

# Cisco 12000 Series internetrouterarchitectuur: Packet-switching

## Inhoud

[Inleiding](#)

[Voorwaarden](#)

[Vereisten](#)

[Gebruikte componenten](#)

[Conventies](#)

[Achtergrondinformatie](#)

[Packet-switching: Overzicht](#)

[Packet-switching: Engine 0 en Engine 1 lijnkaarten](#)

[Packet-switching: Engine 2 lijnkaarten](#)

[Packet-switching: Switching Cells over Fabric](#)

[Packet-switching: Verzendpakketten](#)

[PacketFlow-overzicht](#)

[Gerelateerde informatie](#)

## [Inleiding](#)

Dit document onderzoekt de belangrijkste architecturale elementen van Cisco 12000 Series Internet Router — die pakketten overschakelt. Switching pakketten zijn radicaal anders dan elk van het gedeelde geheugen of de bus-gebaseerde Cisco-architecturen. Door een dwarsbalkstof te gebruiken, biedt Cisco 12000 zeer grote hoeveelheden bandbreedte en schaalbaarheid. Bovendien gebruikt de 12000 virtuele uitvoerwachtrijen om het sluitstuk van de lijn in het switch te verwijderen.

## [Voorwaarden](#)

### [Vereisten](#)

Er zijn geen specifieke vereisten van toepassing op dit document.

### [Gebruikte componenten](#)

De informatie in dit document is gebaseerd op de volgende hardware:

- Cisco 12000 Series internet-router

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 de potentiële impact van

elke opdracht begrijpen.

## Conventies

Zie de [Cisco Technical Tips Convention](#) voor meer informatie over documentconventies.

## Achtergrondinformatie

(Het overschakelingsbesluit op een Cisco 12000 wordt uitgevoerd door de lijnkaarten (LCs). Voor sommige LC's switches een speciale Application-Specific Integrated Circuit (ASIC) daadwerkelijk de pakketten. Distributed Cisco Express Forwarding (dCEF) is de enige beschikbare switchmethode.

**Opmerking:** De motoren 0, 1 en 2 zijn niet de meest recente motoren die door Cisco ontwikkeld zijn. Er zijn ook Engine 3, 4 en 4+ lijnkaarten, en er kan meer volgen. Engine 3 lijnkaarten kunnen Edge-functies met lijnsnelheid uitvoeren. Hoe hoger de Layer 3 motor, hoe meer pakketten in hardware worden geschakeld. U kunt nuttige informatie over de verschillende lijnkaarten die beschikbaar zijn voor Cisco 12000 Series router en de motor waarop zij gebaseerd zijn op [Cisco 12000 Series Internet Router: Veelgestelde vragen](#).

## Packet-switching: Overzicht

De pakketten worden altijd doorgestuurd door de ingang lijnkaart (LC). De stress LC voert alleen de uitgaande Quality of Service (QoS) uit die in de rij afhankelijk is (bijvoorbeeld Weighted Random Early Detection (WRED) of Committed Access Rate (CAR)). De meeste pakketten worden geschakeld door LC die gedistribueerd Cisco Express Forwarding (dCEF) gebruikt. Alleen de controle pakketten (zoals routingupdates) worden naar de Gigabit routeprocessor (GRP) verzonden voor verwerking. Het Packet-switching pad is afhankelijk van het type switchmotoren dat op de LC wordt gebruikt.

Dit gebeurt er wanneer een pakje inkomt:

1. Een pakket wordt geleverd in de Physical Layer Interface Module (PLIM). Hier gebeuren verschillende dingen: Een transceiver zet optische signalen in elektrische signalen (de meeste CSR-lijnkaarten hebben glasvezel connectors) L2-framing wordt verwijderd (SAN's, Asynchronous Transfer Mode (ATM), Ethernet, High-Level Data Link Control (HDLC)/Point-to-Point Protocol - PPP) ATM-cellen worden opnieuw geassembleerd Pakketten die niet voldoen aan de controle van de cyclische redundantie (CRC) worden weggegooid
2. Aangezien de verpakking wordt ontvangen en verwerkt, wordt het Direct Memory Access tot een klein (ongeveer 2x maximum transmissie unit (MTU)-geheugen, genaamd "First In, First Out (FIFO) burst memory" (First In, First Out (FIFO)). De hoeveelheid van dit geheugen is afhankelijk van het type LC (van 128 KB tot 1 MB).
3. Nadat het pakket volledig in FIFO-geheugen is opgeslagen, contacteert een applicatiespecifiek geïntegreerd circuit (ASIC) op de PLIM het Buffer Management ASIC (BMA) en vraagt een buffer om het pakket in te zetten. De BMA wordt verteld welke maat het pakje is en deelt dienovereenkomstig een buffer toe. Als de BMA geen buffer van de juiste grootte kan krijgen, wordt het pakje ingetrokken en wordt de "negeren" teller verhoogd op de inkomende interface. Er is geen terugvalmechanisme zoals bij sommige andere platforms. Terwijl dit gebeurt, kan de PLIM nog een pakje in het FIFO burst memory ontvangen,

