

10Gbps 라인속도 지원을 위한 개방형 플랫폼 구성 방법

박준석, 권경인, 김영준
LG전자 정보통신사업본부 시스템연구소

An Approach to Open Platform Architecture Design for 10Gbps Wire-speed

Joon-Seok Park, Kyoung-In Kwon, and Young-Joon Kim
LG Electronics

Abstract - 광대역통합망 환경에서 에지 라우터에 요구되는 여러 가지 기능 및 성능항목들 중에서 차세대 개방형 시스템 플랫폼을 적용하면서도 10Gbps 라인속도를 제공할 수 있는 방안을 연구하고 그에 따른 상위수준의 시스템 구조를 제안한다. 10Gbps 인터페이스에 대해서 라인속도를 제공하기 위해서는 스위치 카드와 라인카드간의 물리적 연결에 최소한 10Gbps의 두배에 해당하는 20Gbps 대역폭 할당이 필요하며 이를 위해서 개방형 플랫폼의 패브릭 인터페이스에서는 라인카드당 8개의 연결 링크를 사용하는 듀얼-듀얼 성형 토폴로지를 사용하여야 한다.

1. 서 론

통신과 방송 인터넷이 융합된 품질보장형 광대역 멀티미디어 서비스를 언제 어디서나 끊임없이 안전하게 이용할 수 있는 차세대 통합네트워크를 광대역통합망이라 통칭한다[1]. 광대역통합망은 기존의 다양한 액세스망을 융합하고 품질보장이 지원되는 새로운 프리미엄망 구축을 통하여 가입자별 또는 응용 별로 차별화된 서비스를 제공할 수 있는 새로운 개념의 차세대 망이라고 할 수 있다.

광대역통합망에서는 데이터 서비스 뿐만 아니라 음성과 영상 그리고 방송을 포함한 다양한 서비스들이 단일화된 패킷 전달망을 통하여 각각의 특성과 요구사항에 맞게 처리 될 수 있어야 하며 이를 위해서는 시스템이 다기능 및 고성능을 제공하여야 한다.

또한 폭발적으로 늘어나는 정보통신 단말과 홈네트워크로 대변되는 정보가전의 발전으로 인하여 통신에 참여하는 개체수가 기하급수적으로 증가함에 따라 이들을 수용할 수 있는 새로운 주소체계인 IPv6의 도입이 필수적이라고 할 수 있다.

그리고 망이 통합되면서 통신 사업자들은 기존의 다양한 액세스망에 품질보장형 망을 추가로 구축하여 효율적으로 운영하기를 원하고 있으며 이를 위해서는 표준화된 개방형 시스템 플랫폼의 사용이 요구되고 있다

본 논문은 광대역통합망 환경의 전달망 요소인 에지 라우터에 요구되는 많은 기능 및 성능항목들 중에서 160Gbps 이상의 시스템 용량을 제공하면서 개방형 시스템 플랫폼을 적용할 수 있는 방법과 10Gbps 라인속도를 구현하기 위한 방안을 도출하고 그에 따른 상위수준의 시스템 구조를 제안한다.

광대역통합망에 적합한 다기능 IPv6 에지 라우터의 구조에 관한 본 논문의 구성은 다음과 같다. 2장에서는 광대역통합망 환경의 전달망 장비가 갖추어야 할 필수 기능 및 구조를 광대역통합망의 요구사항을 통해서 살펴보고 3장에서는 시스템의 핵심 칩인 스위치 패브릭과 네트워크 프로세서에 요구되는 성능을 예측한다.

4장에서는 3장에서 분석된 칩셋의 성능 요구 조건을 바탕으로 광대역통합망을 위한 다기능 에지 라우터의 상위수준 시스템 구조를 설명하고 5장에서 결론을 맺는다.

