

# 고속 패브릭 연결망 기반 메모리 중심 컴퓨팅 기술 동향

## Trends in High Speed Fabric-Interconnect-Based Memory Centric Computing Architecture

차승준 (S.-J. Cha, seungjunn@etri.re.kr)

석성우 (S.-W. Sok, swsok@etri.re.kr)

권혁제 (H.J. Kwon, heavenwing@etri.re.kr)

김영우 (Y.W. Kim, bartmann@etri.re.kr)

김진미 (J. Kim, jinmee@etri.re.kr)

김학영 (H.Y. Kim, h0kim@etri.re.kr)

고광원 (K.-W. Koh, kwangwon.koh@etri.re.kr)

김강호 (K.-H. Kim, khk@etri.re.kr)

고성능컴퓨팅시스템연구실 선임연구원

고성능컴퓨팅시스템연구실 책임연구원

고성능컴퓨팅시스템연구실 책임연구원

고성능컴퓨팅시스템연구실 책임연구원

고성능컴퓨팅시스템연구실 책임연구원

고성능컴퓨팅시스템연구실 연구전문위원

고성능컴퓨팅시스템연구실 책임연구원/실장

초성능컴퓨팅연구본부 책임연구원/본부장

### ABSTRACT

Applications such as artificial intelligence continue to grow in complexity and scale. Thus, the demand for scalable computing is increasing for achieving faster data processing and improved efficiency. This requirement has led to the development of memory-centric computing and high-speed fabric interconnection technologies. Memory-centric computing reduces the latency and enhances the system performance by shifting the focus from the central processing unit to the memory, whereas high-speed fabric interconnects enable efficient data transfer across various computing resources. Technologies such as Gen-Z, OpenCAPI, and CCIX have been integrated into the CXL (Compute Express Link) standard since 2019 to improve communication and cache coherence. Ethernet-based interconnects such as RoCE, InfiniBand, and OmniXtend also play a crucial role in providing high-speed data transfer and low latency. We explore the latest trends and prospects of these technologies, highlighting their benefits and applications.

**KEYWORDS** CXL, OmniXtend, 고속 패브릭 연결망, 메모리 중심 컴퓨팅

\* DOI: <https://doi.org/10.22648/ETRI.2024.J.390510>

\* 본 연구는 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임[No. 2018-0-00503, 메모리 중심 차세대 컴퓨팅 시스템 구조 연구].



본 저작물은 공공누리 제4유형

출처표시+상업적이용금지+변경금지 조건에 따라 이용할 수 있습니다.

©2024 한국전자통신연구원

## 1. 서론

최근 인공지능을 비롯한 다양한 응용 프로그램의 대형화와 복잡화가 가속화되면서 더 빠른 데이터 처리와 향상된 효율성을 제공하는 확장 가능한 컴퓨팅 성능이 요구되고 있다. 이러한 요구는 단일 컴퓨팅 노드로는 감당할 수 없는 수준에 이르렀으며, 이에 따라 메모리 확장성과 효율적인 자원 활용을 위한 기술들이 필요하게 되었다. 이러한 변화는 데이터센터의 대규모 고성능 컴퓨팅 시스템부터 온디바이스 형태의 소규모 기기까지 언제 어디서나 실행 가능한 고성능 컴퓨팅 환경을 필요로 한다. 이에 따라 메모리 중심 컴퓨팅과 고속 패브릭 연결망은 혁신적인 기술로 자리매김하게 되었다.

메모리 중심 컴퓨팅은 데이터 집약적 응용 프로그램, 특히 인공지능, 머신러닝, 빅데이터 분석의 보편화로 인해 빠르고 효율적인 메모리 접근의 필요성이 증가하면서 기존의 CPU 중심 아키텍처에서 벗어나 메모리가 중심 역할을 하게 되어 대기 시간을 줄이고 시스템 전반의 성능을 개선한다. 이러한 환경에서는 분산된 메모리 자원을 사용하여 메모리 용량을 확장하려면 고속 패브릭 연결망 기술이 필요하다.

고속 패브릭 연결망은 여러 컴퓨팅 자원(CPU, 메모리, 스토리지 등)을 고속 네트워크 패브릭을 통해 연결하여 데이터를 빠르고 효율적으로 전송하는 기술이다. 빠른 데이터 전송 속도와 낮은 대기 시간을 제공하며, 메모리 중심 컴퓨팅을 지원하는 데 필수적이다. 또한, 증가하는 데이터 수요에 맞추어 확장성과 고속 데이터 이동을 제공하여 자원 활용을 최적화하고 성능을 향상할 수 있다. 하지만, 이러한 기술을 활용하기 위해서는 서로 다른 하드웨어와 소프트웨어 구성요소 간의 복잡성을 극복하여 호환성을 보장해야 하며, 하드웨어뿐만 아니라 시스템 기

능을 향상하기 위해 메모리 확장, 메모리 분리, 메모리 공유를 지원해야 하며, 빠르고 효율적인 데이터 처리를 제공하는 전문적이고 차별화된 시스템 소프트웨어가 필요하다[1].

