Goldilocks™

“1,000Cores이상 Scale Out 가능한 클러스터 데이터베이스”

Pure In-Memory DBMS for High Performance

2018

SUNJESOFT
(주)선재소프트

Document Ver. : 3.1.2

Goldilocks™

1. 선재소프트 소개
2. Goldilocks 소개
3. 구축사례
별첨 : Goldilocks DB 구성의 특장점
Q&A

© 2017 SUNJESOFT CO., LTD. All rights reserved.

선재소프트 소개

Pure In-Memory DBMS for High Performance

대한민국
SW발전을 위해
혁신을 이루었습니다.

Goldilocks DBMS TPC위원회의 국제공인 인증을 획득하고
공식 홈페이지에 게시되었습니다.

2017년 대한민국
SW 품질 대상 수상

암호화폐거래소
"빗썸, 빗썸Pro"메
인 매매체결
DBMS로 선정

유진초저온 차세
대시스템 (ERP,
물류시스템 등)
기반 DBMS로
선정

차세대 DBMS
Goldilocks 발표

2010 | 2013 | 2016 | 2017 | 2017.05 | 2017.06 | 2017.11 | 2017.12

삼성증권 차세대 주문시스템에 원장 DBMS로 채택
- HTS / 법인주문 / 딜러주문 / LP 등 모든 주문에 사용
- 서울과 수원간 원거리 Active-Active 구현
- 60개 증권사 중 주문속도가 가장 빠름
- 기존 DB : IBM DB2원장 DB를 이관

중국 차이나유니콤 CBSS 집중화 시스템 전체 DBMS 납품

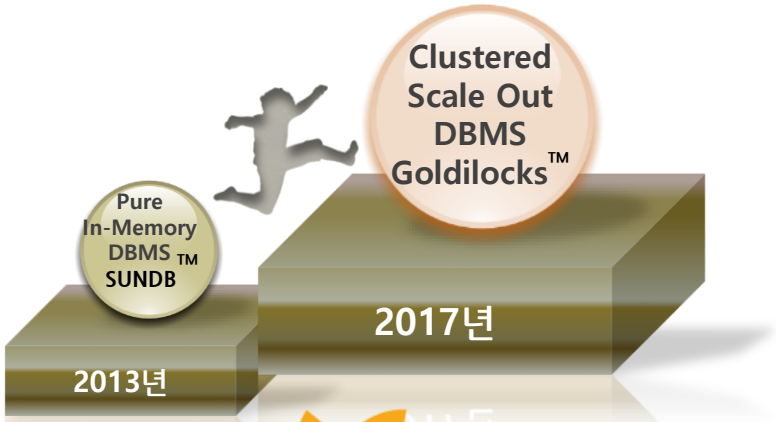
한국거래소 차세대 매매체결 시스템 개발 참여(계약규모 : 83억)
- DBMS , 메세징 M/W, Framework 등

선재소프트 창업
세계 최고의 혁신적인 DBMS 제품개발을 목표(Not Fast Follow , First Mover)

© 2017 SUNJESOFT CO., LTD. All rights reserved.

선재소프트 연혁

선재소프트는 In-Memory Computing 전문회사로서 2010년 설립이래 In-memory기반의 DBM, DBMS, Messaging Middleware분야의 기술을 선도하고 있으며 미래지향적 데이터 처리 플랫폼을 위하여 연구를 거듭하고 있는 기술력 중심의 회사입니다.



SUNJESOFT
(주)선재소프트

선재소프트는
In-memory computing 기술력을 바탕으로
" Future Focus In-Memory Scale Out DBMS Platform"
핵심 기술력 회사로 성장 할 것입니다.

2017	• 유진초저온 통합물류 정보시스템(물류/ERP) Goldilocks 수주 • 한국전력 차세대 배전지능화 기반 MMDB 수주 • 금융결제원 빅데이터사업 Goldilocks 납품 • 세계 1등 암호화폐거래소 빗썸거래소 매매체결시스템용 Goldilocks 납품 • 초록우산어린이재단 통합정보시스템 및 데이터분석시스템 구축 수주 • 중국 허난모바일 CRM 시스템 Goldilocks 납품 • 한국전력 빅데이터 통합 플랫폼 사업 SUNDB 납품 (내부 수집서버) • 전력연구원 In-Memory 기반 전력데이터 고속 분석 인프라 Goldilocks 납품(4TB) • 경찰청 안전Dream 시스템 Goldilocks 납품
2016	• NH투자증권 기관/상품 주문 시스템 납품 • 한국마사회 One card total 서비스 지원 시스템 구축 • 경찰청 정보전산시스템 성능개선사업 납품(BI 분석용) • 건강보험심사평가원 의약품 유통정보 분석모니터링시스템 납품 • 한국철도기술연구원 실시간 철도안전관제 납품 • SK증권 / 하나금융투자 FIX-OMS시스템 납품 • SK증권 FDS 시스템 구축사업 DBMS 납품
2015	• 미래창조과학부 "SW컴퓨팅산업원천기술개발사업 GCS 과제 "선정 • 한전KDN 스마트 배전 시스템 SUNDB 납품 • 중국 차이나유니콤 cBSS 집중화 시스템 SUNDB 납품 • 중국 차이나모바일 (하남성) CRM권한시스템 SUNDB 납품 • 스포츠토토 판매점관리 시스템 구축 • 삼성증권 차세대 플랫폼 및 주문 시스템 사업 수주 • 하이투자증권 투자정보 시스템 구축 사업 수주 • 유안타증권 FDS 시스템 구축 사업 수주
2014	• 한국거래소 차세대 시스템 (EXTURE+) 성공적 오픈 및 운용계약 • KTNET 전자수입인지 시스템 구축완료
2013	• 코스콤 통합시세정보 1단계 구축(KBBO-Korea Best Bid and Offer) • 코스콤 PowerBase+프로젝트에 GOLDILOCKS납품(25copy) • 한국거래소 신파생 시장감시시스템 SUNDB 납품(DB사이즈 2.8TB)
2012	• 한국거래소 EXTURE+ 본개발(업무개발부문) 사업 수주 • 한화증권 고속주문시스템 SUNDB납품 • 코스콤 금융데이타센터 구축사업 SUNDB납품
2011	• 한국거래소 EXTURE+ 선도개발 사업 수주
2010	• 한국거래소 "매칭엔진 파일럿시스템 개발 용역" 사업 수주 • 선재소프트 설립

© 2017 SUNJESOFT CO., LTD. All rights reserved.

2. Goldilocks Cluster™ 소개

© 2017 SUNJESOFT CO., LTD. All rights reserved.

What is the Goldilocks Cluster?

