

40G/100G 이더넷을 위한 PCS 송수신부 설계 및 기능 검증

정회원 한 경 은*, 김 승 환*, 안 계 현*, 김 광 준*

Design and Verification of PCS Transmitting and Receiving Module for 40/100 Gigabit-Ethernet

Kyeong-Eun Han*, Seung-Hwan Kim*, Kye-hyun Ahn*, Kwangjoon Kim* *Regular Members*

요 약

본 논문에서는 40G/100G 이더넷의 구조적 특징을 고려하여 PCS(Physical Coding Sublayer) 송수신부를 설계하고 로직 시뮬레이션을 통하여 각 모듈의 기능을 검증한다. 이를 위하여 PCS 송수신부의 기능 모듈 및 입출력 신호를 정의하고 HDL 프로그래밍 언어를 사용하여 각 기능을 구현하였다. 설계한 PCS 송수신부는 64B/66B 인코딩과 디코딩, 동작 모드에 따른 스크램블링과 디스크램블링, 송수신 프레임의 유효성 여부 판단 기능, 다중 레인 분배 및 정렬 기능을 제공한다. 설계한 PCS 송수신부의 기능 검증을 위하여 ModelSim 시뮬레이터를 이용하였으며, 입력 데이터에 따른 모듈 동작 및 타이밍 관점에서 기능을 확인하였다. 이를 통하여 설계한 PCS 송수신부의 각 기능 모듈이 올바르게 동작함을 확인할 수 있었다.

Key Words : 40G/100G Ethernet, PCS, optical networks, logic simulation

ABSTRACT

In this paper, we design the PCS(Physical Coding Sublayer) transmitting and receiving module for 40G/100G Ethernet and verify the performance of it through logic simulation. In this work, we defined each function module and internal/external control signals and implemented them using HDL programming language. We also designed 64B/66B encoding/decoding, scrambling/descrambling including operation mode, detection of invalid frames, and multi-lane based distribution/arrangement. It was simulated using ModelSim and verified in terms of the operation and timing according to input data. The simulation result shows that all designed modules in 40G/100G Ethernet are correctly performed.

I. 서 론

최근 인터넷 트래픽이 급격히 증가하고 트리플 플레이 서비스(Triple Play Service: TPS)와 같은 광대역-고품질 기반의 다양한 어플리케이션이 출현함에

따라 이를 수용하기 위한 전송 링크 및 망의 고속화 방안이 요구되고 있다¹⁻³⁾. 이러한 고속화 방안으로 2007년 12월 IEEE 802.3ba Task Force(TF)는 로컬 서버와 인터넷 기간망이라는 서로 다른 두 시장의 요구를 모두 수용할 수 있는 40G/100G 이더넷 표준을

※ 본 연구는 지식경제부 및 정보통신연구진흥원의 IT핵심기술개발사업의 일환으로 수행하였음[2008-F017-01, 100Gbps급 이더넷 및 광전송 기술 개발].

* 한국전자통신연구원 광인터넷연구부 (kehan@etri.re.kr, sammy@etri.re.kr, khyun@etri.re.kr, kjk@etri.re.kr)

논문번호: KICS2010-04-158, 접수일자: 2010년 4월 8일, 최종논문접수일자: 2010년 8월 20일

제안하여 2010년 표준화 완료를 목표로 하고 있다^{4,6)}.

표준에서 제안된 40G/100G 이더넷은 기존 802.3 이더넷 MAC(Media Access Control)과 프레임 구조를 그대로 수용하는 한편 대용량 데이터 전송을 위하여 저가의 고속 데이터 전송이 가능한 다중 레인을 사용한다는 특징을 갖는다. 이를 위하여 표준에서는 40G/100G 이더넷을 위한 공통 물리 계층을 정의하고, 40G 이더넷과 100G 이더넷 각각에 대하여 다수개의 PCS 가상 레인과 링크 당 물리적인 레인의 개수를 명시하고 있다. 따라서 기존 10G 이더넷과는 달리 40G/100G 이더넷 구현을 위해서는 다중 레인의 특성이 반영되어야 하며, 이를 위하여 40G/100G 이더넷 물리 계층, 특히 라인 코딩 및 데이터 송수신을 담당하는 PCS(Physical Coding Sublayer)는 다중 레인 기반의 데이터 전송 및 정렬 기능, 다중 레인을 고려한 동기화, 얼라인먼트 처리 기능을 포함해야 한다.

40G/100G 이더넷 PCS는 현재의 구현 기술 및 환경을 고려하여 다양한 방법으로 구현될 수 있다. 현재 FPGA 및 다양한 장비 업체에서 40G/100G 이더넷 상용화 제품 출시를 준비 중인 반면에 40G/100G 이더넷 관련 연구는 기초 단계에 머물러 있는 실정이다. 따라서 본 논문에서는 40G/100G 이더넷 설계를 위한 고려 사항, 주요 기능 모듈 설계 과정 및 파라미터를 명시함으로써 40G/100G 이더넷 구현을 위한 기초 자료를 제시한다.

본 논문에서는 40G/100G 이더넷의 구조적 특성을 고려하여 PCS의 주요 기능 블록을 설계하고 로직 시뮬레이션을 통한 기능 검증을 수행한다. 이를 위하여 PCS의 주요 기능 모듈 및 입출력 신호를 정의하고 HDL 프로그래밍 언어를 사용하여 각 기능을 구현하였다. 설계한 PCS 송수신부는 64B/66B 인코딩과 디코딩, 모드에 따른 스크램블링과 디스크램블링, 송수신 프레임의 유효성 여부 판단 기능, 다중 레인 분배 및 정렬 기능을 제공한다. 이때 각 기능 모듈은 레인별 제어 블록과 데이터 블록을 각각 분리하여 처리하도록 설계되었으며, 내부 레지스터를 이용하여 각 모듈에서의 처리 지연에 따른 데이터의 출력 시점을 일정하게 유지하도록 하였다.