연결망 기술의 관점에서 기존의 컴퓨팅 아키텍처는 CPU 중심 구조로 메모리와 데이터 입출력 사이의 대역폭이 제한되어 성능 병목현상이 발생한다. 이를 해결하기 위해 고성능 컴퓨팅 시스템에서는 PCIe(Peripheral Component Interconnect Express)나 이더넷(Ethernet)과 같은 고속연결망 기술을 도입하였지만, 대규모 데이터 전송 요구를 충족시키기에는 부족하다. 이에 따라 데이터센터와 클라우드 컴퓨팅 환경에서 더욱 효율적인 데이터 전송과 메모리 확장을 지원하는 기술들을 연구하여 적용하고 있다 [2].

대표적으로 발표된 고속 패브릭 연결망 기술로는 Gen-Z, OpenCAPI(Open Coherent Accelerator Processor Interface), CCIX(Cache Coherent Interconnect for Accelerators)가 있었으나, 2019년 발표된 CXL(Compute Express Link)에 점진적으로 통합되었다. CXL은 PCIe 인터페이스 Gen5 이상의 환경을 기반으로 하는 연결망 표준 기술로, CPU와 메모리, 가속기 등 다양한 컴퓨팅 자원 간의 캐시 일관성 문제를 해결하는 등 효율적인 통신이 가능하다.

이더넷 기반의 연결망 기술은 호환성과 확장성을 제공한다. RoCE(RDMA over Converged Ethernet), Infiniband, iWARP의 기술은 기존의 이더넷 인프라를 활용하여 높은 데이터 전송 속도와 낮은 대기 시간을 제공한다. RISC-V 진영에서는 이더넷 기반으로 캐시 일관성의 상호 연결망을 제공하는 표준인 OmniXtend 프로토콜 표준을 제공한다[3].

이러한 다양한 고속 패브릭 연결망 기술들은 데이터 전송의 병목현상을 해결하고, 시스템 성능을 극대화하며, 데이터센터의 자원 활용도를 향상시키

는 데 중요한 역할을 한다. 본고에서는 고속 패브릭 메모리 연결망 기술의 최신 동향과 실현 전망에 대해 분석하고, 각 기술의 장단점 및 적용 사례를 통해 향후 발전 방향을 알아보려고 한다.

II장에서는 고속 패브릭 연결망 기술의 전체 동향과 PCIe와 이더넷 기반 고속 패브릭 메모리 연결망 기술 동향을 설명한다. III장에서는 RISC-V 진영에서 제공하는 고속연결망 프로토콜인 OmniXtend를 설명한다. 마지막으로 IV장에서는 고속 패브릭 연결망 기반 기술의 활용 사례와 시사점을 중심으로 본고의 결론을 제시한다.

## II. 고속 패브릭 메모리 연결망 기술

### 1. 고속 패브릭 메모리 연결망 동향

인공지능과 빅데이터 분석 등 데이터 집약적 응용 프로그램의 확산으로 메모리 중심 컴퓨팅과 고속 패브릭 연결망 기술의 중요성이 대두되었다. 그림 1과 같이 2014년에 인텔과 HPE에서 연구를 시작하여 현재까지 다양한 산업표준이 발표되고, 또한 통합되고 있다.

인텔은 2014년에 랙스케일 디자인(RSD: Rack Scale Design)에서 프로세서, 메모리, 스토리지, 네트워크 등 컴퓨팅 자원을 필요에 따라 조합하여 노드를 구성하는 환경을 제시하고, 이를 통해 컴퓨팅 자원 활용의 효율성을 향상시킬 수 있는 기술을 제안하였

다. 하지만, 메모리는 프로세서 자원과 통합되어 있어서 메모리 자원만에 대한 유연한 구성은 불가능하여 더 이상 연구가 진행되지 않았다[4]. 하지만 이 기술은 추후 스토리지 솔루션과 데이터센터 기술에 영향을 주었다. HPE는 메모리 중심 차세대 컴퓨터의 구조 연구를 통해 2016년 더 머신(The Machine)을 소개하였다. 더 머신의 프로토타입은 기존의 프로세서 중심의 폰노이만 구조를 벗어나서, 160TB 규모의 메모리 풀을 중심으로 광네트워크로 구성된 패브릭을 통해 CPU, 가속기 등 컴퓨팅 프로세서들이 메모리 풀을 공유하는 구조이다[5]. 초기 발표 이후 개발의 어려움으로 프로젝트가 중단되었지만, 새로운 메모리 기술과 메모리 중심 아키텍처 아이디어는 HPE의 다른 제품과 기술에 통합되었다.

메모리 풀 접근을 위한 연결망 표준을 선점하기 위한 관련 업계의 표준화 기술을 연구하여 발표했다. 사용하는 연결망 물리계층에 따라 PCIe 기반의 CCIX, OpenCAPI가 있으며, 이더넷 기반으로는 RoCE, Infiniband, OmniXtend가 있다. 또한, 독립적인 연결망 기술로 Gen-Z가 발표되었다. 이들은 기반이 되는 시스템 버스의 제약에 따라 메인보드 내, 혹은 샤시 내부에서 패브릭 메모리 연결을 지원한다. 특히, 이더넷 기반의 연결망은 노드나 랙 경계를 넘어서 다양한 토폴로지의 연결망 구성을 지원한다.