Goldilocks Cluster is a class of modern RDBMS that seek to provide NewSQL

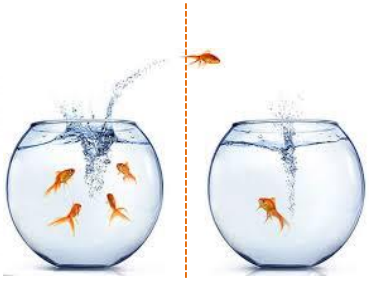
☑ Define (Michael StoneBraker)

	Old SQL	NoSQL	NewSQL
Relational	Yes	No	Yes
SQL	Yes	No	Yes
ACID transactions	Yes	No	Yes
Horizontal scalability	No	Yes	Yes
Performance / big volume	No	Yes	Yes
Schema-less	No	Yes	No

☑ This Category

VOLTDB

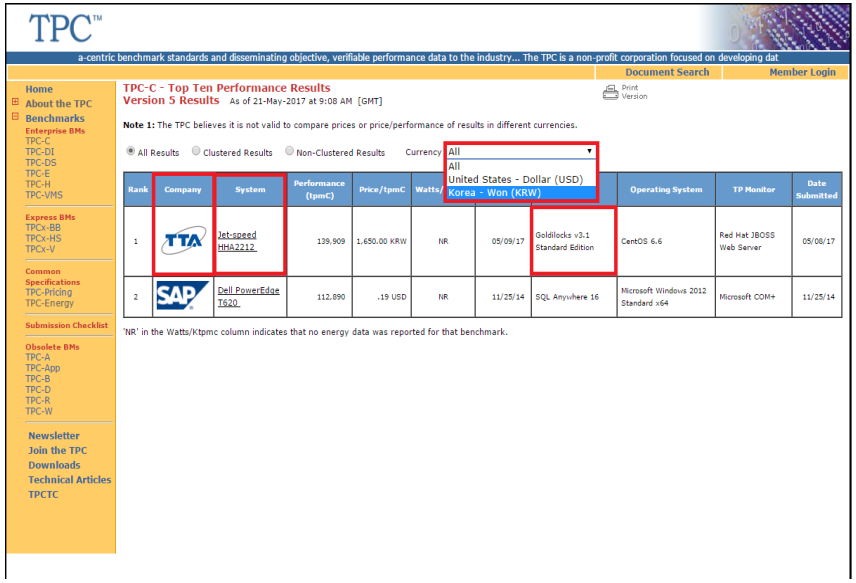
memsql



Clustrix

SAP HANA

☑ TPC-C 국제공인인증 획득(2017/05)
www.tpc.org



© 2017 SUNJESOFT CO., LTD. All rights reserved.

Goldilocks Cluster 제품 개발 배경

IT서비스 환경의 변화와 사용자 요구에 부응하는 고성능 Scale Out 데이터베이스 플랫폼을 개발함.

- 실시간 고속트랜잭션 처리 및 분석처리에 최적화
- 데이터베이스 시스템 운영환경의 혁신 (No Back-Up, Active-to-Active DR)

IT 패러다임의 변화

비즈니스 인프라 : 클라우드, 빅데이터, IoT, 모바일 서비스에 적합한 인프라 요구

데이터 제공 : 공급자 중심 ➔ 사용자 참여 중심 저비용, 고성능, 무중단 서비스 요구

차세대 DB 아키텍처 고민

고성능. 고효율 Scale Out 아키텍처를 지원하는가?

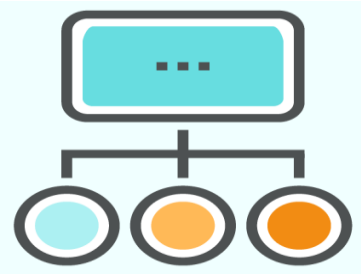
폭증하는 데이터에 대한 처리?

24*365서비스 위한 가용성 확보 방안?

표준 SQL지원으로 개발의 편리성. 생산성을 보장?

안정적인 운영 환경을 지원하는가?

차세대 Scale Out Cluster DBMS



- 데이터와 트랜잭션의 폭증에 순응할 수 있는 Scale Out 아키텍처 채택
- 비즈니스 연속성을 보장하는 고가용 기술 및 무 정지 서비스 환경 제공
- SQL 지원 및 완벽한 ACID 보장 등 기존 DBMS 기능 지원

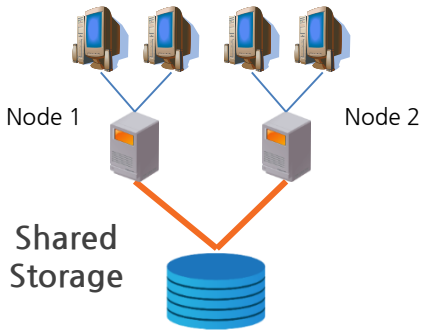
© 2017 SUNJESOFT CO., LTD. All rights reserved.

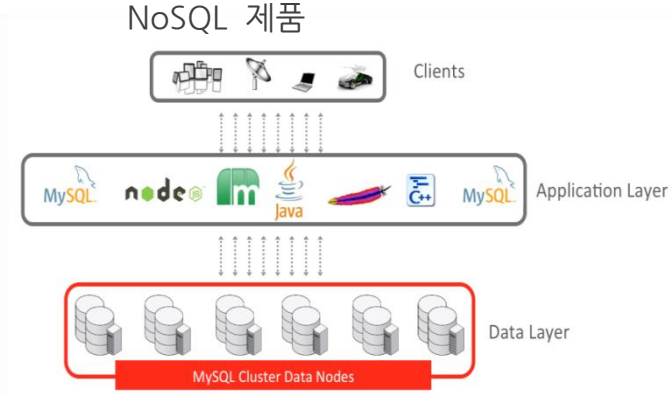
기존 기술과의 차별성/혁신성

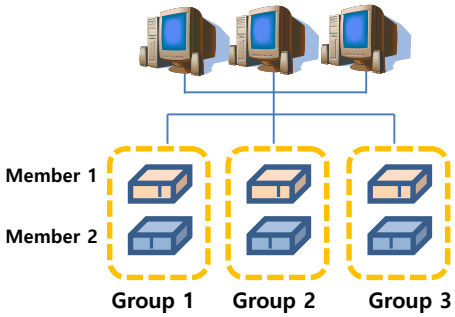
Legacy SQL 및 Transaction을 지원하면서 동시에 유연한 Scale Out Cluster 구조로 설계되었습니다.

