

AI 메모리 반도체 분야 주요 기술 및 시장 동향

Key Technologies and Market Trends in Artificial Intelligence (AI) Memory Semiconductors

홍아름 (A.R. Hong, areumh@etri.re.kr) 미래전략연구실 선임연구원
이재성 (J.S. Lee, jaeseonglee@etri.re.kr) 기술경제연구실 선임연구원
연주원 (J.W. Yeon, juwon@etri.re.kr) 반도체연구정책센터 연구원

ABSTRACT

As artificial intelligence (AI) technology advances, memory semiconductors supporting AI accelerators are becoming increasingly important, alongside improvements in computational performance. The demand for AI-optimized memory with high bandwidth and low latency is growing rapidly, and the memory market continues to expand substantially on this basis. AI memory refers broadly to memory architectures designed for AI computing workloads. A representative example is high-bandwidth memory (HBM), which delivers an ultrahigh memory bandwidth. Additional memory types are utilized according to their specific application requirements: graphics double data rate (GDDR) memory, primarily used in graphics processing and mid-range AI accelerators; double data rate (DDR) memory, serving as main memory in server environments; and low-power double data rate (LPDDR) memory, deployed in on-device environments owing to its low-power characteristics. Several emerging technologies have recently attracted considerable attention, including compute express link (CXL)-based memory pooling, high bandwidth flash (HBF), which overcomes the capacity limitations of HBM by vertically stacking nonvolatile memory, and three-dimensional DRAM (3D DRAM), which addresses the scaling limitations of conventional DRAM. This study reviews the technological and market trends of these AI memory technologies, and derives implications for navigating competition in the AI and semiconductor industries.

KEYWORDS AI, CXL, DDR, GDDR, HBF, HBM, LPDDR, Memory, 3D DRAM

I. 서론

최근 생성형 AI(Artificial Intelligence) 등 다양한 AI 기술이 산업 전반으로 빠르게 확산되면서, AI 가속기 연산 성능뿐만 아니라 이를 뒷받침하는 메모리

의 중요성이 크게 부각되고 있다. 현재 적용되고 있는 LLM(Large Language Model), LMM(Large Multimodal Model), LAM(Large Action Model) 등은 방대한 양의 데이터를 실시간으로 처리해야 하며, 이에 따라 고대역폭·저지연 특성을 갖춘 AI 특화 메모리에 대한

* DOI: <https://doi.org/10.22648/ETRI.2026.J.410302>

* 본 연구는 한국전자통신연구원 연구운영지원사업의 일환으로 수행되었음[26ZF1130, ICT 국가기술전략 정책연구].

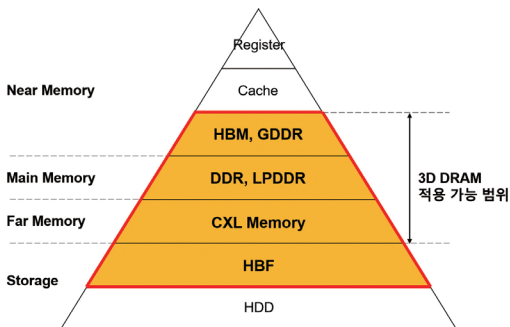


수요가 급증하고 있다. 이러한 수요 확대에 힘입어 메모리 시장은 빠르게 성장하고 있으며, 시장조사 업체 트렌드포스는 2026년 메모리 시장 규모가 최대 5,516억 달러(약 806조 원)에 이를 것으로 전망하였다[1]. 특히, AI 특화 메모리의 대표주자인 HBM (High Bandwidth Memory)은 공급 부족 현상을 겪고 있으며, HBM뿐만 아니라 서버용 DRAM 메모리 등도 공급-수요 격차로 인해 급격한 가격 상승과 매출 증가 추세를 보인다.

이처럼 AI 메모리의 중요성은 계속 강화되고 있으며, 단순한 보조 부품이 아니라 AI 시스템의 성능과 효율성을 좌우하는 핵심 요소로 자리매김하고 있다. 앞으로 AI 및 반도체 산업에서의 경쟁에 대응하기 위해서는 이러한 AI 메모리에 대한 이해와 동향 파악이 중요한 상황이며, 이에 따라 본고에서는 AI 메모리의 기술 및 시장 동향을 파악하고 시사점을 도출하고자 한다.

II. AI 메모리 기술 개념

AI 메모리 기술은 전통적인 메모리 계층 구조(Register-Cache-RAM-Storage)를 바탕으로 정의된다. AI 연산 환경에서 요구되는 고대역폭 데이터 처리, 대용량 메모리 제공 및 메모리 확장성에 주목하여,



출처 ETRI ICT전략연구소 기술전략연구본부 작성

그림 1 AI 메모리 계층 구조

표 1 AI 메모리 기술 주요 특징

세부 영역	특징
Near Memory	(HBM, GDDR) GPU/AI 가속기에 직접 연결되어 대규모 병렬 연산에서 데이터 이동 병목을 최소화
Main Memory	(DDR, LPDDR) 서버·모바일 장비에서 CPU/SoC에 연결되는 메인 메모리로, 연산에 필요한 데이터의 저장과 고속 접근을 담당
Far Memory	(CXL) CPU·가속기와 외부 메모리를 고속·일관성 기반으로 연결하는 인터커넥트 기술로, 메모리 확장 및 공유를 통해 자원 활용 효율을 제고
Storage	(HBF) NAND 기반 비휘발성 메모리로, AI 추론용 대규모 KV 캐시 및 데이터 저장을 지원

그림 1과 같이 해당 기능을 담당하는 메모리 계층을 중심으로 AI 메모리 기술 범위를 도식화하였다. 또한, 이와 같이 정의된 범위는 연산장치와의 거리와 기능적 역할을 기준으로 연산 인접 영역(Near), 시스템 내부 기본 메모리(Main), 시스템 외부 확장 영역(Far), 그리고 데이터 저장 영역(Storage)으로 구분하였다(표 1).

III. 주요 AI 메모리 기술 동향

1. HBM

1.1 개념

HBM(High Bandwidth Memory)은 여러 개의 DRAM 다이(Die)를 TSV(Through-Silicon Via) 기술을 이용해 수직으로 적층한 3D 스택 메모리 구조로, 신호 경로 단축과 병렬 데이터 처리 구조를 통해 대규모 데이터 전송 대역폭을 제공하는 고성능 메모리 기술이다.

1.2 기술 동향

전 세계 고성능 컴퓨팅(HPC: High Performance Computing)과 AI 기술의 발전으로 데이터 처리 성능에 대한 수요가 급증하는 가운데, HBM은 메모리 대역폭을 확장해 시스템 병목을 완화할 수 있는 핵

심 기술로 평가된다.

이 기술은 불과 2년 주기로 세대가 발전해 현재 4세대에 이르렀으며, 메모리 기술 중에서도 가장 빠른 혁신 속도를 보여준다. 세대가 진화할수록 대역폭·용량은 약 2배, 데이터 전송 속도는 약 50% 증가하는 추세이다.

