

Arquitetura do Roteador de Internet do Cisco 12000 Series: Switching de pacotes de informações

Contents

[Introduction](#)

[Prerequisites](#)

[Requirements](#)

[Componentes Utilizados](#)

[Conventions](#)

[Informações de Apoio](#)

[Switching de pacotes de informações Overview](#)

[Switching de pacotes de informações Placas de linha Engine 0 e Engine 1](#)

[Switching de pacotes de informações Placas de linha do Engine 2](#)

[Switching de pacotes de informações Comutação de células em toda a estrutura](#)

[Switching de pacotes de informações Transmitindo pacotes](#)

[Resumo de fluxo de pacote](#)

[Informações Relacionadas](#)

[Introduction](#)

Este documento examina os elementos arquitetônicos mais importantes do Cisco 12000 Series Internet Router — comutação de pacotes. Pacotes de switching são radicalmente diferentes de qualquer uma das arquiteturas Cisco com base em memória compartilhada ou barramento. Usando uma estrutura de barras cruzadas, o Cisco 12000 oferece grandes quantidades de largura de banda e escalabilidade. Além disso, o 12000 usa filas de saída virtual para eliminar o Head of Line Blocking dentro do Switch Fabric.

[Prerequisites](#)

[Requirements](#)

Não existem requisitos específicos para este documento.

[Componentes Utilizados](#)

As informações neste documento são baseadas no seguinte hardware:

- Cisco 12000 Series Internet Router

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. If your network is live, make sure that you understand the potential impact of any command.

Conventions

Para obter mais informações sobre convenções de documento, consulte as [Convenções de dicas técnicas Cisco](#).

Informações de Apoio

(A decisão de switching em um Cisco 12000 é feita pelas placas de linha (LCs). Para algumas LCs, um ASIC (Circuito integrado específico do aplicativo) é capaz de comutar os pacotes. O Distributed Cisco Express Forwarding (dCEF) é o único método de switching disponível.

Lembrete: Os mecanismos 0, 1 e 2 não são os mecanismos mais recentes desenvolvidos pela Cisco. Também existem as placas de ingresso Mecanismo 3, 4 e 4+, com mais a seguir. As placas de linha do Engine 3 são capazes de executar recursos de extremidade na taxa de linha. Quanto maior o Layer 3 Engine, mais pacotes são comutados no hardware. Você pode encontrar algumas informações úteis sobre as diferentes placas de linha disponíveis para o Cisco 12000 Series Router e o mecanismo no qual elas se baseiam no [Cisco 12000 Series Internet Router: Perguntas frequentes](#)

Switching de pacotes de informações Overview

Os pacotes sempre são encaminhados pela placa de entrada (LC). O LC de saída executa somente a Qualidade de Serviço (QoS) de saída que depende da fila (por exemplo, Detecção Antecipada Aleatória Ponderada (WRED - Weighted Random Early Detection) ou Taxa de Acesso Consolidada (CAR - Committed Access Rate). A maioria dos pacotes é comutada pelo LC usando o Cisco Express Forwarding distribuído (dCEF). Apenas os pacotes de controle (como, por exemplo, atualizações de roteamento) são enviados para o Gigabit Route Processor (GRP) para processamento. O caminho de switching do pacote depende do tipo de mecanismo de switching usado no LC.

Isso é o que acontece quando um pacote entra:

1. Um pacote entra no PLIM (módulo de interface da camada física). Várias coisas acontecem aqui: Um transceptor transforma sinais óticos em elétricos (a maioria das placas de linha CSR têm conectores de fibra) O enquadramento L2 é removido (SANE, Asynchronous Transfer Mode (ATM), Ethernet, High-Level Data Link Control (HDLC)/Point-to-Point Protocol - PPP) As células ATM são remontadas Pacotes que falham na verificação de redundância cíclica (CRC) são descartados
2. À medida que o pacote é recebido e processado, é a memória direta acessada em uma pequena (aproximadamente 2 x memória intermédia MTU) chamada de memória de intermitência FIFO (First In, First Out). A quantidade desta memória depende do tipo de LC (de 128 KB a 1 MB).
3. Quando o pacote estiver completamente na memória FIFO, um ASIC (application-specific integrated circuit, circuito integrado específico da aplicação) no PLIM entra em contato com o BMA (Buffer Management ASIC, ASIC de gerenciamento de buffer) e solicita um buffer para colocar o pacote. O BMA recebe a informação de qual o tamanho do pacote e aloca um