CXL은 2019년에 발표된 인텔이 주도하는 차세대 고속연결망 기술로, 메모리와 컴퓨팅 자원의 효율적인 공유를 목표로 한다. 기존의 CCIX, OpenCAPI, Gen-Z와 같은 표준들이 CXL로 통합되면서, 업계 전반에서 단일 표준을 채택하는 추세가 강화되고 있다. 이러한 통합은 다양한 하드웨어 간의 호환성을 높이고, 데이터센터의 효율성을 극대화하는 데 기여한다.

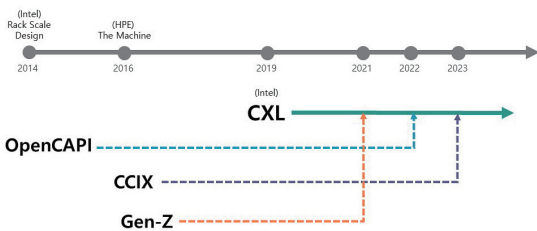


그림 1 고속 패브릭 메모리 연결망 기술의 발전

## 2. PCIe 연결망 기술 동향

PCIe 연결망은 비영리 기관인 PCI-SIG(Peripheral Component Interconnect Special Interest Group)를 통하여 개발되고 있는 대표적인 업계 표준 규격이며, 인텔, AMD와 같은 프로세서 제조사를 중심으로 다양한 주변장치 제조사들의 요구를 바탕으로 규격을 발전시키고 있다[6]. PCIe는 2003년 기존의 병렬 입출력 버스인 PCI와 PCI-X를 대체하는 기술로, 컴퓨팅 시스템의 프로세서와 입출력을 연결하는 버스 기술로 시작하여 최근에는 CXL의 기본 인프라로 사용하는 등 컴퓨팅 노드 간의 시스템 연결망 기술로 활용 및 발전하고 있다.

PCIe 기술은 점대점 직렬 통신에 기반한 기술로 그림 2와 같이 PCIe Gen1의 레인당 2.5GT/s의 속도를 시작으로 하여, 2019년 레인당 32GT/s, 2022년 레인당 64GT/s 속도의 Gen5, Gen6 규격을 발표하였으며, 2024년 현재 레인당 128GT/s 속도의 PCIe Gen7 규격의 개발이 진행 중이다[6,7].

PCIe 장치는 크게 전송계층, 데이터-링크 계층, 물리계층으로 구성하며, PCIe Gen3 이후로 데이터-링크, 물리계층의 규격화를 통하여 다양한 규격의 전송계층 도입이 가능하다. 이를 통하여, CCIX, CXL 규격과 같은 새로운 기술 규격의 기본 연결망으로 활용이 가능하다.

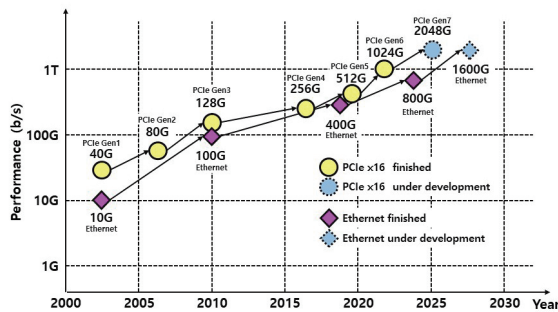


그림 2 PCIe 및 Ethernet 기술 발전 추이

- Gen2: Peer-to-Peer 데이터 전송 도입
- Gen3: 128b/130b 코딩 도입
- Gen5: CXL 규격 지원
- Gen6: PAM-4, 242B/256B 코딩 도입

특히 PCIe Gen5 이후로는 시스템 수준의 메모리 확장(메모리 풀링 및 셰어링)을 위한 인텔 주도의 CXL 규격을 위한 기본 직렬 버스 규격으로 사용되면서, PCIe 기술이 기존 입출력 버스에서 메모리 시스템 버스 분야의 확장을 주도한다. 또한, PCIe Gen6 이후의 규격은 기존의 전통적인 NRZ(Non-Return Zero) 코딩 방식에서 PAM-4(Pulse Amplitude Modulation) 코딩을 도입함으로써 보다 빠른 속도의 시그널링을 지원한다.

### 가. OpenCAPI

OpenCAPI 기술은 IBM Power 프로세서에 기반하여 가속기 장치의 고성능화를 위하여 개발한 CAPI (Cache Coherent Interconnect for Accelerators) 기술에 기반한다. 초기의 CAPI 기술 규격은 PCIe Gen3에 기반한 IBM 독자 기술로 시작하였으며, 2016년에 AMD, 구글, IBM, Mellanox 등의 다양한 업체의 컨소시엄으로 발전하여 기술 공개를 통한 OpenCAPI 규격으로 발전하였다[8].

OpenCAPI는 3.0에서 NVIDIA사의 NVLink 2.0을 지원하는 형태로 발전하였으며, 4.0부터 내부에 PSL(Process Service Layer)의 캐시 지원 유닛을 두어 메모리 접근에 대해 캐시일관성을 제공한다[9]. 2022년 CXL 컨소시엄에 관련 기술을 공여하여 CXL 기술 발전에 기여하였으며, 현재는 활동이 종료되었다[10].

