

40G/100G 이더넷 표준 기반의 라인카드 기술 동향

Trends of Line Card Technology Based on 40G/100G Ethernet Standard

네트워크 기술의 미래 전망 특집

양종열 (C.R. Yang)	광전송기술연구팀 책임연구원
안계현 (K.H. Ahn)	광전송기술연구팀 선임연구원
김승환 (S.H. Kim)	광전송기술연구팀 책임연구원
고제수 (J.S. Ko)	광전송기술연구팀 책임연구원
김광준 (K. Kim)	광전송기술연구팀 팀장

목 차

-
- I. 서론
 - II. 40G/100G 이더넷 표준 기반의 라인카드 개발 동향
 - III. 결론

* 본 연구는 지식경제부 및 정보통신연구진흥원의 IT 핵심기술 개발사업의 일환으로 수행하였음. [2008-F017-01, 100Gbps급 이더넷 및 광전송 기술개발]

UCC, 트위터 등 멀티미디어 콘텐츠 증가, 유틸리티 컴퓨팅과 같은 다양한 신규 서비스의 급증, IPTV 등 높은 대역폭을 요구하는 애플리케이션의 증가, 가상화 데이터 센터의 등장과 함께 40G/100G 이더넷 기술이 차세대 광대역 서비스 대역폭 요구에 대한 장기적 해결방안의 하나로 제시되고 있는 가운데 세계적으로 40G/100G 이더넷으로의 네트워크의 진화가 시작되고 있다. 본 고에서는 최근 세계적으로 뜨거운 쟁점이 되고 있는 차세대 인프라 40G/100G 이더넷 표준을 기반으로 하는 디바이스 및 프로덕트의 출시 동향을 살펴보고 현재 사용 가능한 상용 칩을 이용한 40G 이더넷 라인카드의 구조와 향후 구현 가능한 100G 이더넷 라인카드의 구조 그리고 40G/100G 이더넷 상의 OTN 네트워크 응용에 대해 고찰한다.

I. 서론

최근 국가 그린 IT 전략에 따라 추진 목표로서 그린 IT 제품 개발 및 수출 전략화와 IT 서비스 그린화, 10배 빠른 안전한 네트워크 구축, 즉 초광대역 융합망 구축 및 핵심기술 확보, 고효율 액세스 네트워크 구축 등이 설정되었다. 이에 부합하는 기술 중의 하나가 최근 고속 이더넷 기술의 추세로 세계적으로 뜨거운 쟁점이 되고 있는 차세대 40G/100G 이더넷이다. 이는 40GE/100GE 기술이 국가 그린 네트워크 및 차세대 광대역 서비스 요구대역폭 장기적 해결방안의 하나로 전망되기 때문이다.

비디오 기반 애플리케이션으로 인해 네트워크 대역폭의 필요성이 더욱 증가하고 있는 가운데 40G와 100G 고속 이더넷이 이를 충족하면서 나아가 고성능 컴퓨팅, 비즈니스 연계, 가상, VOD, 저장장치, 비디오 감시 및 VoIP 같은 대역폭 집약적 애플리케이션을 크게 성장시킬 것으로 전망된다. 특히, 데이터 센터는 지속적으로 증가하는 애플리케이션과 이를 활용한 다양한 서비스를 제공하기 위해 성능, 효율성 및 확장성이 업체들의 요구사항이 되고 있으며, 가상화 기술을 도입함으로써 빠른 속도로 변화하는 고객 요구에 신속하게 대응할 수 있는 유연한 IT 인프라 환경을 제공할 수 있을 것이다.

최근 미국 스탠포드 대학의 연구 결과[1]에 따르면, 2000년에서 2005년 사이 전 세계 데이터 센터의 에너지 사용량은 연간 710억 kWh/yr에서 1500억 kWh/yr 이상으로 2배가 넘게 증가했으며, 특히 한국을 포함한 아시아 지역의 데이터 센터 전력사용 증가율이 가장 높은 것으로 알려지면서 2009년 데이터 센터용 40G/100G 이더넷 라인카드의 필요성이 급부상하고 있다.

궁극적으로 40G/100G 이더넷을 제공을 위해서

40G/100G 트렁크 그룹을 이용하여 링크 집성(agggregation)을 통하여 제공하는 방법도 가능하나 싱글 40G/100G 이더넷은 싱글 40G/100G 포트와 케이블만 관리하므로 네트워크와 호환성이 높고 관리가 용이하며 그린 에너지, 즉 전력 효율 또한 매우 높은 장점을 갖는다. 이러한 이유로 40G/100G 이더넷 기술이 데이터 센터 네트워크 집성업 링크로 인한 속도 및 트래픽 처리용량의 증대를 해결하기 위해 요구되고 있다[2].

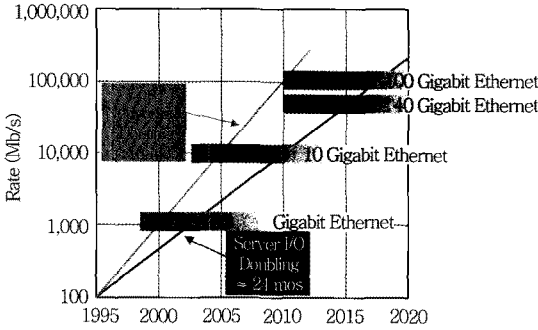
경제성에 대하여 CIR의 보고서에 따르면, 40G와 100G 이더넷은 2016년까지 SONET을 대체하는 등 네트워크의 대부분이 40G/100G 이더넷으로 대체될 것으로 전망되고 있다. 40G 이더넷 시장이 100G 이더넷 시장보다 더 클 것으로 기대하고 있으며 2016년까지 세계시장이 40G 이더넷은 31억 달러, 100G 이더넷은 12억 달러로 전망하고 있으며 2010년 국내 데이터 시장은 7억 원 규모로 전망하고 있어 국가 주도로 선도하여 시장 경쟁력을 확보하고 국내 유선통신제조업의 신규 시장을 적극 창출할 필요가 있다[3].

본 고에서는 I장 서문에 이어, II장에서 IEEE 802.3ba 이더넷 PHY 기반의 40G/100G 고속 이더넷 표준화 동향과 이를 기반으로 하는 주요 세계 통신회사의 프로덕트의 개발 동향을 살펴보고, 현재 출시되어 있는 상용 칩을 이용한 40G 이더넷 라인카드의 구조, 향후 구현 가능한 100G 라인카드 구조 및 OTN 네트워크 응용에 대해 고찰하고, III장에서 결론을 맺고자 한다.

