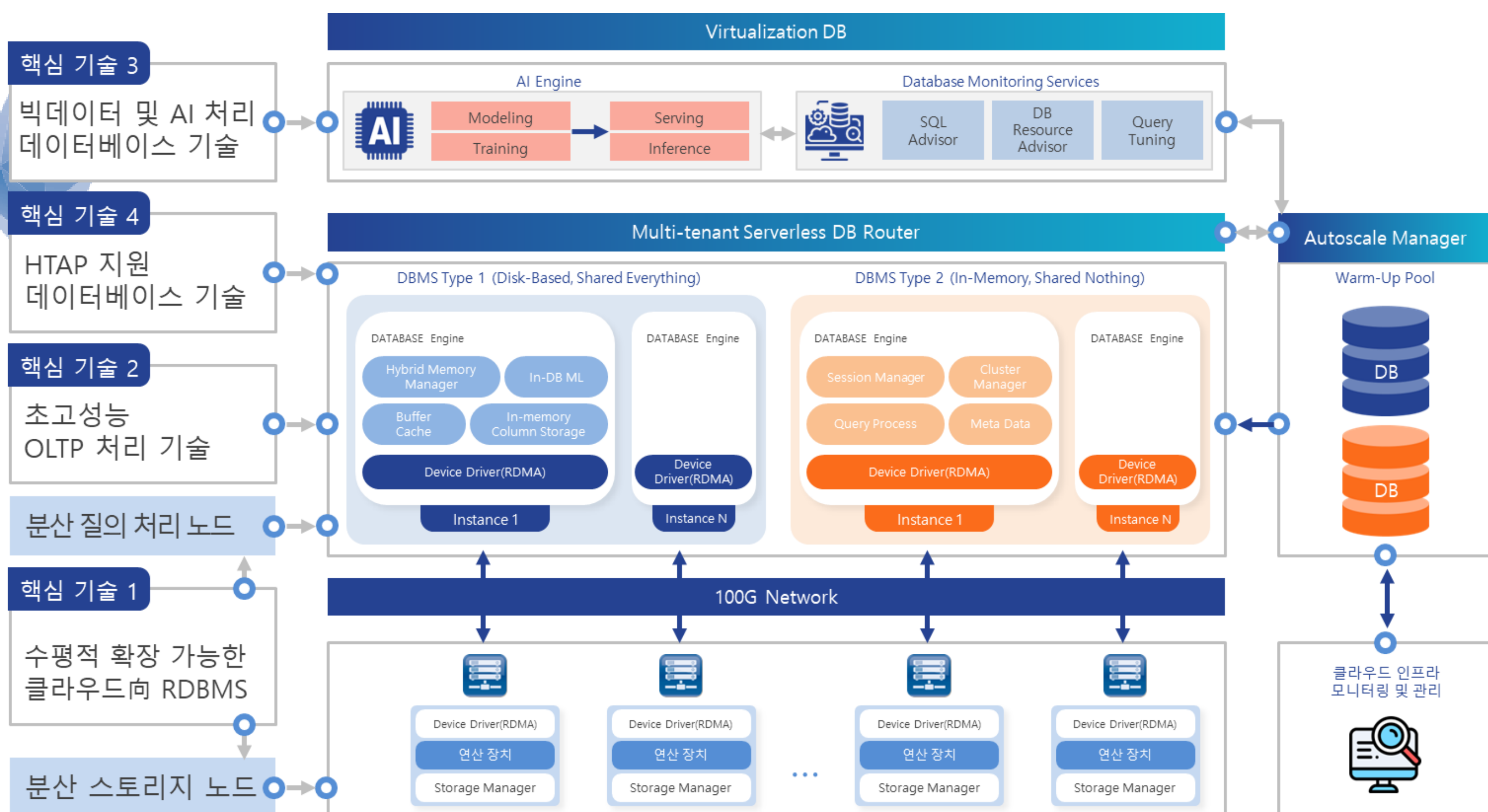


고성능 HW 기반 OLTP, 빅데이터 및 AI 데이터를 효율적으로 저장 관리하는 차세대 Cloud-Native DBMS 기술 개발

총괄/1세부

TmaxTibero

연구개발 목표

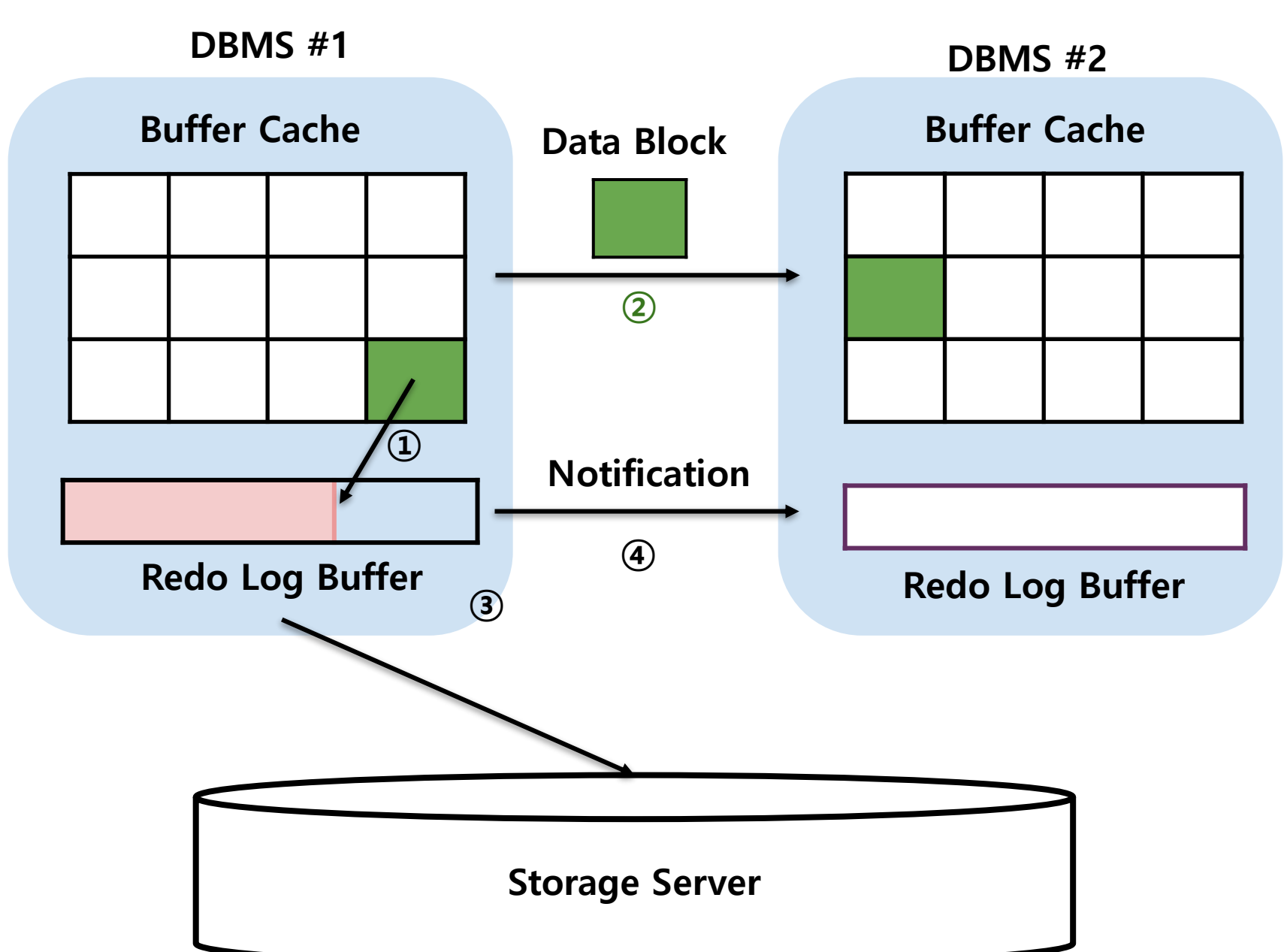


- ▶ 초고성능 OLTP 데이터베이스 최적화 기술 개발
 - ❖ 고성능 하드웨어를 이용한 분산 데이터베이스 (完)
 - ❖ 신뢰성, 가용성, 동시성 제어, 성능 향상, 복제 관리 기술 (完)
 - ❖ 개발한 기술을 활용한 OLTP 데이터베이스 최적화 기술 (完)
- ▶ 차세대 Computing Resource 기반의 빅데이터 및 AI 처리 데이터베이스 기술 개발
 - ❖ 빅데이터 및 AI 처리를 위한 AI 엔진 연동 (未)
 - ❖ In-Database ML 기술과 분석질의 처리 과정에서 필요한 대용량 임시 데이터 저장 및 관리 (未)
 - ❖ 컬럼 기반 저장과 압축 기술 (完)
 - ❖ Auto-scaling 및 저지연 Cold Start 기술 (未)
- ▶ HTAP 지원 데이터베이스 기술 개발
 - ❖ OLTP 및 OLAP을 통합한 HTAP 데이터베이스 최적화 기술 (中)
 - ❖ HTAP 질의처리를 위한 In-memory 컬럼 기반 처리 기술 개발 (完)
- ▶ 타 세부 기술과의 연동을 위한 기술 개발
 - ❖ Virtual Database 기술을 통해 타 DBMS의 Data Object에 대한 CRUD 및 조회 기능 개발 (中)
 - ❖ 연산 스토리지에 Query Predicate를 Offloading 하는 기능 개발 (中)