waardoor het 2xMTU groot is.

4. Als er een vrije buffer beschikbaar is in de rechterrij, wordt het pakket opgeslagen door de BMA in de gratis rijlijst van de juiste grootte. Deze buffer wordt in de rij gezet, die door de Salsa ASIC of de R5K CPU wordt onderzocht. De R5K CPU bepaalt de bestemming van het pakket door de lokale DCEF-tabel in Dynamic RAM (DRAM) te raadplegen, en verplaatst vervolgens de buffer van de wachtrij naar een ToFabric-wachtrij die overeenkomt met de doelsleuf. Als de bestemming niet in de CEF-tabel is, wordt het pakje ingetrokken. Als het pakket een controlepakket is (bijvoorbeeld, het routeren van updates), wordt het aan de rij van het GRP gevraagd en door het GRP verwerkt. Er zijn 17 ToFab wachtrijen (16 unicast, plus 1 multicast). Er is één toFab wachtrij per lijnkaart (dit omvat de RP). Deze rijen staan bekend als "virtuele uitvoerrijen" en zijn belangrijk zodat het blokkeren van de kop-van-lijn niet plaatsvindt.
5. The ToFab BMA snijdt het pakket in 44 byte stukken op, die de lading voor wat uiteindelijk als "Cisco Cells" zal worden bekend zijn. Deze cellen krijgen een kop van 8 bytes en een bufferheader van 4 bytes door de frFab BMA (de totale gegevensgrootte is tot nu toe 56 bytes), om vervolgens te worden nagekeken in de juiste ToFab wachtrij (op welk punt komt de #Qelem-teller in het zwembad waar de buffer vandaan kwam, met één omlaag en de ToFab-wachtrij-teller wordt verhoogd). De "beslisser" is afhankelijk van het type schakelmotoren: Op motor 2+ kaarten wordt een speciale ASIC gebruikt om de manier te verbeteren waarop de pakketten worden geschakeld. Normale pakketten (IP/Tag, geen opties, checksum) worden direct door de Packet Switching ASIC (PSA) verwerkt, dan passeert u de ruwe wachtrij/CPU/SSA-combinatie en worden rechtstreeks op de toFab-wachtrij nagezocht. Alleen de eerste 64 bytes van het pakket worden door de Packet Switching ASIC doorgegeven. Als het pakket niet door de PSA kan worden geschakeld, wordt het pakket nagekeken naar de RawQ die door de CPU van de LC moet worden verwerkt zoals eerder uitgelegd. Op dit moment is de switching-beslissing genomen en is het pakket nagezocht naar de juiste ToFab uitvoerwachtrij.
6. Met de toFab BMA DMA's (Direct Memory Access) kunnen de cellen van het pakket in kleine FIFO-buffers in de fabric-interface ASIC (FIA) worden opgeslagen. Er zijn 17 FIFO-buffers (één per ToFab-wachtrij). Wanneer het FIA een cel van het toFab BMA krijgt, voegt het een CRC van 8 bytes toe (totale celgrootte - 64 bytes); 44 bytes payload, 8 bytes-celheader, 4 bytes-bufferheader) Het FIA heeft seriële lijneinterface (SLI) ASIC's die vervolgens 8B/10B-codering op de cel uitvoeren (zoals de Fibre Distributed Data Interface (FDDI) 4B/5B) en zich voorbereiden op verzending via het weefsel. Dit kan veel overhead lijken (44 bytes data worden omgezet in 80 bytes over het weefsel!), maar het is geen probleem omdat de capaciteit van het weefsel dienovereenkomstig is verdeeld.
7. Nu een FIA klaar is om te verzenden, vraagt de FIA toegang tot de stof van de momenteel actieve kaartserver en klok (CSC). De CSC werkt op een vrij complex eerlijke algoritme. Het idee is dat geen enkele LC het recht heeft om de uitgaande bandbreedte van een andere kaart te monopoliseren. Merk op dat zelfs als een LC gegevens uit één van zijn eigen poorten wil verzenden, dit nog steeds door het weefsel moet. Dit is belangrijk omdat als dit niet zou gebeuren, één haven op een LC alle bandbreedte voor een bepaalde haven op die zelfde LC zou kunnen monopoliseren. Het zou ook het omschakelingsontwerp ingewikkelder maken. De FIA stuurt cellen in het switch naar hun uitgaande LC (gespecificeerd door gegevens in de Cisco Cell header die wordt geplaatst met de switchingmodule). Het billijkheidsalgoritme is ook ontworpen voor optimale matching; indien kaart 1 naar kaart 2 wil worden verzonden en kaart 3 tegelijkertijd naar kaart 4 wil verzenden, gebeurt dit parallel. Dat is het grote verschil tussen een switch stof en een busarchitectuur. Zie het als analoog