- 노드 추가에 따른 선형적인 성능 배가와 확장
- 무중단 데이터 서비스 제공

	Shared Disk Cluster	Shared Nothing Cluster	Shared Nothing / Scale Out Cluster
Key Idea	스토리지 공유 방식	수평적 확장 개념	수평적 확장 개념 완벽한 SQL, Transaction 지원
Strength	정합성 유연성 (SQL 지원)	제한 없는 확장성	선형확장, SQL지원, 정합성 ACID, 고속 트랜잭션
Weak Point	고비용 구조 백업, DR 부대비용 제한적인 확장성	낮은 정합성 (No ACID, Poor Transaction) SQL 호환성 낮음 (Key/Value)	구현 기술 난이도 높음
Examples	Oracle RAC, Tibero TAC	VoltDB, Redis MySQL Shard, MongoDB 등 NoSQL 제품	**Goldilocks Cluster 플랫폼**







© 2017 SUNJESOFT CO., LTD. All rights reserved.

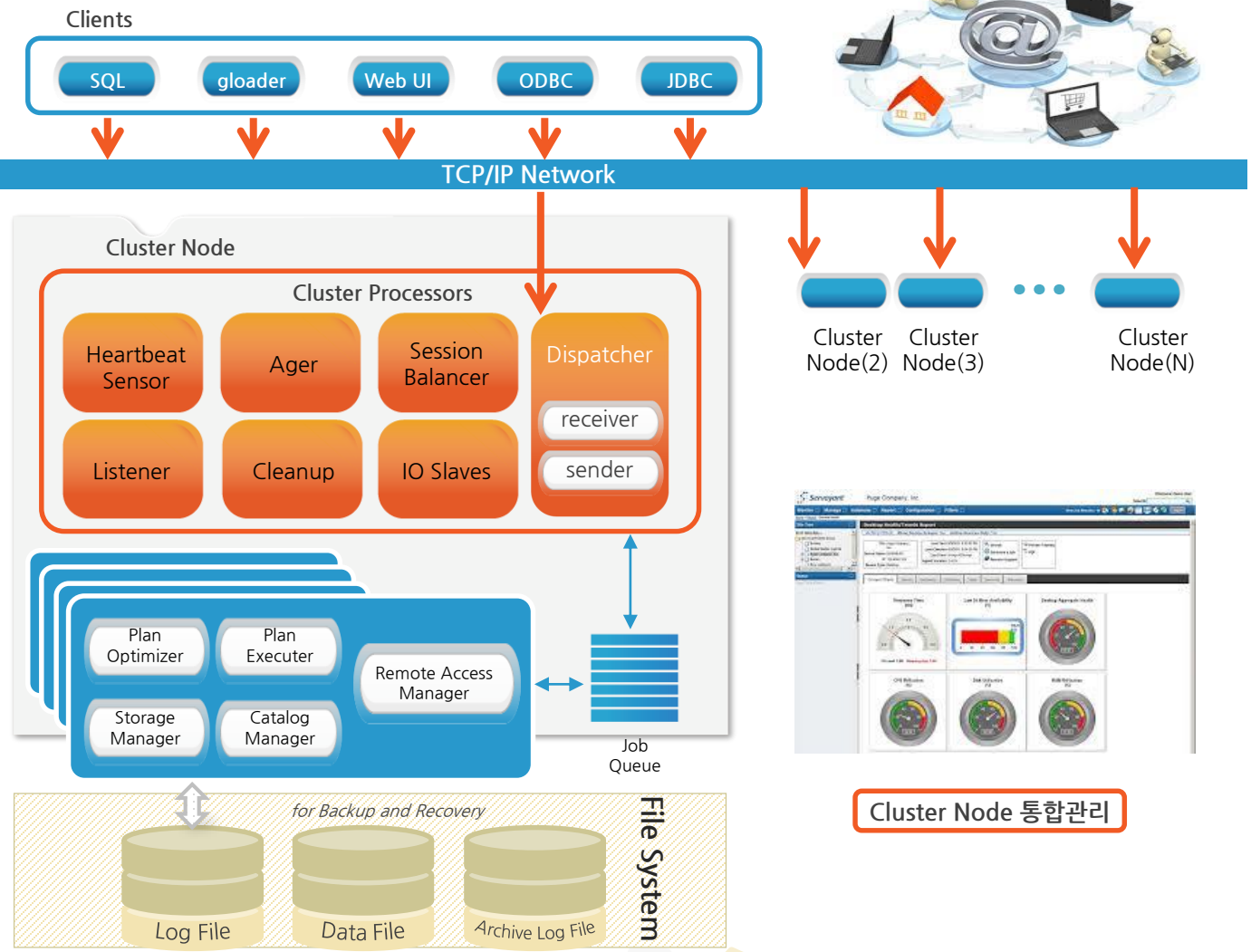
Goldilocks Cluster 아키텍처

새로운 IT환경에 부응하기 위해 고성능 인메모리 기반의 Scale Out 방식 클러스터 데이터베이스를 최적화된 Engineered 아키텍처로 개발하였습니다.

In Memory Centric RDBMS

Commodity X86 및 Any H/W

융합 기술을 이용한 HW와 SW 최적화



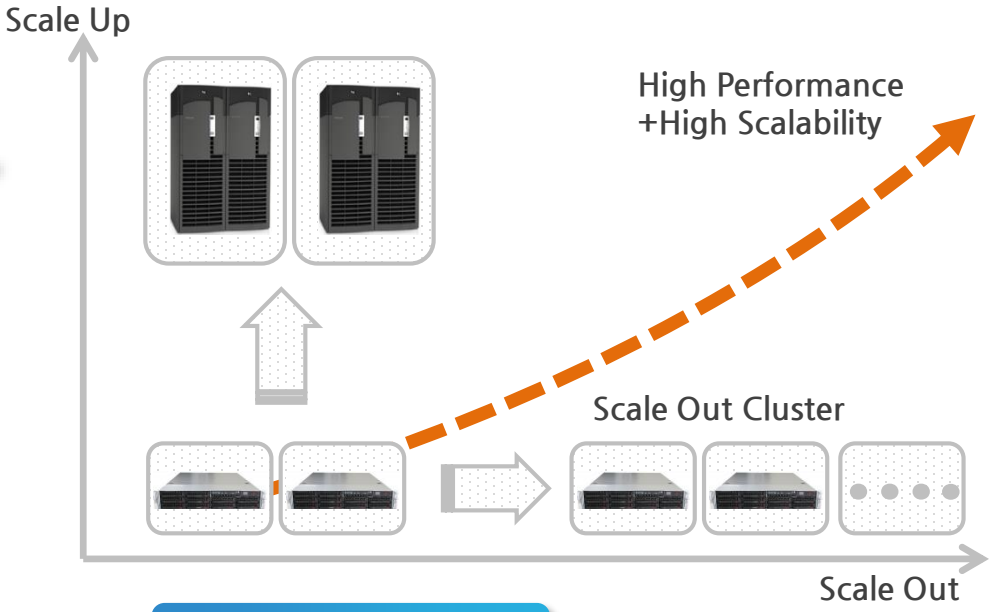
Cluster Node 통합관리

© 2017 SUNJESOFT CO., LTD. All rights reserved.