이러한 기술 진화의 흐름 속에서 삼성전자는 2026년 2월 세계 최초로 HBM4를 양산 출하에 성공하였다고 밝혔다. 해당 제품은 JEDEC(Joint Electron Device Engineering Council) 표준(8Gbps)을 약 46% 상회하는 11.7Gbps의 데이터 처리 속도를 확보하였으며, 단일 스택 기준 총 메모리 대역폭도 HBM3E 대비 약 2.7배 향상된 3.3TB/s에 달한다[2]. 이는 AI 모델 확산에 따라 심화되는 데이터 병목 문제를 구조적으로 개선할 수 있는 수준으로 기대된다.

1.3 시장 동향

HBM 시장 규모 및 발전 추이

시장조사기관 Yole에 따르면, 글로벌 HBM 매출은 2024년 170억 달러에서 2030년 980억 달러로 증가할 것으로 전망되며, 연평균 성장률은 33%에 달한다. 또한, DRAM 시장 매출에서 HBM이 차지하는 비중은 2024년 18%에서 2030년 50%까지 확대될 것으로 전망된다[3].

수요 구조를 살펴보면, 2025년 기준 엔비디아 중심의 GPU가 전체 HBM 수요의 63.6%를 견인하며 압도적인 비중을 차지하고 있다[4]. 그러나 골드만 삭스에 따르면, 향후 특정 AI 알고리즘에 최적화된 맞춤형 ASIC(Application-Specific Integrated Circuit) 반도체에 대한 HBM 수요가 급증해 전체 시장의 3분의 1 수준까지 확대될 것으로 전망된다[5]. 이는 엔비디아의 범용 GPU(Graphics Processing Unit)가 주도

표 2 주요 기업별 HBM 관련 생산 및 투자 현황

기업명	투자 규모 및 주요 내용
삼성전자	<ul style="list-style-type: none"> 삼성디스플레이로부터 천안 공장(약 28만m²)을 임대하여 HBM 전용 첨단 패키징 라인을 구축 엔비디아와 협력해 5만 개 이상의 블랙웰 GPU를 기반으로, 생산 전 주기를 AI가 분석·예측·제어하는 반도체 AI 팩토리 구축
SK 하이닉스	<ul style="list-style-type: none"> 청주 및 용인 반도체 클러스터에 약 20조 원 규모 투자를 통해 HBM 제조 역량 강화 미국 인디애나 주에 차세대 HBM 생산을 위한 첨단 패키징 공장 건설
마이크론	<ul style="list-style-type: none"> 약 70억 달러를 투자해 2026년 가동을 목표로 싱가포르 내 HBM 패키징 공장을 착공 500억 달러를 투입한 미국 아이다호 공장에서 2027년부터 HBM 생산 착수 예정
CXMT	<ul style="list-style-type: none"> 중국 정부의 지원하에 HBM 개발에 주력 중이며, 약 300억 위안 규모 자금을 조달하여 관련 설비 확장 추진

출처: 사레별 자료 참고하여 저자 작성[6-11].

하던 HBM 수요가 구글, 아마존 등 빅테크 기업들이 자체 개발하는 맞춤형 AI 칩(ASIC) 영역으로 크게 확대되고 있음을 시사한다.

HBM 생산·투자 방향

HBM 시장의 가파른 성장세에 대응하여 주요 기업들은 주도권 선점을 목표로 대규모 투자 및 설비 확장을 통한 생산 능력 확대를 추진 중이다(표 2).

2. GDDR

2.1 개념

GDDR(Graphics Double Data Rate)은 GPU의 연산 속도에 맞춰 방대한 데이터를 빠르게 처리하기 위해 설계된 그래픽 전용 고성능 메모리를 말한다. 일반적인 시스템 메모리(DDR: Double Data Rate)와 달리 짧은 시간 내에 대량의 데이터를 한꺼번에 실어나르는 “대역폭” 확보에 특화되어 있어 영상 처리와 연산 가속에 최적화되어 있다. 이를 바탕으로 고해상도 그래픽 출력뿐만 아니라, 최근에는 초거대

AI 모델의 추론 성능을 극대화하는 핵심 솔루션으로 쓰인다.

2.2 기술 동향

GDDR6는 NRZ(Non-Return to Zero) 신호 방식과 QDR(Quad Data Rate) 커맨드 버스를 기반으로 제어 효율을 개선하면서, 2023년 업데이트된 JESD250D 규격에 이르기까지 16~24Gbps급 전송 속도의 기술적 기틀을 다졌다[12]. 이후 마이크론과 엔비디아의 협력으로 등장한 GDDR6X는 JEDEC 공식 표준은 아니나, 그래픽 메모리 최초로 PAM4(Pulse Amplitude Modulation) 신호 방식을 채택해 신호 전압을 4단계로 세분화함으로써 물리적 주파수 증가 없이 전송 효율을 이론상 두 배 향상시켰다[13,14]. 이는 차세대 표준인 GDDR7의 신호 체계 논의에 핵심적인 영향을 미쳤다.

2024년 3월 공식 발표된 GDDR7(JESD239)은 JEDEC 표준 DRAM 최초로 PAM3 신호 인터페이스를 도입했다[15]. PAM3는 2주기에 3비트를 전송하는 방식으로, NRZ 대비 전송 효율을 50% 개선하면서도 PAM4보다 설계 복잡도와 노이즈 민감도를 낮춰 실질적인 구현 안정성을 확보한 것이 특징이다[15,16]. 또한, 단순 속도 향상을 넘어 클라우드 컴퓨팅 및 AI 연산 환경을 겨냥한 RAS(Reliability, Availability, and Serviceability) 요구사항을 표준에 포함시켰다[15]. 2026년 현재 GDDR7은 초기 32Gbps에서 최대 48Gbps 수준의 속도 확장을 타진하며 고성능 컴퓨팅 생태계의 표준 솔루션으로 자리 잡고 있다[16].

2.3 시장 동향

GDDR 메모리 시장은 기존 게이밍 GPU 중심의 수요 구조를 넘어, 최근 AI 추론 전용 가속기 시장으로 그 영역을 빠르게 확장하고 있다. 2025년 하반기

표 3 GDDR 관련 기업 동향

기업	동향
삼성전자	<ul style="list-style-type: none"> • 업계 최초 GDDR7 개발 성공(24)에 이어, '25년 말부터 24Gb(3GB) 고용량 모듈 양산 돌입 • 24Gb 제품은 기존 16Gb 대비 용량을 50% 늘려 AI 워크스테이션과 데이터센터의 실장 효율 극대화 • 28~32Gbps급 제품을 중심으로 엔비디아의 블랙웰 아키텍처 기반 GPU에 주요 공급사로 참여 • '26년 초 기준 GDDR7 시장 점유율 약 70%를 상회할 것으로 추산
SK하이닉스	<ul style="list-style-type: none"> • HBM에서의 리더십을 GDDR로 전이시키기 위해 속도 경쟁 집중 • 2025년 1분기 GDDR7 양산 시작 • ISSCC 2026에서 업계 최고 수준인 48Gbps급 초고속 GDDR7 기술을 공개하며 기술적 우위 과시
마이크론	<ul style="list-style-type: none"> • 엔비디아와의 장기적 협력 관계를 바탕으로 GDDR7 초기 공급망에서 안정적인 입지를 확보 • GDDR6X부터 쌓아온 고속 신호 전송 노하우를 GDDR7 표준에 이식하며, 게이밍용 RTX 50시리즈에 칩을 공급 중

출처: 사레별 자료 참고하여 저자 작성[13,19-23].