buffer adequadamente. Se o BMA não obtiver um buffer do lado direito, o pacote será descartado e o contador "ignorar" será incrementado na interface de entrada. Não há mecanismo de recuo como em algumas outras plataformas. Enquanto isto está acontecendo, o PLIM poderia estar recebendo outro pacote na memória de intermitência FIFO, por isso é que tem o tamanho 2xMTU.

4. Se houver um buffer livre disponível na fila da direita, o pacote será armazenado pelo BMA na lista de filas livres do tamanho apropriado. Esse buffer é colocado na Fila Bruta, que é examinada pela Salsa ASIC ou pela R5K CPU. A CPU R5K determina o destino do pacote, consultando sua tabela local dCEF na RAM dinâmica (DRAM); em seguida, desloca o buffer da fila bruta para uma fila ToFabric correspondente ao slot de destino. Se o destino não estiver na tabela CEF, o pacote será descartado. Se o pacote for um pacote de controle (por exemplo, atualizações de roteamento), ele será colocado na fila do GRP e processado pelo GRP. Há 17 filas ToFab (16 unicast mais 1 multicast). Há uma fila toFab por placa de linha (isso inclui o RP). Essas filas são conhecidas como "filas de saída virtuais" e são importantes para que o bloqueio de head-of-line não ocorra.
5. O ToFab BMA corta o pacote em pedaços de 44 bytes, que são o payloads do que posteriormente serão conhecidas como células Cisco. A essas células é fornecido pelo BMA frFab um cabeçalho de 8 bytes e um cabeçalho de buffer de 4-bytes (tamanho total dos dados até agora = 56 bytes) e, em seguida, elas são enfileiradas na própria fila ToFab (neste ponto, o contador #Qelem no conjunto do qual o buffer é proveniente diminui em um ponto e o contador de fila ToFab aumenta em um ponto). O "responsável pelas decisões" depende do tipo de mecanismo de switching: Nas placas Engine 2+, um ASIC especial é usado para melhorar a maneira como os pacotes são comutados. Pacotes normais (IP/Tag, nenhuma opção, soma de verificação) são processados diretamente pelo ASIC de Switching de Pacotes (PSA), em seguida, ignoram a combinação de fila bruta/CPU/Salsa e são colocados diretamente na fila toFab. Somente os primeiros 64 bytes do pacote são transferidos pelo ASIC de switching de pacotes. Se o pacote não puder ser comutado pelo PSA, ele será colocado na fila RawQ para ser tratado pelo CPU do LC, como explicado anteriormente. Nesse momento, a decisão de switching terá sido tomada e o pacote terá sido enfileirado na fila de saída ToFab adequada.
6. O toFab BMA faz um DMA (Acesso Direto à Memória) das células em buffers FIFO pequenos na interface da tela ASIC (FIA). Existem 17 buffers FIFO (um por fila ToFab). Quando o FIA obtém uma célula do BMA toFab, adiciona um CRC de 8 bytes (tamanho total da célula - 64 bytes; payload de 44 bytes, cabeçalho de célula de 8 bytes, cabeçalho de buffer de 4 bytes). O FIA contém ASCIs de interface de linha serial (SLI) que então efetuam uma codificação 8B/10B na célula (como Interface de Dados Distribuídos de Fibra Ótica, FDDI, 4B/5B) e se prepara para transmiti-la sobre a estrutura. Isso pode parecer uma grande sobrecarga (44 bytes de dados são transformados em 80 bytes pela estrutura!), mas não é um problema, pois a capacidade da estrutura foi provisionada de acordo.
7. Agora que um FIA está pronto para transmitir, o FIA solicita acesso à tela do programador de cartão e relógio (CSC) atualmente ativos. O CSC funciona em um algoritmo de clareza bastante complexo. A ideia é que nenhum LC tem permissão para monopolizar a largura de banda de saída de qualquer outra placa. Observe que mesmo que um LC queira transmitir dados de uma de suas próprias portas, ele ainda precisa passar pela estrutura. Essa característica é importante porque, não fosse por ela, uma porta em um LC poderia monopolizar toda a largura de banda de outra porta nesse mesmo LC. Também tornaria mais complicado o projeto de switching. O FIA envia células através da matriz de comutação para o LC de saída (especificado pelos dados no cabeçalho da célula da Cisco colocado lá