2. 본 론

2.1 광대역통합망의 일반적 요구사항

유무선을 포함하여 통신과 방송이 융합되는 광대역통합망 환경에서는 다양한 측면에서 여러 가지 요구사항이 있을 수 있다. 먼저 사용자는 편리하면서도 사용하기에 익숙한 고품질 통신을 저렴한 서비스 이용요금으로 제공 받기를 원하고 있으며 통신사업자는 고객의 요구사항을 만족시키기 위해 망의 기능향상과 효율적인 망운용을 통하여 이를 충족시키고자 한다.

정보통신부에서 발표한 광대역통합망 구축 기본계획에 따르면 그림 1과 같이 광대역통합망에서 전달망 계층의 요구사항을 크게 품질보장(QoS), 보안(security), IPv6, 그리고 개방형 API의 네 가지 범주로 나누고 있다[1]. 본 장에서는 광대역통합망에서 요구되는 주요 기능들을 간략히 살펴봄으로써 다기능 에지 라우터의 요구사항을 도출하고 이를 시스템 상위구조설계에 반영하고자 한다.

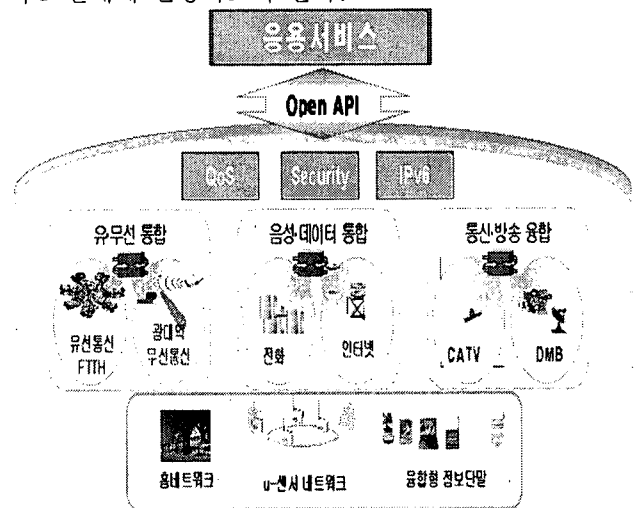


그림 1. 광대역통합망 기본 개념도

2.1.1 품질보장 요구사항

광대역통합망에서 트래픽 엔지니어링 기술과 이를 통한 서비스 품질보장 기능은 인터넷망을 포함하여 기존의 데이터망과 차별화시킬 수 있는 가장 중요한 부분이다. 특히 기존의 회선망을 이용한 음성서비스가 광대역통합망에서는 단일화된 패킷망을 통하여 서비스

스되기 때문에 이들을 서비스하고 전달할 전달망 장비의 품질보장 기능은 더욱 중요하게 된다.

따라서 다기능 에지 라우터와 같은 광대역통합망 환경의 전달망 장비에서는 장비로 유입되는 트래픽을 가입자별, 또는 응용별로 분류하고 각 서비스의 특성에 맞게 처리함으로써 품질보장 요구사항을 만족시킬 수 있어야 한다. 또한 광대역통합망 환경의 다기능 에지 라우터는 DiffServ(2)의 기능은 물론 시스템 내에서 64,000개 이상의 트래픽 플로우를 지원하고 향후 최대 1,000,000개의 트래픽 플로우를 수용할 수 있도록 설계함으로써 다양한 수준의 차등화 서비스 및 품질보장 서비스를 지원해야 한다.

2.1.2 보안 요구사항

광대역통합망의 신뢰성과 안전성을 위해서는 그 구성요소인 전달망 장비에서 보안기능을 제공하여야 한다. 대표적인 보안방법으로는 장비로 유입되는 트래픽에 대해서 지속적으로 감시 함으로써 이상 트래픽을 감지하고 그에 따라 적절하게 대응하는 것이다.

광대역통합망에서는 네트워크 보안에 관해 3단계로 계획하고 있다[1]. 1단계에서는 과다 트래픽의 발생을 감시할 수 있는 보안 모니터링 체계를 구축하고 2단계에서는 유해 트래픽 차단 등의 침해 대응 체계를 구축하며 마지막으로 3단계에서 비정상 트래픽 제어 등 능동보안체계를 구축 하는 것으로 규정하고 있다.