aan een Ethernet switch versus een hub; op een switch, als port A naar haven B wil sturen en haven C wil praten met haven D, dan gebeuren die twee stromen onafhankelijk van elkaar. Op een hub zijn er half-duplexproblemen zoals aanrijdingen en terugkoppeling en het opnieuw proberen van algoritmen.

8. De Cisco-cellen die uit de stof komen, worden door de SLI-bewerking verwerkt om de 8B/10B-codering te verwijderen. Als er hier fouten zijn gemaakt, verschijnen deze in de bestandsindeling van het showcontroller als "celpariteit". Zie [Hoe u de uitvoer van de bestandsindeling van de showcontroller leest](#) voor meer informatie.
9. Deze Cisco Cells zijn DMA's in FIFO's op de FrFab FIA's, en vervolgens in een buffer op de FrFab BMA. FrFab BMA is degene die de herassemblage van cellen in een pakje doet. Hoe weet de FrFab BMA welke buffer om de cellen in te zetten voordat ze opnieuw in elkaar worden gezet? Dit is een ander besluit van de binnenkomende lijnkaartwisselmachine; omdat alle wachtrijen op het gehele vakje van dezelfde grootte en in dezelfde volgorde zijn, heeft de switchmachine alleen de Tx LC het pakket in de zelfde rij van het aantal geplaatst waar het de router binnendrong. De frFab BMA SDRAM wachtrijen kunnen worden bekeken met de opdracht van de showcontroller van de frStation op de LC. Zie [Hoe de uitvoer van het controllerformaat van de show lezen | De Opdrachten van de tofab wachtrij op een Cisco 12000 Series Internet Router](#) voor meer informatie. Dit is in principe hetzelfde idee als de BMA-output van Fab. Pakketten komen binnen en worden in pakketten geplaatst die van hun respectieve vrije rijen worden gedewachtrij geplaatst. Deze pakketten worden in de rij van het uit-stof geplaatst, nagezocht op of de interfacerij (er is één rij per fysieke poort) of ruweQ voor uitvoerverwerking. Er gebeurt niet veel in het ruweQ: multicast-replicatie per poort, Modified Deficit round robin (MDRR) - hetzelfde idee als Distributed Weighted Fair Queuing (DWFQ) en output CAR. Als de rij voor verzenden vol is, wordt het pakket verbroken en wordt de uitrolteller verhoogd.
10. FrFab BMA wacht tot het TX-gedeelte van de PLIM klaar is om een pakje te verzenden. De frFab BMA voert de eigenlijke MAC-herschrijven uit (gebaseerd, onthoud, op informatie in de Cisco Cell header) en DMA's maken het pakket over naar een kleine (opnieuw 2xMTU) buffer in het PLIM-circuit. De PLIM voert de ATM SAR en SONET in, indien van toepassing, en geeft het pakket door.
11. ATM-verkeer wordt opnieuw gemonteerd (door de SAR), gesegmenteerd (door de tofab BMA), opnieuw gemonteerd (door de fromfab BMA) en opnieuw gesegmenteerd (door de fromfab SAR). Dit gebeurt heel snel.