II. 40G/100G 이더넷 표준 기반의 라인카드 개발 동향

1. 40G/100G 이더넷 표준화 동향

이더넷 가입자 수가 증가하고 가입자단의 고대역



<자료>: An Overview: Next Generation of Ethernet - IEEE 802HSSG_Tutorial_1107

(그림 1) 40G/100G 컴퓨팅 및 네트워킹

화가 증가하며 다양한 인터넷 서비스의 폭발적인 증가로 고대역 인터페이스의 필요성이 제기되었고, 2002년 6월 저렴한 가격 및 신속한 개발을 목표로 10G 이더넷 표준화가 완료(802.3ae)되었다. 이후 (그림 1)에서 보는 바와 같이 서버와 컴퓨팅 트래픽이 2년마다 2배씩 증가하고 네트워킹은 18개월마다 2배씩 증가하여 40G/100G 이더넷 인터페이스의 필요성이 확인되었다. 특히 10GE 표준화에 미 포함되었던 OTN 기능으로 SDH/SONET의 장점을 수용하고 IEEE와 ITU-T와의 상호협조가 진행중이다.

2007년 7월 IEEE 802.3ba HSSH가 시장 요구에 맞추어 표준화하기 위해 시작되었고, 12월에 새로운 표준에 대해 연구하기 위해 공식적으로 TF가 구성되었다. 속도 및 확장 관점에서 100G가 부적합으로 결정되어 서버 I/O에 따른 적합성 및 시장성을 근거로 100G 외에 40G를 포함하게 되었으며 대역기반 서비스를 위해 40G, 애플리케이션 수용을 위해 100G가 제안되었다. 40G/100G 이더넷 MAC/PHY 표준이 IEEE 802.3ba SG에 의해 2010년 6월 최종 승인되었다[4],[5].

2. 40G/100G 이더넷 개발 동향

2010 인터롭 컨퍼런스에서 40G 이더넷이 뜨거운

쟁점이 되었고 익스트림에서 초기 40G 이더넷 공급업체로 지정되었다.

Mellanox Technology는 데이터 센터와 네트워크 스토리지 시스템용 40GE NIC(Connect EN 40G)를 2009년 9월에 출시하였고, Spirent/Ixia/EXFO/JDSU 등에서 40G/100G 이더넷 패킷 분석장비를 2009년 말부터 출시하고 있다. 또한, Ixia는 2009년 12월 인터롭에서 40GE/100GE CFP MSA 광트랜시버 모듈을 통해 10기가 10포트의 100G 이더넷 트래픽을 송수신하였고, 또 JDSU는 100G 이더넷 테스트 프로덕트를 발표하였으며, 알카텔 루슨트는 100GE 서비스 인터페이스 장비(7450)를 2010년 상용화 출시하였다. 쥘니퍼 네트워크, 시스코는 표준화에 부합하는 100GE 장비(7000 업그레이드)를 준비중이다[3].

Network World[6]에 의하면 Brocade Communication사에서 데이터 센터를 위한 100G 이더넷 서비스를 준비하고 있으며 기존 대비 데이터 센터 에너지는 37%까지 감소한다고 발표하였다. 데이터 센터용 40GE/100GE를 2010년 말까지 출시할 예정이며 2015년까지 1테라급 시스템이 예상된다고 발표하였다. 그러나 40GE의 경우 10GE 가격의 7~8배, 100G 이더넷의 경우 10G 이더넷 가격의 20배 가량 가격이 비싸므로 저가화가 필요한 상황이다. 100G 이더넷 서비스 시장은 기업의 대역폭 요구량과 대부분의 트래픽 문제를 해결하며 오랫동안 지속될 것으로 예상된다.

최근 데이터 센터의 서버로 40G 이더넷 서비스가 준비되고 있으며, 고성능 컴퓨팅 같은 대역폭 집약적인 애플리케이션이 고속 이더넷의 필요성을 주도할 것이다. 데이터 센터는 이미 성능을 향상시키기 위해 10G 업 링크를 집성하고 있어 곧 40G 이더넷이 서버 시장에서 가시화될 것으로 보이며, 40G와 100G 이

더넷은 대형 라우터, 1테라급 이더넷 스위치 시스템, 코어 전송 및 백본 네트워크에 요구되고 있다.

ETRI는 40G/100G 이더넷 연구 프로젝트를 수행 중이며 참여할 기업을 찾고 있다. 저전력 기능을 갖는 10G 4 레인의 데이터 센터용 40G 이더넷 라인카드, 40인치 백플레인에서 40G 데이터를 전송하기 위한 적응 이퀄라이저, 그리고 4개의 10G 신호를 CWDM 방식으로 광전송하는 기능을 수행하는 2포트 40G 이더넷 광트랜시버, 그리고 10G/40G/100G 이더넷 신호를 백본 망에서 수용하여 2.5G/10G/40G/100G OTN 신호로 인캡슐레이션 및 디캡슐레이션 하여 전송할 수 있는 OTH 프레이머 칩 등을 개발하고 있다.

3. 40G 이더넷 라인카드 개발

가. 40G 이더넷 요구사항

UCC 및 멀티미디어 콘텐츠 증가, 유틸리티 컴퓨팅과 같은 다양한 신규 서비스의 급증, IT 아웃소싱의 확대 등으로 데이터 센터에 대한 수요가 증가하고 있는 가운데 2009년에 ‘클라우드 컴퓨팅(cloud computing)’이라는 개념이 새로 등장하면서 데이터 센터는 설계, 구축, 운용 등 모든 면에서 근본적인 변화를 요구 받고 있고, 그린 네트워크 기술이 필수적으로 요구된다.

