

# Cisco 12000 Series 인터넷 라우터 아키텍처:패킷 스위칭

## 목차

[소개](#)

[사전 요구 사항](#)

[요구 사항](#)

[사용되는 구성 요소](#)

[표기 규칙](#)

[배경 정보](#)

[패킷 스위칭:개요](#)

[패킷 스위칭:엔진 0 및 엔진 1 라인 카드](#)

[패킷 스위칭:엔진 2 라인 카드](#)

[패킷 스위칭:패브릭 간 셀 스위칭](#)

[패킷 스위칭:전송 패킷](#)

[패킷 흐름 요약](#)

[관련 정보](#)

## 소개

이 문서에서는 Cisco 12000 Series Internet Router의 가장 중요한 아키텍처 요소(패킷 스위칭)에 대해 살펴봅니다.스위칭 패킷은 공유 메모리 또는 버스 기반 Cisco 아키텍처와 근본적으로 다릅니다. Cisco 12000은 크로스바 패브릭을 사용하여 매우 많은 양의 대역폭과 확장성을 제공합니다.또한 12000은 가상 출력 대기열을 사용하여 스위치 패브릭 내에서 Head of Line Blocking을 제거합니다.

## 사전 요구 사항

### 요구 사항

이 문서에 대한 특정 요건이 없습니다.

### 사용되는 구성 요소

이 문서의 정보는 다음 하드웨어를 기반으로 합니다.

- Cisco 12000 Series 인터넷 라우터

이 문서의 정보는 특정 랩 환경의 디바이스를 토대로 작성되었습니다.이 문서에 사용된 모든 디바이스는 초기화된(기본) 컨피그레이션으로 시작되었습니다.현재 네트워크가 작동 중인 경우, 모든 명령어의 잠재적인 영향을 미리 숙지하시기 바랍니다.

## [표기 규칙](#)

문서 규칙에 대한 자세한 내용은 [Cisco 기술 팁 표기 규칙](#)을 참조하십시오.

## [배경 정보](#)

(Cisco 12000에 대한 스위칭 결정은 라인 카드(LC)에 의해 이루어집니다. 일부 LC의 경우 전용 ASIC(Application-Specific Integrated Circuit)가 실제로 패킷을 전환합니다.dCEF(Distributed Cisco Express Forwarding)가 유일한 스위칭 방법입니다.

**설명:**엔진 0, 1, 2는 Cisco에서 개발한 최신 엔진이 아닙니다.Engine 3, 4 및 4+ 라인 카드도 있으며, 추가 정보가 제공됩니다.엔진 3 라인 카드는 라인 레이트로 Edge 기능을 수행할 수 있습니다.레이어 3 엔진이 높을수록 하드웨어에서 더 많은 패킷이 스위칭됩니다.Cisco 12000 Series 라우터에 사용할 수 있는 다양한 라인 카드와 [Cisco 12000 Series 인터넷 라우터](#)의 기반 엔진에 대한 유용한 정보를 찾을 수 있습니다.[FAQ](#).

## [패킷 스위칭:개요](#)

패킷은 항상 인그레스 라인 카드(LC)에 의해 전달됩니다. 이그레스 LC는 대기열 종속(예: WRED(Weighted Random Early Detection) 또는 CAR(Committed Access Rate)인 아웃바운드 QoS(Quality of Service)만 수행합니다. 대부분의 패킷은 분산형 Cisco Express Forwarding(dCEF)을 사용하여 LC에 의해 스위칭됩니다. 제어 패킷(예: 라우팅 업데이트)만 처리를 위해 GRP(Gigabit Route Processor)로 전송됩니다.패킷 스위칭 경로는 LC에서 사용되는 스위칭 엔진 유형에 따라 달라집니다.

패킷이 들어올 때 다음과 같은 상황이 발생합니다.