Dat is de levenscyclus van een pakje, van begin tot eind. Als je wilt weten hoe een SR er aan het einde van de dag uit ziet, lees dan deze hele krant 500.000 keer!

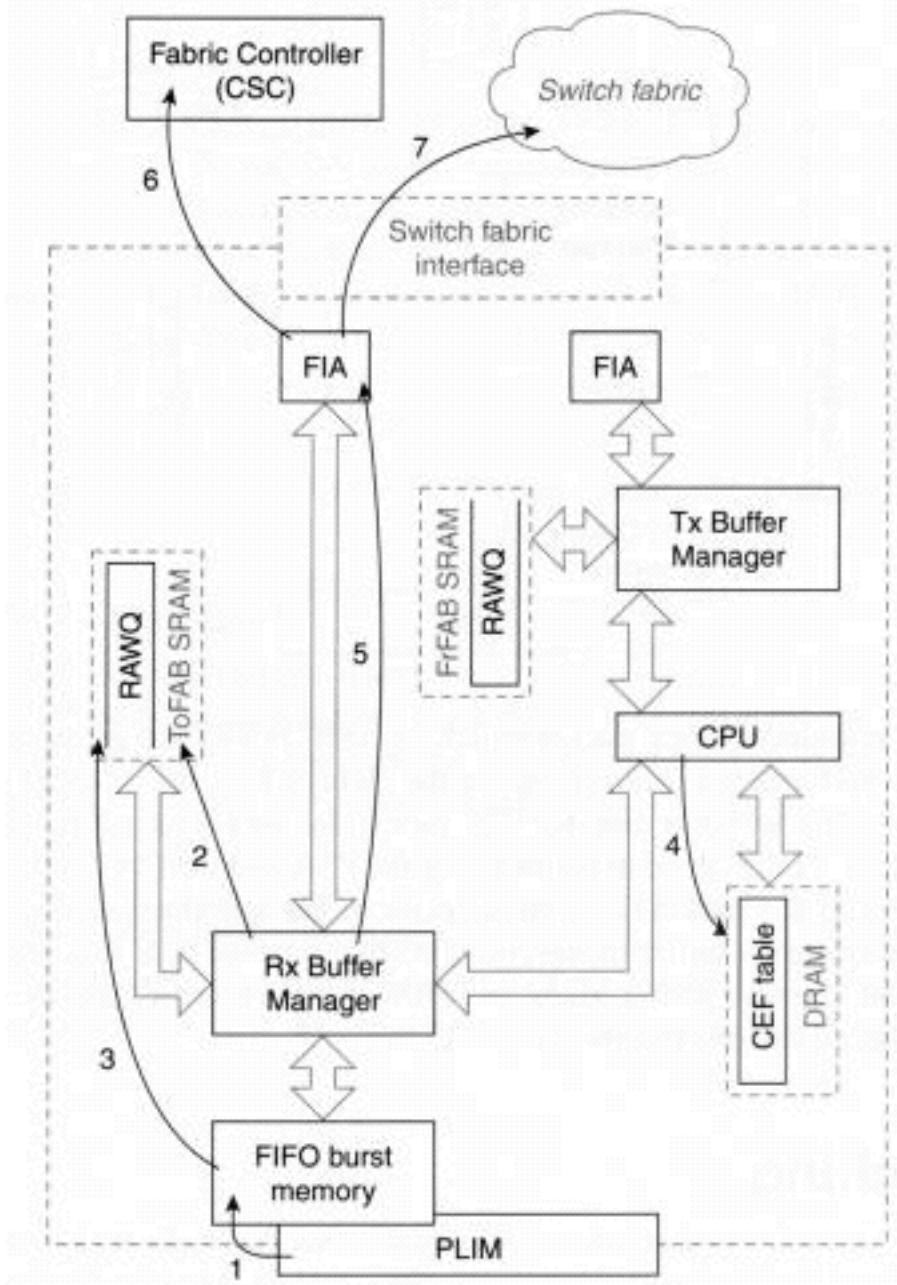
Het pakketswitchpad op de GSR is afhankelijk van het type verzendingsmotor op de LC. Nu gaan we door alle stappen voor Engine 0, Engine 1 en de twee LC's.