pelo mecanismo de comutação). O algoritmo de compatibilidade também foi concebido para uma correspondência ótima; Se a placa 1 desejar transmitir para a placa 2 e a placa 3 desejar transmitir para a placa 4 ao mesmo tempo, isto acontece em paralelo. Essa é a grande diferença entre uma matriz de comutação e uma arquitetura de barramento. Pense nisso como análogo a um switch Ethernet versus um hub; em um switch, se a porta A desejar enviar à porta B e a porta C quiser se comunicar com a porta D, esses dois fluxos acontecem independentemente um do outro. Em um hub, há problemas half-duplex como colisões e algoritmos de backoff e de repetição.

8. As células Cisco que saem da estrutura passam pelo processamento SLI para remover a codificação 8B/10B. Se houver algum erro aqui, eles aparecerão na saída do comando `show controller fia` como "paridade de célula". Consulte [Como ler a saída do comando show controller fia](#) para obter informações adicionais.
9. Essas células Cisco são DMAs em FIFOs nos frFab FIAs e depois em um buffer no frFab BMA. O frFab BMA é o que realmente faz a remontagem de células em um pacote. Como a BMA frFab sabe em que buffer colocar as células antes de remontá-las? Esta é outra decisão tomada pelo mecanismo de switching da placa de linha recebida; como todas as filas em toda a caixa são do mesmo tamanho e na mesma ordem, o mecanismo de comutação apenas tem a LC Tx que colocou o pacote na fila de número igual a partir da qual ele entrou no roteador. As filas FrFab BMA SDRAM podem ser visualizadas com o comando `show controller frfab queue` no LC. Consulte [Como ler a saída do comando show controller frfab Comandos | tofab queue em um Cisco 12000 Series Internet Router](#) para obter detalhes. Essa é basicamente a mesma idéia da saída toFab BMA. Os pacotes chegam e são colocados em pacotes que são retirados de suas respectivas filas livres. Esses pacotes são colocados na fila da estrutura de partida, enfileirados na fila da interface (há uma fila por porta física) ou na fila bruta para processamento de saída. Pouca coisa acontece na cruzQ: replicação multicast por porta, Modified Deficit Round Robin (MDRR): mesmo conceito apresentado pelo Distributed Weighted Fair Queuing (DWFQ) e pelo CAR de saída. Se a fila de transmissão estiver cheia, o pacote será descartado e o contador de queda de emissor será incrementado.
10. O frFab BMA espera até que a parte TX do PLIM esteja pronta para enviar um pacote. O frFab BMA faz a regravação real de MAC (com base, lembre-se, nas informações contidas no cabeçalho da célula da Cisco) e DMAs o pacote em um pequeno buffer (novamente, 2xMTU) no circuito PLIM. O PLIM encapsula o SAR e o SONET do ATM, quando apropriado, e transmite o pacote.
11. O tráfego ATM é remontado (pelo SAR), segmentado (pelo BMA tofab), remontado (pelo BMA fromfab) e segmentado novamente (pelo SAR fromfab). Isto ocorre muito rápido.

Esse é o ciclo de vida de um pacote, do início ao fim. Se quiser saber como é um GSR no final do dia, leia este artigo inteiro 500.000 vezes!