연구 개발 항목

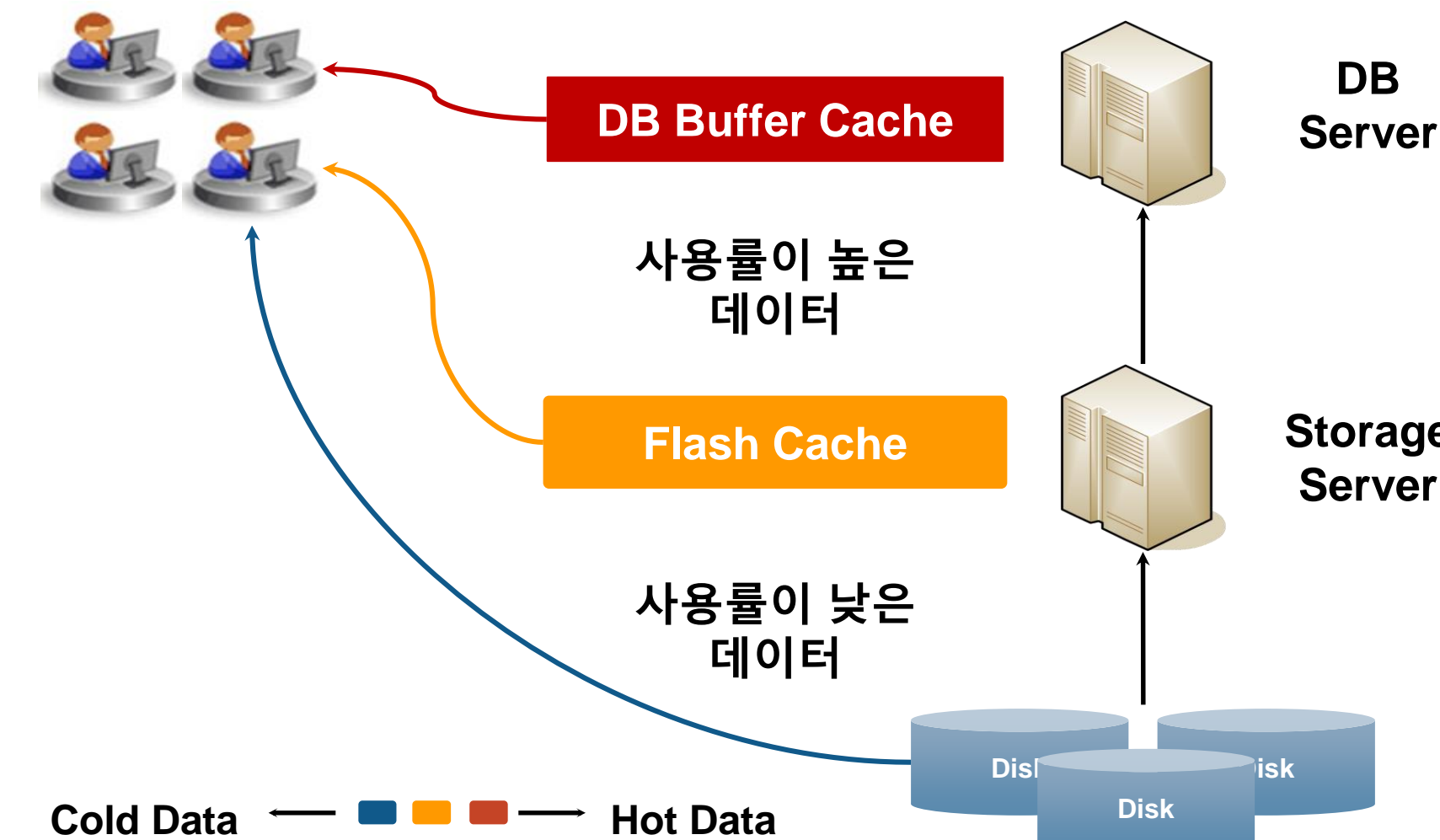
초고성능 OLTP 데이터베이스 최적화 기술

1. Redo Log FLush 없이 Current/Consistent-read(CR) Block 전송하는 기술 개발



- 세부 설명**
- ▶ 동작 방식
 - ① Data Block 수정 및 Redo Log 생성
 - ② Redo Log를 Flush하기 전에 Data Block 전송
 - ③ Storage에 Redo Log Flush 수행
 - ④ Redo Log Flush 완료 정보는 별도로 전송
 - ⑤ Data 수정 및 참조 작업을 마치고 비동기적으로 Redo Log Flush 완료 여부 체크
 - ▶ 성능 효과
 - ❖ Buffer Cache를 빠르게 공유할 수 있음
 - ❖ 다수의 Node에서 동일 Data를 수정 및 참조하는 상황에서 부하 처리 성능 향상 가능

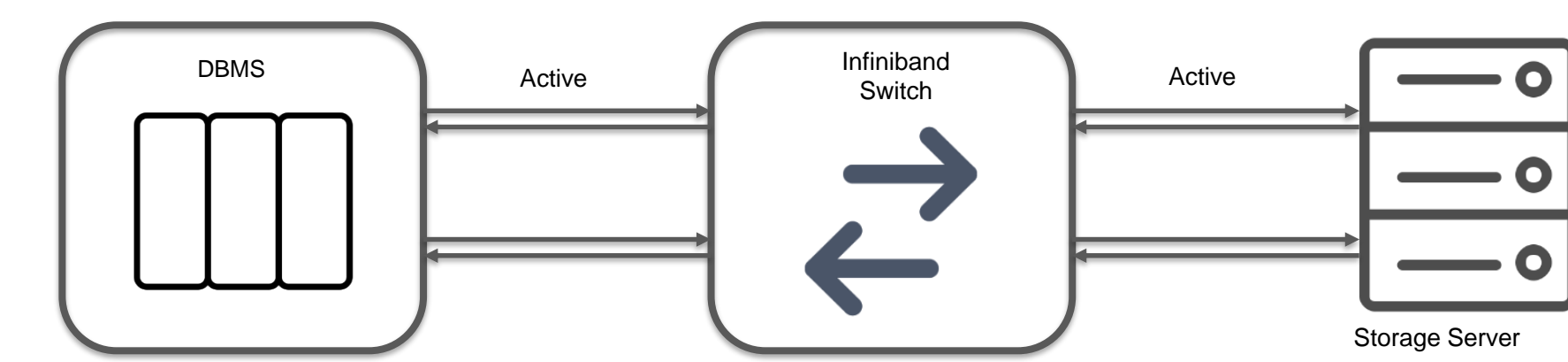
2. 스마트 스토리지 서버의 NVM-SSD 기반 디스크 캐시 기술 개발



- 세부 설명**
- ▶ 스토리지 서버 구성 시 NVM-SSD를 추가해 디스크에 대한 Write-back 캐시로 사용하는 기능
 - ❖ Flash Cache를 이용하여 Random I/O 응답시간 향상
 - ❖ 상대적으로 자주 사용되는 Hot Data를 Flash Caching

