

IDC 네트워크에 적합한 OIF 800ZR 기술 소개

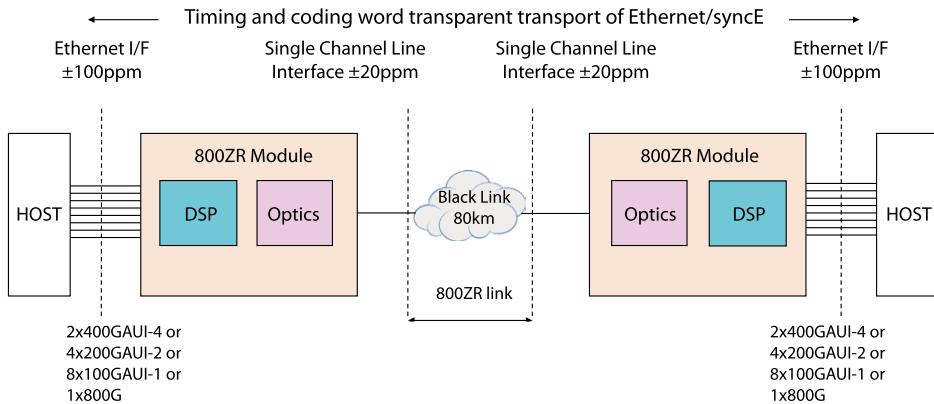
전경규 광전송(PG201) 부의장, 한국전자통신연구원 광네트워크연구실 연구전문위원

1. 머리말

인터넷 트래픽이 증가함에 따라 이를 저장하고 처리하는 인터넷 데이터센터(IDC, Internet Data Center) 용량이 증가하고, IDC 간에 주고 받는 트래픽량 또한 증가하기 때문에 전송 용량이 큰 전달 장치가 필요하다.

시내국간용 메트로 전송망과 시외국간 백본

전송망과 관련된 전달망 장치 표준은 ITU-T에서 개발되어 현재 800G급 표준이 진행되고 있다. 건물내 혹은 캠퍼스와 같이 비교적 짧은 구간에서는 이더넷 기술 표준이 적용되어 IEEE 802.3에서는 현재 800G급과 1.6T급 MAC/PHY 기술 개발이 진행되고 있다. ITU-T에서 개발된 OTN(Optical Transport Network) 표준 기반 장치는 최근 급증하는 IDC간 트래픽 전



*출처 : Draft Working Document: oif2021.144.12

[그림 1] 800ZR 참조 다이어그램

달에 비효율적이고, IEEE 802.3에서 개발 중인 이더넷 기술로는 80km 이상 전송할 수 없어 IDC 네트워크에 적합하지 않다. 2020년 3월 OIF에서는 IDC망에 적합한 400G 신호 전송 장치 규격 OIF-400ZR-01.0 'Implementation Agreement 400ZR'을 발간하여 400G급 이더넷 신호를 최대 120km까지 전달할 수 있도록 하였다. 이어서 800G급 신호 전달용 규격 'Implementation Agreement for 800ZR Coherent Interface' 문서가 2023년 4분기 최종 승인되어 발간될 예정이다.

본 동향서는 OIF에서 현재 개발 중인 800ZR 기술을 소개하기 위해 시스템의 참조 모델, 프레임 구조, 오버헤드 기능, 다중화 방법, 시스템 블록 다이어그램에 대해 기술하고 맺음말로 끝을 맺는다.

2. 800ZR 참조 다이어그램

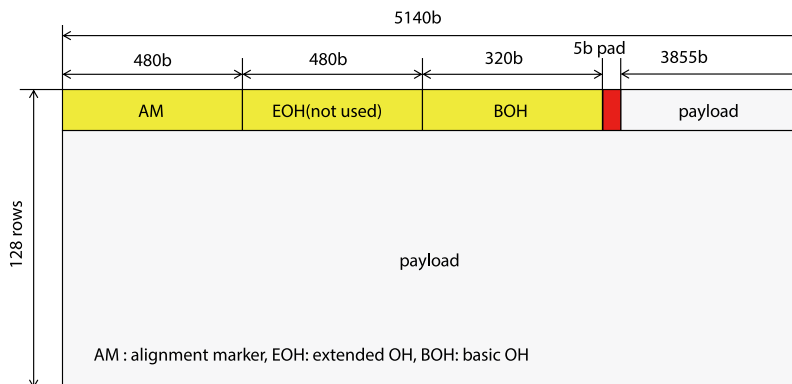
[그림 1]은 800ZR의 외부 인터페이스, 800ZR의 주요 기능 모듈, 인터페이스에 따라 요구되는 클럭 허용 오차, 광링크가 포함된 800ZR 참