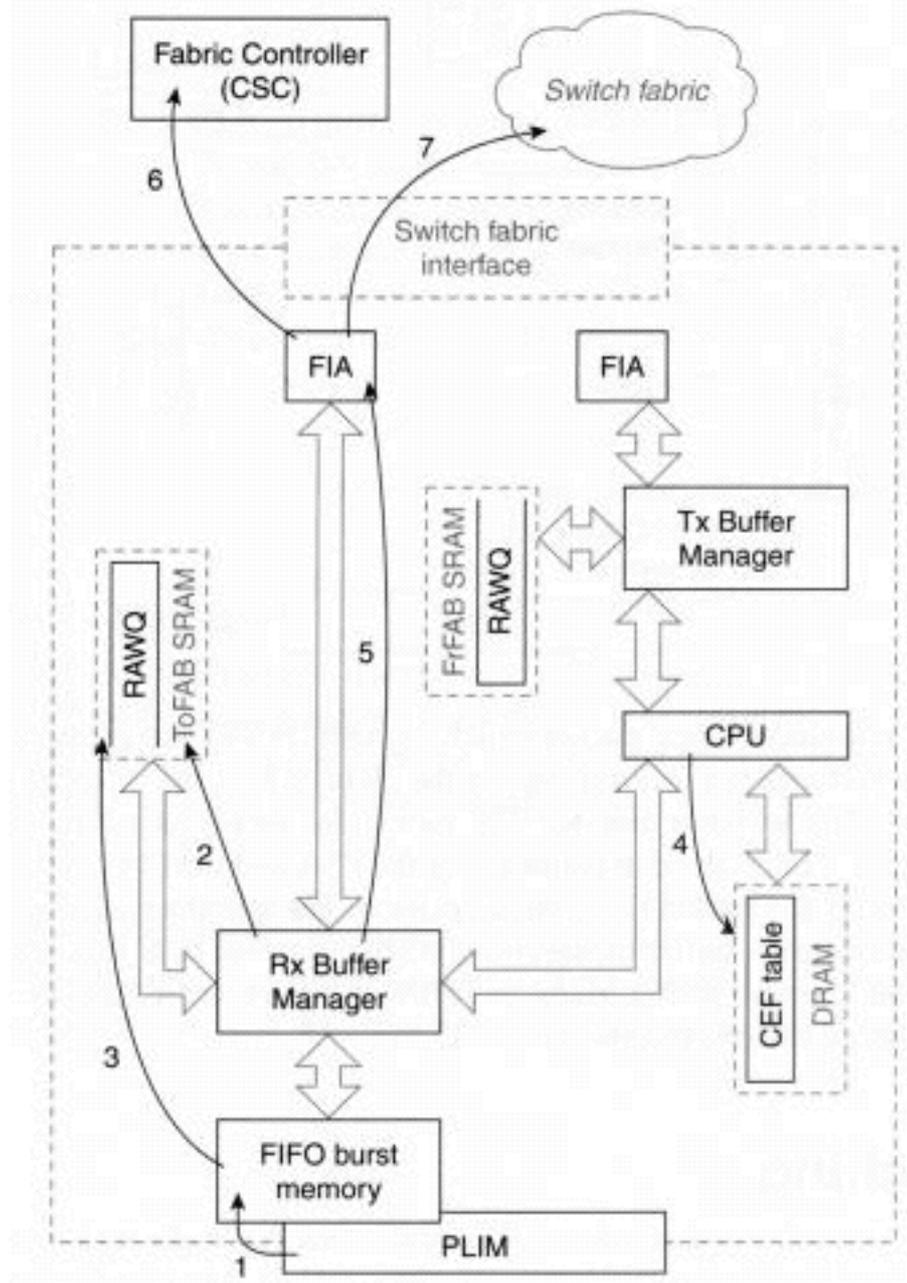
O caminho de switching do pacote no GSR depende do tipo de mecanismo de encaminhamento da LC. Agora, revisaremos todos os passos do Engine 0, Engine 1 e as duas LCs.