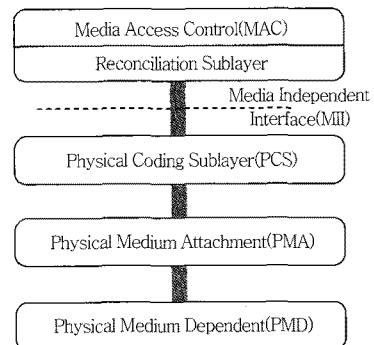
최근 상용화되고 있는 상용 디바이스들은 저전력 규격 요구조건을 기본적으로 생산하고 40G/100G 이더넷 라인카드 설계를 위한 요구사항을 부합하기 위해 최적화되어 있다. 따라서 40G 이더넷 라인 카드는 IEEE 802.3ba 표준을 만족하는 상용 칩을 최적 조건에서 조합하여 고밀도, 고성능 및 저전력을 위한 시장 요구를 만족하고 고속 설계 요구조건을 부합함으로써 낮은 지터 성능의 40G 광트랜시버 성능을 요구하는 40G 설계가 가능하다.

40G 이더넷 라인카드를 위한 사용자 요구사항 [3]은 다음과 같다.

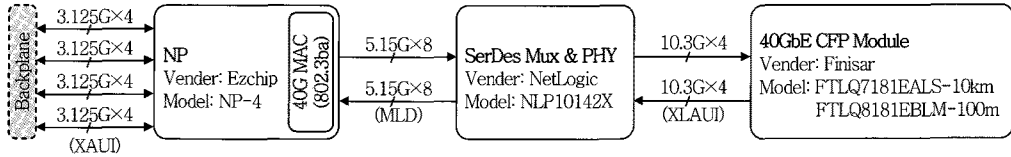
- 40G 이더넷 라인카드는 802.3ba 40GE MAC/PCS 기능을 포함하여야 한다.
- 40G 이더넷 라인카드는 40GE 광트랜시버 기능을 포함하여야 한다.
- 40G 이더넷 라인카드의 PCS 데이터 레이트는 40Gb/s이어야 한다.
- 40G 이더넷 라인카드의 PMA는 4 레인×10G를 지원하여야 한다.
- 40G 이더넷 라인카드는 사용자 요구에 따라 광대역서비스 기능을 제공하기 위한 기능을 지원하여야 한다.
- 40G 이더넷 라인카드는 고성능, 대용량 및 확장성은 물론 공간 효율적인 고밀도 시스템과 가상화 클라우드 컴퓨팅을 위한 차세대 시스템에 적합하도록 설계되어야 한다.

나. 40G 이더넷 PHY 표준

40G 이더넷 물리 인터페이스 규격은 표준 IEEE 802.3ba 프로토콜 스택 구조에 따라 10G×4 레인으로 제공된다. (그림 2)는 PCS 계층 및 상급 40G/100G 이더넷 요구조건에 관한 표준 IEEE 802.3ba 프로토콜 스택의 개념도이다.



(그림 2) 표준 IEEE 802.3ba 프로토콜 스택

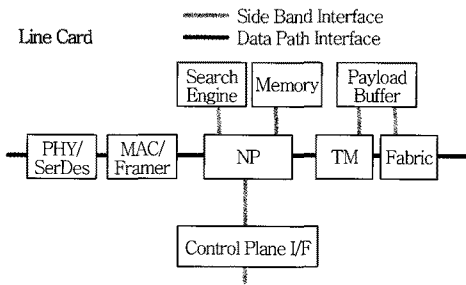


(그림 3) 상용 칩을 이용한 40G 이더넷 PHY

(그림 3)은 2010년 상반기 출시된 상용 칩을 이용하여 IEEE 802.3ba 40G 이더넷 PHY를 수용하는 10G 4레이의 인터페이스로 40G 이더넷 라인카드를 구현하기 위한 구조이다.

다. 40G 이더넷 라인카드 구조

(그림 4)는 IEEE 802.3ba에 규정하는 40G 고속 이더넷 표준[4]을 부합하기 위한 인터페이스 구조이다. 기능별로 크게 외부 네트워크로부터 들어오는 트래픽 패킷 데이터를 수용하여 목적지로 스위칭하여 전송하기 위한 데이터 경로 인터페이스(data path interface)와 제어 기능을 처리하기 위한 대역 외 인터페이스(side band interface)로 구분된다.



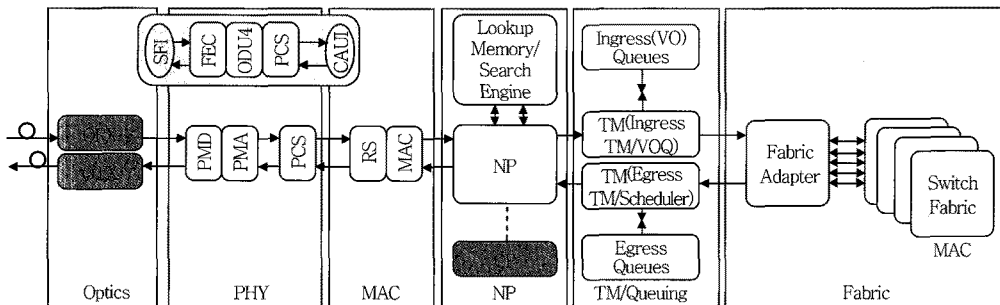
(그림 4) 40G 이더넷 라인카드 인터페이스 구조

(그림 5)에 IEEE 802.3ba 이더넷 표준에 부합하는 공통 라인카드(universal line card) 구조를 나타내었다. 이 구조에서 PHY 블록에 40G 이더넷 PHY가 채용되면 40G 이더넷 라인카드 하드웨어가 되고, OTN PHY(회색 부분)이 채용되면 OTN 라인카드 하드웨어가 된다.

사용된 주요 상용 칩은 2010년 상반기에 출시되었으며 <표 1>은 40G 이더넷을 위한 40G PHY 상용 칩으로 40G CFP MSA 광트랜시버, 넷로직의 40G SerDes Mux, 10G급 SerDes 내장형 알테라 FPGA 그리고 40G MAC을 내장하고 있는 EZ NP-4 네트워크 프로세서[7] 등이 있다.