조 다이어그램이다. 800ZR은 100G, 200G, 400G, 800G 이더넷 신호를 종속 신호로 수용할 수 있다. 즉, 8개의 100G, 4개의 200G, 2개의 400G 그리고 1개의 800G급 이더넷 신호를 수용할 수 있으나 이종의 이더넷 신호를 혼합 수용하지 않는다. 두 장치간 80km 이상 신호를 전송하기 위해서는 클럭 허용 오차가 $\pm 20\text{ppm}$ 이하가 되어야 하므로 800ZR 프레임 구조는 TDM(Time Division Multiplex) 형태가 되어야 한다. 클럭 허용 오차가 $\pm 100\text{ppm}$ 인 이더넷 신호는 GMP(Generic Mapping Procedure)를 통해 800ZR 프레임으로 매핑되어 TDM화 된다. 이렇게 함으로써 패킷인 이더넷 신호가 TDM프레임에 실려 중계기 없이 80km 이상 전송 가능하게 된다.

3. 800ZR 구조 및 오버헤드 기능

3.1 800ZR 프레임 구조

[그림 2]는 ITU-T G.709.1 FlexO 프레임을 기반으로 하는 800ZR에 적용되는 100G ZR 인스턴스의 프레임 구조이다. 여기서 8개의 100G ZR



*출처 : Draft Working Document: oif2021.144.12

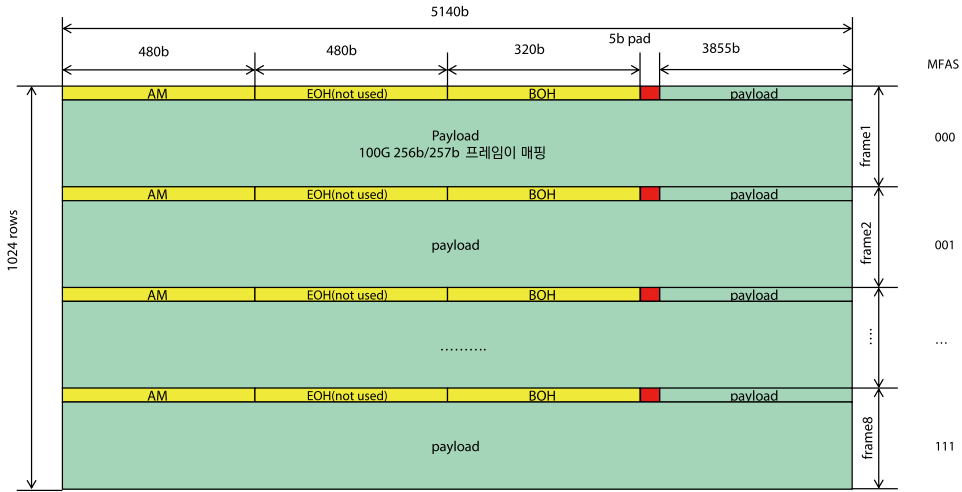
[그림 2] 100G ZR 인스턴스 프레임 구조

인스턴스가 순차적 인터리빙 방식으로 다중화되어 하나의 800ZR 프레임이 만들어진다. 100G ZR 인스턴스 프레임은 가로 5,140비트, 세로 128열로 구성되고 오버헤드 비트 수는 1,280비트로 FlexO와 동일하다. 그러나 FlexO 페이로드 영역에 5 비트의 패드를 추가하여 프레임을 구성하는 것이 FlexO와 다르다.