본 논문의 구성은 다음과 같다. 먼저 2장에서는 40G/100G 이더넷 물리 계층의 구조와 PCS 송수신부 설계 요구 사항에 관하여 기술한다. 3장에서는 40G/100G 이더넷 PCS 송수신부 설계를 위한 각 기능 모듈 및 입출력 신호를 정의하고 HDL 프로그래밍 언어를 기반으로 구현한다. 4장에서는 설계한 PCS의 기능 검증을 위하여 ModelSim 환경에서 시뮬레이션

을 수행하고 그 검증 결과를 분석한다. 마지막으로 5장에서 결론을 맺는다.

II. 40G/100G 이더넷 물리계층

2.1 구조

40G/100G 이더넷은 기본적으로 광케이블에서 전이중 방식을 지원하며 802.3 이더넷 MAC 프레임 송수신, 64B/66B 라인 코딩, 그리고 링크 오류 검출 및 표시 기능을 제공해야 한다. 그림 1은 802.3ba 표준에서 정의된 40G/100GBASE-R 이더넷의 물리 계층 구조와 이에 따른 인터페이스를 나타낸다.

IEEE 802.3ba 표준에서 정의된 40G/100G 이더넷 물리 계층은 크게 RS(Reconciliation), PCS, PMA(Physical Medium Attachment), PMD(Physical Medium Dependent)로 구성되며, FEC(Forward Error Correction)와 AN(Auto-Negotiation)은 선택 사항이다. 또한 40GBASE-R과 100GBASE-R 각각에 대하여 RS와 PCS간 인터페이스는 XLGMII(40Gbps Media Independent Interface)와 CGMII(100Gbps Media Independent Interface)와 PMA와 PMA간 인터페이스는 XLAUI(40Gbps Attachment Unit Interface)와 CAUI(100Gbps Attachment Unit Interface)로 명시하고 있다. 이더넷 물리 계층의 RS는 링크 장애 감지 및 복구 기능과 MAC에서 순차적으로 전송한 데이터 스트림을 XLGMII/CGMII 병렬 데이터 경로에 맞게 변경하여 전달하는 기능을 수행한다. XLGMII/CGMII는 RS와 PCS를 연결하는 논리적 인터페이스로 40Gbps와 100Gbps의 전송 속도 및 높은 확장성을 제공한다. 또한 전이중 방식을 지원하며 양방향에 대하여 각각 64-비트 폭의 독립적인 데이터 경로를 제공한다. 40G/100GBASE-R PCS는 MAC과

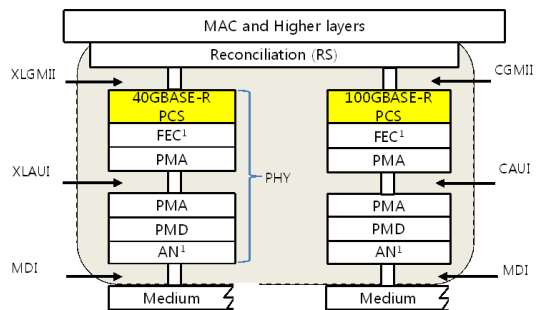


그림 1. 40G/100GBASE-R 이더넷 물리 계층 구조
Fig. 1. Physical layer architecture of 40G/100GBASE-R Ethernet

PMA간 데이터 송수신, 신호 부호화 및 복호화, 얼라인먼트, 동기 검출, 디스큐(deskew) 기능을 수행함으로써 XLGMII/CGMII와 PMA간 전송 속도를 조절하고 전송의 효율성과 신뢰성을 제공한다. 40GBASE-R PCS의 경우 4개의 PCS 가상 레인을 통해 PMA로 전송되며, PMA 간 인터페이스는 4개의 물리적 레인으로 구성된다. 100GBASE-R PCS의 경우에는 20개의 PCS 가상 레인을 통해 PMA로 전송되며 PMA간 인터페이스는 10개의 물리적 레인으로 구성된다. PMA는 PCS 라인 수에 따라 전송된 신호들을 하위 부계층의 가상 또는 물리적 라인 수에 맞게 적절히 조정하는 기능과 비트 기반의 다중화(bit level multiplexing) 기능을 제공한다. 또한 선택적으로 테스트 신호와 루프백(loopback) 기능을 제공한다. 40G/100G 이더넷은 PCS와 PMD 레인의 수와 전송 속도 또는 칩간(chip-to-chip) 인터페이스에 따라 한 개 이상의 PMA 부계층을 사용할 수 있도록 정의하고 있다.

2.2 40G/100G PCS 설계 시 고려 사항

40G/100G 이더넷 PCS는 기존의 이더넷과 다르게 XLGMII/CGMII로부터 수신한 데이터의 비트 정렬을 위하여 64-비트 단일 전송기(transfer)를 사용하고 기어 박스를 포함하지 않는다. 수신한 64 비트 데이터들은 인코딩을 거쳐 66 비트 블록 단위로 스크램블링된 후 각 PCS 가상 레인을 통해 PMA로 전송된다. 라인 별 데이터 전송은 라운드-로빈(Round-Robin) 방식을 기반으로 하며, 라인 별로 수신한 데이터의 정렬(alignment), 디스큐(deskew) 및 동기화 기능은 각 라인 정보에 따라 병렬적으로 처리된다. 또한 PCS는 라인 별 디스큐 및 정렬 기능을 위하여 16,383개의 66 비트 블록마다 얼라인먼트 마커(marker)를 주기적으로 삽입한다. 얼라인먼트 마커는 각 라인마다 고유한 값을 가지며 매 주기마다 모든 레인으로 동시에 삽입된다.