다기능 에지 라우터는 3계층 또는 그 이상의 계위에서 패킷의 헤더를 분석하고 패킷 감시 및 필터링 기능을 수행함으로써 비정상적인 트래픽 흐름에 대해서 고속으로 감지, 종합분석하여 실시간으로 대응해야 하며 IPv4에 대해서는 8,000개, IPv6에 대해서는 2,000개의 ACL(access control list)을 지원 함으로써 보안요구사항을 만족시켜야 한다.

2.1.3 IPv6 요구사항

인터넷 호스트, 이동전화(IMT-2000, WCDMA, PCS 단말 등), 스마트 정보가전(인터넷 TV, 인터넷 냉장고, 등) 등의 증가로 IPv4 주소로는 급증하는 주소자원의 수요를 충족시킬 수 없으므로 IPv6 주소체계의 도입이 절실히 요구된다. 특히 광대역통합망을 통하여 유비쿼터스 환경을 실현하기 위해서는 망 요소 장비에서 새로운 주소체계인 IPv6의 지원이 무엇보다도 중요하다.

광대역통합망 구축계획에서는 IPv6의 도입을 4단계로 나누어 추진한다[1]. 1단계에서는 일부 단말 및 가입자망에 IPv4/IPv6 듀얼스택을 적용하고 IPv4 전달망을 터널링하여 격리된 IPv6망을 연동한다. 2단계는 IPv6를 가입자망에 확대적용하고 더불어 전달망에 IPv4/IPv6 듀얼 스택을 도입함으로써 IPv6망간 연동을 제공하는 것이다.

IPv6기능이 필수 요구사항인 광대역통합망에서 다기능 에지 라우터는 IPv4/IPv6 라우팅 및 패킷 전달 기능을 수행하게 된다.

2.1.4 개방형 플랫폼 요구사항

시스템에 표준화된 개방형 하드웨어 플랫폼을 적용함으로써 서로 다른 시스템 제조사간 또는 서로 다른 기능 모듈간에 원활한 상호연동성을 제공할 수 있다. 시스템 제조사로서는 시스템의 모든 기능 모듈을 독자 개발하지 않아도 되기 때문에 시스템의 개발기간을 단축하고 특정 기능 모듈에 개발역량을 집중함으로써 시

스템의 안정성 및 신뢰성을 확보함과 동시에 개발 비용을 절감할 수 있다. 뿐만 아니라 적기에 시장진입을 시도할 수 있는 장점을 가지게 된다. 또한 통신사업자로서는 기능 모듈별로 제조사를 달리하여 시스템을 구축할 수 있기 때문에 시스템 전체를 특정 제조사에 의존하게 되는 비합리성을 제거하고 표준화된 단일 플랫폼을 사용할 수 있어 지속적인 시스템의 업그레이드 및 운용의 효율성을 취할 수 있다.

다기능 에지 라우터는 시스템 플랫폼에 대한 광대역 통합망의 요구사항을 수용하여 차세대 개방형 표준 플랫폼 구조를 가질 수 있도록 설계해야 한다.

2.2 주요 칩의 성능 요구사항

2.2절에서는 광대역통합망 환경의 전달망 요소인 다기능 에지 라우터에 요구되는 성능항목들 중에서 10Gbps 라인속도를 구현하기 위한 방안을 스위치 패브릭과 네트워크 프로세서의 성능 관점에서 도출한다.