1. 패킷은 PLIM(물리적 레이어 인터페이스 모듈)으로 들어갑니다. 여기에서 다양한 일이 발생합니다.트랜시버는 옵티컬 신호를 전기로 바꿉니다(대부분의 CSR 라인 카드에는 파이버 커넥터가 있음).L2 프레이밍 제거(SANE, ATM(Asynchronous Transfer Mode), 이더넷, HDLC(High-Level Data Link Control)/Point-to-Point Protocol - PPP)ATM 셀 재조합CRC(cyclic redundancy check)에 실패한 패킷은 삭제됩니다.
2. 패킷이 수신 및 처리되면 "First In, First Out (FIFO) burst memory"라는 작은 (약 2 x MTU(Maximum Transmission Unit) 버퍼에 직접 메모리가 액세스됩니다. 이 메모리의 양은 LC의 유형(128KB에서 1MB)에 따라 달라집니다.
3. 패킷이 완전히 FIFO 메모리에 있으면 PLIM의 ASIC(Application-specific integrated circuit)가 BMA(Buffer Management ASIC)에 연결하여 패킷을 넣을 버퍼를 요청합니다.BMA는 패킷의 크기를 알려주고 그에 따라 버퍼를 할당합니다.BMA가 올바른 크기의 버퍼를 가져올 수 없으면 패킷이 삭제되고 수신 인터페이스에서 "ignore" 카운터가 증가합니다.다른 플랫폼과 달리 대체 메커니즘은 없습니다.이 과정이 진행되는 동안 PLIM은 FIFO 버스트 메모리에서 다른 패킷을 수신할 수 있습니다. 따라서 PLIM의 크기는 2xMTU입니다.
4. 오른쪽 대기열에 사용 가능한 버퍼가 있는 경우 해당 패킷은 BMA가 해당 크기의 사용 가능한 대기열 목록에 저장합니다.이 버퍼는 Salsa ASIC 또는 R5K CPU에서 검사하는 Raw Queue에 배치됩니다.R5K CPU는 DRAM(Dynamic RAM)에서 로컬 dCEF 테이블을 참조하여 패킷의 대상을 결정한 다음, Raw Queue에서 대상 슬롯에 해당하는 ToFabric 대기열로 버퍼를 이동합니다.대상이 CEF 테이블에 없으면 패킷이 삭제됩니다.패킷이 제어 패킷(예: 라우팅 업데이트)인 경우 GRP의 대기열에 추가되고 GRP에서 처리됩니다.17개의 ToFab 큐(16개의 유니캐스트 및 1개의 멀티캐스트)가 있습니다. 라인 카드당 하나의 toFab 대기열이 있습니다(여기에는

RP가 포함됨). 이러한 대기열은 "가상 출력 대기열"이라고 하며 OOB(Head-of-Line) 차단이 발생하지 않도록 해야 합니다.