Goldilocks Cluster Feature(1/3)

Shared Nothing Cluster-분산 데이터베이스

- Shared Nothing Cluster Database
  - No Leading Vendor
  - Oracle : Shared Disk Cluster (Scale Out 제한적)
- 현재 시장 기술 수준
  - Poor Performance
  - Lots of Limitations



Core Technology

- High Throughput
  - Sharding을 통한 데이터 분산
  - 서버 추가에 따른 total throughput 향상
- High Availability
  - 데이터 다중화를 통한 고가용성 확보
  - No Single-Point Failure
- Online Scale Out
  - Server추가 및 제거 시 서비스 중단 없는 Scale Out

주요 관련 기능

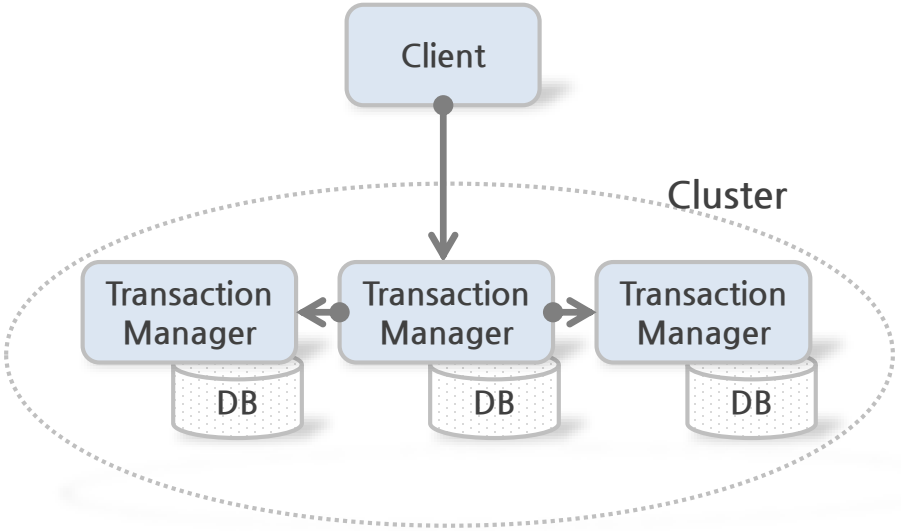
- High Throughput
  - Sharding strategy
  - Inter-Node Parallel DML
- High Availability
  - Active-Active Replication
  - Auto Data Re-balancing
- Online Scale Out
  - Online Add/Remove Nodes
- Node Failure
  - No Single-Point Failure
  - Connection fail-over
  - DML fail-over

© 2017 SUNJESOFT CO., LTD. All rights reserved.

Goldilocks Cluster Feature(2/3)

Global Transaction 준수

- Global Transactional in Cluster Database
  - 2Phase Commits Processing
  - Multi-Node Transaction Support
- Scale Out DBMS의 Core Technology
  - ACID Compliant With MVCC
- 현재 시장 수준
  - Not /Poor Support Global Transaction



Core Technology

- Stand-Alone 과 동일한 ACID
  - Atomicity
  - Consistency
  - Isolation
  - Durability
- 분산 환경의 특성
  - Network deadlock
  - Node Failure

주요 관련 기능

- ACID in Shared Nothing Cluster
  - 2-phase commit protocol
  - Global statement level consistency
  - Global View SCN
- Networked DML
  - Inter-Node Parallel DML

© 2017 SUNJESOFT CO., LTD. All rights reserved.

Goldilocks Cluster Feature(3/3)

Legacy SQL in Scale Out Cluster

- Less Legacy SQL function in Shared Nothing Cluster Database
  - Application Modify needs
- 현재 시장 수준
  - Complex SQL not Support

Core Technology

- Legacy DBMS의 SQL Power
  - Transparent SQL in shared nothing cluster
  - Better Performance than Stand-Alone DBMS
  - Obstacle in shared-nothing cluster prototypes
- Query
  - No Limitation of Join, Grouping, Aggregation etc
- DML
  - Transparent DML w/ sharding strategy
  - Less Performance Down with many replica
- SQL-Object
  - User, Schema, Tablespace
  - Table, Index, View, Sequence, Constraint

주요기능

- SQL Shipping
  - Generated SQL
  - Query decomposition
  - Node selection
  - Inter-Node parallelism
- Join Power
  - shard condition X shard condition
  - non-shard condition X shard condition
  - non-shard condition X non-shard condition
- DML
  - Inter-Node Parallel DML
  - Non-deterministic DML
- Dictionary Synchronization

© 2017 SUNJESOFT CO., LTD. All rights reserved.