시스템 수준 또는 보드수준에서 목표 성능을 얻기 위해서는 먼저 사용하는 주요 칩과 인터페이스 자체의 성능이 이를 만족하여야 한다. 차세대 다기능 에지 라우터를 비롯하여 전달망 계위 라우터 시스템의 주요 핵심 구성요소로는 스위치 패브릭 칩과 네트워크 프로세서를 들 수 있으며 이들의 성능이 시스템 전체의 성능을 좌우한다고 해도 과언이 아니기 때문에 본 논문에서는 스위치 패브릭 칩과 네트워크 프로세서의 성능에 초점을 맞추어 분석한다

2.2.1 스위치 패브릭 칩

그림 2는 공유메모리 구조를 갖는 스위치 패브릭 칩의 일반적인 논리 블록도를 나타낸 것이다. 스위치 패브릭 칩은 라인카드의 스위치 패브릭 인터페이스와의 연결을 위한 여러 개의 SerDes(serializer/deserializer)와 스위치 패브릭 내부로 인입된 데이터 세그먼트를 적절한 또다른 SerDes로 스위칭 하기위한 공유메모리 스위치부로 구성된다. 현재의 대다수 스위치 패브릭 칩의 SerDes는 2.5Gbps 또는 3.125Gbps 속도의 완전 전이중(full duplex) 시리얼 링크이다. 이러한 구성을 갖는 스위치 패브릭 칩을 일단 또는 다단으로 구성함으로써 원하는 용량과 성능의 스위치 카드부를 설계할 수 있다.

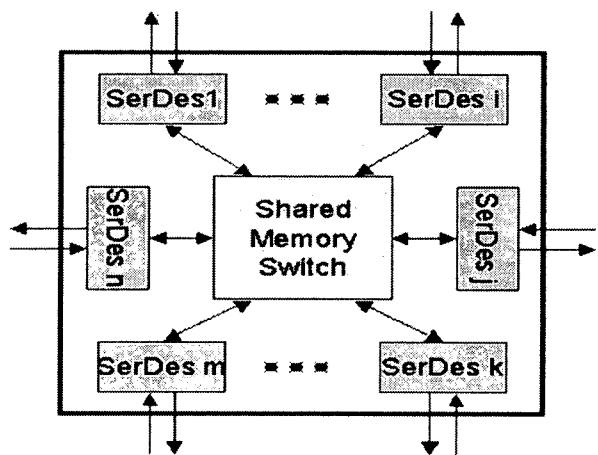


그림 2. 스위치 패브릭의 논리 블록도

차세대 다기능 에지 라우터에 사용 가능한 스위치 패브릭 칩의 성능비교는 칩이 제공하는 SerDes의 개

수와 각 SerDes의 속도 그리고 이들 SerDes를 라인 카드의 스위치 패브릭 인터페이스와 어떤 토폴로지 및 메커니즘으로 구성하느냐에 따라 크게 세 가지 종류 (A, B, C)로 분류하여 성능을 분석한다.

첫 번째 스위치 패브릭 칩(A)은 성능면에서 모든 패킷 크기 및 인터페이스 종류에 대해서 라인속도 스위칭 성능을 제공하며 가용도 측면에서는 스위칭 경로 상으로 트래픽을 분산하여 처리하는 부하 균형 기능을 제공함과 동시에 데이터 경로상에 오류가 발생할 경우 이를 동적으로 복구할 수 있는 기능을 가진다. 또한 최대 수십 Tbps까지 스위칭 능력을 확장할 수 있고 10Gbps 인터페이스에 대해 라인속도를 지원하면서 차세대 개방형 하드웨어 플랫폼 규격을 준수하여 시스템을 설계할 수 있는 구조를 가지고 있다.

그림 3은 스위치 패브릭 칩 A에서 라인카드당 8개의 SerDes를 할당하였을 경우 10Gbps속도의 모든 패킷 크기에 대해서 라인속도를 제공할 수 있음을 보인 것이다. 그림에서 처럼 10Gbps POS(packet over SONET) 트래픽의 요구 대역폭이 보드가 제공하는 최대 가능 대역폭 및 실효 보드 대역폭 보다 작으므로 모든 패킷 크기에 대해서 라인속도 제공이 가능함을 알 수 있다.