오버헤드는 AM(Alignment Marker)용으

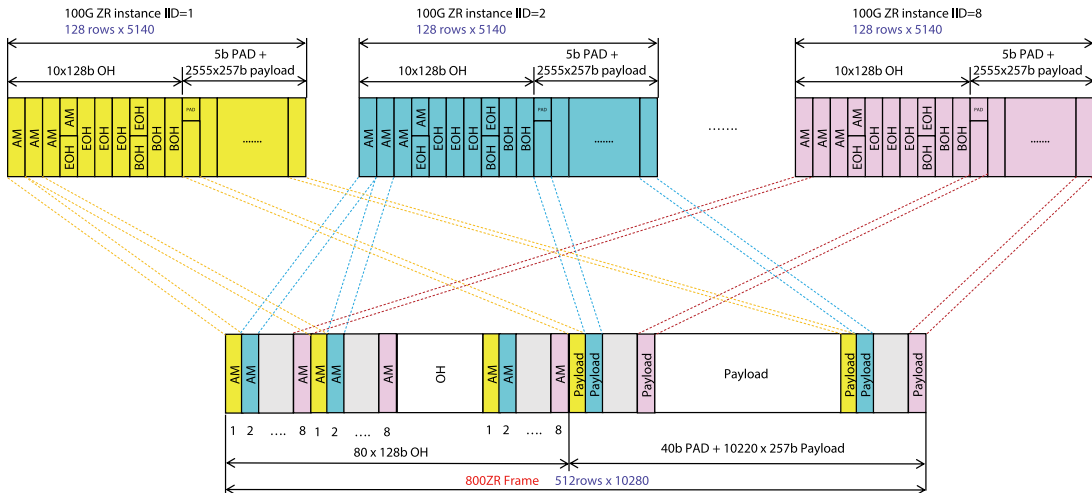
로 480비트, EOH용으로 480비트, BOH용으로 320비트가 각각 할당되나 800ZR에서는 EOH(Extended Overhead)는 사용하지 않고 BOH(Basic Overhead)만 사용한다.

[그림 3]은 연속된 8개의 100G ZR 인스턴스로 구성된 멀티 프레임 구조로서 128개의 열 8개가 합쳐져서 1,024열이 된다. 매 프레임의 MFAS(Multi Frame Alignment Signal)오버



*출처 : Draft Working Document: oif2021.144.12

[그림 3] 100G ZR 인스턴스의 멀티 프레임 구조



*출처 : Draft Working Document: oif2021.144.12

[그림 4] 800ZR 프레임 구조

헤드에 저장된 값에 의해 프레임 순서가 구분되고, 송신단에서는 보내는 순서에 따라 '000'에서 '111'까지 값을 각각 저장한다. 이렇게 멀티프레임 방식을 취하는 이유는 풍부한 오버헤드를 사용하기 위함이다.

3.2 800ZR 다중화 구조

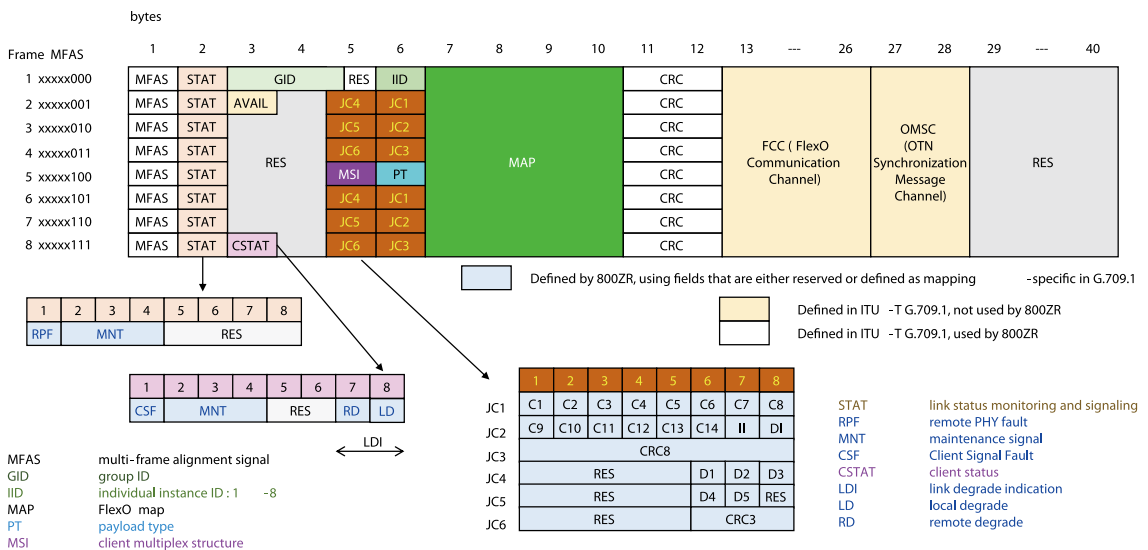
800ZR 프레임은 8개의 100G ZR 인스턴스 프레임이 128비트 단위로 순차적 인터리빙 (Interleaving) 방식에 의해 하나의 프레임이 생성된다. [그림 4]와 같이 1280비트의 오버헤드 8개가 128비트 단위로 인터리빙되고 이어서 페이로드가 128비트 단위로 행해진다. 그 결과 128 row × 5,140비트의 8배인 512 row × 10,280비트로 구성된 하나의 800ZR 프레임이 생성된다.