[Switching de pacotes de informações Placas de linha Engine 0 e Engine 1](#)

As seções abaixo são baseadas no livro Inside Cisco IOS Software Architecture (Por dentro da arquitetura Cisco IOS), Cisco Press.

A [Figura 1](#) abaixo ilustra as diferentes etapas durante a comutação de pacotes para um Engine 0 ou Engine 1 LC.

Figura 1: Caminho de switching de Engine 0 e Engine 1



O caminho de comutação para o Engine 0 e Engine 1 LC é essencialmente o mesmo, embora o Engine 1 LC tenha um mecanismo de comutação avançado e um gerenciador de buffer para aumentar o desempenho. O caminho de switching é o seguinte:

- **Etapa 1** - O processador da interface (PLIM) detecta um pacote na mídia de rede e começa a copiá-lo em uma memória FIFO chamada memória de intermitência no LC. A quantidade de memória de intermitência que cada interface tem depende do tipo de LC; as LCs típicas têm de 128 KB a 1 MB de memória de intermitência.
- **Etapa 2** - O processador da interface solicita um buffer de pacote do BMA de recebimento; o pool do qual o buffer é solicitado depende do comprimento do pacote. Se não houver nenhum buffer livre, a interface será descartada e o contador "ignore" (ignorar) da interface será incrementado. Por exemplo, se um pacote de 64 bytes atingir uma interface, o BMA tentará alocar um buffer de pacote de 80 bytes. Se não houver buffers livres no pool de 80 bytes, não

serão alocados buffers do próximo pool disponível.