한편 40G/100G 이더넷 PCS에서는 비트 정렬 구조 및 다중 라인 특성에 따라 인코딩/디코딩 블록 포맷과 PCS 테스트 패턴을 새롭게 정의하고 있다. 기존 PCS는 대부분 구형파(square wave) 및 의사난수(pseudo-random), PRBS(Pseudo-Random Binary Sequence) 31을 PCS 테스트-패턴으로 사용하였으나, 40G/100G 이더넷에서는 라인 별 테스트-패턴 발생 및 검사가 가능하도록 스크램블 된 유희(idle) 패턴을 사용한다.

따라서 40G/100G 이더넷 PCS의 기능 설계를 위한 고려 사항은 다음과 같다.

- 40G/100G 이더넷 표준에서 정의된 PCS 블록 포맷을 기반으로 64B/66B 인코딩 및 디코딩 기능을 제공해야 한다.
- 송수신 프레임의 유효성 여부를 판단하고 에러 검출 기능을 제공해야 한다.
- 스크램블러/디스크램블러 내부에 유희 테스트-패턴 발생기와 검사기를 포함해야 한다.
- 스크램블링된 데이터를 다중 레인으로 분배하고, 디스크램블링된 데이터를 다중 레인으로부터 수신하여 정렬하는 기능을 제공해야 한다.
- 얼라인먼트 기능을 제공하기 위하여 매 16,383개의 66 비트 블록 주기마다 라인 별 고유한 값을 갖는 얼라인먼트 마커를 삽입해야 한다.

2.3 40/100GBASE-R PCS 블록 포맷

그림 2는 40G/100G 이더넷 표준에서 정의한 64B/66B 블록 포맷을 나타낸다^[6]. 각 블록은 2 비트의 동기 비트와 8개의 옥텟(octet)으로 구성된 블록 페이로드로 구성된다. 여기서, D는 데이터, O는 순서-집합(ordered-set) (O/), S는 시작 (S/), T는 종료(terminate) (T/), 그리고 C는 나머지 제어 문자를 나타낸다.

블록 포맷에서 동기 비트는 데이터와 제어 블록의 구분을 위한 것으로 데이터는 ‘01’, 제어 블록은 ‘10’ 값으로 설정된다. 데이터 블록은 8개의 데이터 옥텟으로 구성되고, 일반적인 제어 블록은 8 비트의 블록 타입 필드(block type field)와 8개의 7 비트 제어 코드로 구성된다. 반면에 시작(S/), 순서-집합(O/), 종료(T/)를 나타내는 제어 블록은 블록 타입 필드, 제어 코드, 데이터의 조합으로 구성된다. 40G/100G 이더넷에서는 링크 장애 발생 시 링크 상태나 제어 정보 전송 기능을 확장하기 위하여 순서-집합 제어 블록(O/)을 정의하고 있다. 순서-집합 제어 블록은 O 코드 값에 따라 신호 순서-집합(signal ordered-set)과 순차 순서-집합(sequence ordered-set)으로 구분되나 40G/100G 이더넷에서는 순차 순서-집합 블록만을 사용한다. PCS는 클럭 속도 조정을 위하여 연속으로 수신된 두 개의 순차 순서-집합 블록 중 하나를 제거할 수 있다. 시작 블록과 종료 블록은 각각 패킷의 시작과 끝을 구분하기 위한 것으로 가변 길이를 갖는 이더넷 패킷 특성에 따라 다양한 형태의 종료 블록이 존재한다. 이때, 블록 포맷에서 빗금 처리된 단일 비트는 0 값으로 수신 시 무시된다.

표 1은 제어 문자와 제어 코드 간 대응 값을 나타낸다^[6]. PCS는 MII 제어코드를 기반으로 블록 필드

Input Data	Sync	Block Payload											
Bit Position Data Block Format	0 1	2 65											
D ₀ D ₁ D ₂ D ₃ D ₄ D ₅ D ₆ D ₇	01	D ₀	D ₁	D ₂	D ₃	D ₄	D ₅	D ₆	D ₇				
Control Block Formats		Block Type Field											
C ₀ C ₁ C ₂ C ₃ C ₄ C ₅ C ₆ C ₇	10	0x1e	C ₀	C ₁	C ₂	C ₃	C ₄	C ₅	C ₆	C ₇			
S ₀ D ₁ D ₂ D ₃ D ₄ D ₅ D ₆ D ₇	10	0x78	D ₁	D ₂	D ₃	D ₄	D ₅	D ₆	D ₇				
O ₀ D ₁ D ₂ D ₃ Z ₄ Z ₅ Z ₆ Z ₇	10	0x4b	D ₁	D ₂	D ₃	O ₀	0x000_0000						
T ₀ C ₁ C ₂ C ₃ C ₄ C ₅ C ₆ C ₇	10	0x87					C ₁	C ₂	C ₃	C ₄	C ₅	C ₆	C ₇
D ₀ T ₁ C ₂ C ₃ C ₄ C ₅ C ₆ C ₇	10	0x99	D ₀				C ₂	C ₃	C ₄	C ₅	C ₆	C ₇	
D ₀ D ₁ T ₂ C ₃ C ₄ C ₅ C ₆ C ₇	10	0xaa	D ₀	D ₁			C ₃	C ₄	C ₅	C ₆	C ₇		
D ₀ D ₁ D ₂ T ₃ C ₄ C ₅ C ₆ C ₇	10	0xb4	D ₀	D ₁	D ₂			C ₄	C ₅	C ₆	C ₇		
D ₀ D ₁ D ₂ D ₃ T ₄ C ₅ C ₆ C ₇	10	0xcc	D ₀	D ₁	D ₂	D ₃			C ₅	C ₆	C ₇		
D ₀ D ₁ D ₂ D ₃ D ₄ T ₅ C ₆ C ₇	10	0xd2	D ₀	D ₁	D ₂	D ₃	D ₄			C ₆	C ₇		
D ₀ D ₁ D ₂ D ₃ D ₄ D ₅ T ₆ C ₇	10	0xe1	D ₀	D ₁	D ₂	D ₃	D ₄	D ₅			C ₇		
D ₀ D ₁ D ₂ D ₃ D ₄ D ₅ D ₆ T ₇	10	0xff	D ₀	D ₁	D ₂	D ₃	D ₄	D ₅	D ₆				