Goldilocks Cluster 핵심기술 - Data Sharding Strategy

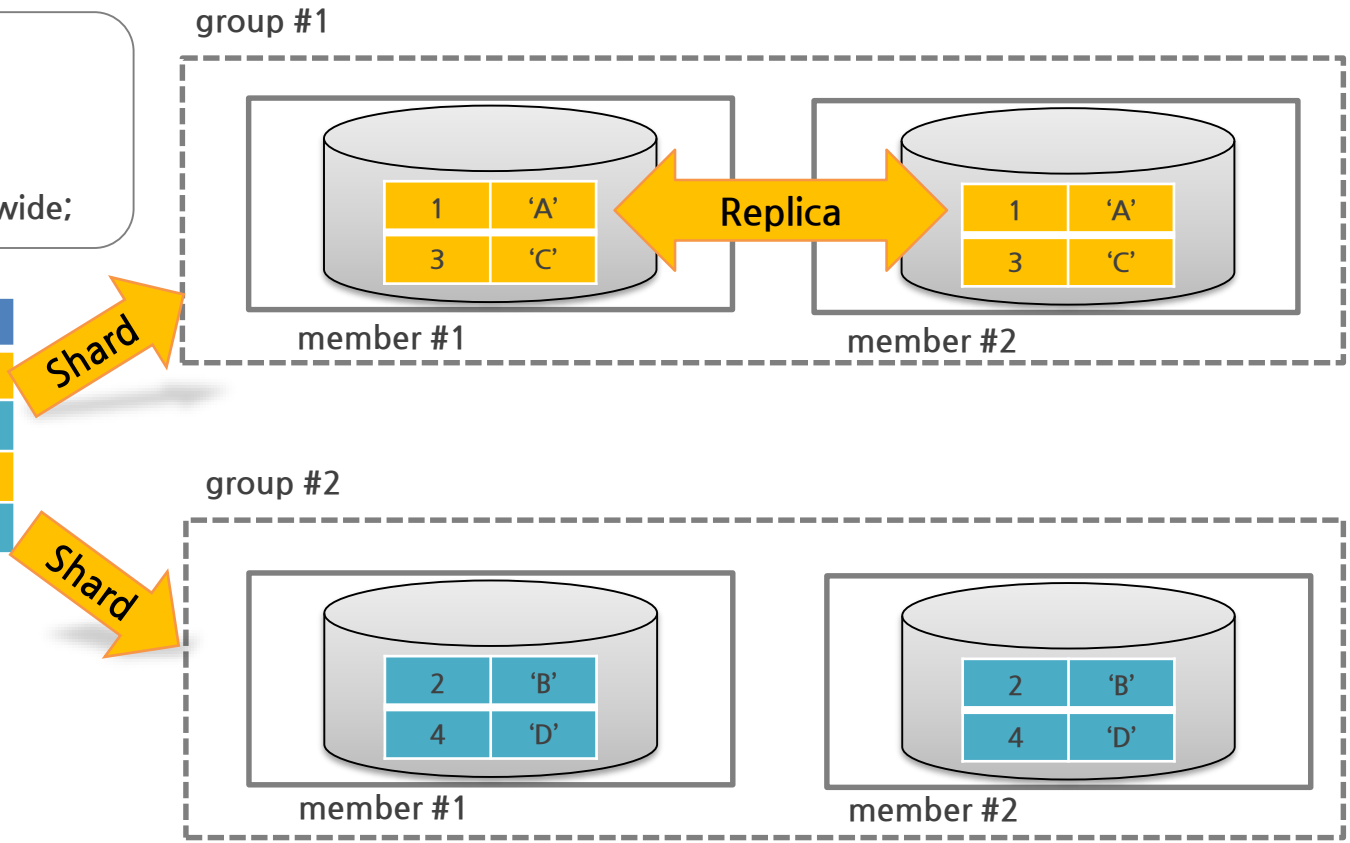
대용량 데이터처리와 부하분산, 병렬처리를 위하여 다양한 Sharding Strategy를 구현 할 수 있습니다

- 수평분할 방식으로 데이터를 분산 저장
- Sharding DB간에 데이터의 Join, Global Transaction처리 가능
- 데이터의 재분배 기능으로 서비스 중지 없이 Scale Up과 Scale Out가능

```
create table t1
(
   id integer,
   name char(10)
) sharding by  hash at cluster wide;
```

id	name
1	'A'
2	'B'
3	'C'
4	'D'

sharded by hash

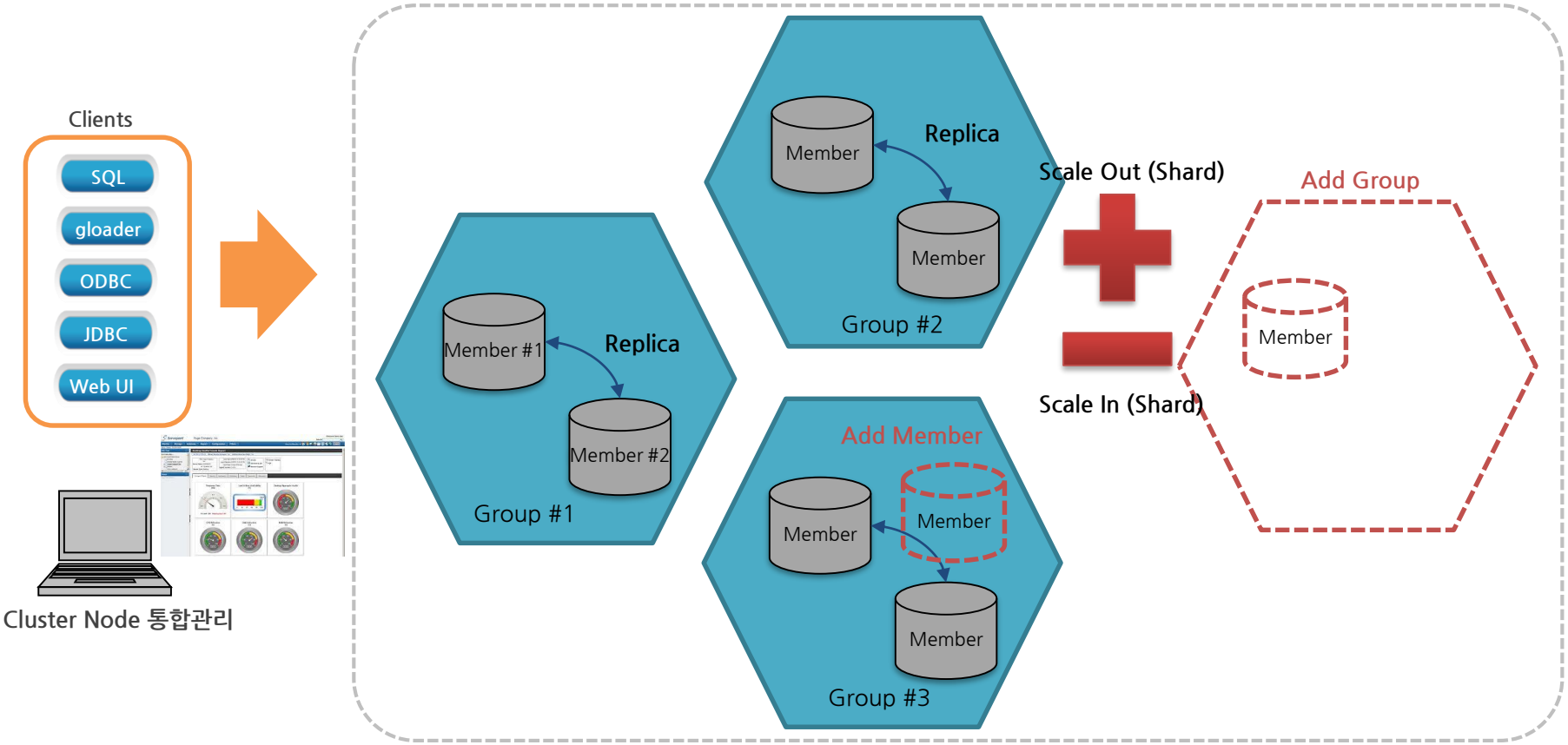


© 2017 SUNJESOFT CO., LTD. All rights reserved.