그림 3. 스위치 칩의 성능(A)

차세대 다기능 에지 라우터의 스위치 패브릭 칩으로 고려 가능한 또다른 형태의 스위치 패브릭 칩 B는 단일 칩으로 160Gbps의 스위칭 용량을 제공하며 4개의 SerDes를 묶어 하나의 포트로 운영할 수 있다. 스피드업(speed up)은 이런 포트를 하나 더 사용하고 셀의 올바른 순서를 지키기 위해 플로우 기반으로 양쪽으로 나누어 스위칭한다. 그리고 양쪽 포트에 전송되는 트래픽을 실시간으로 모니터링하여 양쪽 포트에 트래픽이 균등하게 분배 되도록 부하를 동적으로 관리한다. 이 때 하나의 포트에 스위칭할 수 있는 최대의 페이로드(payload) 대역폭은 셀 오버헤드를 최소로, 그리고 셀 분할 오버헤드는 없는 것으로 가정할 경우 9.41Gbps이다. 그러나 스위치 포트에 10Gbps POS 인터페이스가 연결되고 네트워크 프로세서 및 트래픽 매니저(TM; traffic manager) 헤더를 10바이트로 가정할 경우 대부분의 패킷 크기에서 9.41Gbps를 넘어서 스위치 패브릭 포트에 전달된다. 이 경우 스피드업을 위해 추가한 포트에 트래픽을 나누어 스위칭할 경우 특별한 문제없이 10Gbps 라인속도로 스위칭할 수 있을것으로 예상되나 인가되는 트래픽 전체가 10Gbps 대역폭을 가지는 하나의 연결일 경우에는 부하 균형(load balancing) 기능을 사용할 수 없기 때문에 스위치 패브릭 칩 자체에서 병목현상이 발생하게 된다.

스위치 패브릭 칩 B는 네트워크 프로세서 및 트래픽 매니저와의 연결에 CSIX-L1 인터페이스[3]를 사용하며 라인카드를 위한 패브릭 인터페이스로는 4개의 2.5Gbps 또는 한 개의 10Gbps 속도의 네트워크 프로세서를 지원할 수 있다. 1024개의 유니캐스트 큐와 8개의 멀티캐스트 큐를 지원하는 인입 가상 출력큐(ingress virtual output queue)를 가지고 있으며 효율적인 멀티캐스트를 지원한다.

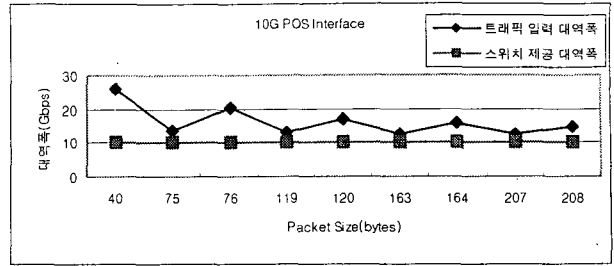


그림 4. 스위치 칩의 성능(B)

그림 4와 그림 5는 각각 스위치 패브릭 칩 B와 C에서 다양한 크기의 POS 패킷에 대해서 요구대역폭을 분석한 그래프이다. 그림에서 처럼 스위치 패브릭 칩 B와 C 모두 특정 패킷 크기에 대해서는 오버헤드를 포함한 트래픽이 요구하는 대역폭을 보드 대역폭이 수용하지 못하는 경우가 발생하기 때문에 10Gbps 라인속도를 지원하지 못함을 알 수 있다

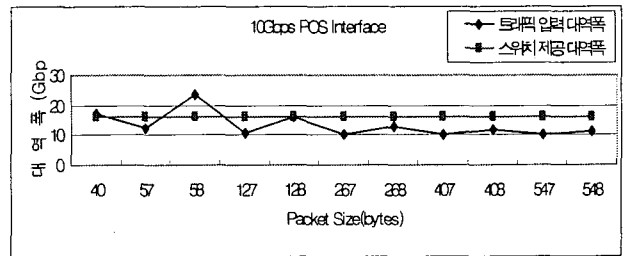


그림 5. 스위치 칩의 성능(C)