기부터 본격화된 GDDR7 양산 체제는 삼성전자, SK하이닉스, 마이크론 등 메모리 3사 간의 기술 고도화 경쟁을 한층 가속화하고 있다[17](표 3). 특히 GDDR7은 HBM 대비 우수한 전력 효율과 가격 경쟁력을 바탕으로, 소비자를 위한 AI 추론 워크로드 구축의 핵심 솔루션으로 부상했다[18].

3. DDR

3.1 개념

DDR(Double Data Rate)은 CPU와 데이터를 주고받으며 시스템 성능을 뒷받침하는 표준 시스템 메모리다. 클럭 신호의 상승·하강 지점에서 모두 데이터를 전송해 기존 대비 전송 속도를 두 배로 높이는 것이 핵심이며, 운영체제 구동부터 복잡한 연산까지 폭넓게 지원한다. 최근에는 전력 효율과 대역폭을 극대화한 DDR5가 서버·클라우드 등 데이터 중심 환경의 필수 자원으로 자리 잡고 있다.

3.2 기술 동향

2020년 7월 발표된 DDR5 최초 규격(JESD79-5)은 단일 채널을 두 개의 독립 하위 채널로 분리하고 모듈 내 PMIC를 직접 탑재하는 구조적 혁신에 집중했다[24]. 이후 2021년 JESD79-5A로 5,600~6,400MT/s급 양산 기틀을 마련했고, 2022년 JESD79-5B에서는 24Gb·32Gb 고밀도 지원 사양을 확립하며 데이터센터 수요를 반영했다[24].

2024년 4월 발표된 최신 규격 JESD79-5C는 전송 속도를 최대 8,800Mbps로 상향하는 동시에, 로우해머 공격을 방어하기 위한 PRAC(Per-Row Activation Counting) 기술을 도입했다[25]. PRAC는 특정 행의 활성화 횟수를 정밀하게 추적해 데이터 탈취를 차단하는 기능으로, 데이터 신뢰성이 중요한 AI 서버 환경에서 필수적인 보안 장치로 평가받는다[26].

차세대 규격인 DDR6는 현재 표준 초안 작성 중이며, 초기 17,600MT/s에서 최대 21Gbps 수준의 속도를 타진하고 있어 DDR5 대비 두 배 이상의 성능이 기대된다[27,28]. 주요 제조사들은 표준 확정 전임을 고려해 신중함을 유지하고 있으나, 차세대 AI 연산의 핵심 솔루션으로서 업계의 기술적 검토는 이미 본격화되고 있다[29].

3.3 시장 동향

현재 글로벌 메모리 시장은 AI 인프라 투자가 가속화되면서 전례 없는 공급 부족과 가격 상승세를 경험하고 있다. 2026년 초 기준, DDR5 주요 규격의 현물 가격은 작년 9월 말 대비 약 다섯 배 수준으로 급등하며 시장 전반의 가격 지표를 흔들고 있다[30]. 이러한 현상은 AI 데이터센터 한 대당 탑재되는 D램 용량이 과거 일반 서버 대비 기하급수적으로 증가함에 따라, 서버용 메모리 조달을 위한 빅테크 기업 간의 확보 경쟁이 치열해진 결과로 분석된다[31].

메모리 제조사들의 생산 전략 변화 또한 시장 수

급 불균형을 심화시키는 핵심 요인으로 작용하고 있다. 삼성전자와 SK하이닉스 등 주요 기업들이 수익성이 압도적으로 높은 HBM 생산 라인을 공격적으로 증설하면서, 상대적으로 일반 DDR5 범용 D램의 생산 가능 물량이 축소되는 “공급 병목” 현상이 발생하고 있다[32]. 업계는 AI 데이터센터발 수요가 꺾이지 않는 한, 램 가격이 과거의 하락 주기처럼 빠르게 하향 안정화되기는 어려울 것이며, 당분간 고공행진을 지속할 것이라는 전망에 무게를 두고 있다[33].

이에 따라 업계 선두 주자들은 늘어나는 수요에 대응하고 원가 경쟁력을 확보하기 위해 10나노급 6세대(1c) 공정 전환에 박차를 가하고 있다. SK하이닉스는 세계 최초 1c DDR5 개발을 통해 기술 우위를 점하려 하고 있으며, 삼성전자는 내년까지 차세대 공정 비중을 전체 생산량의 3분의 1 수준으로 확대해 시장 지배력을 공고히 한다는 계획을 추진 중이다[34,35]. 또한, 중국의 CXMT가 8,000Mbps급 신제품 실물을 공개하며 점유율 5% 선을 돌파하는 등 후발 주자들의 추격도 새로운 변수로 부상하고 있다[36].

4. LPDDR

4.1 개념

LPDDR(Low Power Double Data Rate) 메모리는 스마트폰, 태블릿 등 모바일용으로 활용되는 저전력 DDR DRAM을 의미한다. 낮은 전력 소모와 빠른 속도, 소형 사이즈를 특징으로 하지만, 기존 DDR에 비해 오류정정 기능(ECC: Error Correction Code)이 약해 신뢰성 측면에서 단점이 있다. 초기에는 스마트폰, 태블릿, 노트북 등 모바일 기기에만 주로 활용되었으나, 최근에는 AI PC, 자동차, 데이터센터 등에도 활용되고 있다.

4.2 기술 동향

현재 상용화되어 사용되고 있는 규격은 LPDDR5X이며, JEDEC은 2025년 7월 JESD209-6 LPDDR6 표준을 발표하였다[37]. 이 표준은 모바일 기기뿐만 아니라 AI 등 다양한 용도에서 메모리 속도와 효율성을 크게 향상시키도록 설계되었으며, 2026년 상용화가 예상된다. LPDDR6는 AI 등 고성능 작업을 위해 듀얼 서브채널 아키텍처를 적용하여 효율적이고 유연한 운용이 가능하게 하였다. 또한, 낮은 전압과 저전력 소비가 가능한 VDD2 전원 사용을 통해 전력효율을 더욱 향상시키고, 내장(On-Die) 오류 정정 코드를 포함하여 보안 및 신뢰성을 강화하였다[38]. 데이터 전송속도는 최대 14.4Gbps가 구현 가능할 것으로 전망되고 있으나 [39], 아직까지 양산 제품은 최대 10.7Gbps 수준이다[40,41].