## [Packet-switching: Engine 0 en Engine 1 lijnkaarten](#)

De onderstaande secties zijn gebaseerd op het boek *Inside Cisco IOS Software Architecture*, Cisco Press.

[Afbeelding 1](#) illustreert de verschillende stappen tijdens de pakketswitching voor een motor 0 of motor 1 LC.

**Afbeelding 1: Engine 0 en Engine 1 switchingpad**



Het overschakelingspad voor motor 0 en motor 1 LC is in wezen hetzelfde, hoewel de motor 1 LC over een verbeterde overschakelmachine en bufferbeheerder beschikt voor betere prestaties. Het switchpad is als volgt:

- **Step 1** - De interfaceprocessor (PLIM) detecteert een pakje op de netwerkmedia en begint het te kopiëren naar een FIFO-geheugen dat **barst geheugen** op de LC wordt genoemd. De hoeveelheid barstgeheugen die elke interface heeft, is afhankelijk van het type LC; typische LCs hebben 128 KB tot 1 MB barstgeheugen.
- **Step 2** - De interfaceprocessor vraagt om een pakketbuffer van de ontvangen BMA; het depot waaruit de buffer wordt gevraagd is afhankelijk van de lengte van het pakje. Als er geen vrije buffers zijn, wordt de interface ingetrokken en wordt de "negeren" teller van de interface verhoogd. Als een pakket van 64 bytes bijvoorbeeld in een interface arriveert, probeert de BMA een pakketbuffer van 80 bytes toe te wijzen. Als er geen vrije buffers bestaan in de 80 bytes pool, worden de buffers niet toegewezen in de volgende beschikbare pool.
- **Step 3** - Wanneer een vrije buffer door de BMA wordt toegewezen, wordt het pakket naar de buffer gekopieerd en wordt het nagezocht op de **ruwe wachtrij** (RawQ) voor verwerking door de CPU. Een onderbreken wordt naar de LC CPU verzonden.

- **Stap 4** - De CPU's van de LC verwerken elk pakket in het RawQ zoals het ontvangen is (RawQ is een FIFO), waarbij u de lokale verdeelde tabel van Cisco Express Doorsturen in DRAM raadpleegt om een switching-besluit te nemen.**4.1** Als dit een unicast IP-pakket met een geldig doeladres in de CEF-tabel is, wordt de pakketheader opnieuw geschreven met de nieuwe insluitingsinformatie die uit de nabijheidstabel CEF wordt verkregen. Het geschakelde pakket wordt nagezocht op de virtuele uitvoerwachtrij die overeenkomt met de doelsleuf.**4.2** Als het doeladres niet in de CEF-tabel staat, wordt het pakje ingetrokken.**4.3** Als het pakket een controlepakket is (een routeringsupdate, bijvoorbeeld), wordt het pakket in de virtuele uitvoerwachtrij van het GRP nagezocht en door het GRP verwerkt.
- **Stap 5** - Het ontvangt BMA fragmenten uit het pakket in 64-bytes cellen, en geeft deze af aan het FIA voor transmissie naar de uitgaande LC.