3. 고성능 하드웨어를 이용한 고가용성 기술 개발

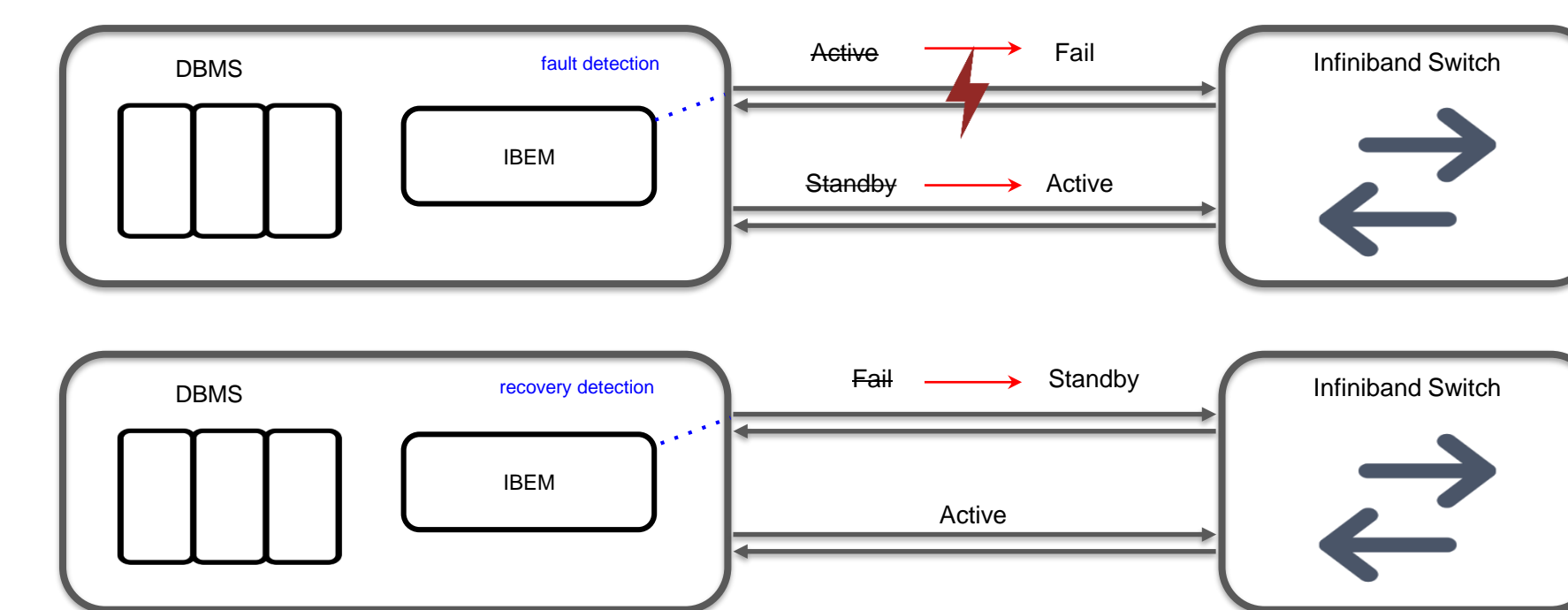
Infiniband Port 이중화 구성



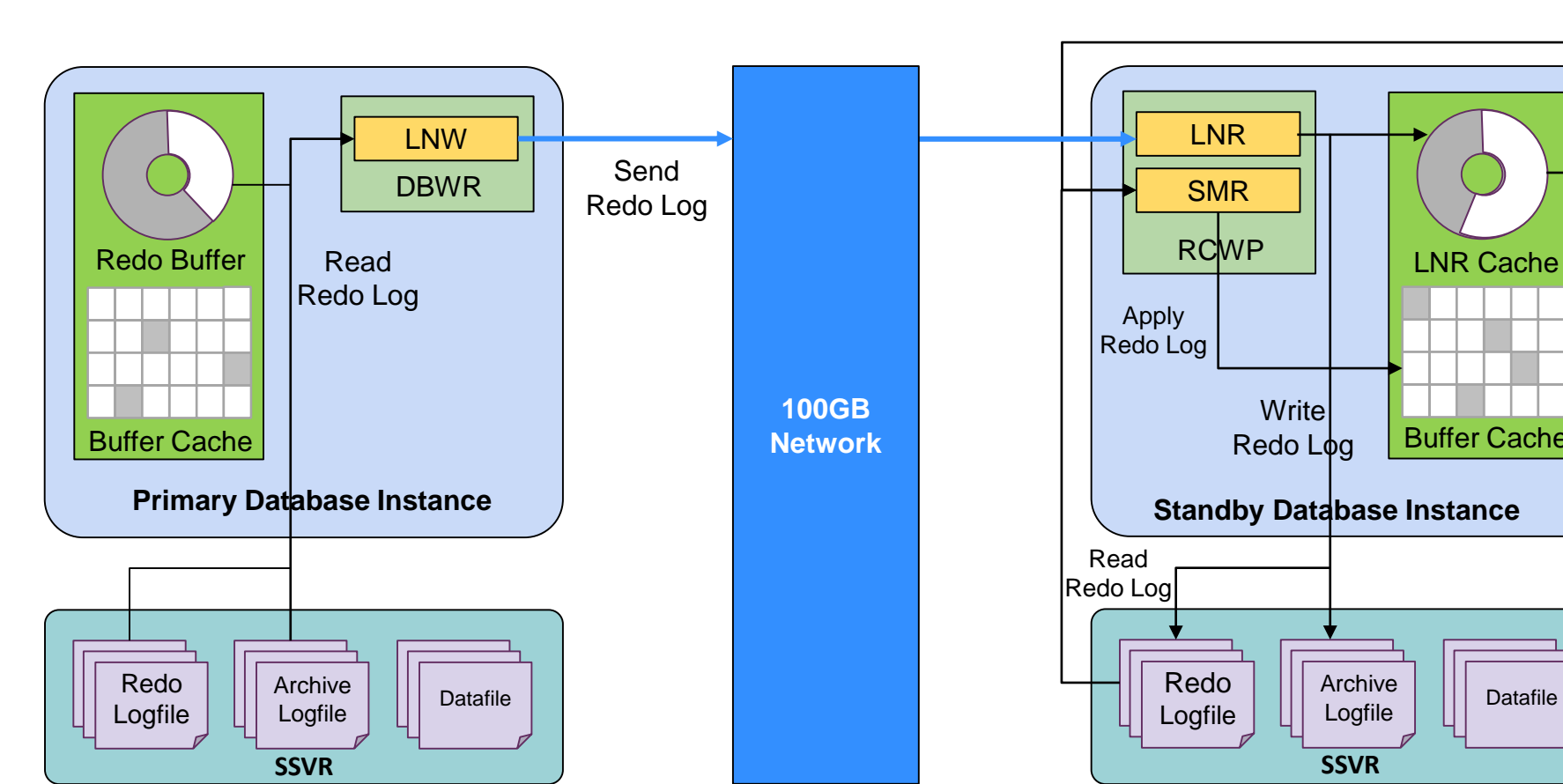
세부 설명

- ▶ Infiniband Port 이중화를 통한 Port Failover 기능
 - ❖ Infiniband Cable 이중화
 - ❖ 최신 Infiniband Switch가 제공하는 Subnet Manager를 활용한 고가용성 기술 지원
- ▶ Infiniband Port Failover 방식
 - ❖ DB와 Storage Server에 Infiniband Event Manager (IBEM)이 Failure 감지를 위해 대기
 - ❖ IBEM이 Port Failure 감지 시 자동으로 Standby Cable을 Active로 격상
 - ❖ Failure 상태의 Port도 IBEM이 복구 시키는 것을 감지하여 Standby로 Rearm