5. ToFab BMA는 패킷을 44바이트 조각으로 자릅니다. 이는 "Cisco Cells"로 알려진 것을 위한 페이로드입니다. 이러한 셀에는 frFab BMA(현재까지 총 데이터 크기 = 56바이트)에 의해 8바이트 헤더와 4바이트 버퍼 헤더가 지정된 다음 적절한 ToFab 큐에 대기됩니다. 이때 버퍼가 가져온 풀의 #Qelem 카운터가 1씩 다운되고 ToFab 큐 카운터가 1씩 증가합니다."의사 결정권자"는 스위칭 엔진의 유형에 따라 달라집니다.Engine 2+ 카드에서는 패킷의 스위치 방식을 개선하기 위해 특수 ASIC를 사용합니다.일반 패킷(IP/Tag, 옵션 없음, 체크섬)은 PSA(Packet Switching ASIC)에 의해 직접 처리된 다음 원시 큐/CPU/Salsa 조합을 건너뛰고 toFab 대기열에 직접 추가됩니다.패킷의 처음 64바이트만 패킷 스위칭 ASIC를 통해 전달됩니다.PSA로 패킷을 전환할 수 없는 경우, 패킷이 RawQ로 대기되어 이전에 설명한 대로 LC의 CPU에서 처리됩니다.이 시점에서 스위칭 결정이 이루어졌고 패킷이 적절한 ToFab 출력 대기열에 넣었습니다.
6. toFab BMA DMA(Direct Memory Access)는 FIA(패브릭 인터페이스 ASIC)에서 패킷의 셀을 소형 FIFO 버퍼로 변환합니다. 17개의 FIFO 버퍼가 있습니다(ToFab 대기열당 하나). FIA가 toFab BMA에서 셀을 가져오면 8바이트 CRC(총 셀 크기 - 64바이트)를 추가합니다.44바이트 페이로드, 8바이트 셀 헤더, 4바이트 버퍼 헤더). FIA에는 SLI(Serial Line Interface) ASIC가 있으므로 셀에서 8B/10B 인코딩을 수행하고(예: FDDI(Fiber Distributed Data Interface) 4B/5B) 패브릭을 통해 전송할 준비를 합니다.이는 많은 오버헤드(44바이트의 데이터가 패브릭에서 80바이트로 전환됨!)로 보일 수 있지만, 패브릭 용량이 적절하게 프로비저닝되었기 때문에 문제가 되지 않습니다.
7. 이제 FIA가 전송될 준비가 되었으므로 FIA는 현재 활성 카드 스케줄러와 클럭(CSC)에서 패브릭에 대한 액세스를 요청합니다. CSC는 다소 복잡한 공정 알고리즘에서 작동합니다.LC는 다른 카드의 발신 대역폭을 독점할 수 없다는 개념입니다.LC가 자체 포트 중 하나에서 데이터를 전송하려는 경우에도 패브릭을 통과해야 합니다.이러한 문제가 발생하지 않을 경우 LC의 한 포트가 동일한 LC에 있는 지정된 포트에 대한 모든 대역폭을 독점할 수 있으므로 중요합니다.스위칭 설계도 더욱 복잡해질 것입니다.FIA는 스위치 패브릭을 통해 셀을 발신 LC로 전송합니다(스위칭 엔진에서 제공한 Cisco Cell 헤더의 데이터에 의해 지정됨).균등성 알고리즘은 최적의 매칭을 위해 설계되었습니다.카드 1이 카드 2로 전송하고 카드 3이 동시에 카드 4로 전송하려는 경우, 이는 병렬로 이루어집니다.스위치 패브릭과 버스 아키텍처의 큰 차이점입니다.이를 이더넷 스위치와 허브의 유사하다고 생각하십시오.스위치에서 포트 A가 포트 B로 전송하고 포트 C가 포트 D와 통신하기를 원할 경우, 이러한 두 플로우는 서로 독립적으로 발생합니다.허브에 충돌, 백오프 및 재시도 알고리즘과 같은 반이중 문제가 있습니다.
8. 패브릭에서 나오는 Cisco 셀은 SLI 처리를 통해 8B/10B 인코딩을 제거합니다.여기에 오류가 있으면 show controller fia 명령 출력에 "cell parity"로 표시됩니다. 자세한 내용은 [show controller fia 명령의 출력을 읽는 방법](#)을 참조하십시오.
9. 이러한 Cisco 셀은 frFab FIA의 FIFO에 DMA를 추가한 다음 frFab BMA의 버퍼에 들어갑니다.frFab BMA는 실제로 셀을 패킷으로 재결합하는 것입니다.frFab BMA는 셀을 리어셈블하기 전에 어떤 버퍼를 넣을 것인지 어떻게 알 수 있습니까?이는 수신 라인 카드 스위칭 엔진에서 결정한 또 다른 결정입니다.전체 상자의 모든 대기열은 크기가 같고 순서가 같기 때문에 스위칭 엔진은 Tx LC가 패킷을 라우터에 입력한 것과 동일한 번호 대기열에 넣기만 하면 됩니다.frFab BMA SDRAM 대기열은 LC에서 show controller frb queue 명령을 사용하여 볼 수 있습니다.show [controller frb의 출력을 읽는 방법 참조](#) 자세한 내용은 [Cisco 12000 Series Internet Router의 tofab queue 명령](#)이것은 기본적으로 toFab BMA 결과와 동일합니다.패킷이 수신되고 해당 여유 대기열에서 제거된 패킷에 배치됩니다.이러한 패킷은 from-fabric 대기열에 배치되며, 인터페이스 대기열(물리적 포트당 하나의 대기열)에 대기열에 추가되거나 출력 처리를 위한 rawQ에 추가됩니다.rawQ에서는 별로 일어나지 않습니다.포트별 멀티캐스트 복제, MDRR(Modified Defense Round Robin) - DWFQ(Distributed Weighted Fair Queuing) 및 출력

CAR과 동일한 개념전송 큐가 가득 차면 패킷이 삭제되고 출력 드롭 카운터가 증가합니다.

10. frFab BMA는 PLIM의 TX 부분이 패킷을 전송할 준비가 될 때까지 기다립니다.frFab BMA는 실제 MAC 재작성(Cisco Cell 헤더에 포함된 정보를 기반으로, 기억)을 수행하고, DMA는 PLIM 회로의 작은(다시 2xMTU) 버퍼로 패킷을 전송합니다.PLIM은 ATM SAR 및 SONET가 캡슐화를 수행하고, 적절하게 패킷을 전송합니다.