Goldilocks Cluster 핵심기술 - Online Scale Out(IN)

비즈니스 업무서비스의 중단 없이 DB서버에 용량과 성능을 확장 및 축소 할 수 있습니다.

- ADD Group, Member Node, Delete Group, Member
- Auto/Manual Data Rebalancing
- 적정 시스템 용량산정으로 비용절감 효과 극대화
- 무중단 시스템 구현(Active-to-Active)

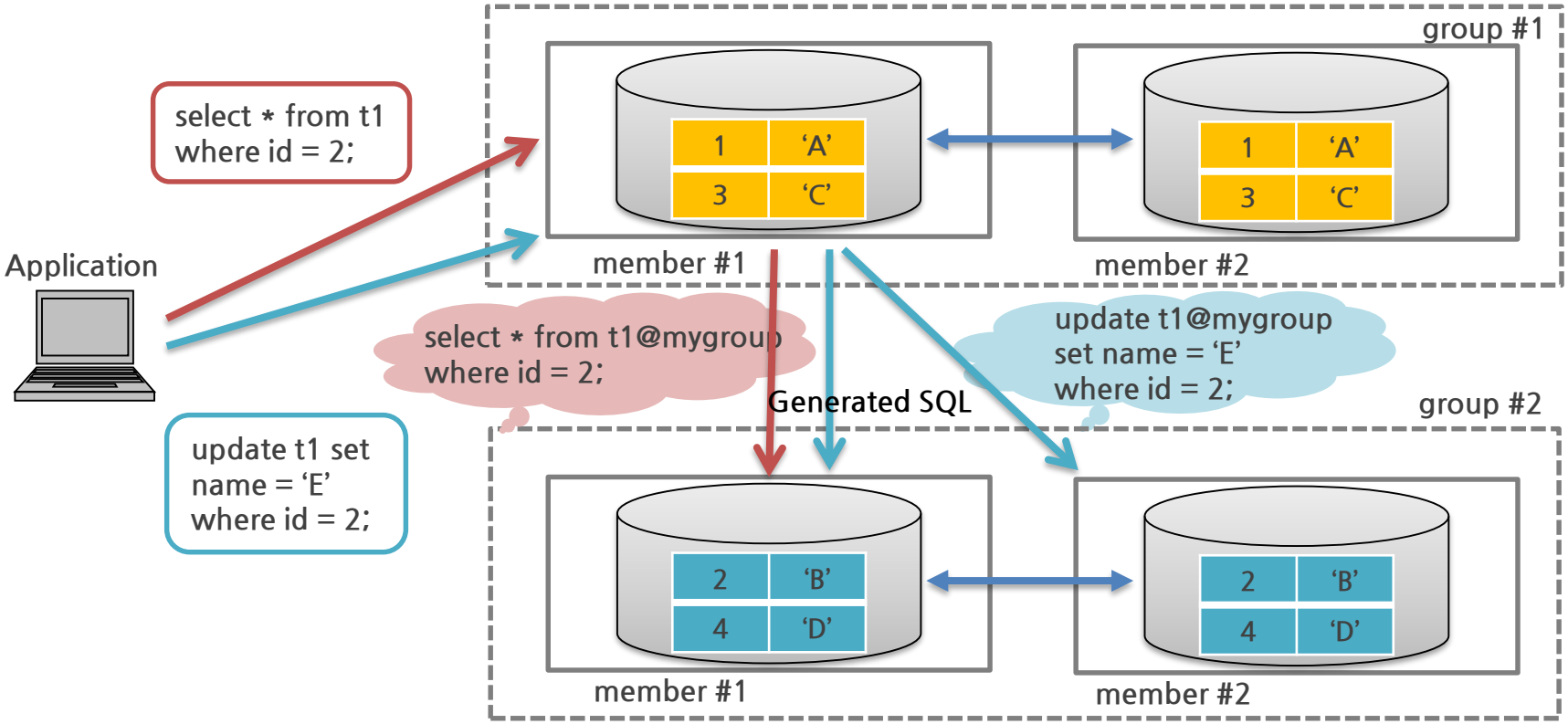


© 2017 SUNJESOFT CO., LTD. All rights reserved.

Goldilocks Cluster 핵심기술 - SQL Shipping

분산 환경에서 데이터 변경 및 데이터 조회 시 노드 간 데이터 이동을 최소화 하여 네트워크 비용을 절감하기 위해 SQL Shipping 기술을 사용합니다.

- 멀티 노드 간에 I/O Overhead 최적화
- SQL Shipping을 통하여 병렬처리 가능 및 고성능 시스템 구현
- 멀티 노드 간에 Data Join처리 가능

