

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

Índice

[Introdução](#)

[Pré-requisitos](#)

[Requisitos](#)

[Componentes Utilizados](#)

[Convenções](#)

[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 Pilhas de comutação através da tela](#)

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

[Resumo de fluxo de pacote](#)

[Informações Relacionadas](#)

[Introdução](#)

Este original examina os elementos arquitetônicos os mais importantes do Cisco 12000 Series Internet Router -- pacotes de switching. Pacotes de switching são radicalmente diferentes de qualquer uma das arquiteturas Cisco com base em memória compartilhada ou barramento. Usando uma tela da barra transversal, o Cisco 12000 fornece muito grandes quantidades de largura de banda e de escalabilidade. Além disso, o 12000 usa filas de saída virtual para eliminar o Head of Line Blocking dentro do Switch Fabric.

[Pré-requisitos](#)

[Requisitos](#)

Não existem requisitos específicos para este documento.

[Componentes Utilizados](#)

As informações neste documento são baseadas no seguinte hardware:

- Cisco 12000 Series Internet Router

As informações neste documento foram criadas a partir de dispositivos em um ambiente de laboratório específico. Todos os dispositivos utilizados neste documento foram iniciados com uma configuração (padrão) inicial. Se a sua rede estiver ativa, certifique-se de que entende o impacto

potencial de qualquer comando.

Convenções

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 motores 0, 1, e 2 não são os motores os mais atrasados desenvolvidos por 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 alguma informação útil sobre as placas de linha diferentes disponíveis para o Cisco 12000 Series Router e o motor em que são baseados no [Cisco 12000 Series Internet Router: Perguntas frequentes](#)

Switching de pacotes de informações Overview

Os pacotes sempre são encaminhados pela placa de entrada (LC). A saída LC executa somente o Qualidade de Serviço (QoS) de partida que é fila-dependente (por exemplo, Weighted Random Early Detection (WRED) ou Committed Access Rate (CAR)). A maioria dos pacotes são comutados pelo LC usando o Distributed Cisco Express Forwarding (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.

Este é o que acontece quando um pacote vem em:

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) Moldar L2 é removida (SÃO, Asynchronous Transfer Mode (ATM), Ethernet, High-Level Data Link Control (HDLC)/protocolo Point-to-Point - o PPP) As células ATM são remontadas Pacotes que falham na verificação de redundância cíclica (CRC) são descartados
2. Porque o pacote é recebido e processado, é memória direta alcançada (em uma memória pequena do buffer da unidade de transmissão máxima aproximadamente 2 x (MTU)) chamou "primeiramente dentro, primeiramente para fora a memória de resposta (FIFO)". A quantidade desta memória depende do tipo de LC (de 128 KB a 1 MB).
3. Uma vez que o pacote está completamente na memória de FIFO, uns circuitos integrados do aplicativo específicos (ASIC) no PLIM contactam o Buffer Management ASIC (BMA) e pedem-no um buffer pôr dentro 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 está na tabela de CEF, o pacote está deixado cair. Se o pacote é um pacote de controle (por exemplo, atualizações de roteamento), está enviado à fila à fila do GRP e estará processado pelo GRP. Há 17 filas ToFab (16 unicast mais 1 multicast). Há uma fila do tofab pela placa de linha (este inclui o RP). Estas filas são sabidas como "filas de saída virtuais", e são importantes de modo que a obstrução do início da linha 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: Em cartões do motor 2+, um ASIC especial é usado para melhorar a maneira que 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 pilha do tofab BMA, adiciona um 8-byte CRC (tamanho de célula total - 64 bytes; 44 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. Isto pode parecer como muitas despesas gerais (44 bytes de dados obtêm transformados em 80 bytes através da tela!), mas não é uma edição desde que a capacidade de fábrica foi fornecida em conformidade.
7. Agora que um FIA está pronto para transmitir, o FIA pede o acesso à tela do planejador do cartão e do pulso de disparo atualmente ativos (CSC). O CSC funciona em um algoritmo de clareza bastante complexo. A ideia é que nenhum LC está permitido monopolizar a largura de banda de saída de todo o outro cartão. Note que mesmo se um LC quer transmitir dados fora de uma de suas próprias portas, ainda tem que atravessar a tela. 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 pilhas através do Switch Fabric a seu LC que parte (especificado por dados no encabeçamento de célula Cisco posto lá pelo mecanismo de switching). O algoritmo de imparcialidade é projetado igualmente para a 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. Aquela é a diferença grande entre um Switch Fabric e uma arquitetura de barramento. Pense dela como análogo a um Switch

Ethernet contra um hub; em um interruptor, se a porta A quer enviar para mover B, e no C da porta quer falar para mover D, aqueles dois fluxos acontecem independentemente de se. Em um hub, há umas edições metade-frente e verso tais como colisões e escritório e experimenta de novo algoritmos.