<표 1> 40G 라인카드용 상용 칩

Device	Function	Part Number	Vendor
Optic Transceiver	40G CFP MSA(1포트)	FTLQ8181-EBLM	Finisar
Optic Transceiver	40G CFP MSA(1포트)	SCF0400LA-xN-ES01	Sumitomo
10G FPGA	40G PCS IP	Stratix IV GT FPGA, EP4S40G5	Altera
40G PHY	SerDes Mux	NLP10142X	NetLogic
EZ NP-4	MAC, Processor	EZ NP-4	EZchip



(그림 5) IEEE 802.3ba 이더넷 표준 라인카드 구조

NP는 패킷 분류, 레이블 생성/스와핑, 어드레스 매핑 및 폴라이싱을 수행하며, CPU에 테이블 관리 및 제어 프로토콜을 포워딩한다. NP의 포워딩 및 마킹을 기반으로 TM이 패킷 큐잉, 셰이핑, 스케줄링을 수행한다. 패브릭은 스위치를 건너 패킷을 스케줄한다. 패브릭은 고속 SerDes를 이용하여 백플레인을 통해 구현된다. IEEE에서 MAC 디바이스용 XLAUI, CAUI, MLD를 정의하고, ITU-T는 OUT-4 표준에 근거한 프레임 구현을 추진하고 있다.

라. 40G 광트랜시버

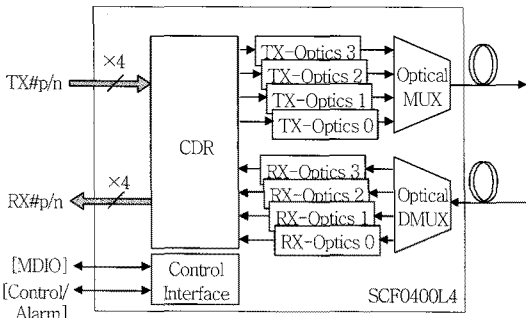
40G (그림 6)의 광트랜시버 모듈은 포트그룹과 물리적 네트워크 인터페이스 사이에서 트래픽 스위칭 역할을 한다. 사용된 광트랜시버 모듈, CFP MSA[8]는 Stratix IV GT FPGA에 직접 10G XLAUI 접속이 가능하므로 외부 PHY 디바이스가 필요 없어 전체 시스템 복잡도가 간단해진다. 40G 광트랜시버 모듈은 4개의 10G 전기신호로 다중화 되며, 각각의 전기신호는 구동 증폭기에서 증폭된 다음 광신호로 변환되고, 광다중화기(Optical Mux)를 통하여 하나의 광섬

유로 다중화된다. 수신부는 광섬유를 통하여 전송된 광신호가 광다중화기를 통하여 4개의 광신호로 분리되고, PIN PD를 거쳐 전기신호로 변환된다. 변환된 전기신호는 증폭되고 4개의 10G 전기신호로 출력된다. 광네트워크 응용을 위해 핫 플러그 방식으로 설계되었다. MDIO 관리 인터페이스를 통해 기능적으로 액세스되는 루프 백(loop back) 기능을 제공한다.

라인카드는 <표 1>에 의거 프로토콜에 따라 40GBASE-LR와 40GBASE-ER 타입, 설계용량에 따라 1포트용과 2포트로 설계될 수 있다. 802.3ba TF의 표준에서는 40G 지원 물리 계층 규격을 SMF에서 10km, OM3 MMF에서 100m, 구리에서 10m, 그리고 백플레인에서 1m로 규정하고 있다.

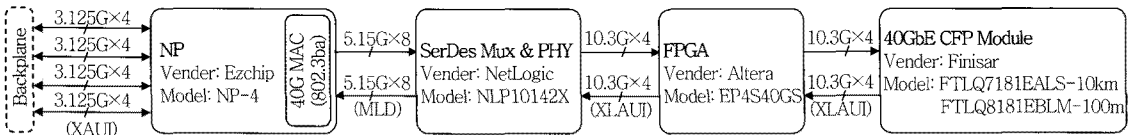
마. 응용서비스 기능 블록 FPGA

저전력 레인 제어, FEC 또는 그 밖의 광대역 응용 서비스를 제공하기 위한 서비스 기능 블록(FPGA)[9]이 CFP MSA 광트랜시버 모듈과 40G SerDes 디바이스와 10G로 인터페이스하기 위해 IP 소프트웨어 패키지 형태로 제공될 수 있다. 이는 IEEE 표준 구조와 독립적으로 제공(option)하기 위한 서비스 기능블록으로서 이 응용서비스 기능블록을 통해 FEC 기능, 그린 에너지 기능 등 차세대 광대역 응용 서비스도 제공할 수 있다. (그림 7)에 10G FPGA를 포함한 구조를 나타내었다. Altera의 STRATX-IV FPGA는 11.3G 트랜시버를 갖고 있어서 client side에서 외부 PHY 디바이스를 사용하지 않고도 바로 CFP 광모듈과 접속이 가능하다. 여기서는 NP와 SerDes/Mux를 통하여 인터페이스 한다.



<자료>: Sumitomo, product Brief NPI-09010B, 2009. 11.

(그림 6) 40G CFP MSA 광트랜시버



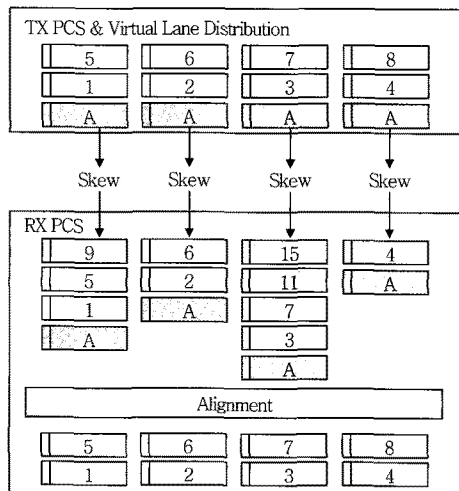
(그림 7) 응용서비스 기능블록을 포함하는 PHY 구조

칩은 작은 로직 용량과 작은 패키지가 접합하며 OTN과 호환되기 위해서 -1, -2, -3 등 3개 speed grade 중에서 11.3G까지 동작이 가능한 speed grade -1이 지원되는 칩이 필요하다.

바. PCS

IEEE 802.3ba 표준은 40G/100G 이더넷을 위한 기술로 물리적인 멀티 레인을 지원하기 위해 이더넷 인터페이스 기능을 수행하는 프로토콜 스택 계층으로 PCS를 규정하고 있다. 이는 가상 레인 데이터 분배 기술[5]로서 40G/100G 이더넷을 한 개의 광선으로 이용할 수 없는 기술적 한계 때문에 표준으로 제안된 기술이다. 가장 중요한 개념은 가상 레인으로서는 40G/100G MAC 프레임이 64B/66B 블록 단위로 블록화 하여 n개의 가상 레인에 라운드 로빈(round robin) 방식으로 할당하고 Rx PCS는 수신한 데이터를 정렬하여 원래의 MAC 프레임 형태로 복원한다.