그림 2. 40G/100G 이더넷을 위한 64B/66B 블록 포맷
Fig. 2. 64B/66B block format for 40G/100G Ethernet

표 1. 제어코드
Table 1. Control codes

Control Character	Notation	MII Control Code	40/100G BASE-R O code	40GBASE-R & 100GBASE-R Control Code
idle	/I/	0x07		0x00
start	/S/	0xfb		Encoded by block type field
terminate	/T/	0xfd		Encoded by block type field
error	/E/	0xfe		0x1e
Sequence ordered_set	/Q/	0x9c	0x0	Encoded by block type 0x4b plus O code, control codes are set to 0x00
Signal ordered_set*	/Fsig/	0x5c	0xf	Encoded by block type 0x4b plus O code, control codes are set to 0x00

* Reserved for INCITS T11 Fibre Channel use

타입과 PCS 블록 포맷을 결정한다. 표 1에서 유희 제어 블록(I/I)은 블록 필드 타입이 '0x00'으로 설정되며 클럭 속도 조절을 위하여 추가되거나 제거될 수 있다. 유희 문자의 추가는 송수신 데이터가 없는 경우 유희 블록 또는 순서-집합 블록 뒤에 가능하며, 종료 문자 뒤에 오는 첫 4개의 유희 문자를 제외하고 제거 가능하다.

III. 40G/100G 이더넷 PCS 송수신부 설계

3.1 PCS 송수신부 기능 블록도

설계한 40G/100GBASE-R PCS 송수신부는 크게 인코딩과 디코딩, 스크램블링과 디스크램블링, 그리고 다중 레인 분배 및 정렬 기능으로 구성되며, 열라인먼트 기능 구현은 포함하지 않는다. 인코딩과 디코딩은 전송 데이터의 클럭 보상, 신뢰성 있는 에러 검출과 효율적인 전송을 제공하기 위한 기능이다. 이를 위하여 전송 에러 검출이 용이하고 높은 전송 밀도와 3%의 낮은 오버헤드(overhead)를 제공하는 64B/66B 코딩 기법이 사용된다. 스크램블링과 디스크램블링은 연속되는 데이터 입력에서 타이밍 정보를 유지하고, 주

기적인 데이터 패턴 반복에 의한 혼변조 현상을 방지하기 위한 기능이다. 이를 위하여 자기-동기화 스크램블러(self-synchronizing scrambler)가 사용되며, 내부에 유휴 패턴 발생기 및 검사기를 포함한다. PCS 라인 분배 송수신부는 스크램블된 데이터를 가상 레인을 통해 분배하고, PMA로부터 수신된 데이터 동기 및 데이터 프레임 조합을 수행한다.

그림 3은 설계한 40G/100G 이더넷 PCS의 기능 블록도 및 데이터 경로를 나타낸다. XLGMII/CGMII 인터페이스를 통해 PCS 송수신부로 전송된 64 비트 데이터(MII_TXD)와 8 비트 제어 코드(MII_TXC)는 66 비트 인코딩 및 스크램블링을 거쳐 PCS 라인 분배부로 전송된다. PCS 라인 분배부는 수신한 66 비트 데이터를 라운드 로빈 방식으로 각 레인에 분배하여 PMA로 전송한다. 반면에 PMA로부터 수신된 66 비트 데이터는 PCS 라인 분배부에서 라인별 동기 및 정렬이 완료된 후 디스크램블링과 디코딩 과정을 거쳐 64 비트 데이터(MII_RXD)와 8 비트 제어 코드(MII_RXC)로 변환되어 XLGMII/CGMII 인터페이스로 전송된다. 이때, PCS 송수신부와 PMA 간 PCS 가상레인의 수는 40G 이더넷의 경우 4개, 100G 이더넷의 경우 20개로 정의된다.

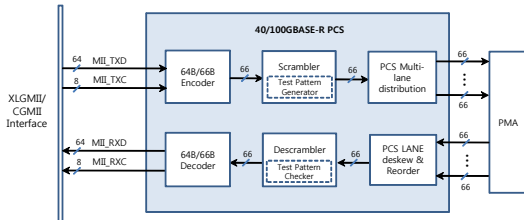


그림 3. 40G/100GBASE-R PCS 블록도 및 데이터 경로
Fig. 3. Block diagram of 40G/100GBASE-R PCS and data path

3.2 PCS 송수신부 기능 모듈 설계

3.2.1 송신부

그림 4는 설계한 40G/100G 이더넷 PCS 송신부 기능 블록 및 입출력 신호를 나타낸다. PCS 송신부는 인코더, 스크램블러, PCS 다중 라인 분배기로 구성되며, 각 모듈의 기능 및 특징은 다음과 같다.

- 인코더: 수신한 데이터를 블록 바이트 별로 처리하여 블록 타입을 결정, XLGMII/CGMII 코드를 PCS 코드로 변환한다. 또한 송신 프레임 블록의 유효성 검사 및 에러 처리를 수행한다.