8. As células Cisco que saem da tela atravessam o processamento SLI remover a codificação 8B/10B. Se lá algum erro aqui, ele apareceria no comando show controller fia output como a "paridade de célula". Veja [como ler a saída do comando show controller fia](#) para a informação adicional.
9. Estas células Cisco são DMA'd em FIFO no frfab FIA, e então em um buffer no frfab BMA. O frfab BMA é esse que faz realmente o remontagem de células em um pacote. Como o frfab BMA sabe que buffer para pôr dentro as pilhas antes que as remonte? Esta é uma outra decisão feita pelo mecanismo de switching entrante da placa de linha; desde que todas as filas na caixa inteira são o mesmo tamanho e na mesma ordem, o mecanismo de switching apenas manda o Tx LC pôr o pacote na mesma fila do número de que inscreveu o roteador. As filas do frfab BMA SDRAM podem ser vistas com o comando show controller frfab queue no LC. Veja [como ler a saída do frfab do controlador da mostra | comandos de fila tofab em um Cisco 12000 Series Internet Router](#) para 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. Estes pacotes são colocados na fila da -tela, enviada à fila na fila de interface (há uma fila pela porta física) ou no RawQ para o processamento de emissor. Não muito acontece no RawQ: replicação de multicast por porta, Rodízio de déficit modificado (MDRR) mesma idéia que o Enfileiramento moderado ponderado distribuído (DWFQ) e 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 parcela TX do PLIM esteja pronta para enviar um pacote. O frfab BMA faz a reescrita do MAC real (baseada, recorde, na informação contida no encabeçamento de célula Cisco), e DMA o pacote sobre (outra vez, 2xMTU) a um buffer pequeno nos circuitos de 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 (eplo 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 você quer conhecer o que um GSR sente como no final do dia, lê este papel inteiro 500,000 vezes!

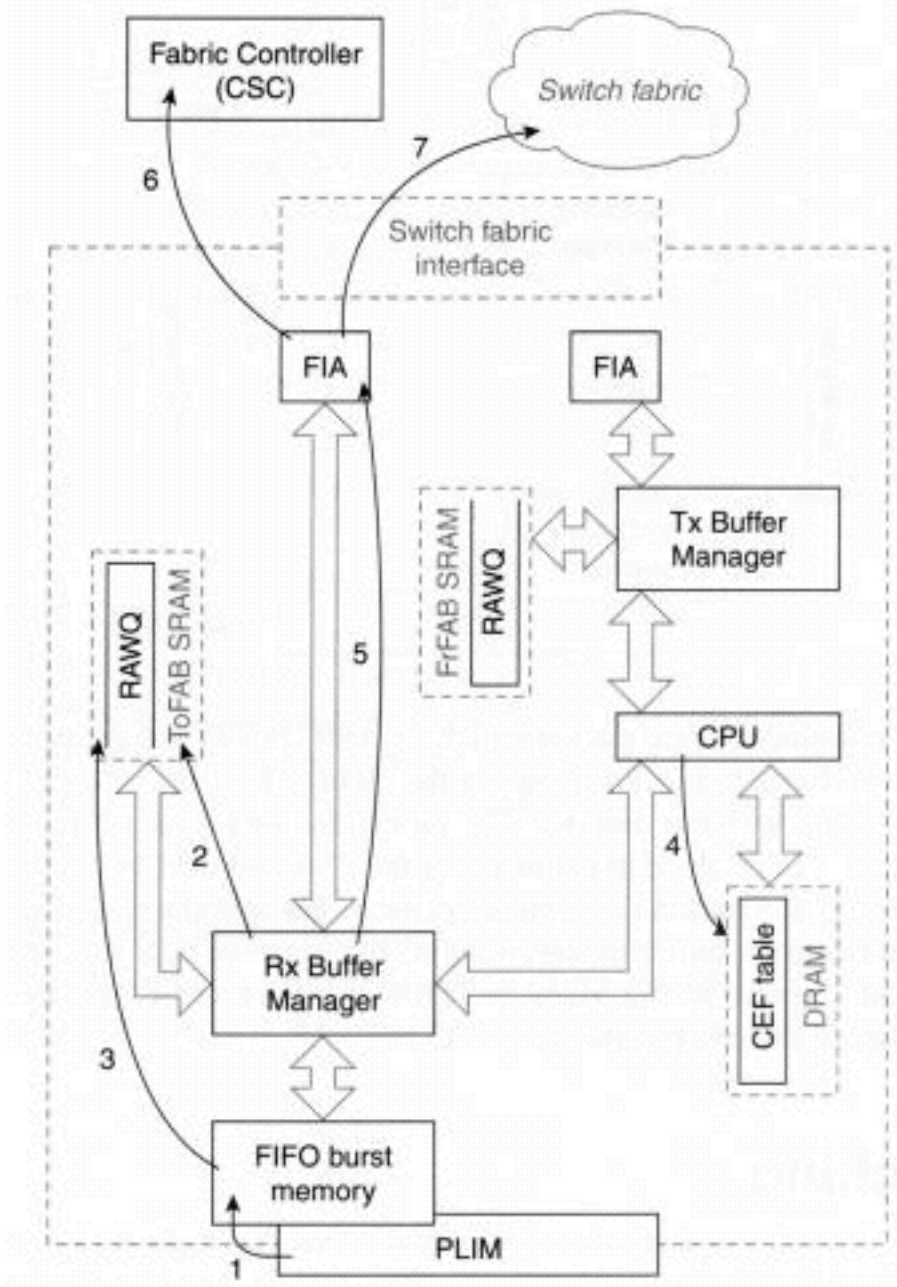
O caminho de switching do pacote no GSR depende do tipo de mecanismo de encaminhamento da LC. Agora, revisaremos todos os passos do Engine 0, Engine 1 e as duas LCs.