### 나. CCIX

CCIX 규격은 ARM, AMD, Xilinx 등과 같이 x86

이 아닌 프로세서 계열의 진영에서 개발한 기술 규격으로, 2018년 1.0 규격을 발표하며 기술 개발이 진행되었다[11]. CCIX는 PCIe의 물리 계층을 활용하여 프로세서, 가속기 등의 장치 간 캐시 일관성을 지원하는 메모리 트랜잭션을 지원한다.

CCIX는 PCIe의 링크 계층을 통해, 또는 전용 링크 계층을 활용하여 원격 메모리를 NUMA(Non Uniform Memory Access) 형태로 접근할 수 있게 한다. 이는 ARM 프로세서의 프로세서 버스와 FPGA 기반 가속기 장치 등에서 활용되어, 컴퓨팅 노드 내의 다양한 컴퓨팅 자원 간의 데이터 전송 효율성을 높인다. CCIX는 높은 대역폭과 낮은 지연 시간을 제공하여, 데이터 병목현상을 줄이고 컴퓨팅 노드 전체의 성능을 극대화한다. 2019년 주요 기능이 CXL에 흡수되어 단일 표준으로 통합되어 2021년 활동을 종료한다고 발표하였다.

### 3. 이더넷 기반 고속연결망 동향

RoCE(RDMA over Converged Ethernet), InfiniBand, iWARP(Internet Wide Area RDMA Protocol)는 고성능 컴퓨팅과 데이터센터 네트워크에서 사용되는 고속 연결망 기술이다. 이 기술들은 모두 원격 직접 메모리 접근(RDMA)을 활용하여 낮은 지연 시간과 높은 대역폭을 제공하며, 각각의 기술은 특정 사용 사례와 요구사항에 맞는 장점과 단점을 가진다.

#### 가. RoCE

RoCE는 InfiniBand 기반의 RDMA 프로토콜을 이더넷 물리 계층을 활용하여 적용한 기술로 다양한 네트워크 구성을 지원하며, RoCEv1과 RoCEv2로 구분된다[12]. RoCEv1은 데이터 링크 계층에서 작동하며, 동일한 서브넷 내의 물리적 네트워크 내에서만 통신이 가능하다. RoCEv2는 네트워크 계층에

서 작동하여 라우팅이 가능하며, 보다 큰 네트워크에서도 활용할 수 있다.

RoCE의 주요 장점은 이더넷 인프라를 활용할 수 있다는 것이며, 기존의 네트워크 인프라를 활용하여 비용을 절감할 수 있다. 다만, 네트워크 혼잡 시 성능 저하가 발생할 수 있어, 이를 해결하기 위한 DCB(Data Center Bridging) 기술이 요구된다[13,14].

#### 나. InfiniBand

InfiniBand는 높은 대역폭과 낮은 지연 시간을 제공하는 고성능 연결망 기술로, 주로 슈퍼컴퓨터와 데이터센터에서 사용된다. InfiniBand는 QoS(Quality of Service), 가상 통로, 자동화된 데이터 전송을 지원하여 네트워크 성능을 극대화한다. 특히, NDR(Next Data Rate) InfiniBand는 최대 400Gbps의 속도를 제공하며, 이는 대규모 데이터 처리가 필요한 환경에 적합하다. InfiniBand는 또한 신뢰성이 높아 데이터 손실이 적고, 고가용성을 제공하는 클러스터 환경에 적합하다[15].

#### 다. iWARP

iWARP는 TCP/IP를 기반으로 하는 RDMA 프로토콜로, WAN 환경에서도 안정적인 RDMA 성능을 제공한다. iWARP는 표준 TCP/IP 프로토콜을 사용하므로 기존 네트워크 인프라와의 호환성이 뛰어나며, 방화벽 및 NAT(Network Address Translation) 환경에서도 원활하게 작동한다. iWARP의 주요 장점은 설치 및 유지보수가 용이하며, RDMA 기능을 WAN 환경으로 확장할 수 있다는 점이다. 다만, TCP/IP의 오버헤드로 인해 다른 RDMA 기술에 비해 상대적으로 높은 지연 시간이 발생할 수 있다[16].

RoCE, InfiniBand, 그리고 iWARP는 각각 고유한 강점과 약점을 가지고 있으며, 특정 사용 사례와 요구사항에 따라 적합한 기술을 선택해야 한다. RoCE

는 비용 효율적인 이더넷 기반 솔루션으로, 대규모 데이터센터에서의 활용이 적합하며, InfiniBand는 높은 성능과 안정성이 요구되는 슈퍼컴퓨터 환경에 최적화되어 있다. iWARP는 WAN 환경에서도 RDMA 기능을 제공하여, 다양한 네트워크 환경에서 활용될 수 있다.

## 4. Gen-Z와 CXL

### 가. Gen-Z

Gen-Z 기술은 고속 패브릭 메모리 연결망을 표방한 최초의 기술로써 2016년 컴퓨팅 시스템 기업과 통신 기업의 협업으로 시작하였다. Gen-Z 컨소시엄은 2018년 Gen-Z 1.0을 발표하였으며, Gen-Z 기술은 PCIe 및 이더넷에 기반한 물리계층을 사용하여 활용도를 높이도록 하였다.