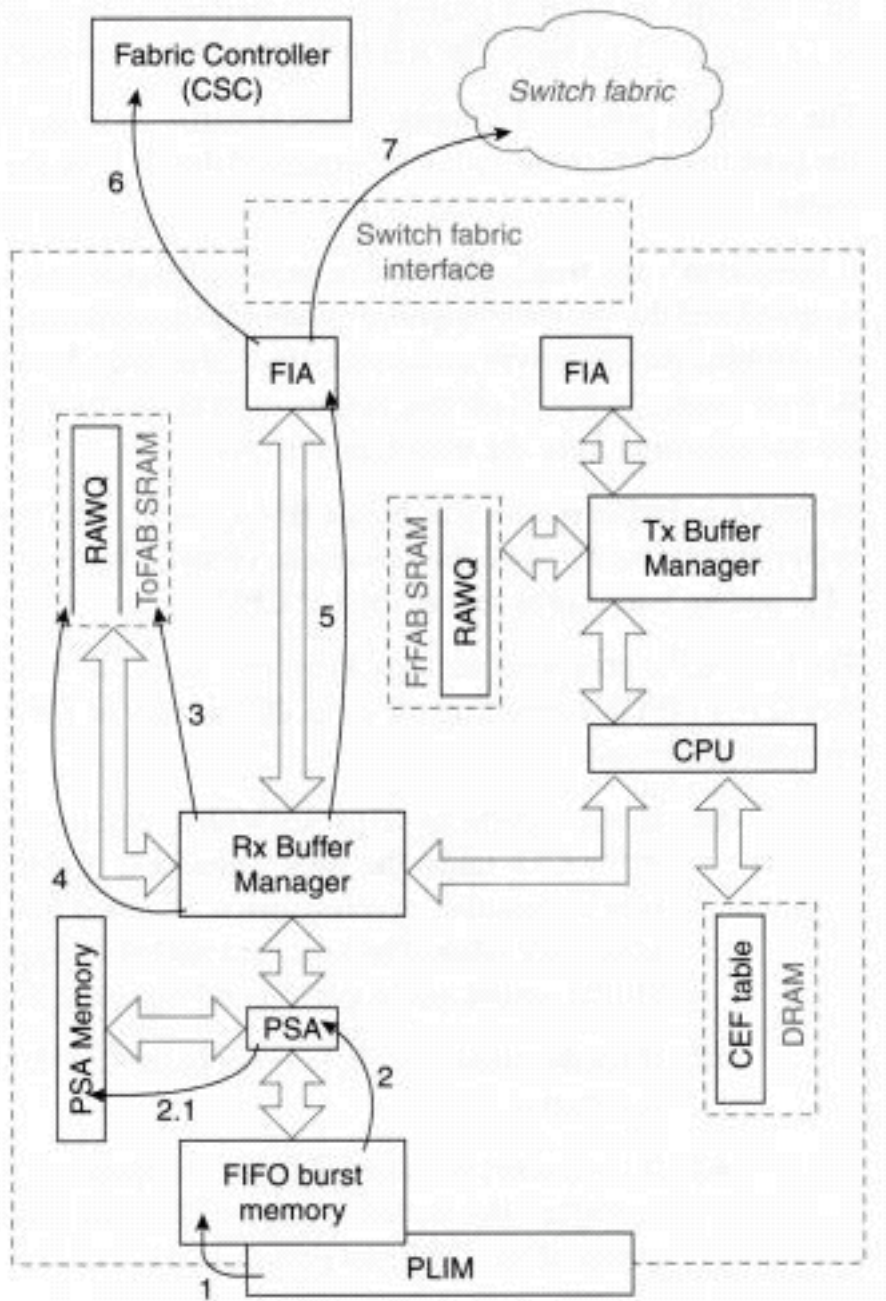
Aan het eind van Stap 5 is het pakket dat in een motor 0/1 LC is aangekomen, geschakeld en klaar om over het switch materiaal als cellen te worden getransporteerd. Ga naar Stap 6 in het gedeelte [Packet Switching: Switching Cells over Fabric](#).

## Packet-switching: Engine 2 lijnkaarten

[Afbeelding 2](#) illustreert het pakketswitchpad wanneer de pakketten in een Engine 2 LC arriveren, zoals in de volgende lijst met stappen beschreven wordt.

**Afbeelding 2: Engine 2 switchingpad**





- **Stap 1** - De interfaceprocessor (PLIM) detecteert een pakje op de netwerkmedia en begint het te kopiëren naar een FIFO-geheugen dat **barst geheugen** op de LC wordt genoemd. De hoeveelheid barstgeheugen die elke interface heeft, is afhankelijk van het type LC; typische LCs hebben 128 KB tot 1 MB barstgeheugen.
- **Stap 2** - De eerste 64 bytes van het pakket, genaamd de header, worden doorgegeven met de Packet Switching ASIC (PSA).**2.1** De PSA switch het pakket door de lokale CEF-tabel in het PSA-geheugen te raadplegen. Als het pakket niet door de PSA kan worden geschakeld, ga dan naar Stap 4. anders gaat u verder met Stap 3.
- **Stap 3** - De Buffer Manager (RBM) accepteert de header van de PSA en kopieert deze naar een gratis bufferkop. Als het pakket groter is dan 64 bytes, wordt de staart van het pakket ook gekopieerd naar dezelfde gratis buffer in het pakketgeheugen en wordt de wachtrij in de uitgaande LC [virtuele uitvoerwachtrij geplaatst](#). Ga naar stap 5.
- **Stap 4** - Het pakket arriveert bij deze stap als het niet door de PSA kan worden geschakeld. Deze pakketten worden op de **ruwe wachtrij** geplaatst (RawQ) en het switchpad is in wezen hetzelfde als voor Engine 1 en Engine 0 LC vanaf dit punt (stap 4 in het geval van motor 0). Merk op dat de pakketten die door de PSA zijn geschakeld nooit in RawQ worden geplaatst