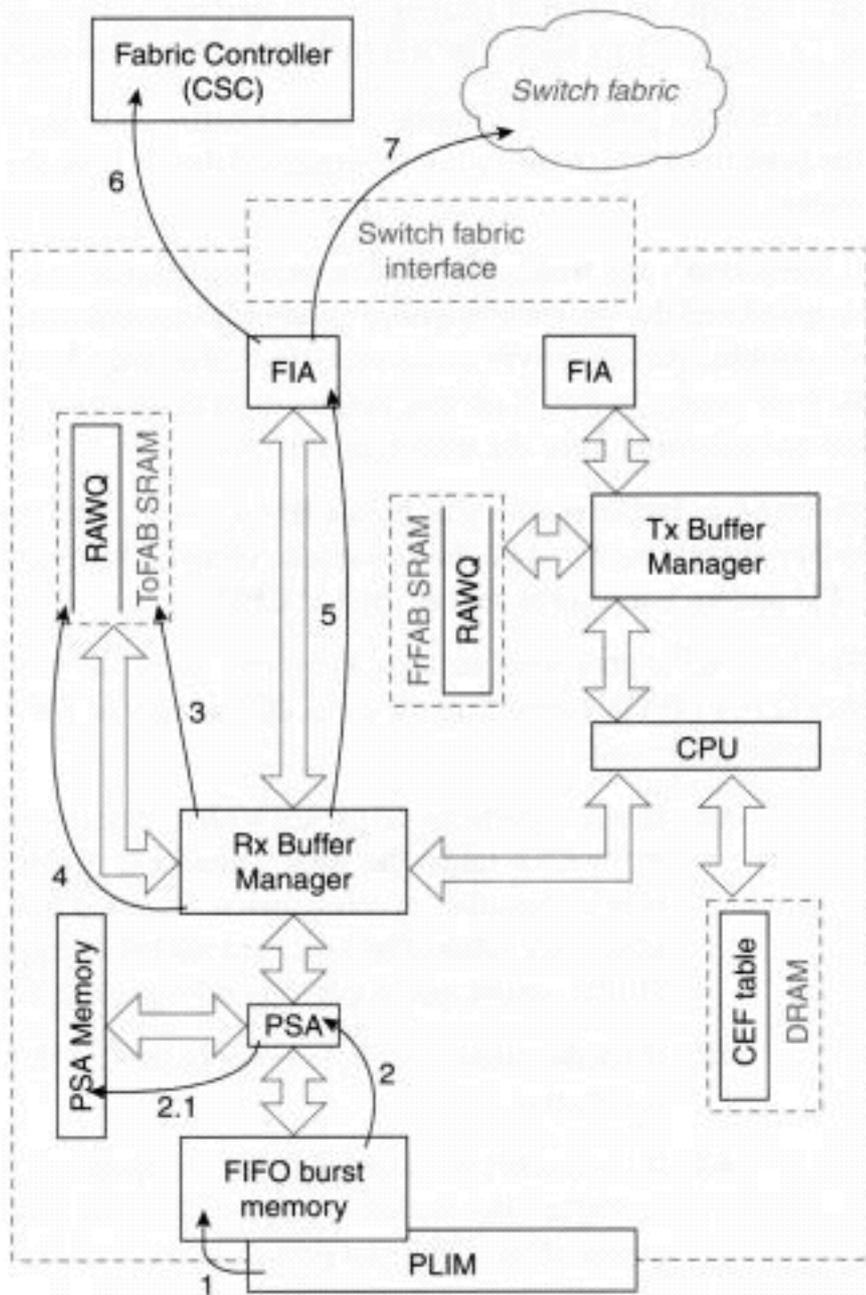
- **Passo 3** Quando um buffer livre é alocado pelo BMA, o pacote é copiado para o buffer e enfileirado na Fila Bruta (RawQ) para processamento pela CPU. Uma interrupção é enviada para a CPU do LC.
- **Etapa 4** - A CPU do LC processa cada pacote no RawQ à medida que ele é recebido (o RawQ é um FIFO), consultando a tabela local distribuída do Cisco Express Forwarding na DRAM para tomar uma decisão de switching.**4.1** Se este for um pacote IP unicast com um endereço de destino válido na tabela CEF, o cabeçalho do pacote será regravado com as novas informações de encapsulamento obtidas da tabela de adjacência CEF. O pacote comutado é enfileirado na fila de saída virtual peer ao slot de destino.**4.2** Se o endereço de destino não estiver na tabela CEF, o pacote será descartado.**4.3** Se o pacote for um pacote de controle (uma atualização de roteamento, por exemplo), o pacote será enfileirado na fila de saída virtual do GRP e processado pelo GRP.
- **Etapa 5** – O BMA de recepção fragmenta o pacote em células de 64 bytes e as transfere para o FIA, para transmissão ao LC de saída.

No final da Etapa 5, o pacote que chegou a um Engine 0/1 LC foi comutado e está pronto para ser transportado através da matriz de comutação como células. Vá para a etapa 6 na seção [Packet Switching: Comutação de células através da estrutura](#).

Switching de pacotes de informações Placas de linha do Engine 2

[A Figura 2](#) abaixo ilustra o caminho de comutação de pacotes quando os pacotes chegam a um Engine 2 LC, conforme descrito na lista de etapas a seguir.

Figura 2: Caminho de switching do mecanismo 2



- **Etapa 1** - O processador da interface (PLIM) detecta um pacote na mídia de rede e começa a copiá-lo em uma memória FIFO chamada memória de intermitência no LC. A quantidade de memória de intermitência que cada interface tem depende do tipo de LC; as LCs típicas têm de 128 KB a 1 MB de memória de intermitência.
- **Passo 2** - Os primeiros 64 bytes do pacote, chamados de cabeçalho, são transmitidos pelo ASIC de switching de pacotes (PSA).
2.1 O PSA comuta o pacote consultando a tabela do CEF local na memória PSA. Se o pacote não puder ser comutado pelo PSA, vá para a Etapa 4; caso contrário, continue com o Passo 3.
- **Etapa 3** – O RBM (Gerenciador de Buffer de Recepção) aceita o cabeçalho do PSA e o copia em um cabeçalho de buffer livre. Se o pacote for maior que 64 bytes, o fim do pacote também é copiado para o mesmo buffer livre na memória de pacotes e é adicionado à fila de saída virtual da LC de saída. Vá para a Etapa 5.
- **Etapa 4** - O pacote chega a esta etapa se não puder ser comutado pelo PSA. Esses pacotes são colocados na fila bruta (RawQ) e o caminho de switching é essencialmente o mesmo para o Engine 1 e Engine 0 LC a partir deste ponto (Passo 4 no caso do Engine 0). Observe que os pacotes que são comutados pelo PSA nunca são colocados no RawQ e nenhuma

interrupção é enviada para a CPU.