Infiniband Port Failover 시나리오



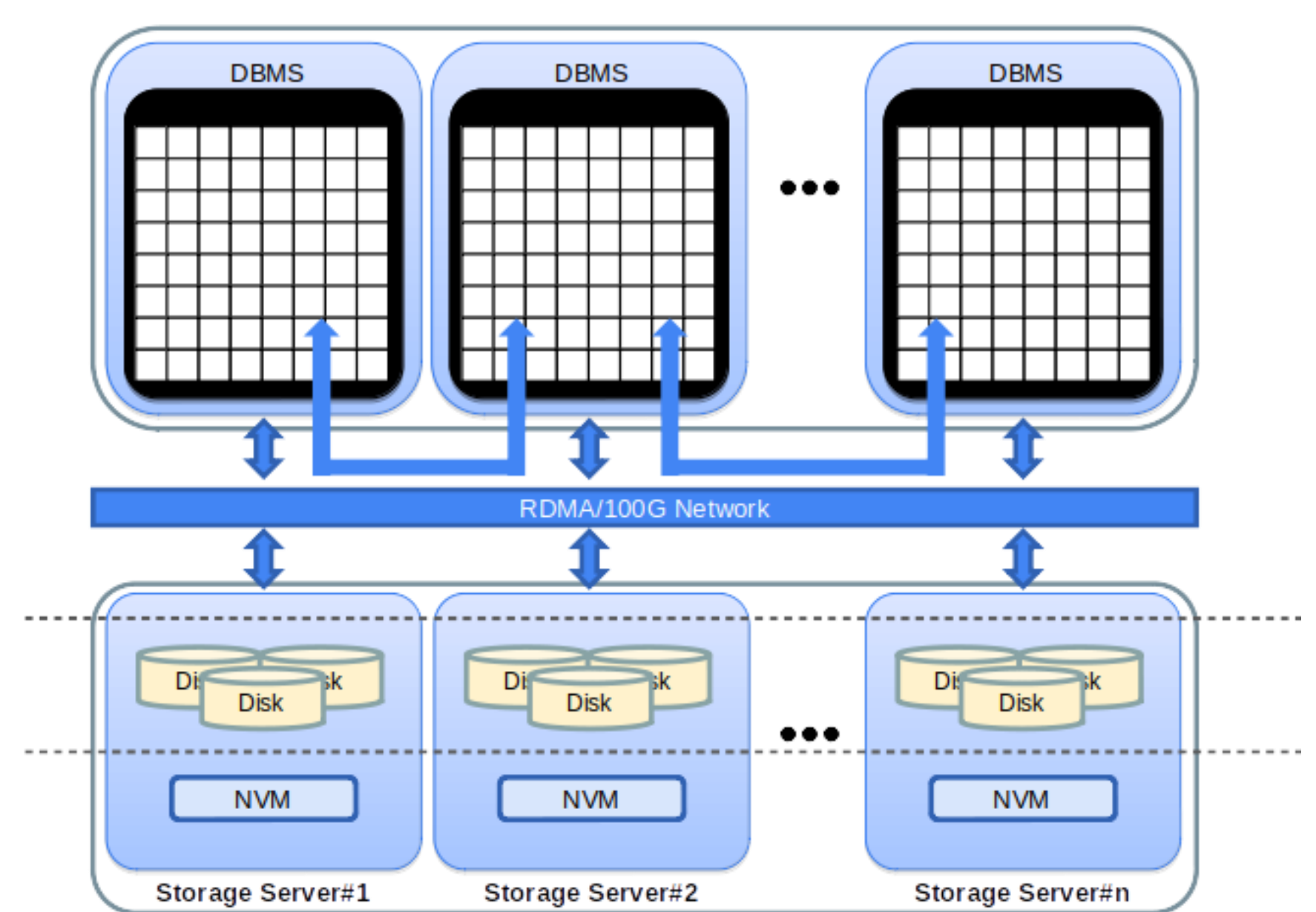
4. 100GB 네트워크를 이용한 저지연 Replica 기술 개발



세부 설명

- ▶ 100GB 네트워크를 이용해 Primary DB에서 Standby로 Redo Log를 고속 전송/복사
- ▶ Standby DB Instance 내 별도 NVM을 Redo Buffer 전용으로 설치 및 활용
 - ❖ Redo Buffer 공간 확장으로 동기화 지연 최소화
 - ❖ 비휘발성 NVM으로 빠른 복구를 통한 동기화 시점 간극 최소화

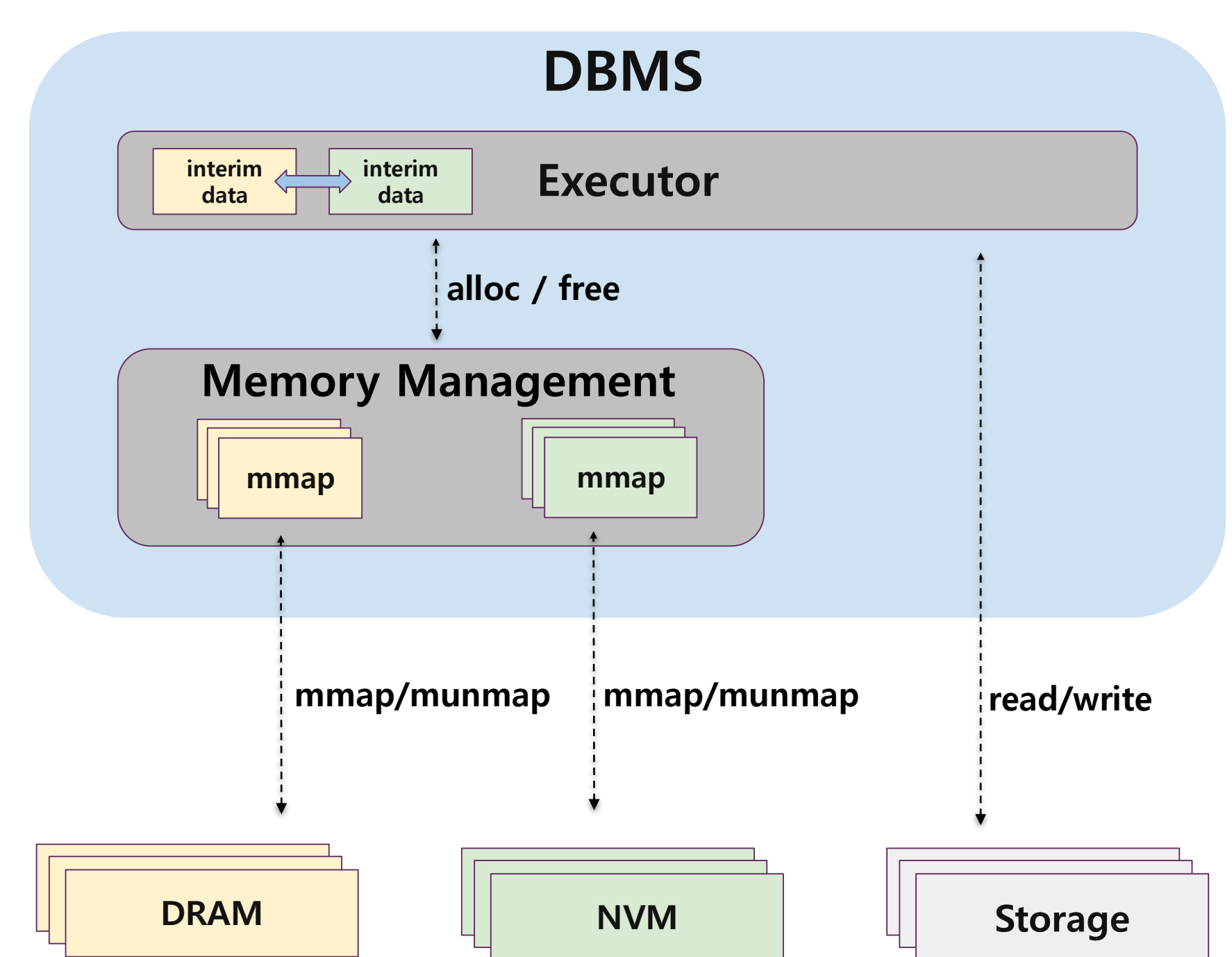
5. 고성능 하드웨어를 이용한 Redo Log Write 성능 향상



세부 설명

- ▶ Internode 통신 과정에서 Infiniband RDMA 활용
 - ❖ Data Block I/O 성능 향상
 - ❖ 안정적인 Buffer Cache Coherence 유지
 - ❖ 연산 스토리지 기능에 RDMA 활용
- ▶ Storage Server에 NVM을 추가하여 성능 향상
 - ❖ NVM에 Redo Log DiskSpace를 따로 구성하여 Redo Flush 성능 향상
 - ❖ Commit 과정 단축 및 전체적인 DBMS 성능 향상