Gen-Z는 이더넷 물리 계층에 기반하여 바이트 수준 접근을 위한 Load/Store 인터페이스, 블록 인터페이스와 메시징 인터페이스 등 다양한 규약을 지원한다[17].

프로세서 수준에서 시스템 연결망 수준까지의 다양한 범위의 활용을 가정하여 개발된 Gen-Z 기술은 초기 HP사를 중심으로 개발이 활발히 이루어졌으며, 국내에서는 한국전자통신연구원이 Gen-Z 1.0 기반의 프론트엔드 브리지, 스위치, Gen-Z 메모리 하드웨어 모듈 개발과 패브릭 관리자 SW를 개발하여 총 4.5TB 규모의 메모리 풀을 지원하는 하드웨어 시스템을 시연하였다[18].

프로세서 수준에서 시스템 수준의 메모리 패브릭의 구성이 가능한 Gen-Z 기술의 상용화는 프로세서 제작사의 적극적 참여가 저조하여 규격 발표 이후 실제 제품화가 늦어지게 되었으며, 2021년 11월에 관련 기술을 CXL에 이관하며 CXL 3.0 규격의 개발에 기여하였다[19].

### 나. CXL

CXL은 프로세서 제조사인 인텔사의 주도로 구성된 CXL 컨소시엄을 통하여 개발하고 있는 시스템 메모리 확장 지원 프로토콜 및 기술 규격이다[20]. CXL은 시스템 수준의 메모리 확장 및 공유를 중심으로 하는 기술이다. 최신의 프로세서는 내부에 수십 개 이상의 코어를 내장하여 성능의 비약적 향상과 인공지능 응용의 발전에 따라 대용량의 메모리가 필요함에 반하여, 프로세서 패키지의 물리적인 제약으로 인하여 프로세서가 필요로 하는 메모리의 확장이 어려운 현실이다. CXL 기술은 기존의 PCIe 시스템 버스를 활용하여, I/O 연결망을 확장하여 추가적인 메모리 연결을 가능케 함으로써 부족한 메모리의 유연한 확장이 가능하다.

CXL 표준은 2019년 3월 1.0이 공개되었고, 2023년 5월 3.1이 공개되었다. 기본적인 특징은 전통적인 I/O(Type 1), 가속기(Type 2), 메모리(Type 3)에 특화되도록 목적에 따른 3가지 프로토콜 유형(CXL.io, CXL.cache, CXL.mem)을 지원하며, CPU가 중심이 되어 전체 패브릭의 캐시 일관성을 유지한다[21].

CXL 1.0은 단일 노드 내에서 다양한 이기종 자원의 연결과 메모리 확장을 위한 기본 규격을 정의하며, CXL 2.0에서는 다중 노드 간의 메모리 풀 제공과 관리, CXL 3.0에서는 CXL 기반의 패브릭 기능을 제공하여 시스템 수준에서의 고속 패브릭 연결망 기반의 메모리 중심 컴퓨팅을 지원하는 규격을 정의한다. 특히 고속 패브릭 연결망 기반의 메모리 중심 컴퓨팅을 위하여는 CXL 2.0 이상의 메모리 풀링과 CXL 3.0 기반의 메모리 공유, 패브릭 확장 및 패브릭 관리 기능이 강화되었으며, 시스템의 유연한 메모리 관리를 위해서는 무엇보다도 고도의 패브릭 관리 기능이 중요하다[22].

CXL은 PCIe, CCIX, OpenCAPI, Gen-Z 등의 선행 기술 배경을 바탕으로 출현한 기술로, 2019년 이

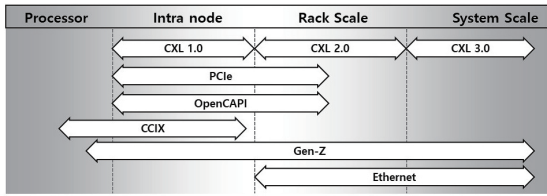


그림 3 패브릭 연결망 기술 적용 범위

후로 2021년 Gen-Z, 2022년 OpenCAPI, 2023년 CCIX 기술을 통합함으로써 기술적인 완성도와 적용 범위를 넓혀왔다. 그림 3은 시스템/패브릭 연결망 기술들의 적용 범위를 표시한 것으로, CXL 기술이 노드 단위에서 시스템 수준까지 확장 가능한 유연한 패브릭 기술임을 확인할 수 있다.

CXL 기술은 프로세서 및 메모리 기술에 의존성이 강하며, 규격 발표 이후 CXL 지원 프로세서 및 관련 장치의 개발은 2023년 이후로 상용 수준의 제품이 출시되고 있다. 프로세서로는 인텔의 사파이어 래피즈(Sapphire Rapids) 이후의 제온(Xeon) 계열과 AMD의 제노아(Genoa) 이후의 에픽(EPYC) 프로세서가 CXL을 지원한다. 국내에서는 삼성과 SK하이닉스가 CXL 2.0을 지원하는 CXL 메모리 장치의 양산을 시작했으며, 관련 SDK도 함께 제공한다[23].