삼성전자와 SK하이닉스는 CES 2026에서 LPDDR6 제품을 공개하였으며[40,41], 삼성전자는 LPDDR6 공개로 CES 혁신상을 수상하였다[41]. SK하이닉스 역시 2026년 3월 LPDDR6 고객사 인증을 완료하였다. 중국의 CXMT는 2026년 LPDDR6를 상용화할 계획이며, 샤오미, 오포, 비보 등 중국 스마트폰 제조사를 수요처로 확보한 상황이다[42].

4.3 시장 동향

온디바이스 AI 디바이스에서의 활용

온디바이스(On-Device) AI 확산에 따라 자동차, XR(eXtended Reality) 기기, 스마트폰, AI PC 등 다양한 기기에서 고성능·저전력 연산 수요가 증가하고 있으며, 이에 따라 LPDDR의 중요성도 더 부각되고 있다. 시장조사기관 Verified Market Reports에 따르면, LPDDR 시장은 2024년 약 100억 달러 규모로 추정되며, 2033년 약 243억 달러까지 성장할 것으

표 4 LPDDR의 AI 서버 활용 관련 동향

기업	동향
퀄컴	<ul style="list-style-type: none"> 2025년 10월 AI 추론용 AI200, AI250 칩 기반 가속기 카드 및 랙 시스템을 공개 HBM 대신 LPDDR을 주요 메모리로 활용하여 비용·전력효율성 극대화
엔비디아	<ul style="list-style-type: none"> LPDDR기반AI서버용모듈형독자규격SOCAMM (SmallOutlineCompressionAttachedMemory Module) 제시 JEDEC은 SOCAMM2 표준화 중 SOCAMM2는 엔비디아 루빈 플랫폼의 Vera CPU에 탑재 예정이며, 메모리 3사는 양산 준비 중
마이크로소프트	<ul style="list-style-type: none"> 자사 AI 데이터센터에 LPDDR을 일부 도입할 예정 자체 개발한 오류 정정 알고리즘을 케이던스의 데이터센터용 LPDDR5X 메모리 시스템 솔루션에 탑재

출처 사레별 자료 참고하여 저자 작성[47-51].

로 예상된다[43].

차량용 메모리의 경우, 삼성전자, SK하이닉스는 차량용 플랫폼에 LPDDR5X를 공급 중이며, 자율주행시스템 안전 표준인 ASIL-D(Automotive Safety Integrity Level D) 등급 인증을 받기도 하였다[44,45]. 또한, 삼성전자는 HBM 활용이 어려운 모바일 및 엣지 서버용으로 LPDDR와 PIM(Processing In Memory)을 결합한 LPDDR-PIM을 개발 중이며, JEDEC 표준 제정을 진행 중이다[46].

AI 데이터센터에서의 활용

LPDDR은 약한 오류 정정 기능과 교체 및 용량 확장의 어려움 등으로 인해 서버에 잘 활용되지 않았으나, 최근 AI 데이터센터의 막대한 총 소유 비용(TCO: Total Cost of Ownership)과 저전력 수요 등으로 인해 AI 서버에서의 LPDDR 활용도 증가하고 있다(표 4).

5. CXL

5.1 개념

CXL(Compute Express Link)은 CPU, GPU, 메모리

간의 고속 연결을 지원하는 차세대 인터커넥트 기술이다. “캐시 일관성”을 핵심으로 여러 장치가 최신 데이터를 실시간으로 공유해 충돌 없이 처리 속도를 높이며, 메모리 풀링과 동적 자원 할당, 계층적 확장을 통해 대규모 데이터 처리 효율을 극대화한다.

5.2 기술 동향

CXL은 2019년 1.0 규격 발표 이후 불과 수년 만에 4.0 세대로 빠르게 진화하며 차세대 메모리·가속기 인터커넥트의 표준으로 자리매김했다. 초기 CXL 1.0/1.1은 PCIe(Peripheral Component Interconnect express) 5.0 기반으로 CPU와 가속기 간 캐시 일관성을 제공해 장치 간 데이터 복사를 생략하고 직접 메모리 접근을 가능케 했다[52]. CXL 2.0(2020년)은 스위칭과 메모리 풀링 기능을 도입해 여러 호스트가 메모리를 동적으로 공유·할당하는 유연한 구조를 갖췄으며, 데이터 암호화(IDE: Integrity and Data Encryption) 등 보안 기능도 표준화했다[52].

CXL 3.0(2023년)은 최대 4,096개의 호스트·장치 간 연결과 다계층 토폴로지를 지원하는 패브릭 아키텍처로 진화하며 데이터센터 전반의 확장성을 확보했다[52]. 2025년 11월 발표된 CXL 4.0은 PCIe 7.0 기반으로 전송 대역폭을 128GT/s(이전 세대 대비 2배)로 끌어올렸으며, 번들 포트(Bundled Port) 기능으로 최대 1.5 TB/s의 대역폭을 구현했다[53-55]. 이를 통해 캐시 일관성을 유지하면서도 100TB 이상 대규모 메모리 풀링이 이론적으로 가능해졌다[56].

5.3 시장 동향

초거대 AI 모델의 급격한 확산은 데이터센터의 연산·메모리 수요를 폭발적으로 증가시켰으며, 이에 따라 운영 효율화 기술이 핵심 과제로 부상하고 있다. 특히 CPU 코어 수는 지속적으로 증가하는 반

면, I/O 단자 수와 전력·발열 등 물리적 한계로 인해 코어당 메모리 대역폭은 오히려 감소하는 성능 불균형이 심화되고 있다[57]. 이러한 제약은 메모리 용량 부족에 따른 데이터 스와핑을 유발해 CPU 유휴 시간을 늘리는 비효율을 초래하며, AI 시대의 처리 데이터량 폭증으로 이 문제는 더 가속화될 전망이다[58]. CXL은 기존 SSD 기반 스와핑 대비 지연 시간을 100~1,000배 이상 단축할 수 있는 핵심 해결책으로 주목받고 있다[59].

CXL 생태계의 기반이 되는 CPU 시장에서는 인텔과 AMD가 지원 규격을 빠르게 확대하며 시장 형성을 주도하고 있다. 인텔은 2024년 서버용 제온 6에 CXL 3.0/3.1 지원을 도입했으며, AMD 역시 에픽 9005 시리즈를 통해 CXL 2.0 생태계를 확장하고 있다[60,61]. 메모리 제조사들의 제품 출시도 잇따르고 있다. 삼성전자는 2025년 메모리 풀링 기반 차세대 솔루션을 공개하며 하드웨어·소프트웨어를 아우르는 CXL 리더십을 강화하고 있고, SK하이닉스는 CXL 2.0 기반 CMM-DDR5(96GB) 제품을 출시해 고용량 메모리 확장 시장을 선점했다[62]. 마이크론 또한 CZ122 확장 모듈을 선보이며 글로벌 메모리 3사 간 CXL 주도권 경쟁이 본격화되고 있다[63].