- **Etapa 5** - O FIM (Fabric Interface Module, módulo de interface de estrutura) é responsável por segmentar os pacotes em [células Cisco](#) e enviar as células para o FIA (Fabric Interface ASIC, módulo de interface de estrutura) para transmissão para o LC de saída.

Switching de pacotes de informações Comutação de células em toda a estrutura

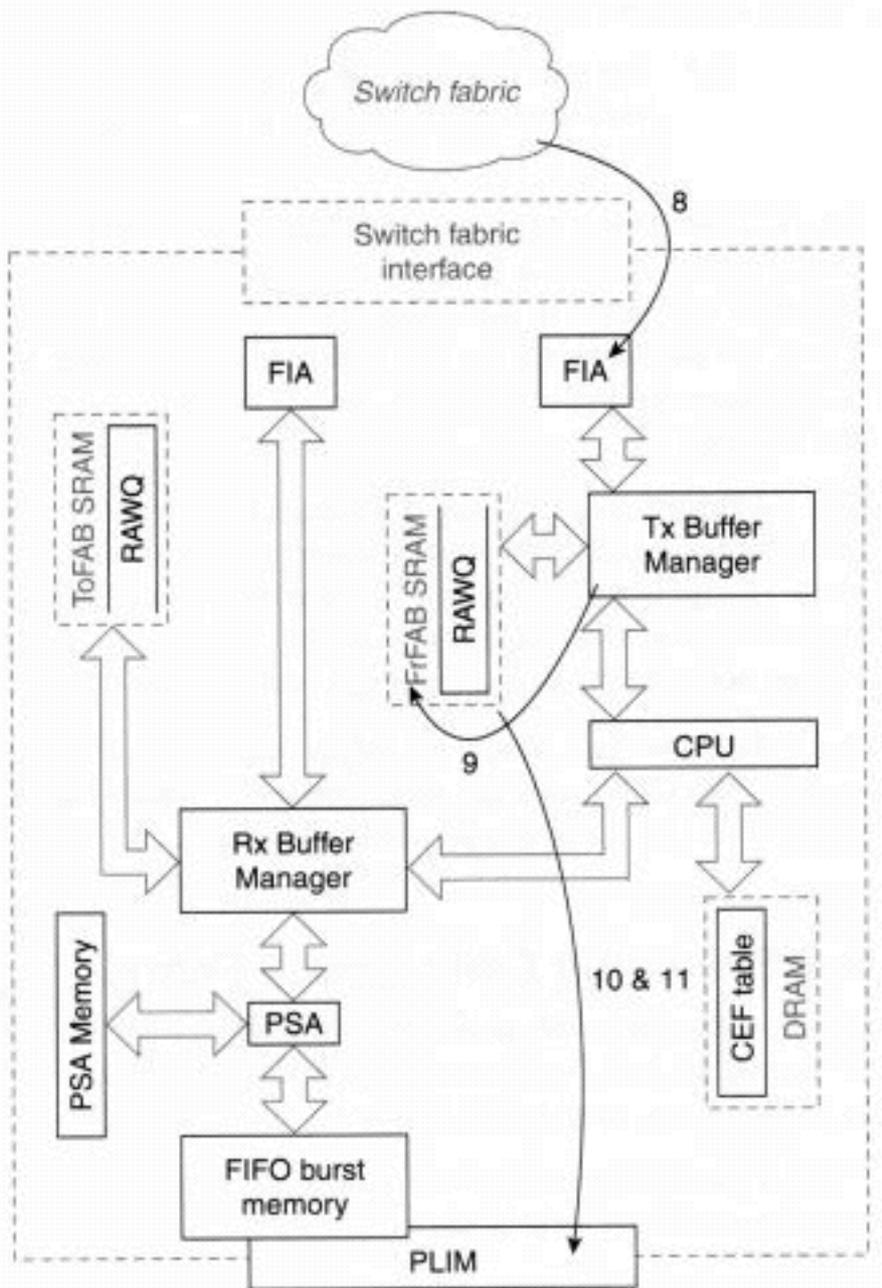
Você chega a este estágio depois que o mecanismo de comutação de pacotes comuta os pacotes. Nesse estágio, os pacotes são segmentados em células Cisco e estão aguardando para serem transmitidos pela tela de switching. As etapas para esse estágio são as seguintes:

- **Step 6** - The FIA sends a grant request to the CSC, which schedules each cell's transfer across the Switch fabric.
- **Etapa 7** - Quando o agendador concede acesso à matriz de comutação, as células são transferidas para o slot de destino. Note que as células podem não ser transmitidas todas de uma só vez; outras células em outros pacotes podem ser intercaladas.

Switching de pacotes de informações Transmitindo pacotes

A Figura 3 abaixo mostra o último estágio da comutação de pacotes. As células são remontadas e o pacote é transmitido na mídia. Isso ocorre na placa de linha de saída.

Figura 3: Comutação de pacotes Cisco 12000: Estágio de transmissão



- Passo 8 - As células comutadas pela tela chegam à placa de linha de destino pelo FIA.
- Passo 9 O gerenciador de buffer de transmissão aloca um buffer da memória do pacote de transmissão e remonta o pacote nesse buffer.
- Passo 10 - Quando o pacote é recriado, o BMA de transmissão coloca o pacote na fila de transmissão da interface de destino no LC. Se a interface transmitir que a fila está cheia (o pacote não pode ser enfileirado), o pacote será descartado e o contador de queda de fila de saída será aumentado. **Observação:** na direção de transmissão, a única vez que os pacotes são colocados no RawQ é quando a CPU do LC precisa fazer qualquer processamento antes da transmissão. Exemplos incluem fragmentação de IP, transmissão múltipla e CAR de saída.
- Passo 11 - O processador de interface detecta um pacote aguardando para ser transmitido, retira o buffer da fila da memória de transmissão, copia-o na memória FIFO interna e transmite o pacote na mídia.

Resumo de fluxo de pacote

Pacotes IP que atravessam o 12000 são processados em três fases:

- Placa de linha de ingresso em três seções: PLIM de entrada (Physical Line Interface Module) - Conversão óptica em elétrica, Rede óptica síncrona (SONET)/SDH (Synchronous Digital Hierarchy), processamento HDLC e PPP. IP Forwarding Decisão de encaminhamento com base na consulta FIB e no enfileiramento em uma das filas de unicast de ingresso ou nas filas de multicast. Gerenciamento de fila de ingresso e interface da tela Processamento da RED (Detecção antecipada aleatória)/WRED (Detecção antecipada aleatória ponderada) nas filas de ingresso e saída em direção à tela para maximizar a utilização da tela.
- Switching de pacotes IP através da tela 12000 da placa de entrada para a placa de saída ou as placas de saída (no caso de multicast).
- Placa de linha de saída em três seções: Interface de estrutura de saída - Remontagem dos pacotes IP a serem enviados e enfileirados em filas de saída; processando pacotes multicast. Gerenciamento da fila de egresso - processamento da RED/WRED nas filas de ingresso e egresso em direção ao PLIM de egresso para maximizar a utilização da linha de egresso. Processamento de saída PLIM - HDLC e PPP, enquadramento SONET/SDH, conversão elétrica para óptica.

[Informações Relacionadas](#)

- [Suporte Técnico - Cisco Systems](#)