### III. RISC-V 진영에서의 고속연결망

#### 1. OmniXtend와 TileLink

OmniXtend[24,25]는 RISC-V 아키텍처의 시스템버스인 타일링크(TileLink[26])를 이더넷을 통해 확장하여 캐시 일관성을 제공하며 노드 간 확장 및 공유 가능한 프로토콜로, 이더넷 전송을 위한 직렬화(TLoE: TileLink over Ethernet) 규격을 제공한다. 칩스 얼라이언스(CHIPS Alliance)에서 개방형 표준으로 기술을 공개하고 있으며 기존의 고속 메모리 패브릭 연결망 표준과는 다르게 누구나 활용할 수 있다는 장점이 있다.

OmniXtend는 기존 x86 시스템의 프로세서 간 IPC(Inter-Processor Communication) 버스인 QPI(Quick Path Interconnect) 혹은 시스템 수준의 메모리 확장을 위한 CXL과 유사한 형태의 시스템 확장을 RISC-V 시스템에 적용할 수 있다[27]. 또한, 개발 자율성이 높은 개방형 RISC-V의 특성을 활용한 다양한 실험이 가능한 특징을 가진다.

타일링크 버스는 RISC-V 시스템인 RocketChip<sup>1)</sup>과 Chipyard 생성자(Generators)에서 사용하는 캐시 일관성 및 메모리 프로토콜이다[28]. 타일링크는 필수 채널 A, D와 캐시 일관성 채널 B, D, E 등으로 구성하는 링크를 형성하는 마스터와 슬레이브 간 병렬 통신 프로토콜이다. 마스터(채널 A) 요청 패킷과 슬레이브(채널 D)의 응답 패킷으로 통신한다. 타일링크는 가장 기본 동작인 읽기(Get)/쓰기(Put) 명령은 정의한 Uncached Lightweight(TL-UL)와 단일(Atomic) 및 복수(Burst) 동작을 정의하고 있는 Uncached Heavyweight(TL-UH)와 캐시 동작을 정의한 Cached(TL-C) 등으로 구분하여 정의한다.

#### 2. 이더넷 프레임과 데이터 전송 메커니즘

OmniXtend는 이더넷 프레임의 페이로드(Payload)에 타일링크 메시지를 포함하여 전송하며, 그림 4와 같이 이더넷 프레임 안에 TLoE 프레임 헤더, 타일링크 메시지, 패딩, TLoE 프레임 마스크로 구성한다.

OmniXtend는 안정적인 데이터 전송을 보장하기 위해 재전송(Retransmission)과 흐름 제어(Flow Control) 기능을 제공한다. 이러한 기능은 특히 대규모 분산 시스템에서 데이터 일관성과 무결성을 유지하는 데 핵심적인 역할을 한다.

1) RocketChip은 U.C Berkeley에서 개발한 In-order RISC-V core이다.

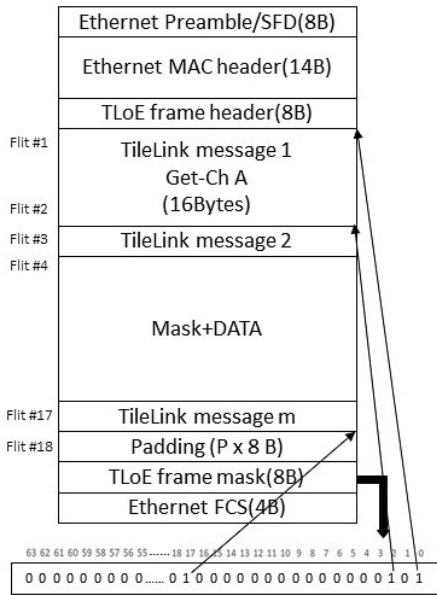


그림 4 TLoE 프레임 구성 예

재전송 기능은 데이터 전송 중 오류가 발생하거나 패킷 손실이 있을 경우, 해당 데이터를 다시 전송하여 정확한 데이터 전달을 보장한다. 이를 위해 패킷 전송 시 순차 번호(Sequence Number)를 부여하여 각 패킷을 식별하고, 수신 측에서 패킷이 정상적으로 수신되었음을 확인하기 위해 수신 확인 순차 번호(Sequence Number Acknowledgment)를 사용한다. 이러한 확인 메커니즘을 통해 패킷 손실이나 오류가 감지되면, 송신 측은 해당 패킷을 재전송함으로써 데이터의 완전한 전송을 보장한다. 재전송은 네트워크 상태에 따라 자동으로 이루어지며, 특정 타임아웃 시간이 지나거나 잘못된 수신이 감지될 경우 트리거된다.

흐름 제어 기능은 송신 측과 수신 측 간의 데이터 전송 속도를 조절하여 네트워크 자원의 효율적 사용을 가능하게 하고, 특히 수신 측에서의 버퍼 오버플로우를 방지한다. OmniXtend의 흐름 제어는 크레딧(Credit) 기반 시스템을 사용하여 구현되며, 각 채널마다 일정량의 크레딧이 할당된다. 송신 측은

크레딧을 소진할 때까지 데이터를 전송할 수 있으며, 수신 측에서 데이터를 처리하고 새로운 크레딧을 발행하면 송신 측은 추가 데이터를 전송할 수 있게 된다. 이 방식은 네트워크의 혼잡을 최소화하고, 전송 속도를 수신 측의 처리 능력에 맞게 조절함으로써 시스템 성능을 최적화하는 역할을 한다.