2.2.2 네트워크 프로세서

차세대 다기능 에지 라우터에는 다양한 종류의 네트워크 프로세서가 사용될 수 있다. 네트워크 프로세서는 복잡한 알고리즘 처리, 상세한 패킷 분석 기능, 트래픽 관리 기능, 그리고 라인속도의 패킷 포워딩 기능 등을 하나의 칩에서 병렬로 수행이 가능한 고성능 프로세서로서 스위치 패브릭 인터페이스 및 프레임머 (framer)와의 연결에 SPI 4.2[4] 또는 CSIX를 주로 사용한다. 그림 6은 일반적인 네트워크 프로세서의 전체적인 논리 블록도를 나타낸 것이다

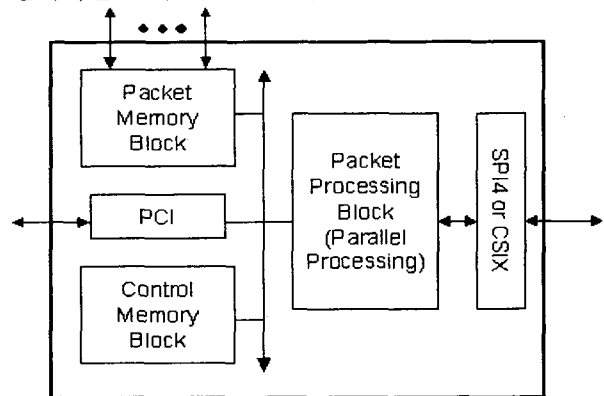


그림 6. 네트워크 프로세서의 논리 블록도

본 논문에서는 네트워크 프로세서가 스위치 패브릭 인터페이스 및 프레임머와의 연결에 SPI 4.2를 사용하는 것으로 가정한다. 따라서 SPI 4.2의 성능 또한 시스템 전체가 10Gbps 라인속도를 지원하는냐의 여

부에 중요한 부분을 담당하게 된다. 그림 7은 SPI 4.2 인터페이스 자체가 제공할 수 있는 성능과 10Gbps 속도의 Ethernet 및 POS에서 요구하는 대역폭을 함께 나타낸 그래프이다. 그림에서 알 수 있듯이 10Gbps의 Ethernet과 POS가 요구하는 대역폭보다 SPI 4.2가 제공하는 대역폭이 크기 때문에 SPI 4.2 인터페이스에서는 병목현상이 발생하지 않음을 알 수 있다.

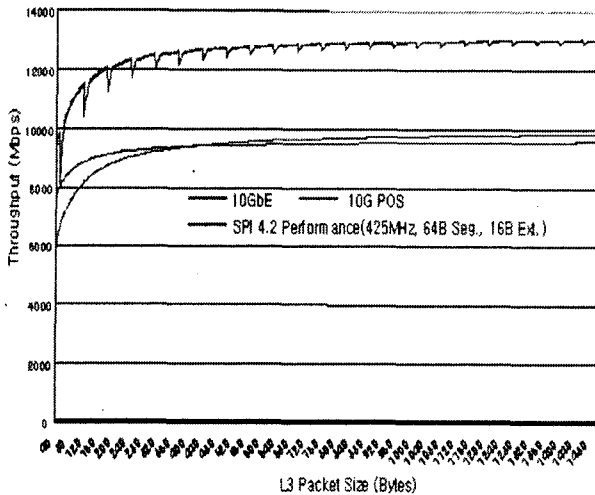


그림 7. SPI 4.2 인터페이스 성능

2.3 시스템 구조

차세대 다기능 에지 라우터는 단일화된 패킷망인 광대역통합망 환경에서 기존의 다양한 가입자 네트워크와 IP/MPLS로 대변되는 코어망 사이에 위치하여 가입자로부터의 트래픽 또는 역방향의 트래픽을 수용하고 정해진 트래픽의 특성에 맞게 처리하여 전달망의 코어 또는 가입자측 장치로 보내는 역할을 한다. 그림 8은 광대역통합망 환경에서 다기능 에지 라우터가 차지하는 위치와 역할을 개념적으로 도식화한 것이다.

