engine 2 LC, comme décrit dans la liste suivante d'étapes.

Figure 2 : Chemin de commutation d'Engine 2



- **Étape 1** - Le processeur d'interface (PLIM) détecte un paquet sur les supports réseau et commence le copiant dans une mémoire FIFO appelée la **mémoire de rafale** sur le LC. La quantité de mémoire de rafale que chaque interface a dépend du type de LC ; les LCS typiques ont 128 KO à 1 Mo de mémoire de rafale.
- **Étape 2** - Les 64 premiers octets du paquet, appelés l'en-tête, sont traversés la commutation de paquets ASIC (PSA). 2.1 La PSA commute le paquet en consultant la table CEF locale dans la mémoire PSA. Si le paquet ne peut pas être commuté par la PSA, passez à l'étape 4 ; autrement, continuez à l'étape 3.
- **Étape 3** - Le gestionnaire de tampons de réception (RBM) reçoit l'en-tête de la PSA et la copie dans une en-tête de mémoire tampon libre. Si le paquet est plus grand que 64 octets, la queue du paquet est également copiée dans la même mémoire tampon libre dans la mémoire de paquet et est alignée sur la [file d'attente de sortie virtuelle](#) sortante LC. Passez à l'étape 5.
- **Étape 4** - Le paquet arrive à cette étape s'il ne peut pas être commuté par la PSA. Ces paquets sont placés sur la **file d'attente de paquets non traités (RawQ)** et le chemin de commutation est essentiellement identique que pour l'engine 1 et l'engine 0 LC de ce point (étape 4 dans le cas d'engine 0). Notez que les paquets qui sont commutés par la PSA ne

sont jamais placés dans le RawQ et aucune interruption est envoyés à la CPU.

- **Étape 5** - Le module d'interface de matrice (FIM) est responsable de segmenter les paquets dans des [cellules de Cisco](#) et d'envoyer les cellules à l'interface de matrice ASIC (la FIA) pour la transmission au LC sortant.

Commutation de paquets : Cellules de changement à travers la matrice

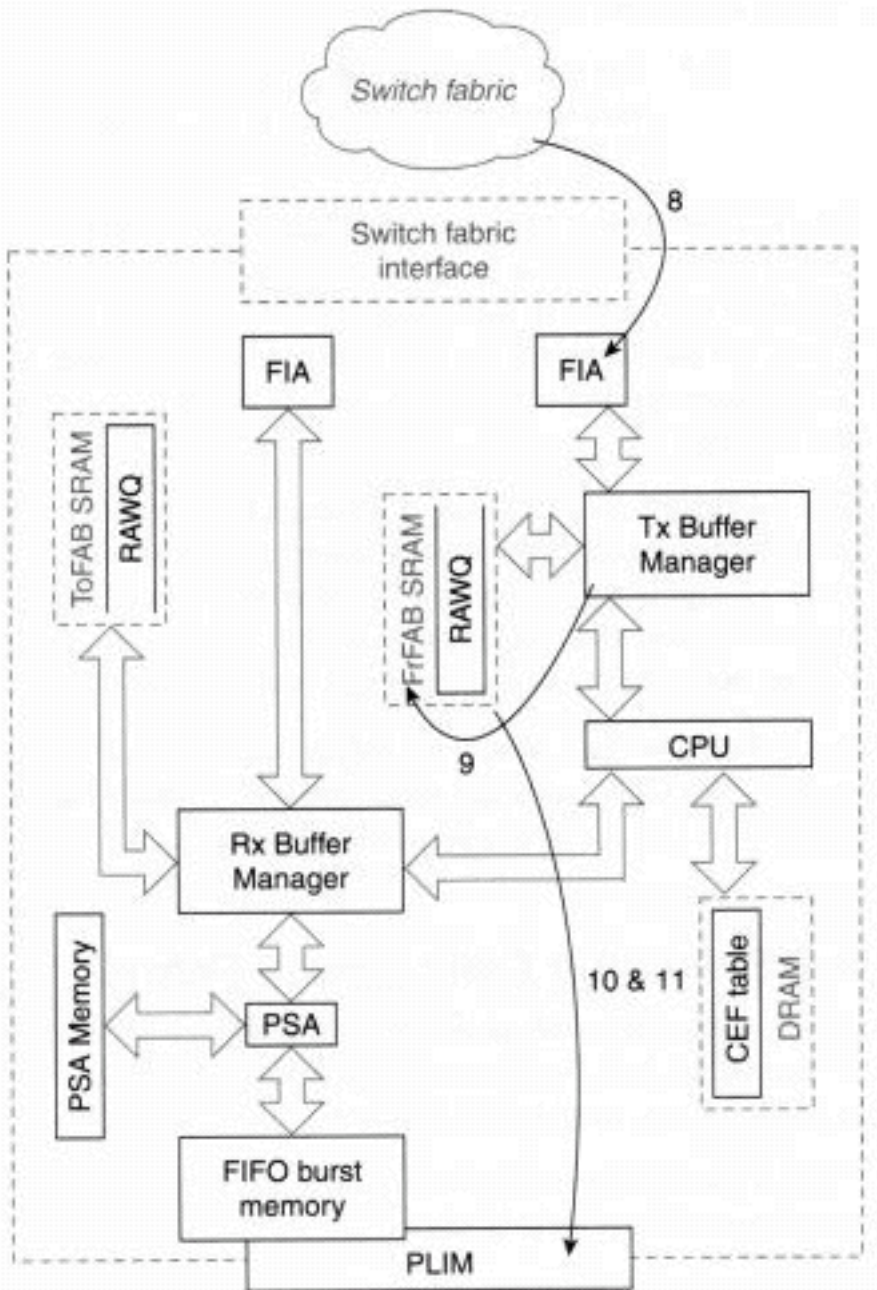
Vous arrivez à ce stade après l'engine de commutation de paquets commute les paquets. À ce stade, les paquets sont segmentés dans des cellules de Cisco et attendent d'être transmis à travers la matrice de commutation. Les étapes pour cette étape sont comme suit :