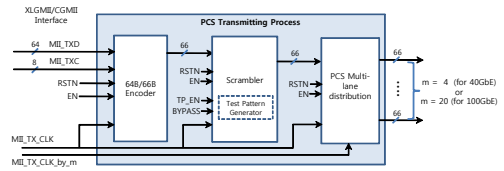


그림 4. 40G/100G 이더넷 PCS 송신부 기능 블록
Fig. 4. Function block diagram of 40G/100GBASE-R PCS transmitting process

- 스크램블러: 제어 신호에 따라 정상 모드(normal mode) 또는 테스트 패턴 모드(test-pattern mode)에서 동작, 바이패스(bypass)기능과 유휴(idle) 패턴을 생성하는 테스트-패턴 발생기를 포함한다.
- PCS 다중 라인 분배기: 스크램블러로부터 전송된 데이터를 라운드-로빈 방식을 기반으로 각 라인으로 분배 전송한다.

인코더는 64비트 MII 데이터(MII_TXD)와 8비트 제어 신호(MII_TXC)를 수신한다. 설계된 인코더는 데이터 타입 디코더 모듈, 블록 타입 생성 모듈, 프레임 생성 모듈로 구성된다. 데이터 타입 디코더 모듈은 수신된 데이터와 제어 신호에 대하여 각 8 비트 서브 블록 단위로 데이터 타입을 판별하고 그 결과를 블록 타입 생성 모듈로 전송한다. 또한 이를 통하여 유효 프레임 여부를 판별하고 에러를 검출한다. 블록 타입 생성 모듈에서는 블록 타입에 따라 블록 타입 필드와 페이로드를 구성하여 PCS 블록을 생성한다. 프레임 생성 모듈은 PCS 블록 타입에 따라 동기 필드를 추가하여 66 비트 PCS 프레임을 생성하고 스크램블러로 전송한다.

인코더로부터 전송된 데이터는 스크램블러에 입력되는 바이패스(bypass) 신호(BYPASS_SCR)에 따라 PCS 다중 라인 분배기로 바이패스(BYPASS_SCR = '1')되거나 스크램블러에서 처리 (BYPASS_SCR = '0')된다. 스크램블러는 스크램블링 모듈과 테스트-패턴 발생 모듈로 구성되며, 테스트-패턴 인에이블(enable) 신호(TP_EN)에 따라 정상 모드와 테스트-패턴 모드에서 동작한다. 정상 모드인 경우(TP_EN = '0'), 스크램블링 모듈은 동기 비트를 제외한 64 비트 페이로드에 대하여 연속적으로 스크램블링을 수행한 후 동기 헤더 비트를 추가하여 66 비트 스크램블된 데이터를 출력한다. 테스트-패턴 모드인 경우(TP_EN = '1'), 테스트-패턴 발생 모듈은 블록 타입이 '0x1e'이고 나머지 비트 값이 모두 '0'인 64 비트 유휴 테스트-패턴을 생성하여 스크램블링 모듈로 전송한다. 스크램

블링 모듈은 유휴 테스트-패턴에 대하여 연속적으로 스크램블링을 수행한 후 동기 비트 '10'을 추가하여 66 비트 스크램블링된 데이터를 출력한다.

스크램블러로부터 바이패스(bypass) 또는 스크램블링된 데이터는 PCS 다중 레인 분배기로 전송된다. PCS 가상 레인은 40GBASE-R의 경우 4개(m=4), 100GBASE-R의 경우 20개(m=20)로 구성된다. 설계한 PCS 다중 레인 분배기는 PCS 초기화 시 40GBASE-R 또는 100GBASE-R로 설정된 값에 따라 제공하는 가상 레인의 개수를 결정한다. 데이터는 66 비트 블록 단위로 각 레인을 통해 동시에 전송되며 블록 별 전송 레인은 라운드-로빈 방식을 기반으로 한다.

PCS 송신부로 입력되는 기본 클럭(MII_TX_CLK)은 40GBASE-R의 경우 625 MHz, 100GBASE-R의 경우 1562.5 MHz가 제공되며, PCS 다중 레인 분배기로 입력되는 클럭(MII_TX_CLK_by_m)은 156.25 MHz로 동일하게 제공된다. 표 2는 40G/100G 이더넷 PCS 송신부 구현을 위하여 정의된 입출력 신호를 나타낸다.

표 2. 40G/100G 이더넷 PCS 송신부 입출력 신호
Table 2. Input and output signals of 40G/100GBASE-R PCS transmitting process

Port	I/O	Description
MII_TX_CLK	Input	Clocks all the encoder and scrambler logic
EN (encoder)	Input	Enables the encoder (active High)
RSTN (encoder)	Input	Asynchronous reset (active High)
MII_TXC[7:0]	Input[7:0]	MII transmit control bus
MII_TXD[63:0]	Input[63:0]	MII transmit data bus
TP_EN	Input	Enables the test-pattern (active High)
RSTN (scrambler)	Input	Asynchronous reset (active High)
EN (scrambler)	Input	Enables the scrambler (active High)
BYPASS_SCR	Input	Indicates the bypass of scrambler
RSTN(PCS multi-lane distribution)	Input	Asynchronous reset (active High)
EN (PCS multi-lane distribution)	Input	Enables PCS multi-lane distribution (active High)
TXD_PCS_n [65:0]	Output [65:0]	Data bus for each PCS lanen (n = 0-3 or 0-19)

3.2.2 수신부

그림 5는 설계한 40G/100G 이더넷 PCS 수신부 기능 블록 및 입출력 신호를 나타낸다. PCS 수신부는 디코더, 디스크램블러, PCS 다중 레인 조정기로 구성되며, 각 모듈의 기능 및 특징은 다음과 같다.