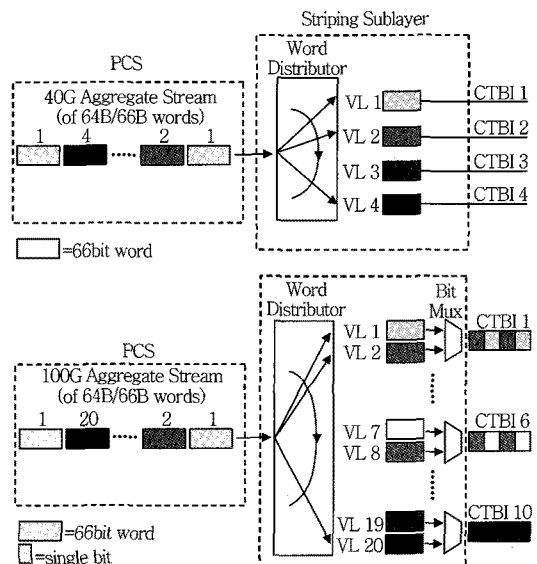
(그림 8)에서 이를 설명한다. PCS 송신단(TX PCS & Virtual Lane Distribution)에서 MAC 프레임을 64B/66B 형태로 처리 후, 4개의 레인으로 분



(그림 8) PCS 가상 레인 분배 기술

산하여 전송한다. 전송된 데이터는 PCS 수신단(RX PCS)에서 다시 정렬하여 원래의 MAC 프레임 형태로 복원되고, 각 레인은 각각 다른 경로를 갖기 때문에 스큐(skew)가 발생하며, 정렬(alignment) 블록에서 이 스큐를 보상한다. 송신단에서는 주기적으로 A를 삽입하여 보내주고, 수신단에서는 정렬을 수행하여 스큐를 보상한다. 이후 분산된 프레임들을 하나의 프레임으로 재조립하면 원래의 프레임으로 복원된다.

(그림 9)는 40G와 100G 가상 레인 개념[10]을 설명한다. 40G는 64B/66B 블록화된 데이터가 4개의 가상 레인으로 분배되고 가상 레인 분배 기술에 의해 4개의 전기적인 레인으로 대응된 다음 전기적 레인은 PMA에 의해 4개의 광 레인으로 대응되는 과정을 나타낸다. 또한, 100G는 64B/66B 블록화된 데이터가 20개의 가상 레인으로 분배되고 가상 레인 분배 기술에 의해 10개의 전기적인 레인으로 대응된 다음 전기적 레인은 PMA에 의해 4개의 광 레인으로 대응되는 과정을 나타낸다. 송신한 데이터와 수신한 데이터가 같음을 알 수 있다.

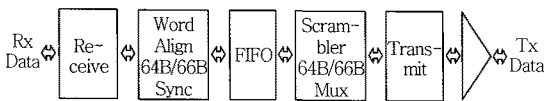


<자료>: WP-01080-1.3, Altera Corporation

(그림 9) PCS 가상 레인 개념

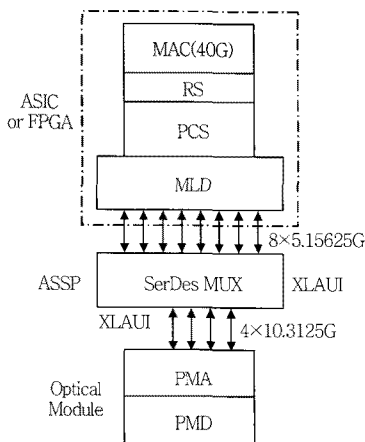
사. SerDes Mux[11]

본 디바이스는 40GBASE-xR4, 100GBASE-xR4 및 100GBASE-xR10(802.3ba) 구현을 지원하는 트랜시버로 10G 이더넷 4 포트로 40G 인터페이스를 지원한다. (그림 10)에 병렬에서 직렬로 변환하여 데이터를 전송하는 경로를 나타내었다. 40G 모드에서 2 라인 인터페이스에서 병렬 데이터를 수신하여 시리얼로 재전송한다. 이 과정 동안 들어오는 데이터는 역시리얼화(deserialized)되고 다중화되어 전송된다. 40G/100G 모드로부터 수신 경로의 첫 컴포넌트는 데이터 입력 수신부이다. 2 디퍼런셜 Tx data 페어에 연결되고, 2 병렬 스트림이 10비트 코드 워드로 변환된다.



(그림 10) 병렬/직렬 변환 전송 경로

(그림 11)에 IEEE 802.3ba PHY 구조를 나타내었다. 그림에서 보는 바와 같이 SerDes Mux는 8개의 5.15625G 인터라켄(Interlaken) 인터페이스를 4 라인의 10G XLAUI 인터페이스로 변환한다. 이후 PMA는 PCS와 PMD 사이에 위치하며 n개의 레인으



(그림 11) IEEE 802.3ba 40G PHY 구조

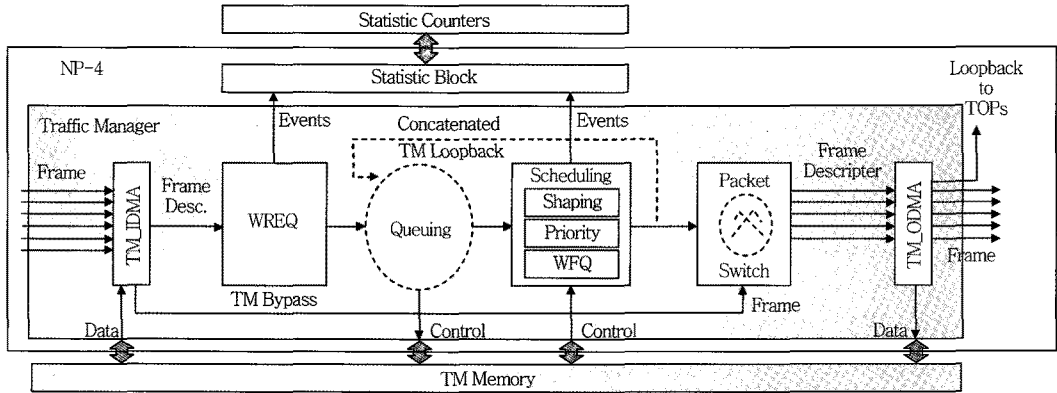
로 구성된 전기적인 인터페이스를 m개의 PMD 레인으로 대응시키는 기능을 한다. 40G/100G 이더넷에서는 PMA가 여러 개로 분리될 수 있으며 비트 레벨 다중/기어박스(gear boxing), 클럭, 데이터 복구 및 클럭 생성 등의 기능을 한다.