3.3 오버헤드 구조 및 기능

800ZR 신호는 DP-16QAM에 의해 단일 캐리어로 800G급 신호가 전달되나, 신호 관리는

100G ZR 인스턴스 단위로 이루어진다.

[그림 5]에서 G.709.1에서 정의되어 있으나 OIF에서 사용하지 않는 FCC, OSMC, AVAIL을 제외한 나머지 오버헤드 기능은 FlexO와 800ZR에서 동일한 방법으로 사용한다. 오버헤드 기능을 간략히 살펴보면, MFAS는 8개로 구성된 멀티프레임을 구분하기 위한 용도로 송신단에서 보내는 순서대로 0에서 7까지 번호를 할당한다. GID는 100G ZR 인스턴스가 속한 800ZR의 아이디를 나타내고 8개의 인스턴스는 동일 GID 값을 갖는다. IID는 100G ZR 인스턴스마다 부여된 고유 번호로서 '1'에서 '8'까지 할당된다. MAP는 8개 각각의 100G ZR 인스턴스가 800ZR 프레임을 생성하기 위해 몇 번째에 위치하는지를 나타내기 위해 사용되나, 256비트 중 마지막 8비트를 제외한 나머지 비트는 '0'으로 설정되고 MAP[255:248]=0xFF로 설정된다. PT는 페이로드 타입을 나타내기 위한 용도이며 클라이언트 신호 즉, 100G, 200G, 400G, 800G에 따



*출처 : Draft Working Document: oif2021.144.12

[그림 5] 100G ZR 인스턴스 오버헤드 구조

라 값이 다르게 설정된다.

MSI는 클라이언트 이더넷 신호가 100G ZR 인스턴스를 차지하고 있는지를 나타내기 위한 용도이다. 차지하고 있으면 MSB를 '1'로, 그렇지 않으면 '0'으로 설정한다. 나머지 7비트는 각 인스턴스 별로 100G인 경우 '0~7'값이 각각 저장되고, 200G인 경우 '0', '0', '2', '2', '4', '4', '6', '6'이, 400G인 경우 '0', '0', '0', '0', '4', '4', '4', '4'이 저장된다. 800G인 경우 모든 인스턴스가 '0'으로 각각 설정된다.

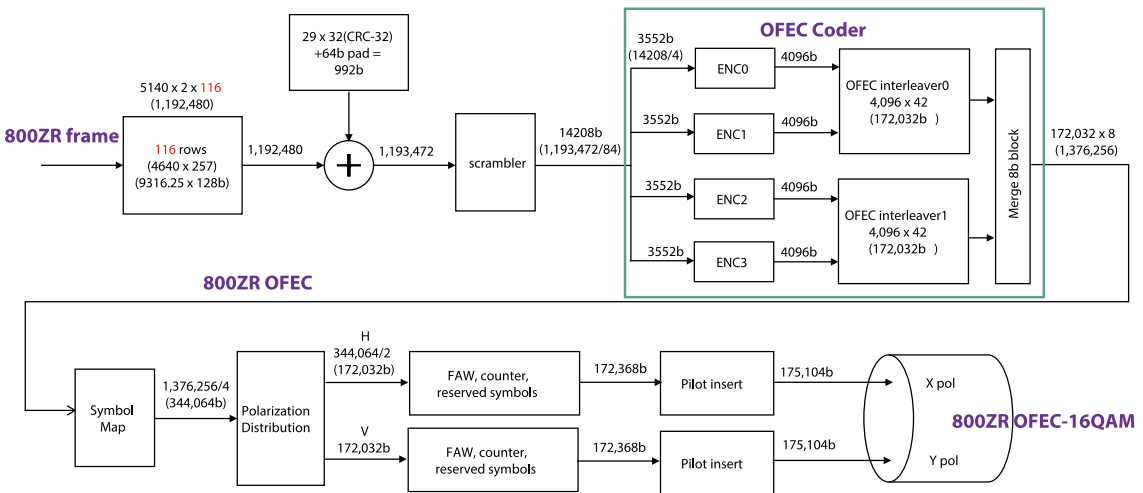
STAT는 100G ZR 인스턴스의 장애상태를 나타내기 위한 용도이다. RPF는 대국 수신장치에서 장애를 검출한 경우를 표시하기 위한 용도이고, MNT는 유지보수를 위한 루프백(loopback)의 현재 수행 여부 상태를 나타낸다. CSTAT는 클라이언트 측, 이더넷 신호의 성능 및 장애 상태를 나타내기 위한 용도로 사용된다. 마지막으로 JC는 100G ZR 프레임내 스템핑(stuffing) 비트의 위치를 수신측에 알려주기 위해 GMP 파라미터에서 페이로드에 실을 데이터 엔터티(entity)

갯수 $C_m(t)$ 와 C_m 나머지 값의 누적치 $\sum C_n D(t)$ 를 저장하기 위한 용도로 각각 사용된다.