11. ATM 트래픽은 SAR에 의해 리어셈블되고(tofab BMA에 의해) 세그멘테이션되고(fromfab BMA에 의해) 리어셈블되고 다시 세그멘테이션됩니다(fromfab SAR에 의해). 이는 매우 신속하게 이루어집니다.

패킷의 라이프사이클은 처음부터 끝까지입니다.GSR이 어떤 기분인지 알고 싶다면, 이 50만 번 신문을 읽으세요!

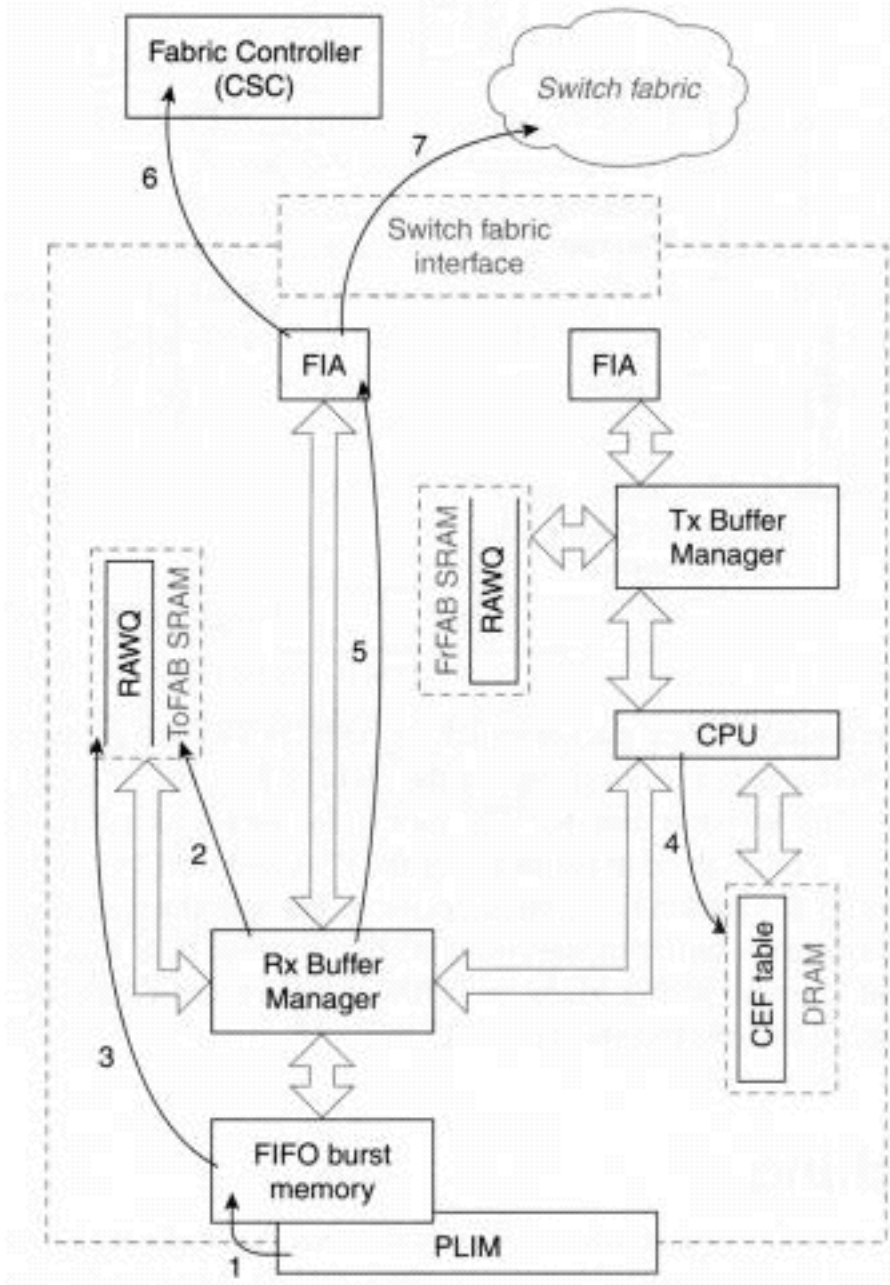
GSR의 패킷 스위칭 경로는 LC의 포워딩 엔진 유형에 따라 달라집니다.이제 엔진 0, 엔진 1 및 두 LC의 모든 단계를 살펴보겠습니다.