- 디코더: 수신한 데이터를 블록 바이트 별로 처리하여 블록 타입을 결정, PCS 코드를 XLGMII/CGMII 코드로 변환한다. 또한 수신 프레임 블록의 유효성 검사 및 에러 처리를 수행한다.
- 디스크램블러: 제어 신호에 따라 정상 모드(normal mode) 또는 테스트-패턴 모드(test-pattern mode)에서 동작, 바이패스(bypass)기능과 유휴(idle) 패턴을 검사하는 테스트-패턴 검사기를 포함한다.
- PCS 다중 레인 조정기: PMA로부터 전송된 데이터를 라운드-로빈 방식을 기반으로 각 레인으로 분배 조정한다.

설계된 디코더는 디스크램블된 66 비트 데이터를 XLGMII/CGMII 데이터(MII_RXD)와 제어 신호(MII_RXC)로 복원하기 위하여 데이터 타입 디코더 모듈, 블록 타입 생성 모듈, 프레임 및 제어 신호 생성 모듈로 구성된다. 데이터 타입 디코더 모듈은 수신한 데이터에 대하여 각 8 비트 서브 블록 단위로 데이터 타입을 판별하고 그 결과를 블록 타입 생성 모듈로 전송한다. 또한 이를 통하여 유효 프레임 여부를 판별하고 에러를 검출한다. 블록 타입 생성 모듈에서는 블록 타입에 따라 블록 타입 필드와 페이로드를 구성하여 XLGMII/CGMII 데이터 블록을 생성한다. 프레임 및 제어 신호 생성 모듈은 결정된 블록 타입에 따라 8 비트 제어 신호를 생성하여 XLGMII/CGMII 인터페이스로 전송한다.

설계한 디스크램블러는 디스크램블링 모듈과 테스트-패턴 검사 모듈로 구성되며, 테스트-패턴 인에이블(enable) 신호(TP_EN)에 따라 정상 모드와 테스트-패턴 모드에서 동작한다. 정상 모드인 경우(TP_EN = '0'), 디스크램블링 모듈은 동기 비트를 제외한 64 비

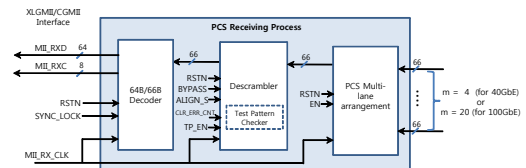


그림 5. 40G/100G 이더넷 PCS 수신부 기능 블록
Fig. 5. Function block diagram of 40G/100GBASE-R PCS receiving process

트 페이로드에 대하여 연속적으로 디스크램블링을 수행한 후 동기 헤더 비트를 추가하여 66 비트 디스크램블된 데이터를 출력한다. 테스트-패턴 모드인 경우 (TP_EN = '1'), 디스크램블링 모듈은 수신한 데이터의 동기 헤더 비트를 제외한 64 비트 페이로드를 디스크램블링하여 테스트-패턴 검사 모듈로 전송한다. 테스트-패턴 검사 모듈은 디스크램블링된 비트들을 검사하여 출력 패턴이 유티 패턴과 일치하는지를 검사하여 판별한다. 유티 테스트-패턴이 아닌 블록에 대해서는 에러로 처리하고 에러 카운터를 증가시킨다. 디스크램블러는 바이패스 신호(BYPASS_DESCR)가 '0'일 때, 디스크램블링 기능을 수행한다.

PMA로부터 전송된 m개의 66비트 데이터는 PCS 다중 레인 조정기에서 라운드-로빈 방식에 따라 순서대로 나열되어 66 비트 단위 블록으로 디스크램블러로 전송된다. 즉, PCS 다중 레인 조정기는 다중 레인을 통해 수신한 264 비트(40GBASE-R) 또는 1,320 비트(100GBASE-R) 데이터를 66 비트 단위로 분리하여 디스크램블러로 전송한다. 표 3은 40G/100G 이더넷 PCS 수신부 구현을 위하여 정의된 입출력 신호를 나타낸다.

표 3. 40G/100G 이더넷 PCS 수신부 입출력 신호
Table 3. Input and output signals of 40G/100GBASE-R PCS receiving process

Port	I/O	Description
MII_RX_CLK	Input	Clocks all the descrambler logic
TP_EN	Input	Enables the test-pattern (active High)
RSTN (scrambler)	Input	Asynchronous reset (active High)
CLR_ERR_CNT	Input	Reset TP_ERR_CNT on MDIO read
ALIGN_S	Input	Indicates completion of synchronization and alignment for all lanes
BYPASS	Input	Indicates the bypass of descrambler
TXD_PCS_n [65:0]	Input	Data bus from each PCS lanen (n = 0-3 or 0-19)
RSTN (decoder)	Input	Asynchronous reset (active High)
SYNC_LOCK	Input	Indicates when data is synchronized (active High)
MII_TXC[7:0]	Output	MII received control bus
MII_TXD[63:0]	Output	MII received data

IV. 기능 검증

설계한 40G/100G 이더넷 PCS 송수신부의 기능 검증을 위하여 ModelSim을 이용한 로직 시뮬레이션을 수행하였다. 이를 위하여 Verilog HDL를 이용하여 각 기능을 구현하고 다음과 같은 검증 시나리오를 기반으로 각 모듈의 기능 검증을 수행하였다.