아. 네트워크 프로세서

40G 데이터 프로세싱을 위하여 full duplex 50G를 지원하는 네트워크 프로세서(EZ NP-4)[12]가 필요하다. EZ NP-4 100G 네트워크 프로세서는 IEEE 802.3ba의 MAC (그림 11)과 같이 ASIC or FPGA(점선부분) 등을 내장하고 있다. 네트워크 프로세서는 TM 블록을 통하여 채널 운용 정보를 CPU에 전달하고 CPU는 이를 TM 메모리에 저장하며 FPGA에서 메모리 룩업하여 가상화를 위하여 채널을 선택적으로 제어하는 기능을 할 수 있다.

(그림 12)에 EZ NP-4 TM 데이터 플로를 나타내었다. TM_IDMA, TM_ODMA는 데이터 경로를 다루고, PFQ 큐잉 블록은 디스크립터 정보를 다룬다. PFQ 큐잉 블록은 모든 서비스 프로비저닝(provisioning), 큐잉(queueing) 및 스케줄링(scheduling)을 책임지므로 TM 블록이기도 하다.

내부 스위치로부터 도착하는 패킷이 TM_IDMA를 통해 TM에 들어가면, TM_IDMA 블록은 링크 테이블을 업데이트하고 패킷 데이터를 TM 메모리에 쓰고, 프레임 포인터(메모리 내 프레임 위치)를 발생한다. 메시지는 IDMA 발생정보(프레임 포인터와 길이)와 TOP 내용(Flow ID, 패킷 길이, 패킷 미터링 컬러, 인터패킷 갭 세팅, 패킷 스위치 ID)을 가진 프레임으로 통과한다. 패킷 디스크립터(descriptor)가 WRED 블록(드롭 또는 큐를 결정)을 통과하고, 플로 ID를 따라 큐잉 블록 속으로 큐된다. 디스크립터와 큐는 패킷 데이터로 TM 메모리에 저장된다. 계층 스케줄러는

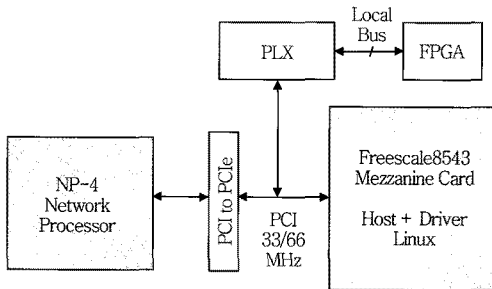


(그림 12) EZ NP-4 TM 데이터 플로우

전 계층에서 폭주, QoS 및 대역폭 보장 상태기반으로 큐로부터 오는 패킷을 스케줄 한다. 계층 스케줄러는 셰이핑, 우선순위 및 폭주관리 블록을 가지고 스케줄링 프로세스를 실행한다. 패킷이 패킷 스위치에 도착하고, 목적지와 서비스 등급별로 출력 큐에 놓인다. 패킷은 TM_ODMA로 보내져 전송되거나 2차 연결(concatenated) TM 플로우를 위해 재순환(recycle) 된다.

자. CPU

40GE 라인카드에 탑재되는 CPU는 MPC8543급이며, (그림 13)에 CPU 블록 구조를 나타내었다. CPU는 PCIe 버스인터페이스를 제공하는 온 보드 디바이스 또는 모듈 타입(Host CPU Mezzanine Card)이 적용 가능하다.

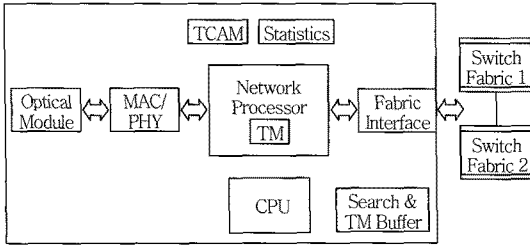


(그림 13) CPU 구조

4. 100G 이더넷 라인카드 개발

가. 100GE 라인카드 구조

(그림 14)는 IEEE 802.3ba 이더넷 표준 기반의 40G/100G 라인카드 기본 형상 구조[1],[4],[13]를 보여준다. 기본 구조는 40G 이더넷 라인카드와 유사하나 라인카드 내 부품 구조의 복잡도와 네트워크 프로세서 등 성능이 달라진다. 100G 연구 시제품은 2007년 독일의 시멘스(Fischer)에 의해 이미 입증된 바 있고, 2008년에는 Comcast(미국)와 시스코에 의해 IEEE 802.3ba 100G 이더넷 표준에 부합하는 시험을 수행한 바 있다. 2010년 6월 IEEE 802.3ba 이더넷 PHY 표준이 승인됨에 따라 통신업체들로부터 2011년부터 표준에 부합하는 100G PHY 디바이스 및 100G 장비 등 프로젝트가 출시될 것으로 예상된다. 100G 전이중(full duplex) 속도를 지원하는 단일 칩 네트워크 프로세서(EZ NP-5)와 100G PHY칩은 2012년 말경에 상용 칩[6],[14]이 출시될 예정이다. 또한, 100G 이더넷 라인카드는 PCS 레인의 다중화 개념을 표준으로 채용하고 있으며 10 레인의 10G와 4 레인의 25G가 필요하다. 이를 위한 디바이스는 연구실험실 수준이고 아직 상용화되지 않고 있다. 따라서 100G 이더넷 라인 인터페이스를 갖는 이더넷 프



(그림 14) 40G/100G 라인카드 기본 형상 구조

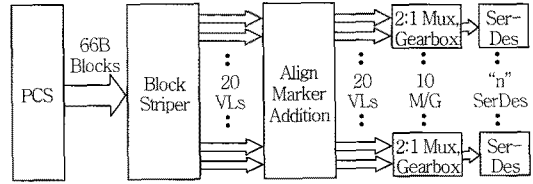
<표 2> 100G 다중레인용 디바이스

10G 4 레인	25G 4 레인
<ul style="list-style-type: none"> 광모듈 접속부인 10G×10 정합 디바이스가 있음(2010년 2분기에 NetLogic 칩 출시 예정) 10G를 지원하는 FPGA가 있음 NP-4(2칩으로 100G 구성) 	<ul style="list-style-type: none"> 현재 광모듈 접속부인 25G×4 정합 25G×4 디바이스가 없음(2011년 2분기 Net-Logic 칩 출시 예정) 현재 25G를 지원하는 FPGA가 없음 NP-4(2칩으로 100G 구성)

로덕트는 2013년~2016년에 본격 상용화될 것으로 전망된다. <표 2>에 10G 10 레인과 25G 4 레인의 PCS 디바이스 전망을 기술한다.