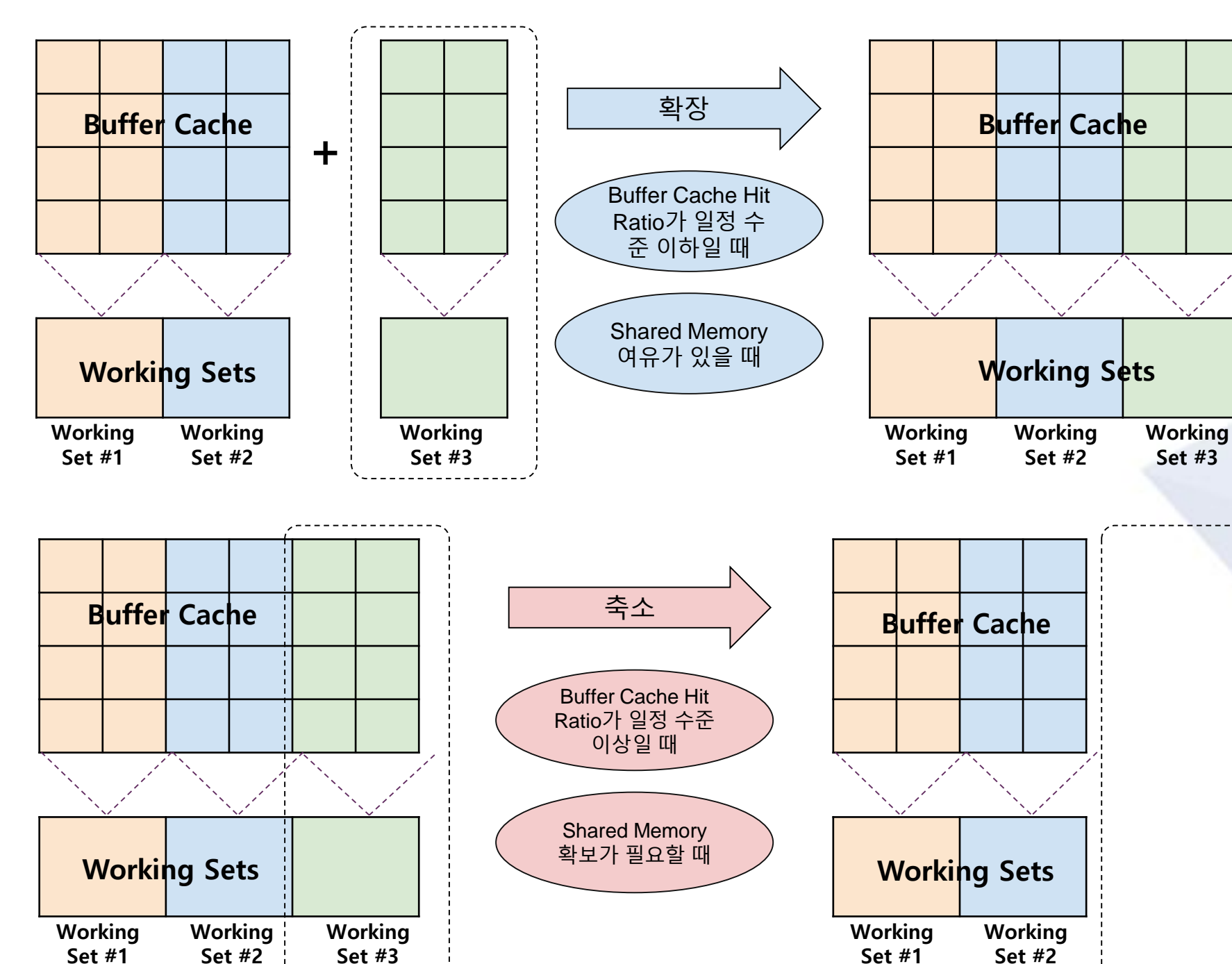
6. DRAM, NVM 하이브리드 메모리를 위한 메모리 관리 기술 개발



세부 설명

- ▶ DRAM/NVM Hybrid Memory Allocator 개발
 - ❖ 용도에 맞게 DRAM과 NVM 영역을 취사 선택하여 메모리 사용 가능
 - ❖ 인스턴스, 프로세스, 세션 별 메모리 사용량 실시간 모니터링을 통해 사용량 제어
- ▶ DRAM이 모자라는 상황에서 기존에 Disk Spill 대신 NVM 영역을 사용하여 비효율 개선
 - ❖ Disk Spill을 위한 Byte Serialization, Deserialization을 NVM 사용 시 불필요
 - ❖ NVM 영역에 있는 데이터와 DRAM 영역에 있는 데이터를 별도 처리 없이 바로 연산 및 비교 가능

7. 실시간 Buffer Cache 사이즈 변화 기능



세부 설명

- ▶ 실시간 Buffer Cache 확장
 - ❖ Buffer Cache Hit Ratio가 일정 수준 이하로 떨어지는 경우 Shared Memory 사용량을 고려하여 Buffer Cache 확장
 - ❖ Buffer Cache에 더 많은 Data Block을 적재하여 Cache Hit Ratio를 높여 성능 향상
- ▶ 실시간 Buffer Cache 축소
 - ❖ Buffer Cache Hit Ratio가 일정 수준 이상이면 Shared Memory 사용량을 고려하여 Buffer Cache를 축소하고 Shared Memory 반납
 - ❖ 반환된 Shared Memory를 다른 용도로 전환하여 성능 향상 (PE, Background Job, etc.)

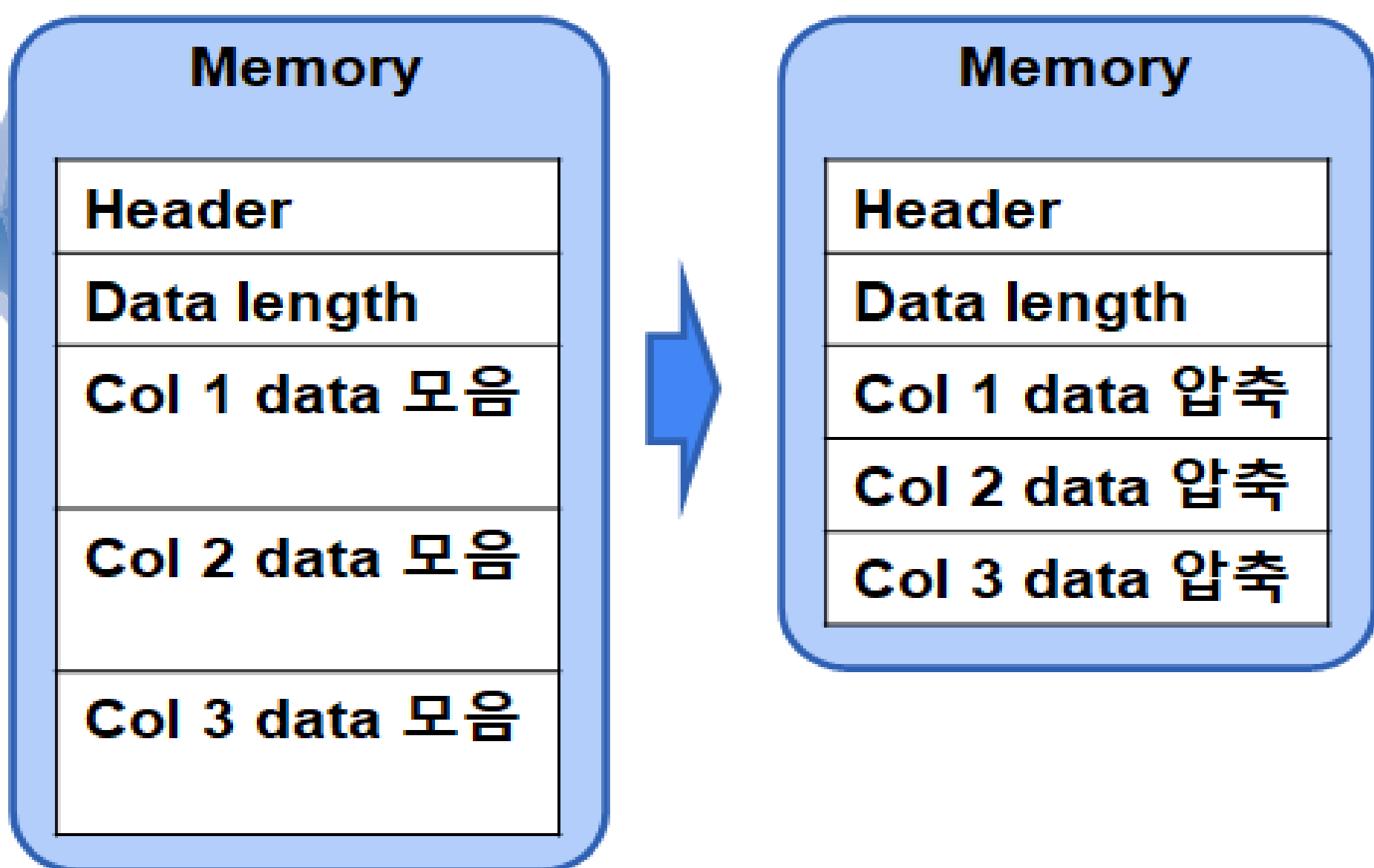
고성능 HW 기반 OLTP, 빅데이터 및 AI 데이터를 효율적으로 저장 관리하는 차세대 Cloud-Native DBMS 기술 개발