## 패킷 스위칭:엔진 0 및 엔진 1 라인 카드

아래 섹션은 *Inside Cisco IOS Software Architecture*, Cisco Press 책을 기반으로 합니다.

아래 [그림 1](#)은 Engine 0 또는 Engine 1 LC에 대한 패킷 스위칭 과정에서 다양한 단계를 보여줍니다

그림 1:엔진 0 및 엔진 1 스위칭 경로



Engine 1 LC는 향상된 스위칭 엔진 및 버퍼 매니저를 통해 성능을 향상시키지만 Engine 0 및 Engine 1 LC의 스위칭 경로는 기본적으로 동일합니다. 스위칭 경로는 다음과 같습니다.

- **1단계** - PLIM(Interface Processor)은 네트워크 미디어에서 패킷을 탐지하여 LC의 버스트 메모리라는 **FIFO 메모리**에 복사하기 시작합니다. 각 인터페이스의 버스트 메모리 양은 LC 유형에 따라 달라집니다. 일반적인 LC에는 128KB~1MB의 버스트 메모리가 있습니다.
- **2단계** - 인터페이스 프로세서가 수신 BMA에서 패킷 버퍼를 요청합니다. 버퍼가 요청되는 풀은 패킷의 길이에 따라 달라집니다. 사용 가능한 버퍼가 없으면 인터페이스가 삭제되고 인터페이스의 "ignore" 카운터가 증가합니다. 예를 들어 64바이트 패킷이 인터페이스에 도착하면 BMA는 80바이트 패킷 버퍼를 할당하려고 시도합니다. 80바이트 풀에 사용 가능한 버퍼가 없는 경우 사용 가능한 다음 풀에서 버퍼가 할당되지 않습니다.
- **3단계** - BMA에서 사용 가능한 버퍼를 할당하면 패킷이 버퍼에 복사되고 CPU에서 처리하기 위해 **원시 대기열(RawQ)**에 대기됩니다. 인터럽트가 LC CPU로 전송됩니다.
- **4단계** - LC의 CPU는 RawQ를 수신하면(RawQ는 FIFO임) RawQ에서 각 패킷을 처리하여 DRAM의 로컬 분산 Cisco Express Forwarding 테이블을 참조하여 스위칭 결정을 내립니다. **4.1** CEF 테이블에 유효한 대상 주소가 있는 유니캐스트 IP 패킷인 경우 패킷 헤더가 CEF 인접성 테이블에서 얻은 새 캡슐화 정보로 재작성됩니다. 스위치드 패킷은 대상 슬롯에 해당하는 가상

출력 대기열에서 대기열에 추가됩니다. 4.2 대상 주소가 CEF 테이블에 없으면 패킷이 삭제됩니다. 4.3 패킷이 제어 패킷(예: 라우팅 업데이트)인 경우 패킷은 GRP의 가상 출력 대기열에 삽입되고 GRP에서 처리됩니다.