나. 100G PCS 채널

(그림 15)에 예로 Altera에서 제공하는 100G PCS 채널 송수신 경로를 나타내었다[14]. 먼저 MAC 데이터가 64/66B의 연속 스트림으로 인코딩되고 스크램블된다. 66비트 스크램블된 데이터는 라운드 로빈



<자료>: 2009 Altera WP

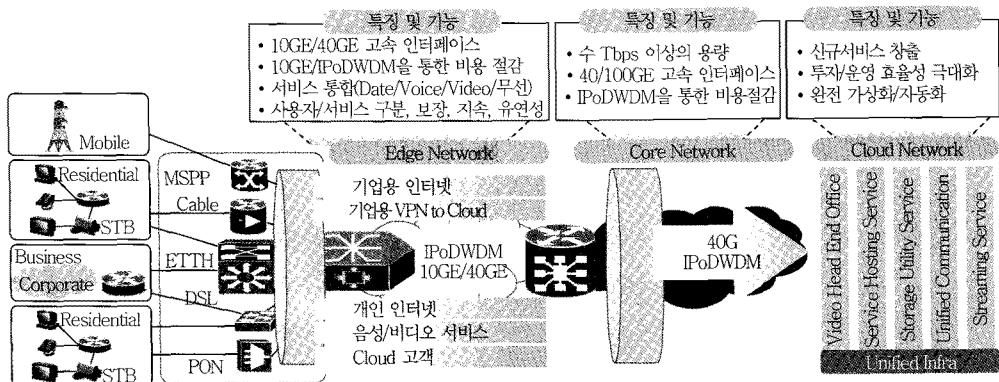
(그림 15) 100G PCS 채널 전송 경로

메커니즘을 통해 20 가상 레인(virtual lane) 위에 스트립(striped) 된다. 마커(marker)가 동시에 가상 라인에 추가된다. 마커는 수신 PCS 블록에서 가상 라인으로부터 데이터를 식별하고, 디스큐(deskew)하고 기록하기 위해 사용된다. 그리고 100G 집합 데이터 스트림을 재조립하기 위해 사용된다. 20개 가상 레인은 궁극적으로 10 레인의 10.3125G PMA로 다중화되고 그 데이터는 역다중된다.

수신 채널은 10 레인으로부터 데이터를 역다중해서 20개 가상 레인에 전달된다. 수신 채널은 디스큐되고 정렬되어 66B 데이터 스트림을 재구축한다.

다. 40G/100G 이더넷 응용

IPTV 등 높은 대역폭을 요구하는 애플리케이션의 증가와 함께 40G/100G 이더넷 기술이 차세대 광대역 서비스 대역폭 요구에 대한 장기적 해결방안의 하나로 제시되고 있다. (그림 16) 시스코의 40G/100G



<자료>: 미래 인터넷 전망과 사업, Cisco

(그림 16) 미래 인터넷 서비스 구조

이더넷의 미래 서비스 전망에 따르면, 40G/100G 고속 이더넷 인터페이스가 에지 네트워크와 코어 네트워크에 필수적으로 요구되고 또한, IPoDWDM(파장 분할 다중화 방식)을 통해 비용절감형 서비스가 가능함을 예측하고 있음을 보여주고 있다.

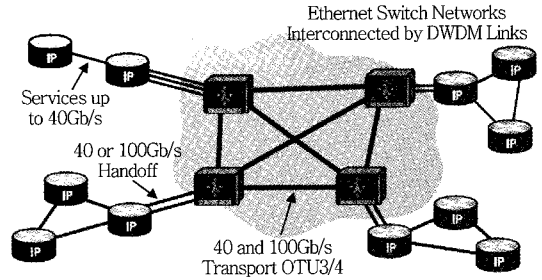
100G OTN 라인카드 구조[15]는 (그림 5)에서 이더넷 PHY 부분이 OTN PHY(진회색 부분)로 대체된다.

IEEE 802.3ba는 2.5G(OTU1)~10G(OTU3) 라인 속도를 갖는 기존 OTN 뿐 아니라 100G OTU4를 통합한 속도에서 OTN 인터페이스를 규정하고 있다. 이를 위해 40G와 100G OTN 표준에서 다음을 목표 [16]로 하고 있다.

- 전이중(full duplex) 동작 지원
- 802.3 MAC 표준을 이용한 802.3ba/이더넷 프레임 포맷 유지
- 802.3ba 표준의 최소 및 최대 프레임 사이즈 유지
- MAC/PCS 서비스 인터페이스에서 10^{-12} BER 지원
- 40G/100G MAC 데이터 속도 지원
- 매체에서 40G/100G 운용을 지원하는 물리계층 규격 지원

100G 이더넷은 (그림 17) 100G 응용[17]에서 보는 바와 같이 OTN 네트워크에 응용 가능하다. OTN은 프레임 헤더 내에 OAM&P 기능의 정보를 전송할 수 있으므로 SONET/SDH가 제공하는 관리 기능을 동일하게 제공한다. 응용분야는 백본 통합, 데이터 센터 네트워크 및 기업 네트워크 그리고 100G 이더넷 전송에서 전송과 이더넷 통합 등에 응용될 수 있다.