총괄/1세부



차세대 컴퓨팅자원 기반의 빅데이터 및 AI 처리 데이터베이스 기술 개발

1. 고속 분석 질의 처리를 위한 컬럼 기반 저장 및 압축 기술 설계

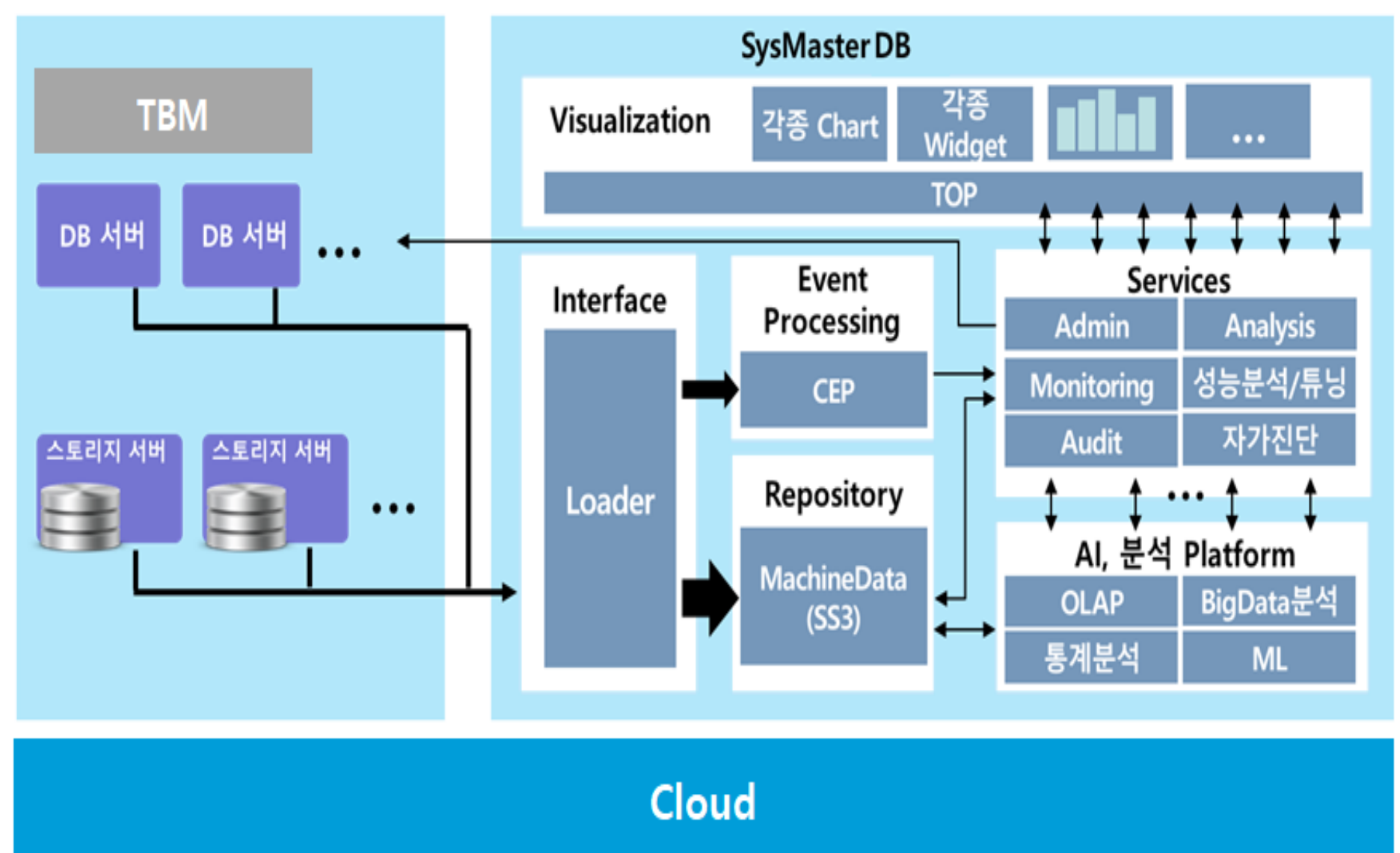


세부 설명

- ▶ 데이터를 Column 기반으로 저장하고 압축함으로써 분석 질의 처리 성능 향상
 - ❖ 압축 대상 Column에 대한 질의 처리 속도 향상
 - ❖ 압축효율을 높여 질의 분석을 위한 I/O 양 축소
- ▶ 기술 상세 개발 내용
 - ❖ 주어진 Row들을 메모리 상에서 Column 별로 모은 후 각 Column 데이터 모음을 압축한다.
 - ❖ 압축된 데이터를 해당 데이터 양에 맞는 만큼의 Data Block에 쪼개어 적재한다.

HTAP 지원 데이터베이스 기술 개발

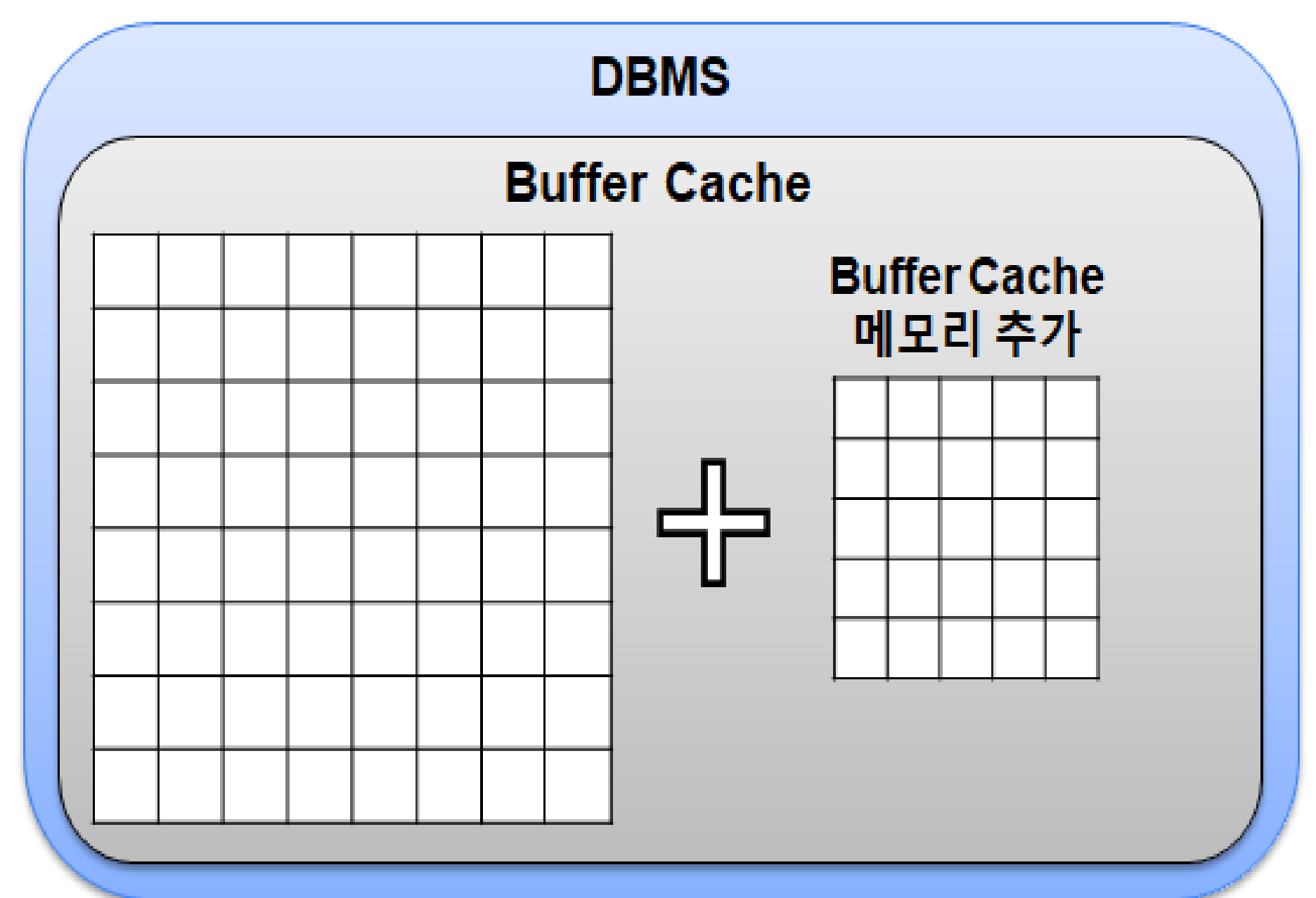
1. ML 기술적용을 통한 데이터베이스 최적화 기술 설계



세부 설명

- ▶ 모니터링 레이어
 - ❖ 부하 지표를 수집하여 ML 엔진에 전달
 - ❖ 일정 주기마다 학습한 모델을 활용하여 새로운 데이터에 대한 부하 여부를 추론
- ▶ ML 엔진 레이어
 - ❖ AutoML 분류 모델 사용
 - ❖ 부하지표에 영향을 줄 수 있는 피처를 정의
 - ❖ 정의된 피처와 타겟 컬럼으로 구성된 학습 데이터셋을 수집

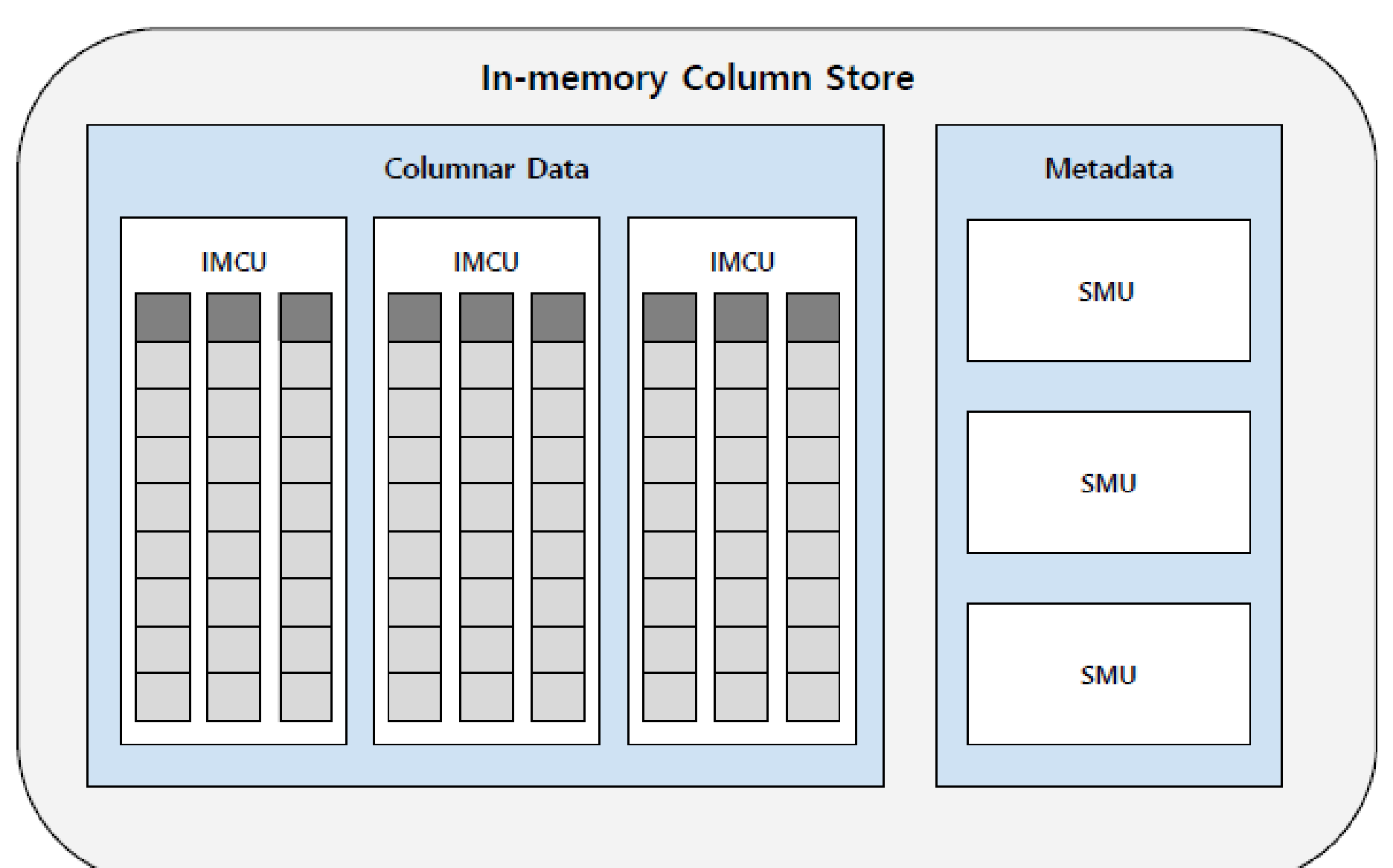
2. 실시간 최적 Buffer Cache 사이즈를 구하고 설정을 변경하는 AI 기능 개발