스위치·컨트롤러 분야에서도 생태계 확충이 이어지고 있다. 엑스콘과 파네시아는 각각 CXL 2.0 및 3.1을 지원하는 고성능 스위치 칩을 공개하며 패브릭 아키텍처 구현의 핵심 동력을 제공하고 있으며, 람버스·퀄리티스반도체는 차세대 컨트롤러 및 PHY(Physical Layer) IP로 설계 기반을 다졌다[64-66]. 중국의 몬타지 테크 또한 CXL 3.1 컨트롤러를 출시하며 기술 추격에 나섰다[67]. 소프트웨어 측면에서는 레드햇과 멤버지가 주요 제조사의 CXL 모듈 인증 및 계층화 솔루션을 통해 기업용 시장의 신뢰성을 확보하며 기술 보편화를 앞당기고 있다[68,69].

6. HBF

6.1 개념

AI 추론 성능 고도화가 확대되면서 연산 성능뿐만 아니라 메모리 저장 용량의 중요성도 함께 확대되고 있다. 생성형 AI는 대화 맥락 유지를 위해 대규모 KV(Key-Value) 캐시 저장이 필요하나, 이를 DRAM 기반의 HBM만으로 처리하기에는 용량 및 비용 측면에서 한계가 존재한다.

이에 따라 낸드 플래시(NAND Flash)를 수직 적층한 HBF(High Bandwidth Flash)가 대안으로 주목받고 있다. HBF는 TSV 기반 3차원 적층 구조를 통해 고용량 낸드를 집적함으로써 HBM의 병렬 데이터 접근 구조를 계승한 고용량 메모리 기술이다.

6.2 기술 동향

2025년 2월 샌디스크가 발표한 로드맵에 따르면, HBF는 차세대 AI 메모리로 상용화를 준비하고 있다. 2026년 상용화를 목표로 하는 1세대 HBF는 HBM과 유사한 수준의 대역폭을 유지하면서도 8~16배에 달하는 대용량 구현을 목표로 한다. 이는 기존 HBM 중심 구조의 용량 한계를 보완하고, AI 추론 워크로드에 적합한 고용량 메모리 계층을 형성하려는 전략으로 해석된다. 이후 2세대에서는 1세대 대비 용량 1.5배 및 대역폭 1.45배 향상을, 3세대에서는 용량과 대역폭을 각각 2배 확대하는 방향으로 단계적 성능 고도화를 추진할 계획이다[70].

한편, SK하이닉스는 2026년 2월 HBM과 HBF를 결합한 하이브리드 메모리 구조인 H3 아키텍처를 제안하였다[71]. H3의 핵심은 GPU 인접 영역에 HBM과 HBF를 단일 인터포저상에 병렬 배치하여 고속 연산과 대용량 저장 기능을 통합하는 데 있다. 엔비디아 블랙웰 B200 GPU 기반 시뮬레이션 환경에서는 8개의 HBM3E와 HBF를 하나의 GPU에 동

시에 연결하는 구성을 제시하였으며, HBF에 대용량 데이터를 저장함으로써 동일 전력 대비 전체 처리량(Throughput per Power)을 기존 HBM 단독 구성 대비 최대 2.69배까지 향상시킬 수 있음을 확인하였다. 이는 차세대 AI 반도체 아키텍처가 단일 고대역폭 메모리 중심 구조에서 벗어나, HBM과 HBF를 결합한 계층형 메모리 구조로 진화하고 있음을 시사한다.

6.3 시장 동향

HBF 시장 규모 및 성장 전망

HBF는 AI 인프라 고도화와 함께 새로운 메모리 계층으로 부상하며 초기 시장을 형성하고 있다. 신영증권에 따르면, HBF 시장은 2030년경 약 120억 달러 규모로 성장할 것으로 전망되며[72], 이는 같은 시기 HBM 시장의 약 10% 수준에 해당한다.

이러한 전망은 HBF가 HBM을 직접 대체하기보다는, AI 시스템 내에서 고용량·고효율 메모리 계층을 담당하는 보완적 시장으로 자리매김할 가능성을 시사한다. 특히 LLM 추론 수요 확대와 함께 전력 대비 성능 및 총소유비용(TCO) 절감이 중요해지면서, 고가의 HBM을 일부 보완할 수 있는 HBF의 전략적 가치가 부각되고 있다.

주요 기업 동향

CES 2026에서 엔비디아 CEO 젠슨 황은 고가의 HBM을 보완하기 위한 새로운 메모리 아키텍처인 ICMS(Inference Context Memory Storage)를 제안하였다. 이는 AI 추론 특화 메모리 아키텍처로, 추론 과정에서 생성되는 KV 캐시 데이터를 낸드 플래시에 저장하여 GPU 메모리의 용량 한계를 보완하는 구조이다. 이를 통해 상대적으로 비용이 많이 드는 HBM 사용량을 줄이면서도 대규모 컨텍스트 데이

터를 효율적으로 관리할 수 있으며, AI 추론 환경에서 메모리 용량 확장과 비용 효율성을 동시에 확보할 수 있는 방식으로 주목받고 있다.

이러한 흐름에 대응하여 주요 기업들도 협력과 기술 개발을 추진하고 있다. SK하이닉스와 샌디스크는 2025년 8월 HBF 기술 표준화를 위한 양해각서(MOU: Memorandum of Understanding)를 체결하였으며, 샌디스크는 2026년 하반기 HBF 메모리 첫 샘플 공급을 목표로 개발을 진행 중이다[73].

7. 3D DRAM

7.1 개념

3D DRAM은 DRAM 셀 어레이 자체를 수직 적층하여 집적도와 면적 효율을 끌어올리는 소자 구조 및 공정 혁신을 의미한다. 현재의 2D DRAM은 1T1C(1 Transistor, 1 Capacitor) 구조로 구성되어 있으며, 커패시터의 부피를 줄이는 방식으로 집적도를 높여왔다. 하지만, 10nm 이하에서는 전류누설, 간섭·저항 증가, 발열 등 미세화에 물리적 한계가 나타나게 되면서 DRAM 구조 변화가 필요하게 되었다. 3D DRAM은 HBM처럼 DRAM을 여러 층으로 쌓는 것이 아니라, 셀 자체를 수직 구조로 새롭게 설계하여 이러한 한계를 극복하고자 하는 기술을 의미한다.

7.2 기술 동향

현재까지는 커패시터를 수평으로 배치하여 수직 적층하는 형태의 구조가 대표적이며, 커패시터-리스(Capacitor-less) 구조(2T0C, 1T0C)의 DRAM도 연구되고 있다.

벨기에의 반도체 연구기관 IMEC은 2025 IEEE VLSI에서 수직 비트라인(VBL: Vertical Bit Line) 구조 3D DRAM에 GAA(Gate-All-Around) 트랜지스터를

통합하는 공정 흐름에 관한 연구를 발표하였다[74]. 또한, 커패시터가 없는 2T0C 구조의 IGZO(Indium Gallium Zinc Oxide)-DRAM 기술도 개발 중이며, 향후 3D DRAM에도 적용 가능하며, 다만 신뢰성 확보가 중요하다고 언급하였다[75].