### 3. 메모리 확장 시스템 구축

한국전자통신연구원에서는 RISC-V 메모리 버스에 외부 확장 메모리를 연결하기 위해 RISC-V를 구성하고 기존 공개 코드인 OmniXtend와 접목하기 위한 브릿지 코드를 작성하여 메모리 확장 시스템을 구축하였다[29]. 그림 5는 구축한 시스템을 호스트 노드와 메모리 노드이다.

호스트 노드는 메모리 노드의 메모리를 특정 메모리 주소에 매핑하여 제공한다. 응용 프로그램이 해당 주소에 접근(Load/Store) 시 이더넷을 통해 메모리 노드를 활용하게 된다. 이를 통해 호스트 노드는 더 큰 메모리가 필요한 응용 프로그램을 동작할 수 있다.

호스트 노드와 메모리 노드는 1:1 연결뿐만 아니라, 이더넷 스위치를 활용한 N:N 연결이 가능하다. 이를 통해 다중 노드 간의 메모리 자원 공유가 가능하며, 이러한 네트워크 구조는 향후 메모리 풀링(Pooling)과 메모리 쉐어링(Sharing) 기능을 지원할 수

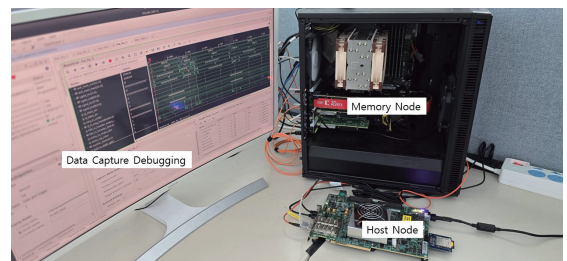


그림 5 외부 확장 메모리를 위한 시스템 구축



있는 유연한 확장성을 제공한다. 시스템 전반의 메모리 자원을 효율적으로 관리하고, 다양한 워크로드에 따라 동적으로 할당함으로써 메모리 활용을 최적화할 수 있다.

### IV. 결론

고속 패브릭 메모리 연결망 기술은 데이터센터 및 클라우드 컴퓨팅 환경에서의 데이터 처리와 자원 활용 효율성을 극대화하기 위한 필수적인 기술이다. 본고에서는 이러한 고속 패브릭 메모리 연결망 기술의 최신 동향을 분석하여 각각의 특징과 장단점을 비교하였다.

특히 CXL(Compute Express Link) 기술은 PCIe 인터페이스를 기반으로 CPU와 메모리, 가속기 등 다양한 자원 간의 효율적인 통신을 가능하게 하여 시스템 전반의 성능을 극대화한다. CXL 기술은 단일 노드에서 다중 노드로의 확장성을 제공하며, 다양한 컴퓨팅 자원을 유연하게 연결하고 관리할 수 있는 단일 표준을 제공함으로써 메모리 중심 컴퓨팅의 핵심 기술로 자리매김하고 있다.

또한, 기존의 고속 메모리 패브릭 연결망 표준과 달리 OmniXtend는 누구나 활용할 수 있는 개방형 하드웨어 플랫폼인 RISC-V를 기반으로 캐시 일관성을 보장하는 패브릭 메모리 연결망 기술을 제공하여, 미래 지향적인 메모리 중심 컴퓨팅 기술의 연구를 가능하게 한다.

고속 패브릭 연결망을 통해 데이터 처리 속도와 효율성을 극대화하여 컴퓨팅 자원의 활용도를 높일 수 있으며, 다양한 기술들의 융합과 표준화는 호환성과 확장성을 제공하여 새로운 응용 프로그램의 개발과 적용을 용이하게 한다. 마지막으로 인공지능, 빅데이터 분석, 고성능 컴퓨팅 등 다양한 분야에서의 혁신을 이끌어낼 수 있다.

따라서, 고속 패브릭 메모리 연결망 기술의 지속적인 연구와 발전은 미래의 컴퓨팅 환경에서 필수적인 요소가 될 것이며, 이를 통해 더 나은 성능과 효율성을 제공하는 혁신적인 컴퓨팅 시스템이 구현될 것이다. 앞으로도 이러한 기술들의 발전 동향을 주의 깊게 살펴보고, 새로운 기술의 도입과 적용을 통해 더욱 발전된 고성능 컴퓨팅 환경을 구축해 나가야 할 것이다.

#### 용어해설

**고속 패브릭 연결망** CPU, 메모리, 스토리지 등 다양한 컴퓨팅 자원을 고속 네트워크로 연결하여 데이터를 빠르고 효율적으로 전송하는 기술로, 낮은 대기시간과 높은 전송속도를 보장

**메모리 중심 컴퓨팅** 메모리 접근을 중심으로 설계한 컴퓨팅 아키텍처로, 메모리가 시스템의 핵심 역할을 수행하여 데이터 전송 병목현상과 대기시간을 단축하여 시스템 성능을 향상시킴

#### 약어 정리

CCIX	Cache Coherent Interconnect for Accelerators
CXL	Compute eXpress Link
DCB	Data Center Bridging
IPC	Inter-Processor Communication
iWARP	internet Wide Area RDMA Protocol
MRIOV	Multi-Root IO Virtualization
NAT	Network Address Translation
NDR	Next Data Rate
NRZ	Non-Return Zero
NTB	Non Transparent Bridge
NUMA	Non Uniform Memory Access
OpenCAPI	Open Coherent Accelerator Processor Interface
PAM	Pulse Amplitude Modulation
PCI	Peripheral Component Interconnect
PCIe	PCI Express
PCI-SIG	Peripheral Component Interconnect Special Interest Group
QoS	Quality of Service
QPI	Quick Path Interconnect

