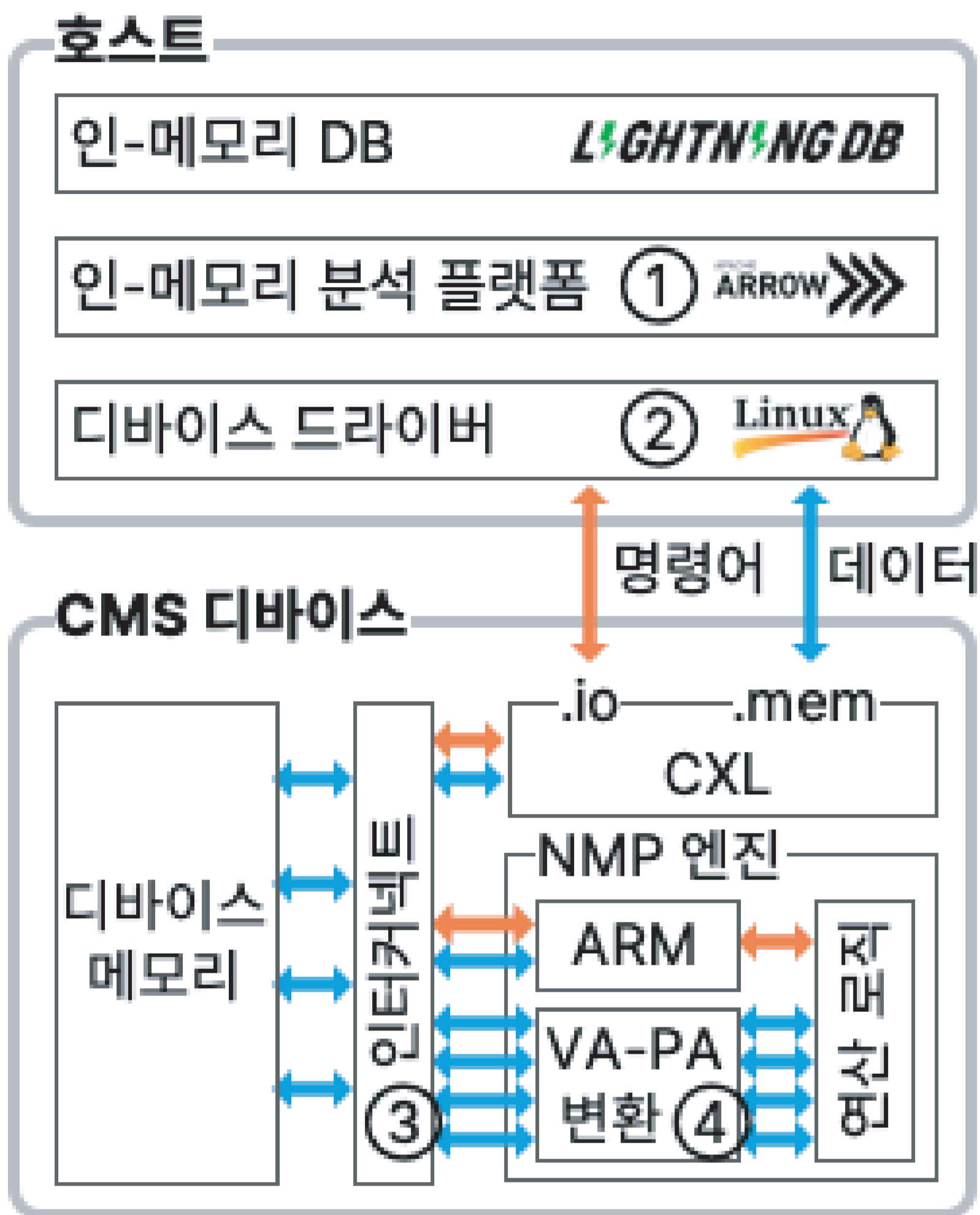


CXL-CMS¹ 2.0

세계 최초 NMP (Near-Memory Processing) 적용 CXL 메모리 솔루션

CPU와 CXL 메모리 간의 데이터 이동 최소화 (Memory-Intensive 워크로드에 최적)
Offloading을 통한 CPU 사용률 최소화 (다른 작업에 CPU 자원 활용 가능)