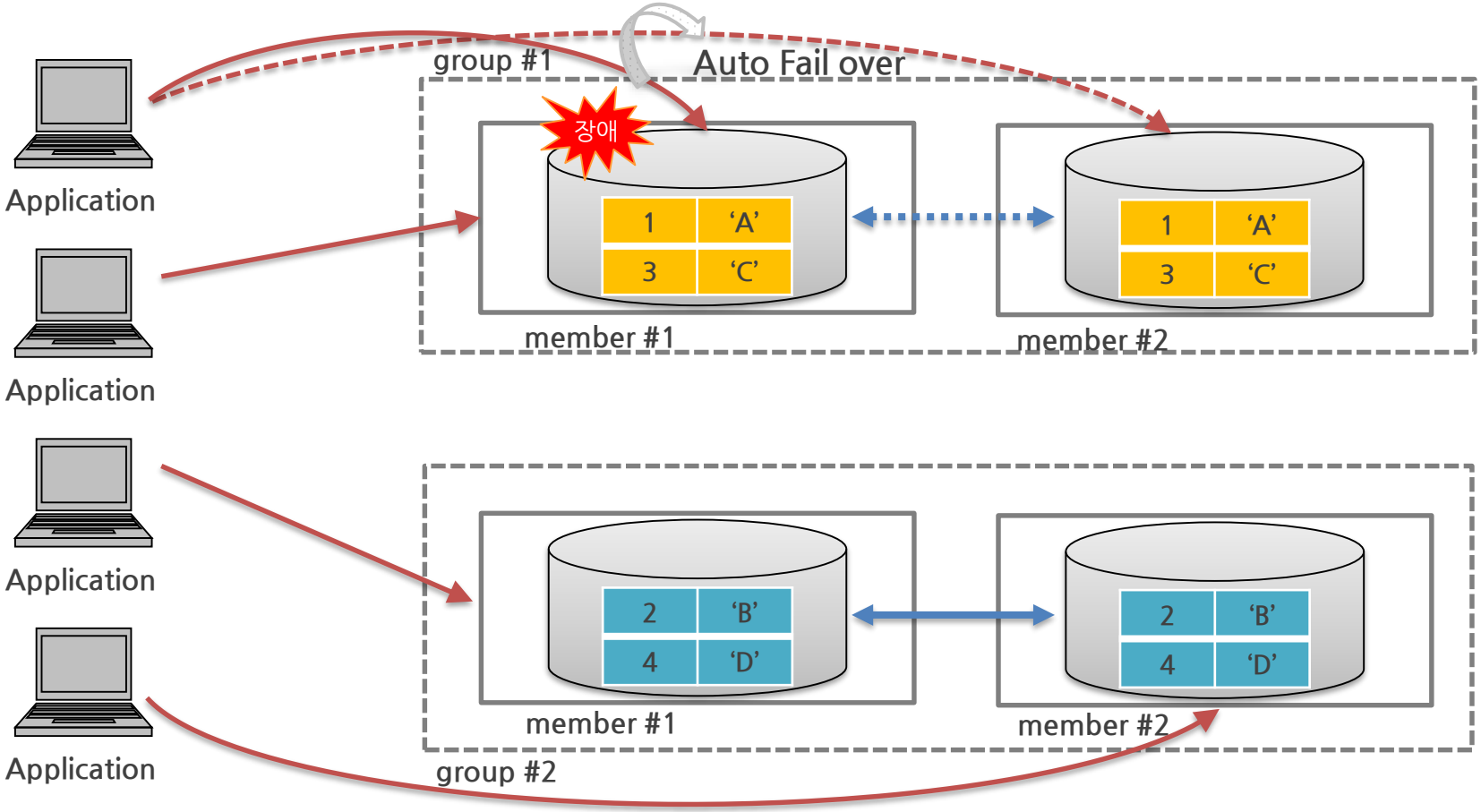


© 2017 SUNJESOFT CO., LTD. All rights reserved.

Goldilocks Cluster 핵심기술 - High Availability

데이터 처리 중인 노드에 장애가 발생하면 자동으로 해당 세션들을 동일 그룹 내 다른 멤버 노드로 이관하여 실패 없이 해당 요청을 처리합니다.

- 장애 노드의 물리적인 복구 완료 후에 Auto/Manual Rebalancing 작업 수행으로 정상 복구

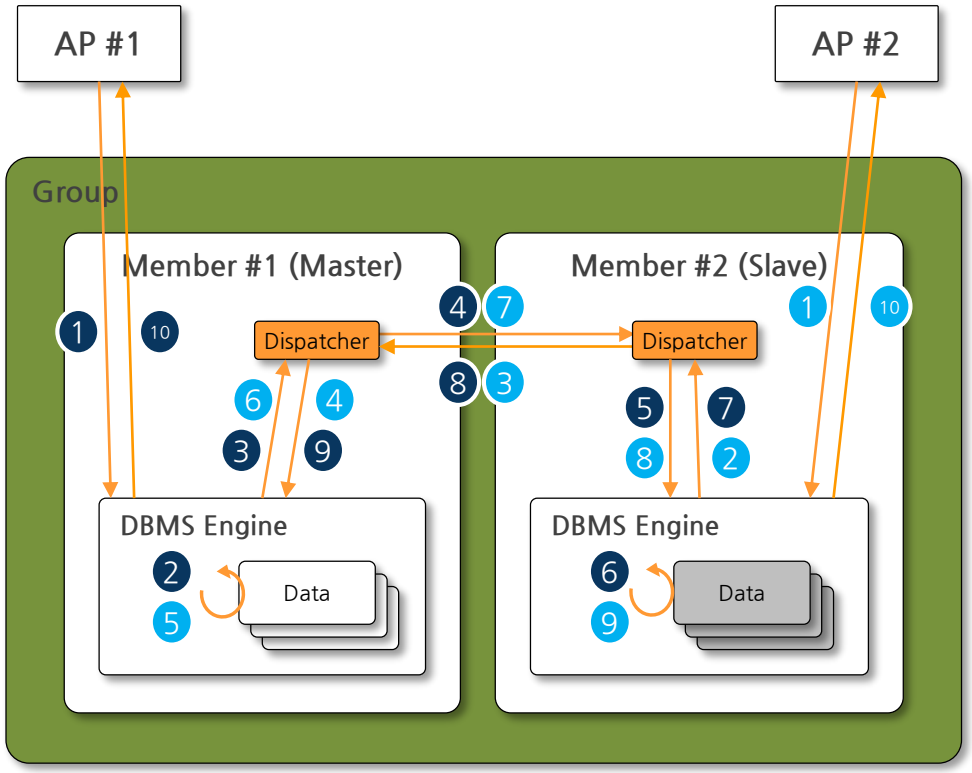


© 2017 SUNJESOFT CO., LTD. All rights reserved.

Goldilocks Cluster 핵심기술 - 데이터 정합성

특정 데이터 노드 장애 시에도 지속적인 서비스가 가능하도록 Replica 구성을 통해 데이터 다중 복제를 함으로써 데이터와 서비스의 연속성을 보장합니다

- 데이터 Layer에서 Data Replication으로 부하 최소화
- SQL Shipping을 통한 N/W 부하 최소화



- **데이터 다중화 구성**
  - 동일 그룹 내 멤버간 데이터 복제
  - 최상의 성능보장을 위해 그룹 내 모든 멤버가 Active상태로 동작
  - 데이터 충돌 방지를 위해 모든 변경 완료는 마스터 노드 완료 후 다른 멤버 노드 완료
  - 각 멤버는 All Master, All Active 구조

- **마스터 노드 변경 요청 처리**
  (AP) -> (Local Session) -> (마스터 노드)

Data 변경에 대한 작업은 로컬 노드에서 먼저 처리 후 원격 노드들의 작업 완료

- **슬레이브 노드 변경 요청 처리**
  (AP) -> (Local Session) -> (슬레이브 노드)