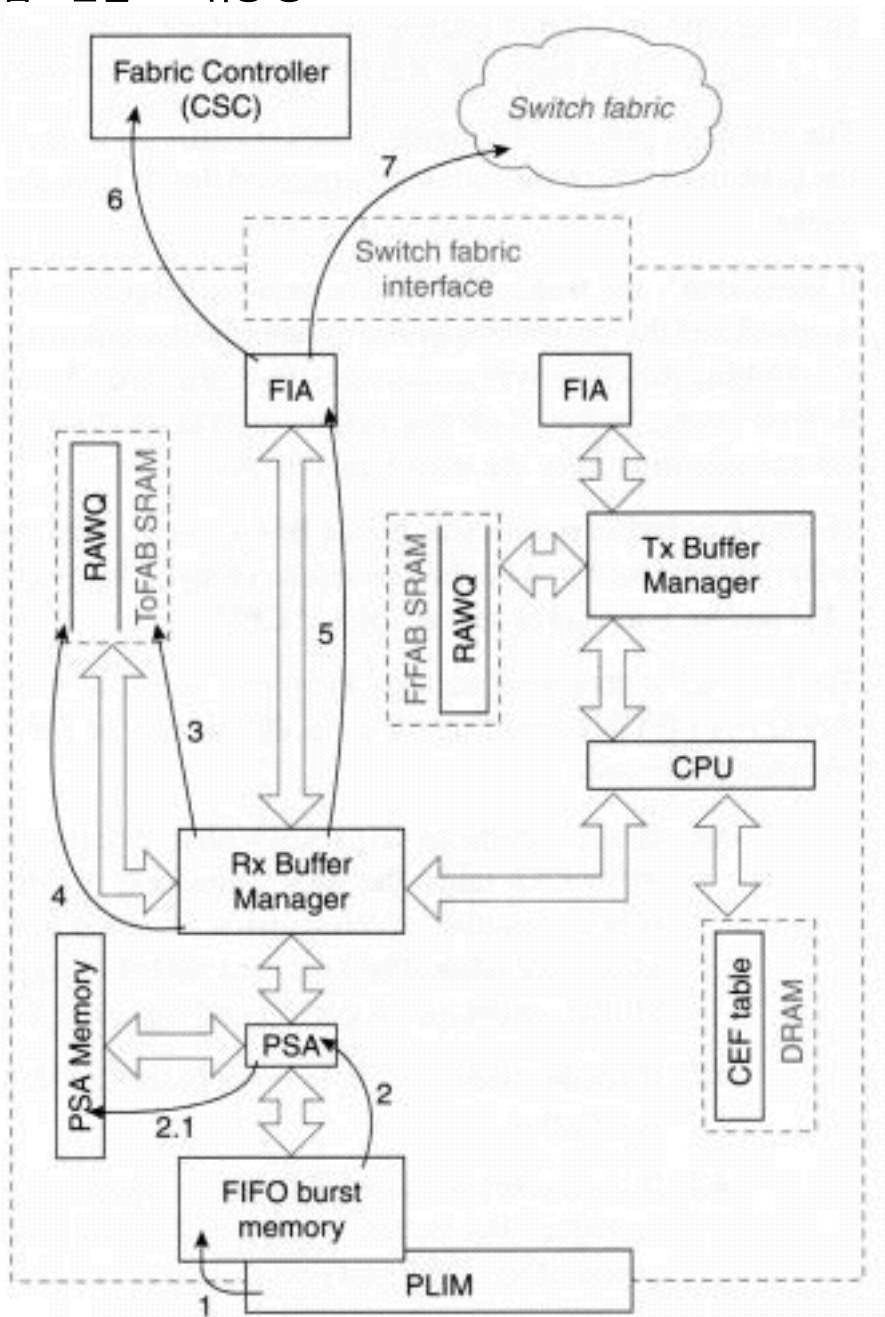
- 5단계 - 수신 BMA는 패킷을 64바이트 셀로 프래그먼트한 다음 아웃바운드 LC로 전송하기 위해 FIA에 전달합니다.

5단계 끝에 Engine 0/1 LC에 도착한 패킷이 스위칭되었으며 스위치 패브릭을 통해 셀처럼 전송할 준비가 되었습니다. [Packet Switching](#) (패킷 스위칭) 섹션의 6단계로 [이동합니다. Fabric을 통해 셀을 스위칭합니다.](#)

## 패킷 스위칭:엔진 2 라인 카드

아래 [그림 2](#)는 다음 단계 목록에 설명된 대로 패킷이 Engine 2 LC에 도착할 때의 패킷 스위칭 경로를 보여줍니다.

그림 2:엔진 2 스위칭 경로



- 1단계 - PLIM(Interface Processor)은 네트워크 미디어에서 패킷을 탐지하여 LC의 버스트 메모

리라는 FIFO 메모리에 복사하기 시작합니다.각 인터페이스의 버스트 메모리 양은 LC 유형에 따라 달라집니다.일반적인 LC에는 128KB~1MB의 버스트 메모리가 있습니다.