- 시나리오-1: 정상적인 포맷을 갖는 프레임을 전송하는 경우
- 시나리오-2: PCS 프레임 형식에 위배되는 비정상 프레임을 전송하는 경우

시나리오-1에서는 검증 데이터로서 40G/100GBASE-R PCS 표준을 따르는 시작 블록, 두 개의 데이터 블록, 종료 블록을 사용한다. 시나리오-2에서는 에러 타입 블록에 대한 검증을 위하여 유효하지 않은 시작 블록, 정상적인 두 개의 데이터 블록과 하나의 종료 블록을 사용한다. 이때, 표 4와 같이 정의된 '시작-데이터-데이터-종료' 패턴은 주기적으로 반복 입력된다.

또한 각 모듈의 기능 검증은 구현된 PCS 송수신부를 각각 루프-백으로 연결하고, 시뮬레이션을 통해 송수신부의 입출력 값을 확인함으로써 수행된다. 그림 6은 기능 검증을 위한 루프-백 블록도를 나타낸다.

그림 7은 MII 데이터 및 제어 신호 입력에 따른 인코딩 결과를 보여준다. MII 인터페이스로부터 '시작-

표 4. XLGMII/CGMII 입력 제어 신호 및 데이터
Table 4. XLGMII/CGMII input control and data

Input data	XLGMII/CGMII TXC	XLGMII/CGMII TXD
Start	8'h01	64'h0e380577200008fb
Start(invalid)	8'h11	64'h0e380555200008fb
Data	8'h00	64'h450008000000008b
Data	8'h00	64'h1b0000661c280000
Terminate	8'hFE	64'h0707070707fd79

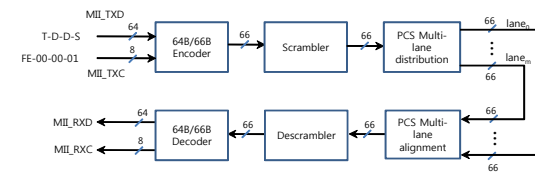


그림 6. 기능 검증을 위한 루프-백 블록도
Fig. 6. Block diagram of loop-back for verification

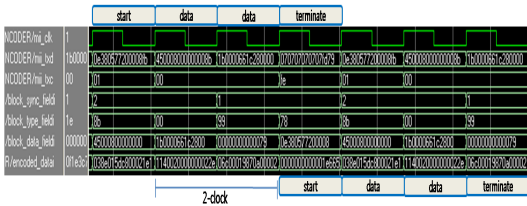


그림 7. MII 데이터 및 제어 신호 입력에 따른 인코딩 결과 (시나리오-1)
Fig. 7. Encoding results of MII input data and control signal (Scenario-1)

데이터-데이터-종료' 블록(mii_txd)과 이에 따른 제어 신호(mii_txc)가 입력되면, 인코더는 2 클럭 후 동기 필드(block_sync_field), 블록 타입 필드(block_type_field), 데이터 필드(block_data_field)를 생성하고, 이를 결합하여 1 클럭 후 66 비트 PCS 블록(encoded_data)을 출력한다. 먼저, 입력된 시작 블록 '0e380577200008fb'은 2 클럭 후, 동기 비트 '1(2 진수 01)', 블록 타입 필드(block_type_field) '0x78', 데이터 필드(block_data_field) '0e380577200008' 값으로 출력된다. 이 값들은 하나의 66비트 PCS 프레임으로 조립되어 1 클럭 후 '038e015dc800021e1' 값으로 출력되는 것을 확인할 수 있다. 또한 그림 7에서 입력된 데이터와 종료 블록 역시 인코딩 예상 값으로 출력되는 것을 확인할 수 있다. 그림 7의 입출력 데이터는 오른쪽이 최하위 비트를 나타낸다. 또한 2진수는 10진수 값으로 16진수는 그대로 표현된다.

그림 8은 디스크램블링 및 디코딩 수행 후 MII 인터페이스로 전송되는 데이터와 제어 신호 출력을 나타낸다. 인코딩 된 데이터(그림 7)는 스크램블링 수행 후 루프-백을 통해 디스크램블러로 전송되고, 디스크램블링 데이터(tx_desc)가 디코더의 입력 데이터가 된다. 디스크램블링된 입력 데이터(tx_desc) '시작(038e015dc800021e)-데이터(1140020000000022e)-데이터(06c00019870a00002)-종료(000000000001e665)' 블록은 디코딩되어 각각 64

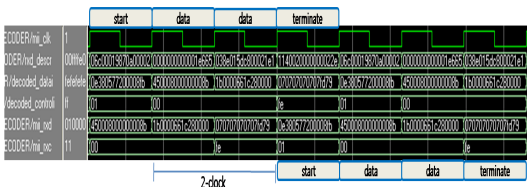


그림 8. 디스크램블링 및 디코딩 수행 후 MII 데이터 및 제어 신호 (시나리오-1)
Fig. 8. MII data and control signal after descrambling and decoding (Scenario-1)

비트 MII 데이터(mii_rxd)인 '0e380577200008fb', '450008000000008b', '1b0000661c280000', '070707070707 fd79' 값으로 출력된다. 또한 이에 대응하는 8 비트 MII 제어 신호(mii_rxc) '00000001', '00000000', '00000000', '11111110' 를 생성하는 것을 확인할 수 있다. 출력된 MII 데이터(mii_rxd)와 제어 코드(mii_rxc)는 2 클럭 후 MII 인터페이스로 전송된다.