[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.

[Figura 1](#) abaixo ilustra as etapas diferentes durante o packet switching para um motor 0 ou o motor 1 LC.

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



O trajeto de switching para o motor 0 e o motor 1 LC é essencialmente o mesmo, embora o motor 1 LC tenha um gerente aumentado do mecanismo de switching e de buffer para o desempenho aumentado. 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 resposta que cada relação tem depende do tipo de LC; as LCs típicas têm de 128 KB a 1 MB de memória de intermitência.
- **Etapa 2** - O processador de interface pede um buffers de pacotes da recepção BMA; o pool de que o buffer é pedido 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 ao LC CPU.

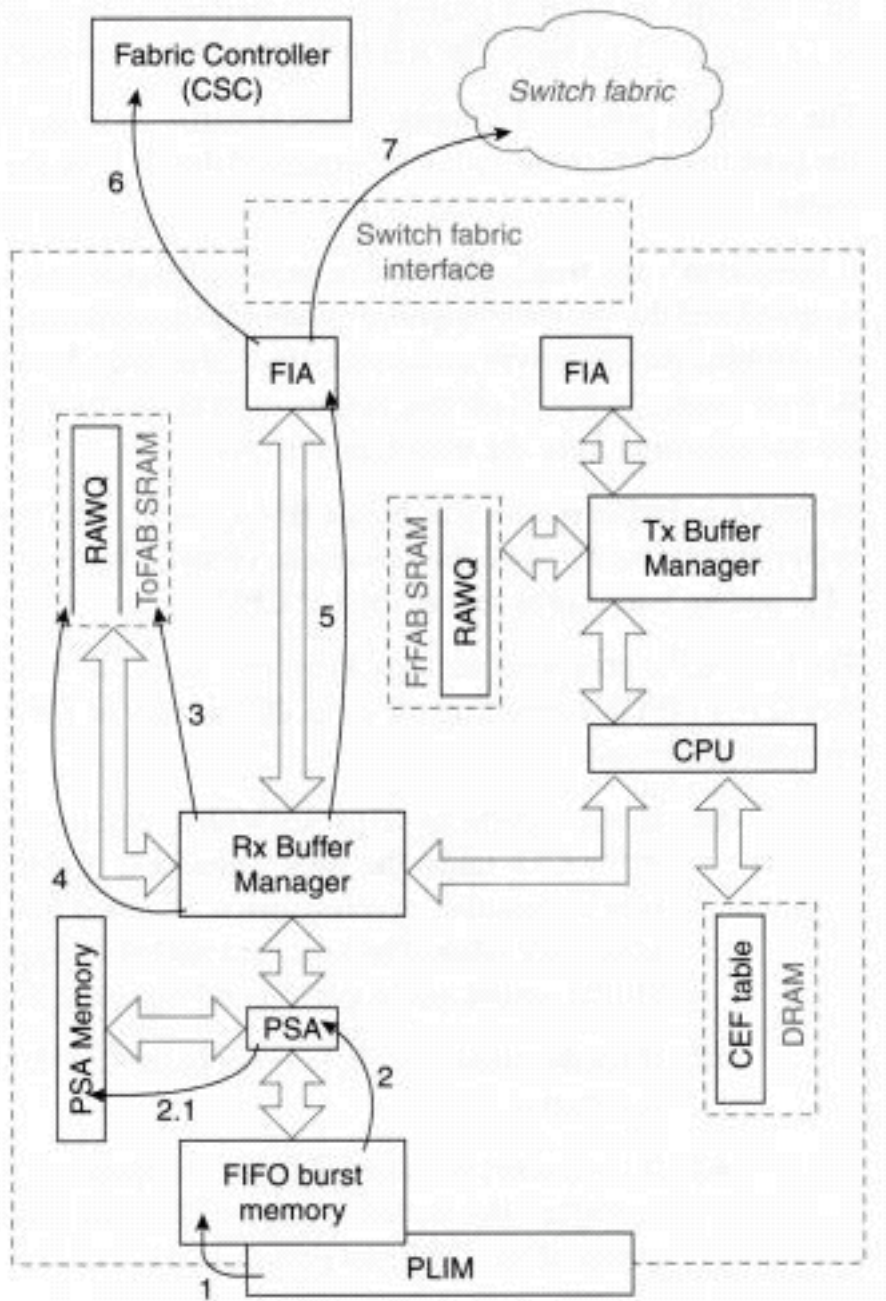
- **Etapa 4** - Os processos de CPU do LC cada pacote no RawQ como é recebido (o RawQ é um FIFO), consultando a tabela local do Distributed Cisco Express Forwarding no DRAM para fazer uma decisão de switching.**4.1** Se este é um pacote IP do unicast com um endereço de destino válido na tabela de CEF, o cabeçalho de pacote de informação está reescrito com a informação de encapsulamento nova obtida da tabela de adjacência de CEF. O pacote comutado é enfileirado na fila de saída virtual peer ao slot de destino.**4.2** Se o endereço de destino não está na tabela de CEF, o pacote está deixado cair.**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 fim da etapa 5, o pacote que chegou em um motor 0/1 LC foi comutado e está pronto para ser transportado através do Switch Fabric como pilhas. Vá pisar 6 no [packet switching da](#) seção: [Pilhas de comutação através da tela](#).

Switching de pacotes de informações Placas de linha do Engine 2

[Figura 2](#) abaixo ilustra o trajeto do packet switching quando os pacotes chegam em um engine 2 LC, como descrito na seguinte lista de etapas.

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 resposta que cada relação 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 pode ser comutado pelo PSA, passe 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 nesta etapa se não pode 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). Note que os pacotes que são comutados pelo PSA estão colocados nunca no RawQ e em nenhuma

interrupção estão enviados ao CPU.

- **Etapa 5** - O módulo de interface de construção (FIM) é responsável para segmentar os pacotes em [células Cisco](#) e enviar as pilhas ao Fabric Interface ASIC (FIA) para a transmissão ao LC de partida.