- **Étape 6** - La FIA envoie une demande de concession au CSC, qui programme le transfert de chaque cellule à travers la matrice de commutateur.
- **Étape 7** - Quand le programmeur accorde l'accès à la matrice de commutateur, les cellules sont transférées vers l'emplacement de destination. Notez que les cellules ne pourraient pas être transmises d'un seul trait ; d'autres cellules dans d'autres paquets pourraient être intercalées.

Commutation de paquets : Paquets de transmission

La figure 3 ci-dessous affiche la dernière étape de la commutation par paquets. Les cellules sont rassemblées et le paquet est transmis sur les medias. Ceci a lieu sur le linecard sortant.

Figure 3 : Commutation de paquets de Cisco 12000 : Transmettez l'étape



- **Étape 8** - Les cellules commutées à travers la matrice arrivent dans le linecard de destination par la FIA.
- **Étape 9** - Le gestionnaire de tampons de transmission alloue une mémoire tampon de la mémoire de paquet de transmission et rassemble le paquet dans cette mémoire tampon.
- **Étape 10** - Quand le paquet est reconstruit, la transmission BMA met le paquet en file d'attente sur la file d'attente de transmission de l'interface de destination sur le LC. Si la file d'attente de transmission d'interface est pleine (le paquet ne peut pas être mis en file d'attente), le paquet est lâché et le compteur de **perte de file d'attente de sortie** est incrémenté. **Note:** Dans la direction de transmission, le seul cas où des paquets sont placés dans le RawQ est quand la CPU LC doit faire le traitement avant transmission. Les exemples incluent la fragmentation IP, la Multidiffusion, et le CAR de sortie.
- **Étape 11** - Le processeur d'interface détecte un paquet attendant d'être transmis, retire la mémoire tampon de la mémoire de transmission, la copie de la file d'attente dans la mémoire interne FIFO, et transmet le paquet sur les medias.

Résumé d'écoulement de paquet

Des paquets IP qui traversent les 12000 sont traités en trois phases :

- Carte de ligne d'entrée dans trois sections :D'entrée PLIM (module d'interface de ligne physique) - Optique à la conversion, au l'ONU-tramage de la hiérarchie du Réseau optique synchrone (SONET) /Synchronous Digital (SDH), au traitement électriques HDLC, et de PPP.Expédition IP - Transmission de la décision fondée sur la consultation de FIB et queue dans une des files d'attente d'unicast d'entrée ou de files d'attente de Multidiffusion.Gestion et interface de matrice de file d'attente d'entrée - Dépistage précoce aléatoire du Détection précoce aléatoire (RED) /Weighted (WRED) traitant sur les files d'attente d'entrée et retirant de la file d'attente vers la matrice afin de maximiser l'utilisation de matrice.
- Paquets IP de changement par la matrice 12000 de la carte d'entrée à la carte de sortie ou aux cartes de sortie (en cas de Multidiffusion).
- Linecard de sortie dans trois sections :Interface de matrice de sortie - Rassemblant les paquets IP pour être envoyé et s'alignant dans des files d'attente de sortie ; traitement des paquets de multidiffusion.Gestion de file d'attente de sortie - RED/WRED traitant sur les files d'attente d'entrée et retirant de la file d'attente vers le de sortie PLIM pour maximiser la ligne utilisation de sortie.De sortie PLIM - HDLC et PPP traitant, tramage SONET/SDH, électrique à la conversion Optique.

Informations connexes

- [Support technique - Cisco Systems](#)