- **2단계** - 헤더라는 패킷의 처음 64바이트가 PSA(Packet Switching ASIC)를 통해 전달됩니다.  
.2.1 PSA는 PSA 메모리의 로컬 CEF 테이블을 참조하여 패킷을 전환합니다.PSA로 패킷을 전환할 수 없는 경우 4단계로 이동합니다.그렇지 않으면 3단계로 진행합니다.
- **3단계** - RBM(Receive Buffer Manager)에서 PSA의 헤더를 수락하고 이를 무료 버퍼 헤더에 복사합니다.패킷이 64바이트보다 크면 패킷 메모리의 동일한 사용 가능한 버퍼에 패킷의 꼬리도 복사되고 나가는 LC [가상 출력 대기열](#)에서 대기됩니다.5단계로 이동합니다.
- **4단계** - PSA로 전환할 수 없는 경우 패킷이 이 단계에 도달합니다.이러한 패킷은 **원시 대기열(RawQ)**에 배치되며 스위칭 경로는 기본적으로 이 지점에서 엔진 1 및 엔진 0 LC와 동일합니다(엔진 0의 경우 4단계). PSA로 전환되는 패킷은 RawQ에 배치되지 않으며 인터럽트가 CPU로 전송되지 않습니다.
- **5단계** - FIM(Fabric Interface Module)은 패킷을 [Cisco Cells](#)로 분할하고 셀을 FIA(Fabric Interface ASIC)로 보내 아웃바운드 LC로 전송합니다.

## 패킷 스위칭:패브릭 간 셀 스위칭

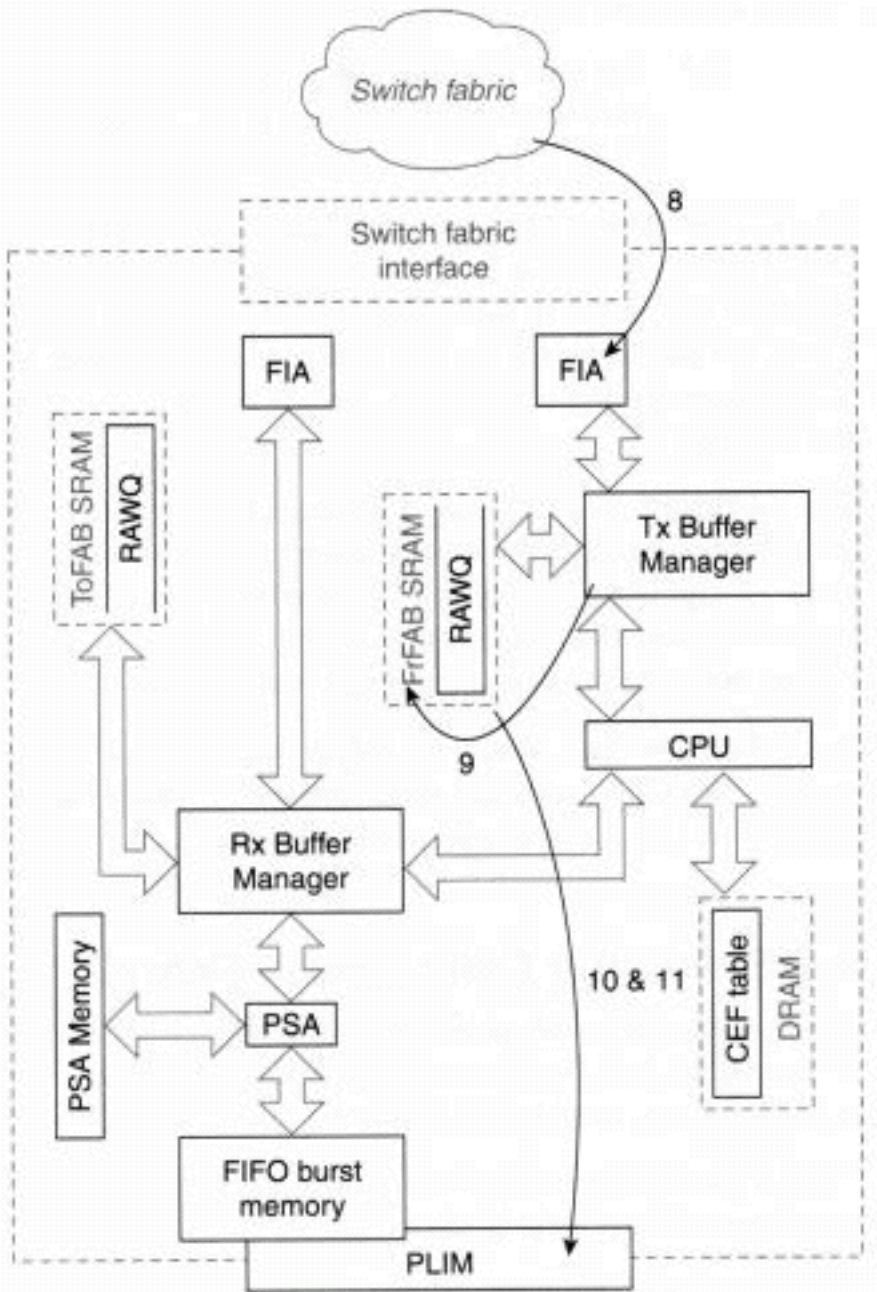
패킷 스위칭 엔진이 패킷을 전환한 후 이 단계에 도달합니다.이 단계에서는 패킷이 Cisco Cells로 분할되며 스위칭 패브릭 간에 전송되기를 기다리고 있습니다.이 단계의 단계는 다음과 같습니다.