Switching de pacotes de informações Pilhas de comutação através da tela

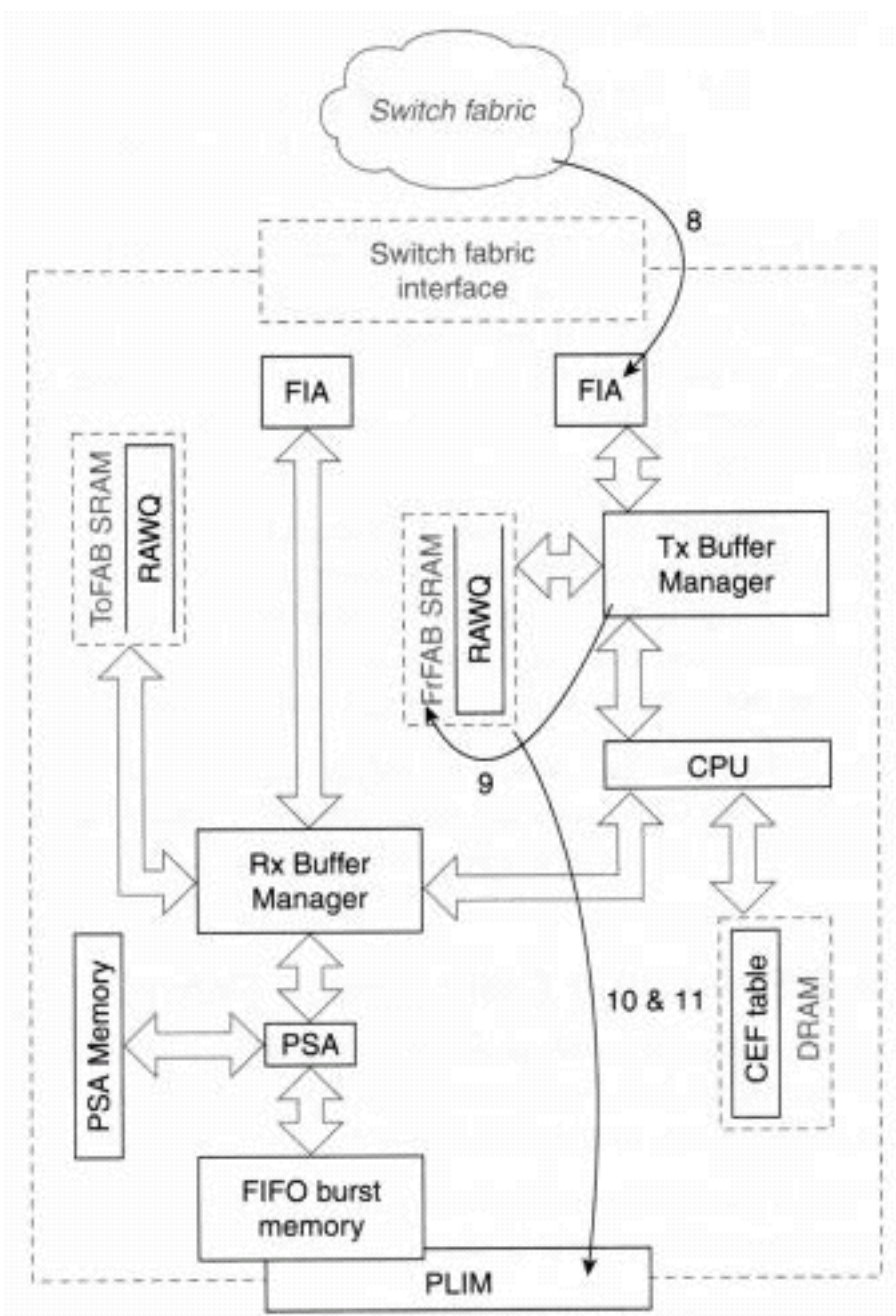
Você chega nesta fase após o motor do packet switching comuta os pacotes. Nesse estágio, os pacotes são segmentados em células Cisco e estão aguardando para serem transmitidos pela tela de switching. As etapas para esse estágio são as seguintes:

- **Etapa 6** - O FIA envia um pedido da concessão ao CSC, que programa transferência de cada pilha através do Switch Fabric.
- **Etapa 7** - Quando o planificador concede o acesso ao Switch Fabric, as pilhas estão transferidas ao slot de destino. Note que as pilhas não puderam ser transmitidas de uma vez; outras pilhas dentro de outros pacotes puderam ser intercaladas.

Switching de pacotes de informações Transmitindo pacotes

Figura 3 abaixo mostra a última fase do packet switching. As células são remontadas e o pacote é transmitido na mídia. Isto ocorre na placa de linha de partida.

Figura 3: Packet switching do 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. **Nota:** Na direção de transmissão, a única vez que os pacotes são colocados no RawQ é quando o a CPU da LC precisa fazer algum 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: Ingresso PLIM (módulo line interface físico) - Ótica à conversão elétrica, à hierarquia digital / Síncrona do Synchronous Optical Network (SONET) (SDH) quemoldam, ao HDLC, e ao processamento de 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 construção da saída - Remontando os pacotes IP para ser enviado e enfileirando-se em filas da saída; processando pacotes de transmissão múltipla. 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](#)