en dat geen onderbreken naar de CPU wordt verzonden.

- **Stap 5** - De Fabric Interface Module (FIM) is verantwoordelijk voor het segmenteren van de pakketten in [Cisco Cells](#) en het verzenden van de cellen naar de Fabric Interface ASIC (FIA) voor transmissie naar de uitgaande LC.

## Packet-switching: Switching Cells over Fabric

Je komt in dit stadium nadat de pakketwisselmachine de pakketten switch. In dit stadium worden de pakketten gesegmenteerd in Cisco Cells en wachten om over het switching materiaal te worden verzonden. De stappen voor deze fase zijn:

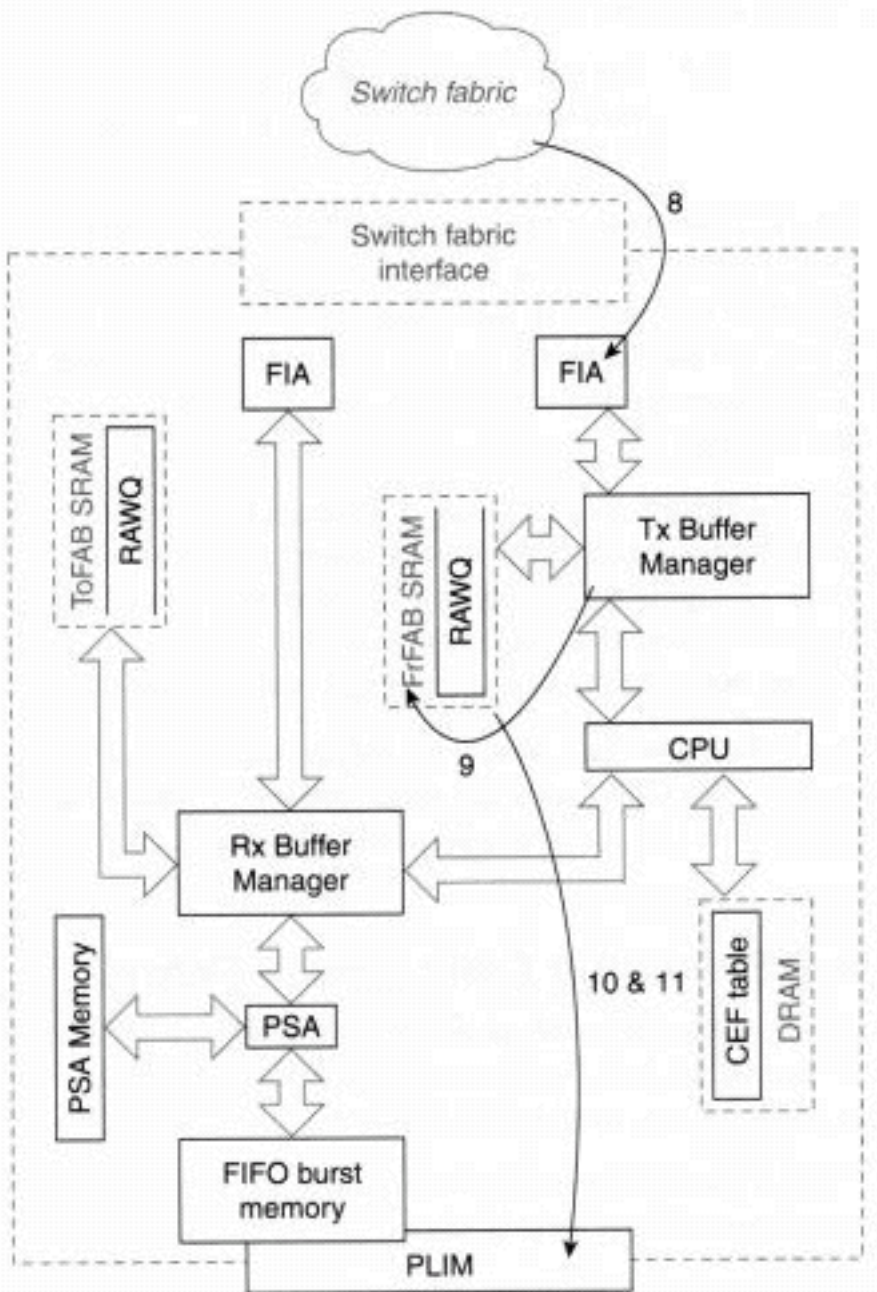
- **Stap 6** - Het FIA stuurt een subsidieverzoek naar de CSC, die de overdracht van elke cel door het switch-weefsel plannt.
- **Stap 7** - Wanneer de planner toegang tot het switch-weefsel verleent, worden de cellen overgebracht naar de doelsleuf. Merk op dat de cellen niet allemaal tegelijk kunnen worden overgedragen; andere cellen in andere pakketten kunnen worden vervormd .