EUV(Extreme Ultra Violet) 노광 장비 구매 및 HK-MG(High-k Metal Gate) 공정 적용 등에 어려움을 겪고 있는 중국에서도 3D DRAM을 활발하게 연구 중이다. 중국과학원 산하 IMECAS(미세전자연구소)에서는 2024년 1월 IGZO 기반의 4F2 2T0C DRAM 셀을 개발하였으며, 3D DRAM에서도 응용될 수 있다고 언급하였다[76]. 북경대학교와 화중과학기술대학교는 협력하여 Science Advances 저널에 IGZO 기반 3D 스택형 2T0C DRAM 어레이 논문을 발표하였다[77].

7.3 시장 동향

현재 메모리 기업에서는 6F2 구조에서 4F2 구조 2D DRAM으로의 변화를 먼저 시도 중이며, 이후 3D DRAM으로 전환할 전망이다. 4F2 구조는 수직 트랜지스터를 통해 6F2 구조보다 사용 공간을 감소시켜 용량·속도·전력 면에서 유리하나, 커플링 문제가 발생할 수 있으며 2D DRAM의 본질적 문제인 미세화 문제 해결이 어려운 등 여러 한계가 있다.

SK하이닉스, 삼성전자, 마이크론, CXMT 등 주요 메모리 기업에서는 수평 커패시터 중심의 연구를 진행 중이며(표 5), 상용화 시점은 2035년경으로 예측된다. 3D DRAM의 기술 난이도가 높고 공정의 큰 변화가 필요하여 양산까지는 상당한 시간이 걸릴 것으로 전망된다.

IV. 결론

AI 기술의 발전으로 인해 AI 메모리의 중요성은

표 5 3D DRAM 관련 기업 동향

기업	동향
SK 하이닉스	<ul style="list-style-type: none"> 2024 IEEE VLSI에서 수직 비트라인 구조 3D DRAM 논문 발표. 아직 5단 수준. 수율은 56.1% 2025 IEEE VLSI 심포지엄에서 10nm 이하의 차세대 노드에 4F² 수직 트랜지스터와 3D DRAM을 채택할 것이라 강조
삼성전자	<ul style="list-style-type: none"> MemCon 2024에서 수직 채널 트랜지스터를 기반으로 한 3D DRAM 개발 로드맵 발표 2030년까지 양산을 목표로 함
마이크론	<ul style="list-style-type: none"> 공식적으로 로드맵 등을 발표하지는 않았으나, 3D DRAM 관련 특허 다수 확보 중
CXMT	<ul style="list-style-type: none"> 2023 IEEE IMW에서 수평 비트라인, 수직 워드라인, 수평 커패시터 구조의 3D DRAM 아키텍처에 대한 논문 발표

출처 사례별 자료 참고하여 저자 작성(78-82).

더욱 높아지고 있으며, 다양한 수요에 맞춰 HBM뿐만 아니라 여러 메모리가 다양한 분야로 영역을 확장해 가고 있다. AI 서버용으로 활용되던 DDR뿐만 아니라 LPDDR 역시 저전력 수요에 맞춰 AI 서버에 활용되기 시작하였으며, 게이밍용으로 활용되던 GDDR 역시 AI 추론용 메모리로서의 활용이 시도되고 있다. 고비용 및 공급 부족으로 인해 고용량 확보가 어려운 HBM을 보완하고자 HBF와 같이 비휘발성 메모리를 활용하는 사례도 등장하였다. 이와 같이 다양한 AI 메모리들이 각 용도에 맞춰 AI 산업에서 널리 활용되고 있다(표 6).

또한, CXMT 등을 중심으로 중국 기업이 메모리 시장의 새로운 플레이어로 등장하고 있으며, 메모리 품귀로 인해 중국산 저가형 메모리에 대한 관심 역시 높아질 것으로 예상된다. 국내 메모리 기업들은 이러한 변화에 대응하기 위해 HBM 등 고부가가치 메모리 기술력 확보, 생산 역량 확보, 표준 주도 등이 필요할 것으로 보이며, 3D DRAM 등 차세대 기술에 관한 연구 및 투자도 지속적으로 유지해야 할 것으로 보인다.

표 6 AI 메모리 기술별 주요 기술·시장 동향 요약

기술	주요 기술·시장 동향
HBM	<ul style="list-style-type: none"> AI 연산 확산으로 AI 시스템의 핵심 메모리로 부상 주요 기업들은 시장 성장세에 대비하여 대규모 투자 및 설비 확장 중 삼성전자는 HBM4 세계 최초 양산('26)
DDR	<ul style="list-style-type: none"> 메모리 기업을 중심으로 DDR6 표준 규격 논의 중 AI 투자 확대로 공급 병목 발생 주요 기업은 수요 대응 및 경쟁력 확보를 위해 10나노급 6세대(1c) 공정 전환 박차
GDDR	<ul style="list-style-type: none"> JEDEC의 GDDR7 표준 발표('24.3) 게이밍 GPU 중심 수요 구조를 넘어 AI 추론용으로 영역 확장
LPDDR	<ul style="list-style-type: none"> JEDEC의 LPDDR6 표준 발표('25.7) 온디바이스AI 확산에 힘입어 모바일 AI 기기 및 차량용 등으로 활용 엔비디아, 퀄컴, 마이크로소프트 등은 저전력 AI 서버용으로 활용 시도
CXL	<ul style="list-style-type: none"> AI 확산에 따른 데이터센터 메모리 병목 해소의 핵심 기술로 부상 CXL 4.0('25.11) 기준 최대 1.5TB/s 대역폭·100TB 이상 메모리 풀링 지원 메모리 3사 및 인텔·AMD 등이 생태계 확장에 참여하며 주도권 경쟁 본격화
HBF	<ul style="list-style-type: none"> HBF 시장은 HBM 시장의 10% 수준의 보완적 시장으로 자리 잡을 것 샌디스크는 '26년 1세대 상용화를 목표로 로드맵 제시('25.2) SK하이닉스, 엔비디아 등 주요 기업은 HBM을 보완하기 위해 HBF를 활용하는 자체 아키텍처 제시('26)
3D DRAM	<ul style="list-style-type: none"> 수평커패시터 구조, 커패시터-리스 구조(2TOC, 1TOC)로 연구 중 주요 기업에서는 수평커패시터를 중심으로 연구 중이며, 상용화는 다소 늦어질 전망(2035년경) EUV 장비 수입이 어려운 중국에서도 적극적으로 연구 중