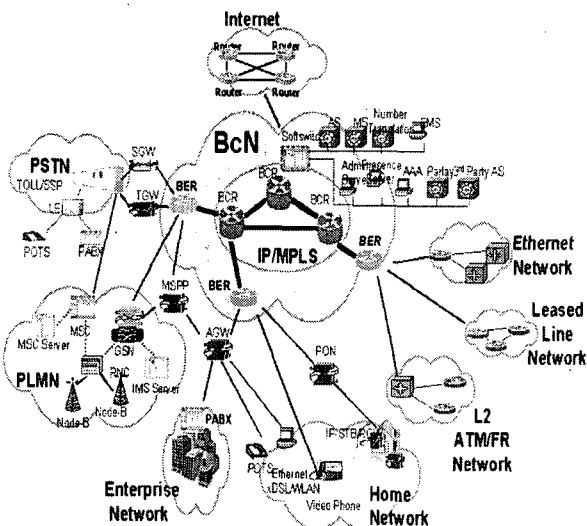


그림 8. 다기능 에지 라우터의 네트워크 모델

본 절에서는 10Gbps 라인카드에서 라인속도를 제공하면서 차세대 개방형 하드웨어 플랫폼을 적용할 수 있는 방법 및 상위 수준의 시스템 구조를 설명한다. 그림 9는 시스템의 전체적인 구성을 논리적으로 그

린 것이다. 다기능 에지 라우터는 일반적인 라우터 시스템과 마찬가지로 크게 제어부, 스위칭부, 그리고 라인카드의 세 부분의 기능블럭으로 구성되어 있다. 제어부는 CPU와 운영체제를 비롯하여 각종 제어모듈과 메모리가 주된 구성요소이며 시스템 전체의 동작과 기능 관리를 수행한다. 라인카드는 네트워크 프로세서를 이용하여 물리 인터페이스로부터 들어오거나 나가는 각종 패킷의 처리 기능을 담당하게 된다. 그리고 스위치 패브릭 칩으로 구성된 스위칭부는 입력포트로 들어온 데이터에 대해서 올바른 목적지 인터페이스로 보내기 위한 스위칭 기능을 수행한다. 그리고 시스템의 이러한 유기적이고 원활한 동작을 위해 각 기능 블럭간에 정보 및 데이터 교환이 필요할 경우 IPC(interprocessor communication)를 통하여 통신하게 된다.

다기능 에지 라우터의 기본형상으로서 통신산업계에 새로운 하드웨어 표준으로 부상하고 있는 차세대 개방형 플랫폼 구조를 적용한다. 차세대 개방형 플랫폼 구조는 새로운 폼팩터(form factor)와 스위치 패브릭 구조에 기반한 차세대 통신장비를 위한 표준화된 개방형 플랫폼의 한 형태이다.

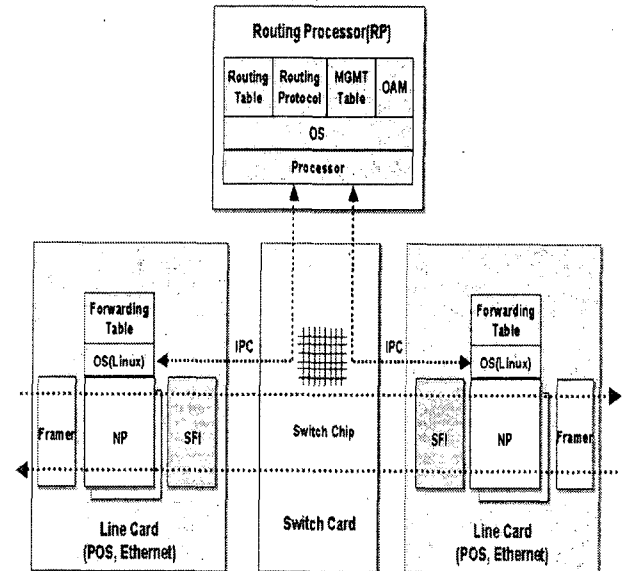


그림 9. 시스템의 논리 구성도

장비 제조사에서는 라우터와 같은 통신시스템에 차세대 개방형 플랫폼 구조를 적용함으로써 개발기간 단축 및 비용을 절감할 수 있는 장점이 있으나 차세대 개방형 플랫폼 구조에도 여러 가지 버전의 규격이 있어 서로 다른 버전을 따르는 시스템간에는 호환성이 없으며 또한 시스템의 용량과 성능의 확장에 대해서는 일부 제약을 갖는 단점이 있다.