4. 800ZR 시스템 구조

[그림 6]은 OFEC(openFEC)과 DP-16QAM (Dual Polarization 16 Quadrature Amplitude Modulation) 기능이 포함된 800ZR 시스템의 송신 기능 블록도이다. 800ZR 프레임이 OFEC 인코더에 인가되기 전 84개의 OFEC블록을 만들기 위한 처리가 수행된다.

수행 방법은 800ZR 프레임에서 116열을 선택하고 선택된 116열에서 매 4열마다 CRC-32를 수행한 후 64비트의 패드를 추가하여 전 프레임에 대해 스크램블링을 한다. 84개의 OFEC 블록은 4개의 인코더 즉, ENCO~ENC3으로 구성된 OFEC 코더에 순차적으로 입력되고 각 인코더에서 계산된 패리티는 인코더의 출력 블록에 추가되어 BCH(239,256) 코드워드가 생성된다. BCH(239,256)은 해밍 거리가 17이므로 239



*출처 : Draft Working Document: oif2021.144.12

[그림 6] 800ZR 시스템 송신 블록도


비트 중에서 8비트 에러 정정이 가능하게 된다. 두개의 인코더 출력을 인터리빙하여 하나의 출력으로 만든 후 다시 8비트 단위로 합한다. 8비트 중 4비트씩 광섬유 편광 제어를 통해 수평축으로 4비트, 수직축으로 4비트씩 각각 보내기 위해 $a+jb$ ($a, b=i^{3/4}1, i^{3/4}3$)값 중 하나가 되도록 DP-16QAM 심볼로 매핑한다.

추가로 DP-16QAM 신호 동기를 위한 FAW(Frame Alignment Word)용 심볼, TS(Training Symbol), PS(Pilot Symbol)등이 추가되어 800G급 800ZR 신호가 단일 파장으로 전송된다.

5. 맺음말

IEEE 802.3에서 개발중인 800G급 이더넷 기술은 40km 이상 전송이 어려워 IDC 망으로 사용할 수 없고, ITU-T에서 개발된 OTN 기술은

수백km 이상 전송이 가능하나 OTH에 기반한 메트로와 백본용 전달 장치에 적합하다. 그러나 OIF에서 개발중인 800ZR 기술 표준은 이더넷 신호 수송을 목적으로 하고 최대 120km까지 전송이 가능하므로 데이터센터간 트래픽 전달용 전송장치에 적합한 표준이다.

OIF는 산업체 중심 멤버십으로 운영되고 이더넷과 OTN 기술 적용이 용이하지 않는 영역에서 현재 상용화 가능한 기술과 국제표준을 적용, 응용하여 새로운 기술 표준을 주도하는 단체이다. 또한 개발 장비간 상호연동성 시험과 시연 등을 통해 국제표준의 위상도 갖고 있다. 현재 IEEE 802.3에서 1.6T급 이더넷 MAC/PHY 기술에 대한 표준도 2026년 목표로 개발을 진행하고 있다. 향후 OIF에서는 이를 수용하고 데이터센터간 연결망으로 사용할 수 있는 1.6T급 기술 표준도 개발될 것으로 예측된다. 

※ 본 연구는 2023년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 [No. 2019-0-00002, 광 클라우드로 네트워킹 핵심기술 개발]

주요용어풀이

- **OTN** (Optical Transport Network) : OTH(Optical Transport Hierarchy) 기반 광전달망
- **OFEC** (Open FEC) : 블록기반 인코더와 반복적 SD 디코더 (Soft Decision decoder)

참고문헌

- [1] OIF Draft Working Document: oif2021.144.12, DRAFT Implementation Agreement for 800ZR Coherent Interface, 2022
- [2] ITU-T G.709.1, Recommendation ITU-T G.709.1/Y.1331.1, Flexible OTN short-reach interface, 2018
- [3] OIF-400ZR-01.0, Implementation Agreement 400ZR, 2020