출처 본문 기반 저자 작성

용어해설

DRAM 전원이 꺼지면 데이터가 사라지는 휘발성 메모리로, 빠른 속도와 저렴한 비용으로 인해 주메모리로 활용됨

JEDEC 반도체 산업에서 메모리 및 전자부품의 규격과 표준을 정하는 국제 표준화 단체

Through Silicon Via 반도체 칩 내부의 실리콘을 관통하는 미세 전극을 형성하여 칩 간을 수직으로 연결하는 3차원 집적 기술

참고문헌

- [1] 김건, "AI 슈퍼사이클에 메모리 질주...올 시장 800조 전망," 디일렉, 2026. 2. 10.
- [2] 삼성전자, "삼성전자, 세계 최초 업계 최고 성능의 HBM4 양산 출하," Samsung Newsroom, 2026. 2. 12.
- [3] Yole Group, "Status of the Memory Industry 2025," Yole Group, 2025.
- [4] Mordor Intelligence, "HIGH BANDWIDTH MEMORY MARKET SIZE & SHARE ANALYSIS - GROWTH TRENDS AND FORECAST (2026 - 2031)."
- [5] SK hynix, "2026 Market Outlook - "Focus on the HBM-Led Memory Supercycle"," SK hynix Newsroom, 2026. 1. 5.
- [6] 김소연, "삼성전자, 천안에 반도체 패키징 공정 증설...HBM 생산 예정," 연합뉴스, 2024. 11. 12.
- [7] 강태우, "SK하이닉스, 'HBM 생산' M15X에 첫 장비반입...연내 준공 순항," 연합뉴스, 2025. 10. 27.
- [8] Micron, "Micron Breaks Ground on New HBM Advanced Packaging Facility in Singapore," Press Release, 2025. 1. 8.
- [9] 김용원, "SK하이닉스 삼성전자 HBM 겨냥한 중국 추격 거세져, CXMT 상장으로 자금 조달," Business Post, 2025. 10. 22.
- [10] 김채연, "마이크론, 아이다호 공장서 내년 중반 HBM 생산 시작," 한국경제, 2026. 2. 18.
- [11] SK하이닉스, "SK하이닉스, 美 인디애나 주와 첨단 후공정 분야 투자협약 체결," 뉴스룸, 2024. 4. 4.
- [12] JEDEC, "Graphics Double Data Rate (GDDR6) SGRAM Standard," 2023. 5.
- [13] Micron, "Micron Samples Next-Gen Graphics Memory for Gaming and AI," Press Release, 2024. 6. 4.
- [14] B. Randolph, "GDDR6X: Memory reimagined," Micron Blog, 2020. 9.
- [15] JEDEC, "JEDEC Publishes GDDR7 Graphics Memory Standard," 2024. 3. 5.
- [16] Rambus, "All You Need to Know About GDDR7," Rambus Press, 2025. 5. 29.
- [17] 고명훈, "GDDR7 본격 개화...삼성·SK, 차세대 그래픽 D램 선점전 '치열'," 시사저널, 2025. 5. 28.
- [18] 문준현, "Samsung's GDDR7 wins Presidential Prize as key memory in AI era," 코리아헤럴드, 2025. 12. 3.
- [19] Samsung, "Samsung Develops Industry's First 24Gb GDDR7 DRAM for Next-Generation AI Computing," Samsung Newsroom, 2024. 10. 17.
- [20] 황정수, "그래픽D램도 잡았다...삼성, GDDR7 시장 70% 장악," 한경 반도체인사이드, 2025. 12. 4.
- [21] Tech Power Up, "SK Hynix to Showcase 48 Gb/s 24 Gb GDDR7 Memory at ISSCC 2026," 2025. 11. 27.
- [22] Tech Power Up, "SK Hynix Targets Q1 2025 for GDDR7 Memory Mass Production," 2024. 6. 12.
- [23] 이나리, "젠스 황 엔비디아, 'RTX' 공개하며...삼성·SK 아닌 '마이크론 탑재'," ZNET Korea, 2025. 1. 7.
- [24] JEDEC, "JEDEC Publishes New DDR5 Standard for Advancing Next-Generation High Performance Computing Systems," 2020. 7. 14.
- [25] JEDEC, "JEDEC Updates JESD79-5C DDR5 SDRAM Standard: Elevating Performance and Security for Next-Gen Technologies," 2024.04.17.
- [26] Tech Power Up, "JEDEC Updates DDR5 Specification for Increased Security Against Rowhammer Attacks, New DDR5-8800 Reference Speed," 2024. 4. 23.
- [27] 이상호, "JEDEC, 2025년 2분기 DDR6 규격 완성 목표," 보드나라, 2024. 5. 22.
- [28] 이지용, "DDR6 D램 곧 시장 개화...메모리 3사, 제품 준비 분주," 뉴시스, 2025. 7. 29.
- [29] 손새로, "DDR5보다 2배 빠른 'DDR6', 언제 나오나?...삼성·하이닉스 "현재 표준 규격 논의 중"," 녹색경제신문, 2025. 1. 15.
- [30] 이성재, "DDR5 램값 고공행진...AI 데이터센터가 수ופן 흔들었다," 금융경제신문, 2026. 2. 9.
- [31] 이우림, "'메모리 없어서 못 만든다' 비명 지르는 빅테크...휘청이는 실적 전망에 中 D램까지 기웃," 중앙일보, 2026. 2. 6.
- [32] 박정환, "HBM 쓸림의 부메랑...'범용 D램 씨 말랐다, 2026년 가격 20% 폭등'," 글로벌이코노믹, 2025. 12. 20.
- [33] 이형수, "AI가 반도체 사이클 바꿨다 침체 짧아지고 호황 길어진다," 연합인포맥스, 2025. 11. 7.
- [34] SK하이닉스, "SK하이닉스, 세계 최초 10나노급 6세대 D램 '1c DDR5' 개발," 뉴스룸, 2024. 8.29.
- [35] 장경윤, "마이크론, '6세대 10나노급' D램 샘플 공급...삼성·SK보다 빨랐다," 지디넷, 2025. 2. 26.
- [36] 이우림, "中 CXMT, 'DDR5'신제품 공개...삼성·SK하이닉스 턱밑 추격," 중앙일보, 2025. 11. 25.
- [37] JEDEC, "JEDEC Releases New LPDDR6 Standard to Enhance Mobile and AI Memory Performance," 2025. 7. 9.
- [38] 박진형, "차세대 저전력 D램 'LPDDR6' 표준 발표...대역폭 20% 개선," 전자신문, 2025. 7. 10.
- [39] S. Knowlton, "LPDDR6 vs. LPDDR5 and LPDDR5X: What's the Difference?," Synopsys Blog, 2026. 2. 10.
- [40] 정일주, "SK하이닉스, 10나노급 6세대 LPDDR6 고객사 인증 완료," 디일렉, 2026. 3. 10.
- [41] <https://www.ces.tech/ces-innovation-awards/2026/lpddr6/>
- [42] 박정환, "중국, 'AI 메모리' LPDDR6 2026년 상용화 선언...삼성·하이닉스 '초격차' 1년으로 좁혔다," 글로벌이코노믹, 2026. 1. 22.