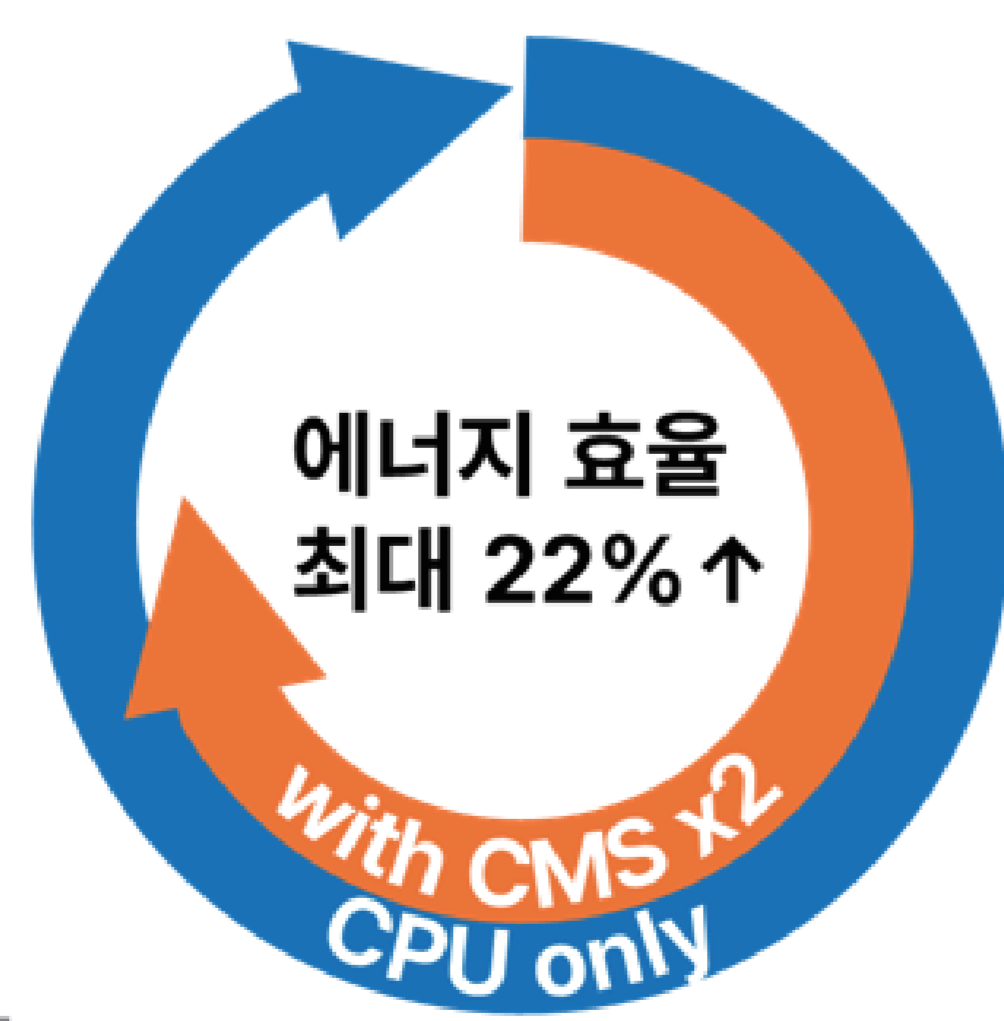
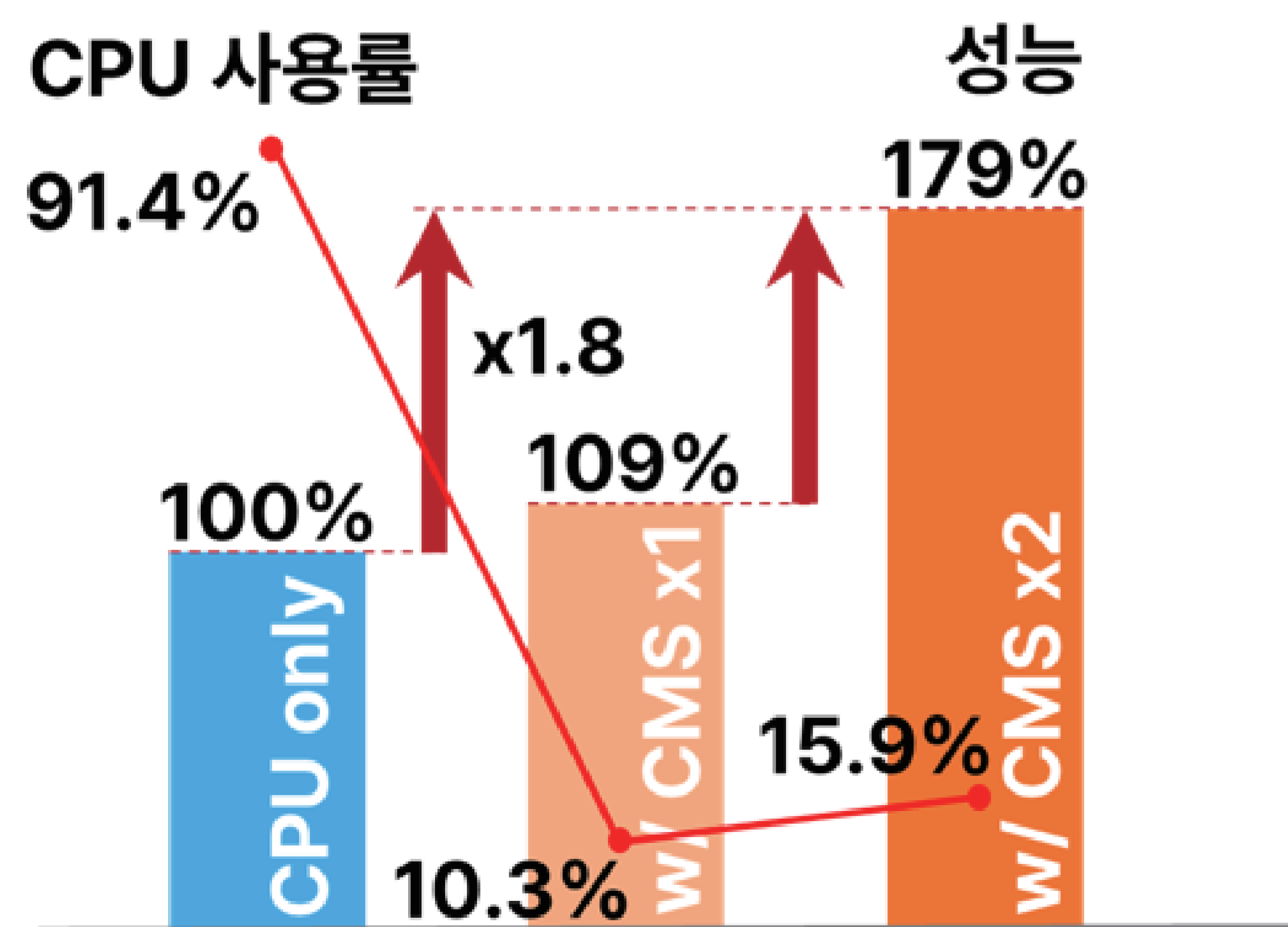
- ① 데이터 분석 APIs 및 라이브러리 제공 (Apache Arrow 지원)
- ② 동적 CXL 메모리 할당 지원
- ③ 채널 Interleaving 지원 (메모리 대역폭 최대 활용)
- ④ VA-PA (Virtual Address-Physical Address) 변환 지원