## Packet-switching: Verzendpakketten

Afbeelding 3 hieronder toont het laatste stadium van pakketswitching. De cellen worden opnieuw geassembleerd en het pakket wordt op de media verzonden. Dit gebeurt op de uitgaande lijnkaart.

**Afbeelding 3: Cisco 12000 Packet-switching: Transmit-stadium**





- **Stap 8** - De cellen die over het weefsel worden geschakeld, komen via het FIA in de doellijnkaart.
- **Stap 9** - De bedenken Buffer Manager wijst een buffer toe uit het geheugen van het verzendpakket en assembleert het pakket in deze buffer.
- **Stap 10** - Wanneer het pakket wordt herbouwd, stuurt het verzenden BMA het pakket naar de verzendwachtrij van de doelinterface op de LC. Als de interface-verzendwachtrij vol is (het pakket kan niet worden nagezocht), wordt het pakje ingetrokken en wordt de **uitvoerrij** bijgewerkt.**N.B.:** In de verzendrichting worden de enige tijdpakketten in RawQ geplaatst wanneer de LC CPU elke verwerking vóór verzending moet uitvoeren. Tot de voorbeelden behoren IP fragmentatie, multicast, en output CAR.
- **Stap 11** - De interfaceprocessor detecteert een pakje dat wacht om te worden verzonden, stelt de buffer van het verzendgeheugen in de wachtrij, kopieert het geheugen in het FIFO-geheugen en geeft het pakje op de media door.

## [PacketFlow-overzicht](#)

IP-pakketten die over de 12000 lopen, worden in drie fasen verwerkt:

- Inloop lijnkaart in drie delen: Ingress PLIM (Physical Line Interface Module) - optisch naar elektrische conversie, synchrone optische netwerkmodule (SONET)/synchrone digitale Hierarchie (SDH) niet-framing, HDLC en PPP-verwerking. IP Forwarding - Forwarding besluit op basis van wachtrijen voor FIB en wachtrijen in één van de ingress unicast-wachtrijen voor multicast. Beheer en fabricinterface van Ingress in wachtrij - willekeurige vroege detectie (RED)/Weighted Random Early Detection (WRED)-verwerking op ingangswachtrijen en dewachtrij voor weefsel om gebruik van stoffen te maximaliseren.
- IP-pakketten via de 12000 fabric-enabled-kaart om te zetten van een toegangkaart naar een opslagkaart of toegangkaart (in het geval van multicast).
- Bovenste lijnkaart in drie delen: uitgaande fabric-interface - opnieuw assembleren van de IP-pakketten die moeten worden verzonden en in wachtrijen worden geplaatst; multicast-pakketten verwerken. Groot rijbeheer - RODE/WRED verwerking in de ingangswachtrijen en wachtrijen naar de res PLIM om het gebruik van de streep-lijn te maximaliseren. Druk op PLIM - HDLC- en PPP-verwerking, SONET/SDH-framing, elektrische to-optische conversie.

## [Gerelateerde informatie](#)

- [Technische ondersteuning - Cisco-systemen](#)