현재로서는 OTN 장비는 OAM 등이 고가로 만들어져 그 필요성이나 존재의 가치에 대해 그다지 부각



<자료>: 참고문헌 [17] JISC, Ciena

(그림 17) 40G/100G 이더넷 응용

되지 않고 있는 실정이다. 그러나 구글 TV 등이 등장하면서 현재 대역은 서로 간에 텍스트에서 영상까지 공유하므로 요구대역폭이 더욱 커지고 OTN의 중요성은 향후 더욱 부각될 것이다. 단, 대역폭만 커질 것이 아니라 네트워크 트래픽의 출렁임(fluctuation) 또한 더욱 심해지기 때문에 네트워크 중간에 IDC, 클라우드 컴퓨팅이 많은 부분 해결방안이 될 수 있다. 당분간 이더넷이나 웹 트래픽을 직접 수용하는 것은 이더넷 망이 될 것이고 일부 구간은 초고속 이더넷 네트워크 자체로 가능할 것이다. 그러나 궁극적으로 이더넷 네트워크와 OTN의 융합 구조가 미래 네트워크의 해결책으로 응용될 수 있다.

III. 결론

본 고에서는 최근 세계적으로 뜨거운 쟁점이 되고 있는 IEEE 802.3ba, 40G/100G 이더넷 표준을 기반으로 하는 디바이스 및 프로덕트의 출시 동향을 살펴보고 현재 사용 가능한 상용 칩을 이용한 40G 이더넷 라인카드의 구조와 향후 구현 가능한 100G 이더넷 라인카드의 구조 그리고 40G/100G 이더넷 상의 OTN 네트워크 응용에 대해 고찰하였다.

싱글 40G/100G 이더넷은 싱글 40G/100G 포트와 케이블만 관리하므로 네트워크와 호환성이 높고 관리가 용이하며 그린 에너지, 즉 전력 효율 또한 매

우 높은 장점을 갖는다.

본 고의 40G 이더넷 라인카드는 사용자 요구에 따라 광 대역서비스 기능을 포함할 수 있는 구조로서 이를 위해 광모듈과 40G 이더넷 PHY간 광대역 응용 서비스 기술을 추가하는 기술(하드웨어적/소프트웨어적), 다양한 패킷 타입에 대해 패킷을 탐색 및 분석하고 모델링하는 마이크로 코드 기술, 이더넷 시험 및 검증 기술, IDC용 100G 이더넷 라인카드 40G/100G 이더넷 광 모듈 정합기술, 데이터 센터 또는 백본 집성 40G 업 링크 등의 100G 이더넷 기반 기술이 축적될 수 있다.

ODMA	Output DMA
OTN	Optical Transport Network
PCS	Physical Coding Sublayer
PD	Photo Diode
PFQ	Per Flow Queuing(TM 이라고도 함)
PHY	Physical
PMD	Physical Media Dependant
SerDes	Serial, Deserial
TCAM	Text CAM
TF	Task Force
TM	Traffic Management(internal block)
TOP	Task Optimized Processor
UCC	User Created Contents
WFQ	Weighted Fair Queuing
WREQ	Weighted Random Early Discard
XLAUI	40G Attachment Unit Interface

● 용 어 해 설 ●

XLAUI/CAUI(40/100 Gigabit Attachment Unit Interface):
 칩간 또는 칩과 모듈간 상호연결에 사용, 40Gbps/100Gbps
 PMA 연결을 연장하기 위한 PMA 서비스 인터페이스

약어 정리

BER	Bit Error Rate
CAM	Content Addressable Memory
CAUI	100G Attachment Unit Interface
CIR	Communications Industry Researchers
DMA	Direct Memory Access block
DMUX	De-multiplexer
FEC	Forward Error Correction
HSSG	High Speed Study Group
IDC	Information Data Center
IDMA	Input DMA
IP	Internet Protocol
IPoWDM	IP over Wavelength Division Multiplexing
IT	Information Telecommunication
LPI	Low Power Idle
MAC	Media Access Control
MDIO	Management Data Input Output
MLD	Multi Lane Distribution
MSA	Multi Source Agreement
NIC	Network Interface Card

참고 문헌

- [1] <http://www.idg.co.kr/newscenter/common/new-CommonView.do?newsId=61578>.
- [2] "An Overview: The Next Generation of Ethernet IEEE 802 Plenary," Atlanta, GA, Nov. 12, 2007.
- [3] "Blade Network Technologies," 40G and 100G Ethernet, Technical Brief, 2008, pp.1-3.
- [4] 신중윤, 안계현, 김승환, 김중호, 양충열, 고제수, "40G/100G 이더넷 기술 및 표준화 동향," 전자통신 동향분석, 제24권 제1호, 2009년 2월, pp.32-42.
- [5] IEEE 802.3ba, Part 3: Carrier Sense Multiple Access with CSMA/CD Access Method and Physical Layer Specifications, 22 June 2010.
- [6] "Brocade Reading 100G Ethernet for Data Centers," http://www.networkworld.com/news/2010/081910-brocade-readying-100g-ethernet-for.html?source=NWWNLE_nlt_datacenter_2010-08-24.
- [7] "NP-4 Network Processor," Document number: 27-7884-14, EZ Chip, 2009.
- [8] CFP MSA Draft 1.0, 23 Mar. 2009.
- [9] "Optical Transport Networks for 100G Implementation in FPGA," WP-01115-1.1, Altera FPGA, July 2010, Altera Corporation.
- [10] "Enabling 100 Gigabit Ethernet Implementing

- PCS Lane,” P/N 915-0909-01 Rev B, IXIA, May 2009.
- [11] “Knowledge-based Processors, 2M, 1M, 512K, 256K, and 128K Recoders,” NL11000RA, Feb. 2010, Netlogic.
- [12] “EZ NP-4 Network Processor Architectural Specifications,” Document number: 27-7884-14, 2009.
- [13] “Using 10-Gbps Transceivers in 40G/100G Applications,” Feb. 2010, WP-01080-1.3, Altera Corporation.
- [14] Raza Rizvi, “100G Ethernet and Beyond: Preparing for the Exabyte Internet,” JISC, July 2009, pp.18-19.
- [15] Christian Hermsmeyer, “Towards 100G Packet Processing; Challenges and Technologies,” Bell Labs Technical Journal, DOI: 10.1002/bij.
- [16] Altera White Paper, “Optical Transport Networks for 100G Implementation in FPGA,” July 2010, Altera Corporation.
- [17] Ciena WP-048 6.2009, “The Road to 100G Networking,” 2008, pp.1-5.