세부 설명

- ▶ 모니터링 레이어
 - ❖ 부하 지표를 수집하여 ML 엔진에 전달
 - ❖ 학습한 모델을 활용하여 현재 Buffer Cache Size와 TPS가 정상범위인지 판단
 - ❖ 비정상인 경우 ML 엔진으로부터 알맞은 Size를 받아와 Tibero에 적용
- ▶ ML 엔진 레이어
 - ❖ AutoML 회귀 모델 사용
 - ❖ 학습한 모델을 활용하여 새로운 데이터에 대한 적정 Buffer Cache Size를 추론

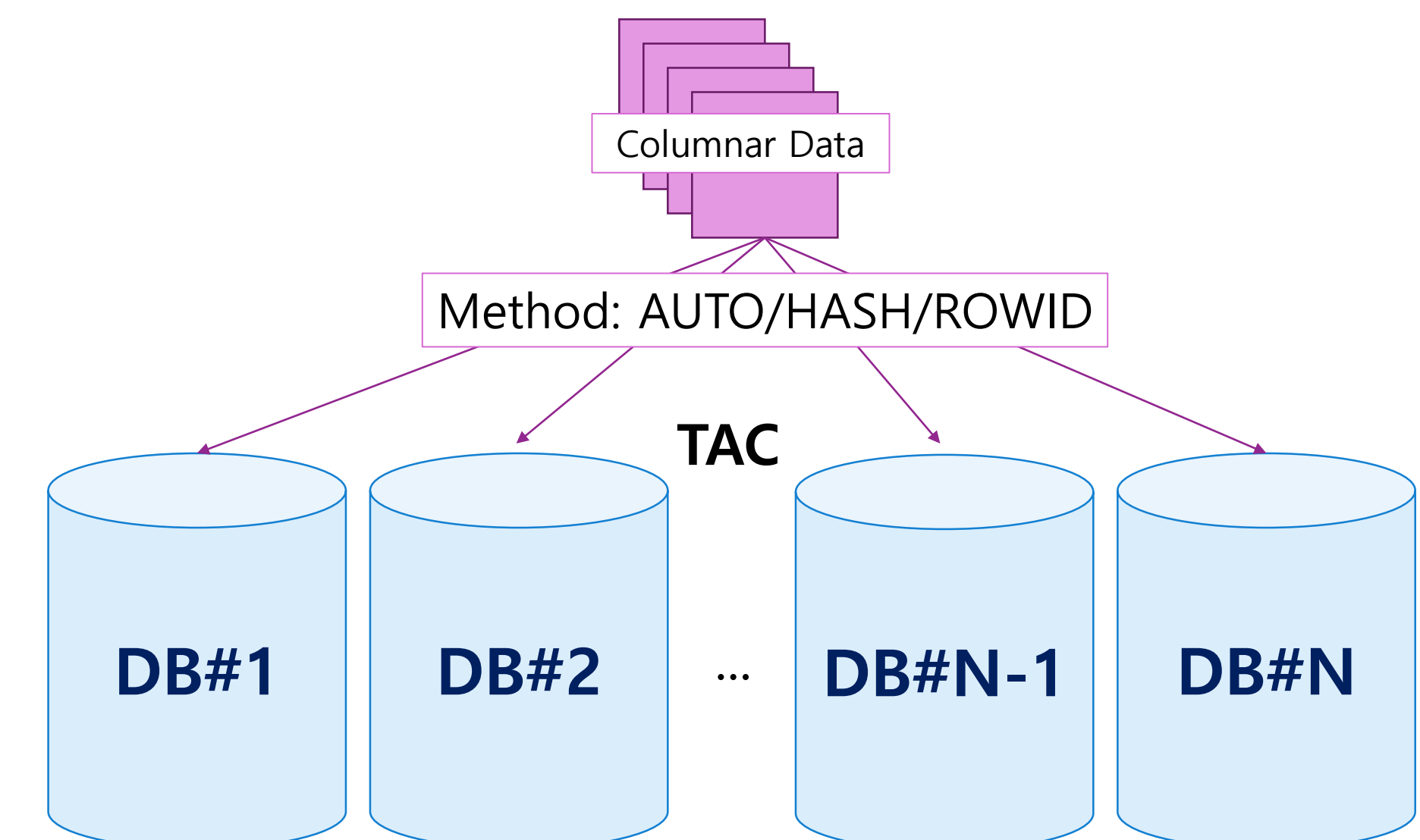
3. HTAP 질의처리를 위한 In-memory 컬럼 기반 처리 기술 개발



세부 설명

- ▶ 지정된 Column을 In-memory로 저장하는 기능
 - ❖ OLAP 쿼리에 최적화
 - ❖ 조회만 가능한 데이터이므로 Concurrency 보장
 - ❖ SIMD를 이용한 Fast Predicate Filtering
- ▶ 두 형태의 데이터 간 정합성 유지를 위한 In-memory 트랜잭션 기능 개발
 - ❖ In-memory로 적재 이후 들어오는 트랜잭션에 대해 저널링을 하여 정합성 유지
 - ❖ IMCS에서 수정된 데이터를 조회할 때는 원본 데이터를 조회하도록 처리