RDMA Remote Direct Memory Access  
RoCE RDMA over Converged Ethernet

### 참고문헌

- [1] 차승준 외, "고속 패브릭 연결망으로 도약하는 메모리 분리 기술 동향과 실현 전망," 주간기술동향, 2133호, 2024, pp. 17-30.
- [2] 김선영 외, "CCIX 연결망과 메모리 확장기술 동향," 전자통신 동향분석, 제37권 제1호, 2022, pp. 42-52.
- [3] M. Radi et al., "OmniXtend: Direct to Caches Over Commodity Fabric," in IEEE Symp. High-Performance Interconnects(Santa Clara, CA, USA), 2019, pp. 59-62.
- [4] Intel, "Intel Rack Scale Design(Intel RSD) Architecture Specification v2.5," Jul, 2019.
- [5] K. Keeton, "The Machine: An Architecture for Memory-centric Computing," in Int. Workshop Runtime Oper. Syst. Supercomput., (Portland, OR, USA), 2015. p. 1.
- [6] PCI-SIG, <https://pcisig.com>
- [7] Ethernet Alliance, <https://ethernetalliance.org>
- [8] 김선영 외, "CXL 인터커넥트 기술 연구개발 동향," 전자통신 동향분석, 제38권 제5호, 2023, pp. 23-33.
- [9] OpenCAPI Consortium, "OpenCAPI 4.0 Transaction Layer Specification," Jun, 16, 2020.
- [10] TechPowerUp, "OpenCAPI Consortium Merges Into CXL," Aug. 2, 2022.
- [11] CCIX Consortium, "An Introduction to CCIX: White Paper," Nov, 2019.
- [12] RoCEv2 Considerations, <https://enterprise-support.nvidia.com/s/article/roce-v2-considerations>
- [13] M. Ahmed and X. Xiang, "Performance Analysis of RoCEv2 in Data Center Networks," IEEE Trans. Netw. Ser. Manag., vol. 17, no. 1, 2020, pp. 55-66.
- [14] N. Kumar and K. Vaddina, "RoCE: Efficient Ethernet-based Data Center Solution," J. Netw. Comput. Applcat., vol. 135, 2019, pp. 1-10.
- [15] A. Gupta and P. Agrawal, "High Performance Computing with InfiniBand: A Review," J. Supercomput., vol. 77, no. 3, 2021, pp. 2312-2330.
- [16] S. Radhakrishnan and S. Bhandarkar, "iWARP: RDMA over TCP/IP for Data Centers," ACM Comput. Sur., vol. 51, no. 6, 2018, pp. 124-137.
- [17] G. Casey, "Gen-Z: High-Performance Interconnect for the Data-Centric Future," OCP Summit 2018, San Jose, CA, USA.
- [18] W.-o. Kwon et al., "Gen-Z memory pool system implementation and performance measurement," ETRI J., vol. 44, no. 3, 2021, pp. 450-461.
- [19] BusinessWire, "CXL Consortium Signs Agreement with Gen-Z Consortium to Accept Transfer of Gen-Z Specifications and Assets," Feb, 10, 2022.
- [20] CXL Consortium, <https://www.computeexpresslink.org>
- [21] K. Lender, "Compute Express Link(CXL): A Coherent Interface for Ultra-High-Speed Transfers," Flash Memory Summit 2019, Santa Clara, CA, USA.
- [22] D.D. Sharma, "Compute Express Link 3.0," CXL Consortium White Paper, Dec, 2023.
- [23] ZDNET Korea, "삼성-SK하이닉스, 이번엔 CXL 기술 경쟁... 제2의 HBM 노린다," 2024. 3. 25.
- [24] M. Radi et al., "OmniXtend: Direct to Caches Over Commodity Fabric," in IEEE Symp. High-Perform. Interconnects, (Santa Clara, CA, USA), 2019.
- [25] OmniXtend 1.0.3 Specification, <https://github.com/chipsalliance/omnixtend/blob/master/OmniXtend-1.0.3/spec/OmniXtend-1.0.3.pdf>
- [26] TileLink Specification 1.8.0, Chipalliance, available at <https://github.com/chipsalliance/omnixtend/blob/master/OmniXtend-1.0.3/spec/TileLink-1.8.0.pdf>, SiFive TileLink Specification, Dec. 3, 2018.
- [27] 김영우 외, "Roofline 모델을 통한 OmniXtend 기반 RISC-V 시스템의 성능 고찰," 대한전자공학회 하계학술대회, 2024, pp. 2595-2598.
- [28] Rocket Chip Generator, Chipsalliance, available at <https://github.com/chipsalliance/rocket-chip>
- [29] 권혁제 외, "원격 메모리 접근 및 확장을 위한 OmniXtend 호스트 시스템 구현 및 실험," 대한전자공학회 하계학술대회, 2024, pp. 2680-2683.