그림 9는 유효하지 않은 프레임 포맷이 입력되었을 때 인코딩 및 디코딩 결과를 나타낸다. 먼저, MII로부터 유효하지 않은 시작 블록(mii_txd)과 제어 신호(mii_txc)로 '0e380555200008fb' 와 '00010001'이 입력되고, 그 뒤로 정상적인 '데이터-데이터-종료' 블록이 입력된다. 인코더는 제어 신호와 프레임 포맷을 검사하여 유효하지 않은 패킷으로 판단, 에러 블록으로 처리한다. 따라서 2 클럭 후에 유효하지 않은 시작 블록은 블록 타입 필드(block_type_field)가 '1e'이고 데이터(block_data_field) 값이 '3c78f1e3c78f1e' 로 인코딩되는 것을 확인할 수 있다. 인코딩된 데이터는 스크램블링, 디스크램블링을 거쳐 디코더로 전송된다. 이때, 디코더는 수신한 블록 타입을 판단하여 에러로 처리하므로 MII 데이터(mii_rxd)와 제어 신호(mii_rxc)를 각각 'fefefefefefefef'와 '11111111'로 디코딩하게 된다. 이는 유효하지 않은 시작 블록이 에러로 처리됨에 따라 일련의 송수신 프레임 순서에서 시작 블록 없이 데이터와 종료 블록이 수신되었으므로 종료 블록도 에러로 처리된 것이다.

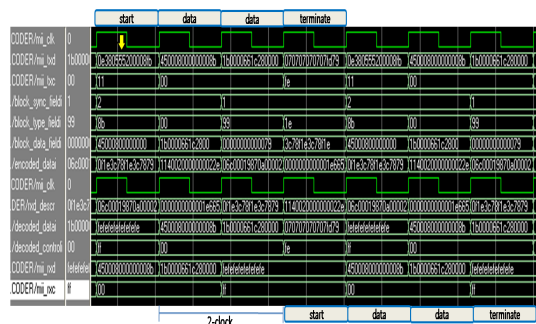


그림 9. 유효하지 않은 데이터 입력에 따른 인코딩 및 디코딩 결과 (시나리오-2)
Fig. 9. Encoding and decoding results according to invalid input data (Scenario-2)

V. 결론

본 논문에서는 40G/100G 이더넷의 구조적 특성을

고려하여 PCS의 주요 기능 블록을 설계하고 로직 시뮬레이션을 통한 기능 검증을 수행하였다. 이를 위하여 PCS의 주요 기능 모듈 및 입출력 신호를 정의하고 HDL 프로그래밍 언어를 사용하여 각 기능을 구현하였다. 설계된 PCS 송수신부는 인코더와 디코더, 테스트-패턴 모드를 지원하는 스캐블러 및 디스크램블러, PCS 다중 레인 분배기 및 조정기로 구성된다. PCS 송수신부에서 MII로부터 입력된 64 비트 데이터는 인코딩과 스캐블링된 후 66 비트 블록 단위로 PCS 다중 레인 분배기를 통해 병렬 전송되고, PCS 수신부에서 다중 레인을 통하여 PMA로부터 수신한 66 비트 데이터는 PCS 다중 레인 조정기에서 일렬로 정렬되고 디스크램블링과 디코딩을 거쳐 MII 인터페이스로 전송된다. 설계된 40G/100G 이더넷 PCS 송수신부의 기능 검증을 위하여 ModelSim 6.0 환경에서 로직 시뮬레이션을 수행하였으며, 시뮬레이션을 통하여 설계한 PCS 송수신부의 각 기능 모듈이 올바르게 동작함을 확인할 수 있었다.

참 고 문 헌

- [1] J. D'Ambrosia, "40 Gigabit Ethernet and 100 Gigabit Ethernet: The Development of a Flexible Architecture," IEEE Communications Magazine, Vol.47, pp.S8-S14, 2009.
- [2] G. Lehmann and R. H. Derksen, "100 Gigabit Ethernet Transmission - Physical Layer Issues," in Proceedings of OFC/NFOEC2007, pp.1-3, March 2007.
- [3] M. Duelk and M. Zirngibl, "100 Gigabit Ethernet - Applications, Features, Challenges," in Proceedings of INFOCOM2006, pp.1-5, April 2006.
- [4] The Road to 100G Networking, Ciena Corp., Linthicum, MD, 2008.
- [5] An Overview of Next-Generation 100 and 40 Gigabit Ethernet Technologies, Ixia Worldwide Headquarters, Calabasas, CA, 2008.
- [6] Part3: Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications, IEEE std 802.3ba TM/D2.0, 2009.

한 경 은 (Kyeong-Eun Han)

정회원

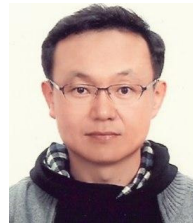


2001년 2월 전북대학교 컴퓨터 공학과
2003년 2월 전북대학교 컴퓨터 공학과 석사
2008년 2월 전북대학교 컴퓨터 공학과 박사
2008년 8월~현재 한국전자통신연구원 연구원

<관심분야> 광통신, 네트워크 프로토콜, Green-IT

김 승 환 (Seung-Hwan Kim)

정회원



1986년 2월 연세대학교 전자공학과 학사
1988년 2월 연세대학교 전자공학과 석사
1988년~1999년 대우전자 반도체연구소 선임연구원
1999년~현재 한국전자통신연구원 책임연구원

<관심분야> 광전송기술, 가입자망기술, ASIC기술

안 계 현 (Kye-hyun Ahn)

정회원



1996년 2월 전북대학교 컴퓨터 공학과
1998년 2월 전북대학교 컴퓨터 공학과 석사
2003년 2월 전북대학교 컴퓨터 공학과 박사
2003년 8월~현재 한국전자통신연구원 선임연구원

<관심분야> 이더넷, 네트워크 프로토콜

김 광 준 (Kwangjoon Kim)

정회원



1981년 2월 서울대학교 물리학과
1983년 2월 서울대학교 물리학과 석사
1993년 6월 미국 Ohio State University 물리학과 박사
1984년 3월~현재 한국전자통신연구원 광전송기술연구팀 팀장 (책임연구원)

<관심분야> 광전송, Optical Networks, WDM