차세대 개방형 플랫폼 구조에서 스위치 카드와 라인카드를 포함한 각 기능 카드간의 연결 토폴로지는 듀얼 성형(dual star), 듀얼-듀얼 성형(dual-dual star), 그리고 메쉬(mesh) 등이 가능하며 이들 토폴로지 중에서 10Gbps 라인속도 지원에 대한 성능문제로 인해 듀얼-듀얼 성형 토폴로지를 사용한다.

메쉬 토폴로지는 분산 스위치 구조로서 10Gbps의 라인속도와 80Gbps 이상의 스위칭 처리율을 지원하는 스위치 패브릭 칩이 없으며, 듀얼 성형 구조는 포트당 4개의 SerDes를 지원하고 현재의 SerDes 성능은 각 SerDes당 2.5Gbps(raw rate 3.125Gbps)

성능을 제공하므로 총 10Gbps의 전송율을 제공할 수 있지만 스위치 패브릭의 분할 오버헤드(segmentation overhead) 및 내부 오버헤드로 인해 모든 패킷 크기에 대해서 10Gbps 라인속도를 지원할 수 없다.

차세대 다기능 에지 라우터의 라인카드는 POS, Ethernet, ATM 인터페이스를 동시에 수용과 함께 10Gbps 라인속도를 제공하며 크게 세 부분의 기능블럭으로 나눌 수 있다. 네트워크 프로세서를 포함하는 NP 모듈이 각각 인입(ingress)과 출구(egress) 트래픽을 처리하고 스위치 카드의 스위치 패브릭 칩과의 연결을 위한 스위치 패브릭 인터페이스 및 트래픽 매니저 부분, 그리고 실제적인 물리 케이블 및 포트가 접속되는 물리카드인 PMC(PCI Mezzaine Card)가 있다.

PMC 는 믹스앤매치(mix and match) 기술을 적용함으로써 라인카드의 물리 인터페이스의 제작 종수를 감소시킬 수 있고 통신사업자는 하나의 라인카드에서 다양한 종류의 인터페이스 및 속도를 제공할 수 있기 때문에 망의 환경에 따라 최적의 시스템 운용을 할 수 있는 장점을 가진다.

3. 결 론

본 논문에서는 광대역통합망 환경에서 전달망 요소인 에지 라우터에 요구되는 기능 및 성능항목들 중에서 차세대 개방형 플랫폼의 적용방법과 10Gbps 라인속도를 구현하기 위한 방안을 도출하고 그에 따른 상위수준의 시스템 구조를 제안하였다.

성능예측을 통하여 10Gbps 라인속도 지원이 가능한 스위치 패브릭 칩과 네트워크 프로세서의 성능 요구 조건을 도출하였으며 스위치 카드와 다른 기능 카드와의 연결 토폴로지는 듀얼-듀얼 성형을 사용한다.

[참 고 문 헌]

- [1] 광대역통합망구축기획단, "정보통신 일등국가 실현을 위한 광대역통합망 구축 기본계획(안)," 정보통신부, Nov. 2003.
- [2] "An Architecture for Differentiated Services," <http://www.ietf.org/rfc/rfc2475.txt>
- [3] "CSIX-L1: Common Switch Interface Specification-L1," Network Processing Forum, Aug. 2000
- [4] "OIF-SPI4-02.0 System Packet Interface Level-4 (SPI-4) Phase 2: OC-192 System Interface for Physical and Link Layer Devices," Optical Interworking Forum