- **6단계** - FIA가 CSC에 보조금 요청을 보냅니다. CSC는 각 셀이 스위치 패브릭에서 전송되도록 예약합니다.
- **7단계** - 스케줄러가 스위치 패브릭에 대한 액세스 권한을 부여하면 셀이 대상 슬롯으로 전송됩니다.셀이 동시에 전송되지 않을 수도 있습니다.다른 패킷 내의 다른 셀은 인터리빙될 수 있습니다.

## 패킷 스위칭:전송 패킷

아래 그림 3은 패킷 스위칭의 마지막 단계를 보여줍니다.셀이 리어셈블되고 패킷이 미디어에 전송됩니다.이는 아웃바운드 라인 카드에서 발생합니다.

그림 3: Cisco 12000 패킷 스위칭:전송 단계



- 8단계 - 패브릭을 통해 전환된 셀이 FIA를 통해 목적지 라인 카드에 도착합니다.
- 9단계 - 전송 버퍼 관리자는 전송 패킷 메모리에서 버퍼를 할당하고 이 버퍼에서 패킷을 리어셈블합니다.
- 10단계 - 패킷이 재구축되면 전송 BMA는 패킷을 LC의 대상 인터페이스의 전송 대기열에 넣습니다. 인터페이스 전송 큐가 가득 차면(패킷을 대기열에 넣을 수 없음) 패킷이 삭제되고 출력 대기열 삭제 카운터가 증가합니다. 참고: 전송 방향에서 LC CPU가 전송하기 전에 처리를 수행해야 하는 경우에만 RawQ에 패킷이 배치됩니다. 예를 들면 IP 프래그먼트화, 멀티캐스트 및 출력 CAR이 있습니다.
- 11단계 - 인터페이스 프로세서는 전송 대기 중인 패킷을 탐지하고, 전송 메모리에서 버퍼를 대기시키고, 내부 FIFO 메모리에 복사한 다음, 미디어에 패킷을 전송합니다.

## 패킷 흐름 요약

12000을 통과하는 IP 패킷은 다음 3단계로 처리됩니다.

- 인그레스 라인 카드(3개 섹션): Ingress PLIM(Physical Line Interface Module) - 옵티컬-전기 변



환, SONET(Synchronous Optical Network)/SDH(Synchronous Digital Hierarchy) 비프레이밍, HDLC 및 PPP 처리.IP 포워딩 - FIB 조회 및 이그레스 유니캐스트 대기열 또는 멀티캐스트 대기열 중 하나로 대기하는 것에 따른 포워딩 결정인그레스 큐 관리 및 패브릭 인터페이스 - 이그레스 큐에서 RED(Random Early Detection)/WRED(Weighted Random Early Detection) 프로세싱을 처리하고 패브릭의 활용도를 극대화하기 위해 패브릭을 향해 대기를 해제합니다.

- 이그레스 카드에서 이그레스 카드 또는 이그레스 카드(멀티캐스트의 경우)로 12000 패브릭을 통해 IP 패킷을 스위칭합니다.
- 이그레스 라인 카드(3개 섹션):Egress Fabric Interface - 전송할 IP 패킷을 리어셈블하고 이그레스 큐로 큐잉합니다.멀티캐스트 패킷을 처리하는 중입니다.이그레스 대기열 관리 - 이그레스 대기열에서 RED/WRED 처리 및 이그레스 PLIM으로 대기열에서 빠져 이그레스 라인 사용률을 극대화합니다.이그레스(egress) PLIM - HDLC 및 PPP 프로세싱, SONET/SDH 프레이밍, Electrical to Optical 변환

## 관련 정보

- [Technical Support - Cisco Systems](#)