- [43] Verified Market Reports, "Global LPDDR RAM Market Size, Industry Trends & Forecast 2026-2034," 2025. 2. <https://www.verifiedmarketreports.com/product/lpddr-ram-market/>
- [44] 최지은, "SK하이닉스 LPDDR5X 차량용 D램, 'ASIL-D' 인증 획득," 머니투데이, 2026. 1. 19.
- [45] 삼성전자, "삼성 12nm급 오토모티브 LPDDR5X: 중앙 집중식 자동차 시스템을 위한 고성능, 고신뢰성 DRAM 솔루션," 삼성전자 테크블로그, 2025. 7. 23.
- [46] 장경윤, "삼성전자 '차세대 메모리 LPDDR6-PIM 표준 거의 완성,'" 지디넷, 2025. 12. 16.
- [47] 장경윤, "퀄컴, AI 추론용 신규 가속기 출시...엔비디아·AMD에 도전장," 지디넷, 2025. 10. 28.
- [48] 장경윤, "엔비디아, SOCAMM 적용 '루빈'으로 연가...삼성·SK도 공급 전략 수정," 지디넷, 2025. 5. 14.
- [49] Kyle Aubrey, "Inside the NVIDIA Vera Rubin Platform: Six New Chips, One AI Supercomputer," NVIDIA Technical Blog, 2026. 1. 5.
- [50] 김소연, "메모리3사 일제히 SOCAMM2 선포...'제2의 HBM' 맞붙," 이데일리, 2025. 10. 23.
- [51] 한주엽, "MS, AI 데이터센터 'LPDDR 시대' 연다," 디일렉, 2026. 1. 27.
- [52] 김선영 외, "CXL 인터커넥트 기술 연구개발 동향" 전자통신동향분석, 제38권 제5호, 2023, pp. 23-33.
- [53] CXL Consortium, "CXL Consortium Releases the Compute Express Link 4.0 Specification," 2025. 11. 18.
- [54] VideoCardz, "CXL 4.0 spec moves to PCIe 7.0, doubles bandwidth over CXL 3.0.," 2025. 11. 24.
- [55] S. Mitchell, "CXL 4.0 doubles bandwidth, introduces bundled ports for data centres," DatacenterNews. 2025. 11. 20.
- [56] B. Crosley, "CXL 4.0과 인터커넥트 전쟁: AI 메모리가 데이터 센터 아키텍처를 재편하는 방법," Introl Blog, 2026. 1. 16.
- [57] C. Petersen, "Software Defined Memory: A Meta perspective," OCP Global Summit, 2021. 11. 8.
- [58] 이동수, "Agentic AI 최적화 이슈에 대한 간단한 소개 - KV cache," LinkedIn, 2025.
- [59] ABI Research, "Opportunities And Challenges For Compute Express Link (CXL)," 2024.
- [60] 미디어어픽, "AMD, 5세대 'AMD EPYC' 임베디드 프로세서 공개," 2025. 3. 11.
- [61] D. Benson, "Intel Launches New Data Center Processor Family at Computex 2024," All About Circuits, 2024. 6. 4.
- [62] SK hynix, "SK하이닉스, CXL 2.0 기반 DDR5 고객 인증 완료 '데이터센터 메모리 혁신 선도'," SK hynix Newsroom, 2025. 4. 23.
- [63] R. Sehgal et al., "Optimizing System Memory Bandwidth with Micron CXLTM Memory Expansion Modules on Intel® Xeon® 6 Processors," Intel White Paper, 2024.
- [64] 파네시아, "파네시아, PCIe 6.4 / CXL 3.2 패브릭 스위치 샘플 공개...업계 최초 포트-기반 라우팅(PBR) 포함 CXL 3.2 표준 전체 기능 구현한 ASIC 스위치 칩," Pannesia Newsroom, 2025. 11. 20.
- [65] Rambus, "CXL 2.0 Controller," Products Catalog. <https://www.rambus.com/interface-ip/cxl/cxl2-controller/>
- [66] 장경윤, "퀄리카스반도체, 메 버리실리콘에 PCIe 6.0·CXL 3.0용 IP 공급," 지디넷, 2025. 1. 9.
- [67] Montage Technology, "Montage Technology Introduces CXL® 3.1 Memory eXpander Controller to Empower Next-Generation Data Center Infrastructure," Press Release, 2025. 9. 1.
- [68] 삼성전자, "삼성전자, 레드햇과 업계 최초 CXL 메모리 동작 검증," Samsung Newsroom, 2023. 12. 27.
- [69] MemVerge, "MemVerge Memory Machine Achieves Red Hat OpenShift Operator and Container Certification," PR Newswire, 2022. 7. 7.
- [70] Sandisk, "Future FWD: Sandisk 2025 Investor Day," <https://investor.sandisk.com/static-files/79481580-ada2-4e08-bdeb-4b440d08f4ab>
- [71] M. Ha et al., "H3: Hybrid Architecture Using High Bandwidth Memory and High Bandwidth Flash for Cost-Efficient LLM Inference," IEEE Comput. Architect. Lett., vol. 25, no. 1, 2026, pp. 49-52.
- [72] 박상욱, "낸드의 계절, HBF가 불러온 혁신," 신영증권, 2025. 9. 22.
- [73] Sandisk, "Sandisk to Collaborate with SK hynix to Drive Standardization of High-Bandwidth Flash Memory Technology," Sandisk Newsroom, 2025. 8. 6.
- [74] N. Rassoul et al., "Process Insights into 3D-DRAM with Vertical Bit Line and Scalable GAA Transistor," in Proc. IEEE Symp. VLSI Technol. Circuits, (Kyoto, Japan), June 2025, pp. 1-3. doi: 10.23919/VLSITechnologyandCir65189.2025.11075009
- [75] A. Belmonte and G.S. Kar, "Disrupting the DRAM roadmap with capacitor-less IGZO-DRAM technology," Nat. Rev. Electr. Eng., vol. 2, 2025, pp. 220-221.
- [76] https://ime.cas.cn/zhxx/zhxw/202401/t20240111_6952284.html
- [77] Q. Li et al., "3D stacked IGZO 2TOC DRAM array with multibit capability for computing in memory applications," Sci. Adv., vol. 11, no. 21, 2025, doi: 10.1126/sciadv.adu4323
- [78] 강해령, "SK하이닉스 '3D D램'의 비밀," 서울경제, 2024. 6. 29.
- [79] SK하이닉스, "SK하이닉스, IEEE VLSI 2025에서 D램 미래 기술 로드맵 발표," SK hynix Newsroom, 2025. 6. 10.

- [80] 임효진, “삼성전자, ‘꿈의 메모리’ 3D D램 올해 공개 예고…AI 시대 승부수?,” 경제일보, 2025. 4. 22.
- [81] KIPOST, “CXMT, 3D D램 직행할수도…3D D램 전환 속도내야,” 2024. 1. 31. <https://www.kipost.net/news/articleView.html?idxno=316201>
- [82] M. Huang et al., “A 3D Stackable 1T1C DRAM: Architecture, Process Integration and Circuit Simulation,” in Proc. 2023 IEEE Int. Memory Workshop, (Monterey, CA, USA), May. 2023, pp. 1-4. doi: 10.1109/IMW56887.2023.10145931.