2.0 주요 특징

- 시스템 메모리 지원 강화
- 연산 오프로딩 지원 강화

폼팩터	Add-In-Card
인터커넥트	CXL 2.0 / PCIe x8
메모리	512GB / 76GB/s
컨트롤러	ARM Cortex [®] -A72 x2

성능 / CPU 사용률 / 에너지 효율



- 성능 1.8배
- CPU 사용률 75%p ↓
- 에너지 효율 22% ↑

데이터 분석 플랫폼

Lightning DB w/ CMS

SK텔레콤의 자체제작 인-메모리 DB로 실시간 데이터 분석 서비스에 활용

Use Case 1 유동인구 분석 (Filter)

- 실시간 유동인구 분석 플랫폼에 적용
- 서버 CPU 급의 데이터 처리 성능 제공
- NMP 컨셉이 실제 서비스 성능 & 에너지 효율 개선에 도움이 됨을 입증한 최초 사례



Use Case 2 Multi Modal AI 응용 (KNN)

- 텍스트를 통한 영상 검색 시스템인 Multi Modal AI 응용에 적용
- Memory-Intensive 워크로드를 CMS에서 처리함으로써 데이터 이동 최소화
- AI형 워크로드에서도 CMS를 통한 성능 & 에너지 효율 개선이 가능함을 보임

벡터 DB

- 벡터와 메타 데이터로 구성
- 유사한 k개의 벡터 검색
- 현재 KNN으로 구현 (DB 내 모든 벡터와 비교)

향후 계획

- 성능 향상을 위해 KNN을 ANN으로 변환 (Graph 기반 Approximation)