4. 클러스터 간 In-memory 컬럼 데이터 분산 저장 기능 개발

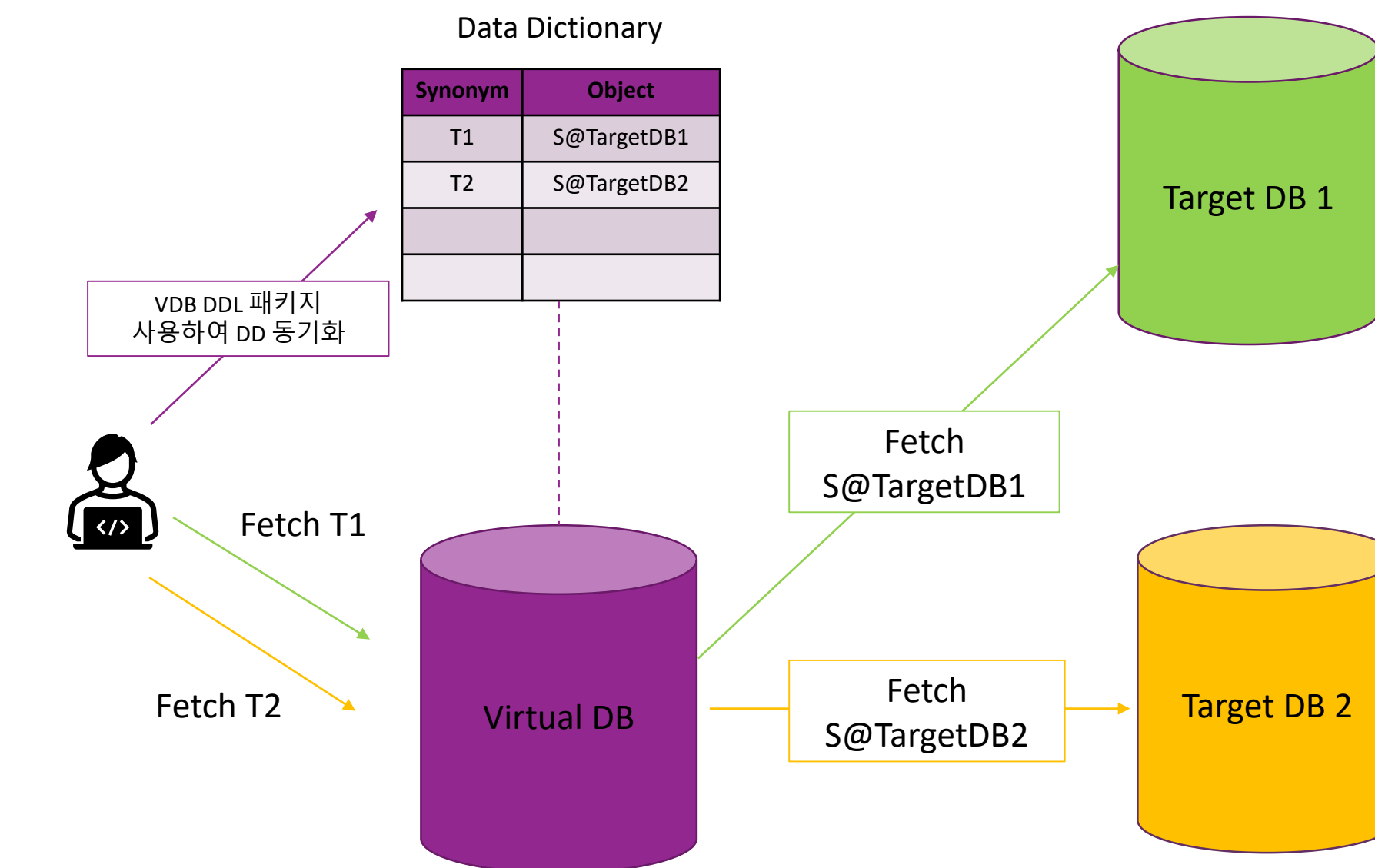


세부 설명

- ▶ 클러스터 간 Columnar Data를 Duplicate 또는 Distribute 형태로 저장 가능하도록 기능 개발
 - ❖ Duplicate: 모든 클러스터 노드가 같은 Columnar Data를 복제하여 가지고 있는 형태
 - ❖ Distribute: Auto/Hash/Rowid 등 규칙을 통해 Columnar Data를 분산 저장하는 형태
- ▶ Columnar Data를 분산 저장 시 각 클러스터 노드에서 연산 처리를 하도록 Query Engine 개발
 - ❖ 간단한 Filter에 대한 Function Offloading 추가

타 세부 기술과의 연동을 위한 기술 개발

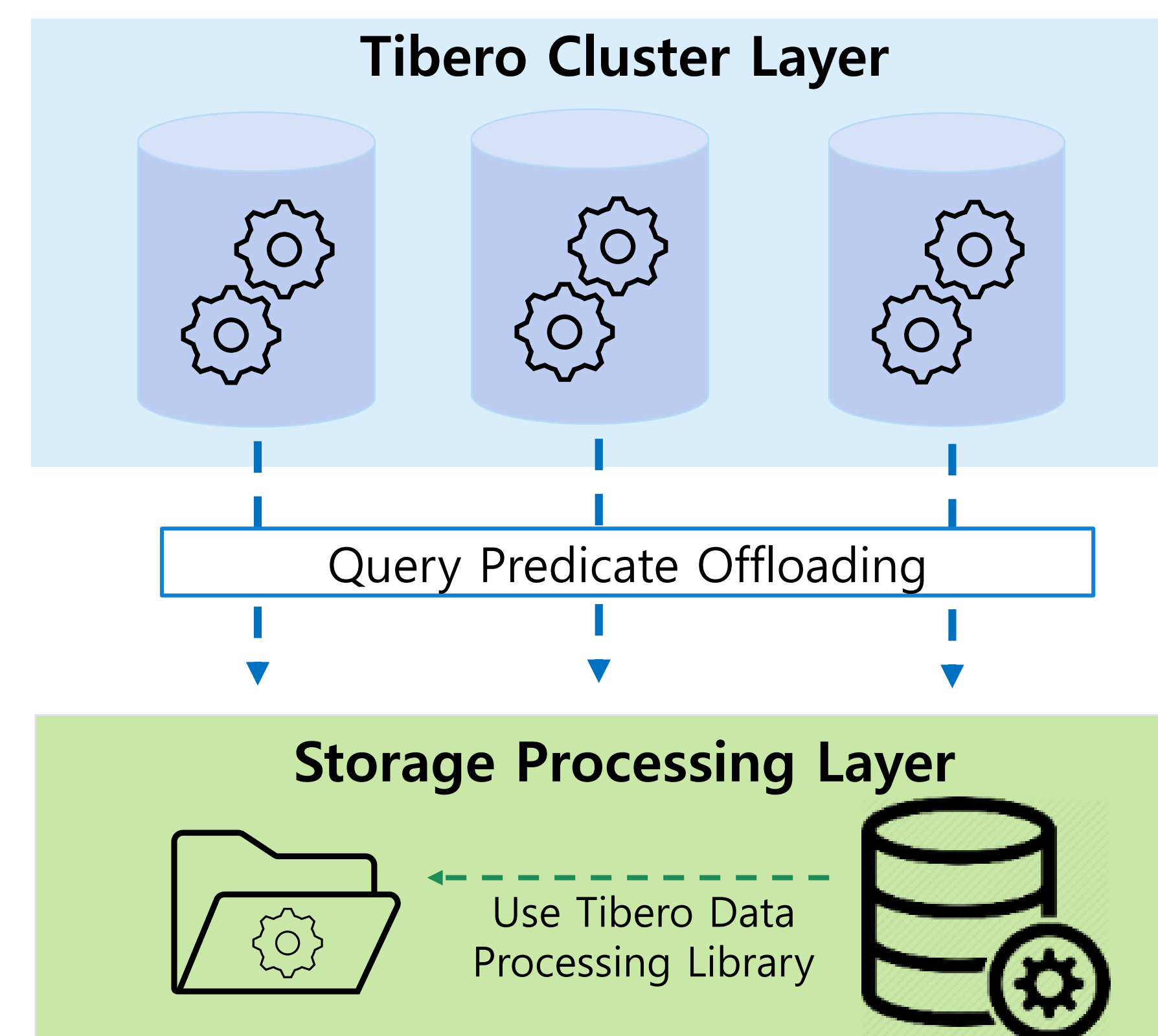
1. Virtual Database 기술을 통해 타 DBMS에 대한 CRUD 및 조회 기능 개발



세부 설명

- ▶ Virtual Database에서 제공하는 VDB Package를 사용하여 Target DB의 Database Object에 대한 CRUD 수행
 - ❖ CRUD 수행 시 Data Dictionary에 Synonym으로서 해당 Table의 접속 정보가 저장
 - ❖ Java External Procedure를 통해 Target Database Object의 메타 정보 획득
- ▶ 조회 시 접속정보가 저장된 Synonym을 사용하여 간편하게 조회 가능

2. 연산 스토리지에 Query Predicate를 Offloading 하는 기능 개발



세부 설명

- ▶ 4세부 연산 스토리지에 Tibero Cluster를 연동하여 Tibero의 데이터를 스토리지에 저장하고 Query Processing 도중 Query Predicate를 연산 스토리지에 Offloading 하여 네트워크 병목을 최소화 하는 기술
- ▶ 연산 스토리지 개발팀에 Tibero Data Processing Library를 제공하여 개발 유연성을 보장하고자 함
 - ❖ Data Block Read/Write Interface
 - ❖ Block Meta Data Interface
 - ❖ Tibero Cluster Networking Protocol Interface
 - ❖ Expression Evaluation Interface
- ▶ 4세부와 지속적인 커뮤니케이션을 통해 개발 및 테스트 진행하여 3분기에 연동 테스트 계획

2차년도 연구 개발 현황

클라우드향 HTAP 개발 현황

주제	상태	진행률
스토리지 서버의 NVM-SSD 기반 하이브리드 디스크 캐시 구현	Complete	100%
고속 분석 질의 처리를 위한 컬럼 기반 저장 및 압축 기술 설계	Complete	100%
In-memory Column Store 기술 설계	Complete	100%
클라우드 환경에서 In-memory Column Store 분산 기술 설계	Complete	100%
Redo Log Flush 없이 Buffer Cache 공유하는 기술 구현	In Progress	55%
실시간 Buffer Cache Size 변화 기능	In Progress	56%
DRAM/NVM Hybrid Memory Allocator 개발	In Progress	85%
DRAM/NVM Hybrid Memory를 이용한 Query Processing 설계 및 구현	In Progress	20%
ML 기술적용을 통한 데이터베이스 최적화 기술 설계	In Progress	20%
실시간 최적 Buffer Cache 사이즈를 구하고 설정을 변경하는 AI 기술 구현	In Progress	20%
Total		66%

Virtual Database 개발 현황

주제	상태	진행률
VDB DDL 동기화 패키지 설계	Complete	100%
Tibero to SunDB DBLink 연동 테스트	Complete	100%
Tibero to SunDB SQL Syntax 테스트	Complete	100%
DDL 동기화 패키지 개발	Complete	100%
DDL 동기화 패키지 테스트	In Progress	10%
TPC-C, TPC-H를 통한 Virtual Database 검증	Scheduled	0%
Total		68%

연산 스토리지 Function Offloading 개발 현황

주제	상태	진행률
Tibero Internal 소개자료 제작 및 공유	Complete	100%
Tibero Data Processing Library API 목록 제작 및 공유	Complete	100%
Tibero Data Processing Library 개발	Scheduled	0%
4세부 연산 스토리지 연동 테스트	Scheduled	0%