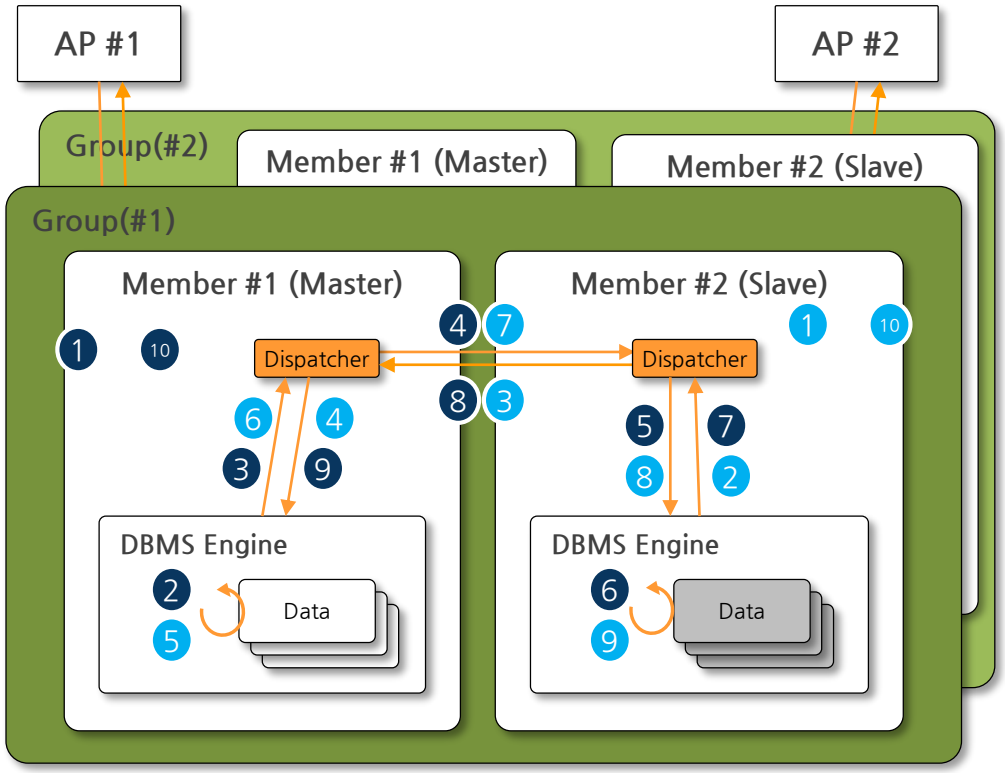
Data 변경 작업은 원격 마스터 노드에서 먼저 처리 후 로컬 노드에서 작업 완료

© 2017 SUNJESOFT CO., LTD. All rights reserved.

Goldilocks Cluster 핵심기술 - Masterless 구조

특정 데이터 노드 장애 시에도 지속적인 서비스가 가능하도록 Replica 구성을 통해 데이터 다중 복제를 함으로써 데이터와 서비스의 연속성을 보장합니다

- 그룹별 Master Node (Coordinator)존재하여 데이터 정합성 문제 해결
- 모든 Member 노드에서 Read/Write가능한 구조



- **데이터 다중화 구성**
  - 동일 그룹 내 멤버간 데이터 복제
  - 최상의 성능 보장을 위해선 그룹 내 모든 멤버가 Active상태로 동작
  - 데이터 충돌 방지를 위해 모든 변경 완료는 마스터 노드 완료 후 다른 멤버 노드 완료

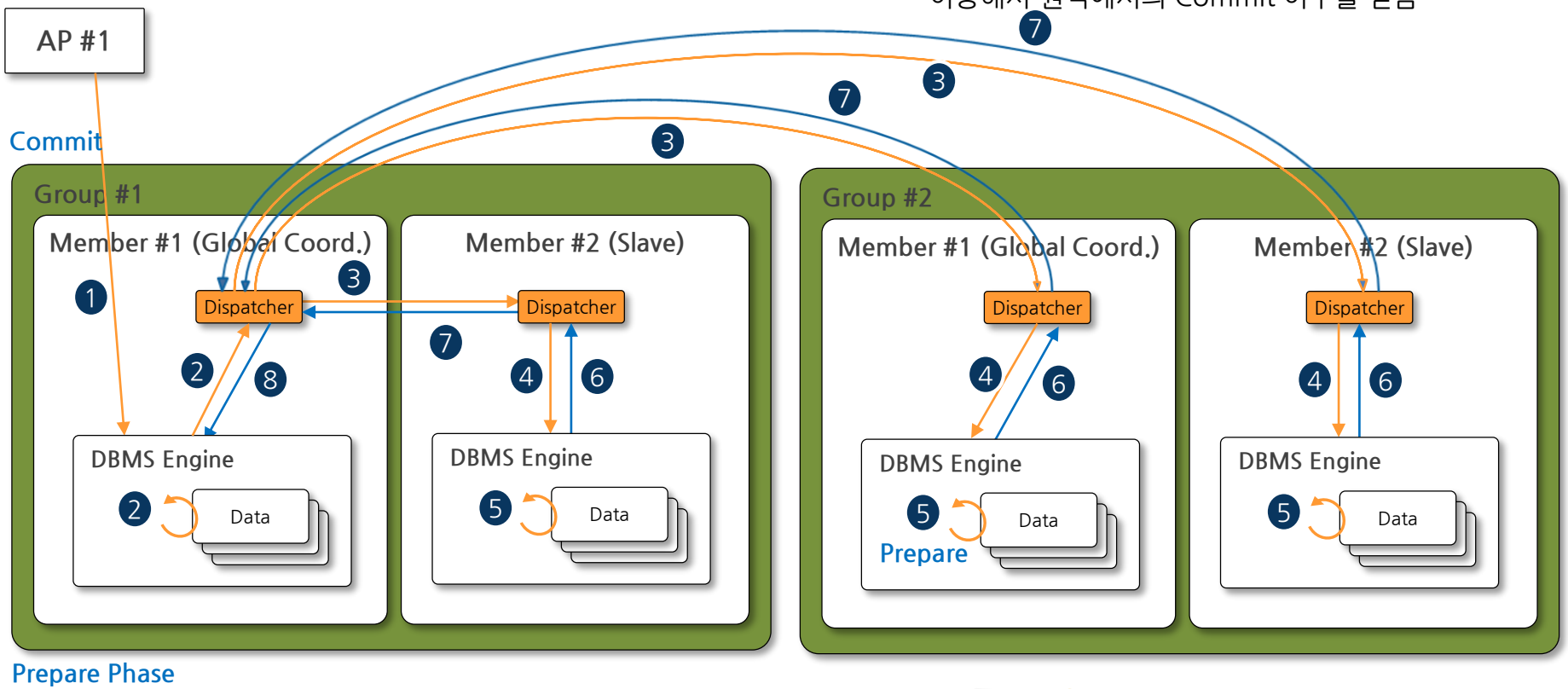
© 2017 SUNJESOFT CO., LTD. All rights reserved.

Goldilocks Cluster 핵심기술 - Global Transaction

Transaction이 변경한 데이터들이 두 개 이상의 Group 에서 이루어진 경우

- 2 phase commit 사용 (Prepare and Commit)
- Transaction Commit 절차가 전체 시스템으로 확장
- 전체 System에 SCN 동기화
- $Global\ Tranaction = \sum_{k=0}^{m}(Local\ transaction\ in\ System)_k$

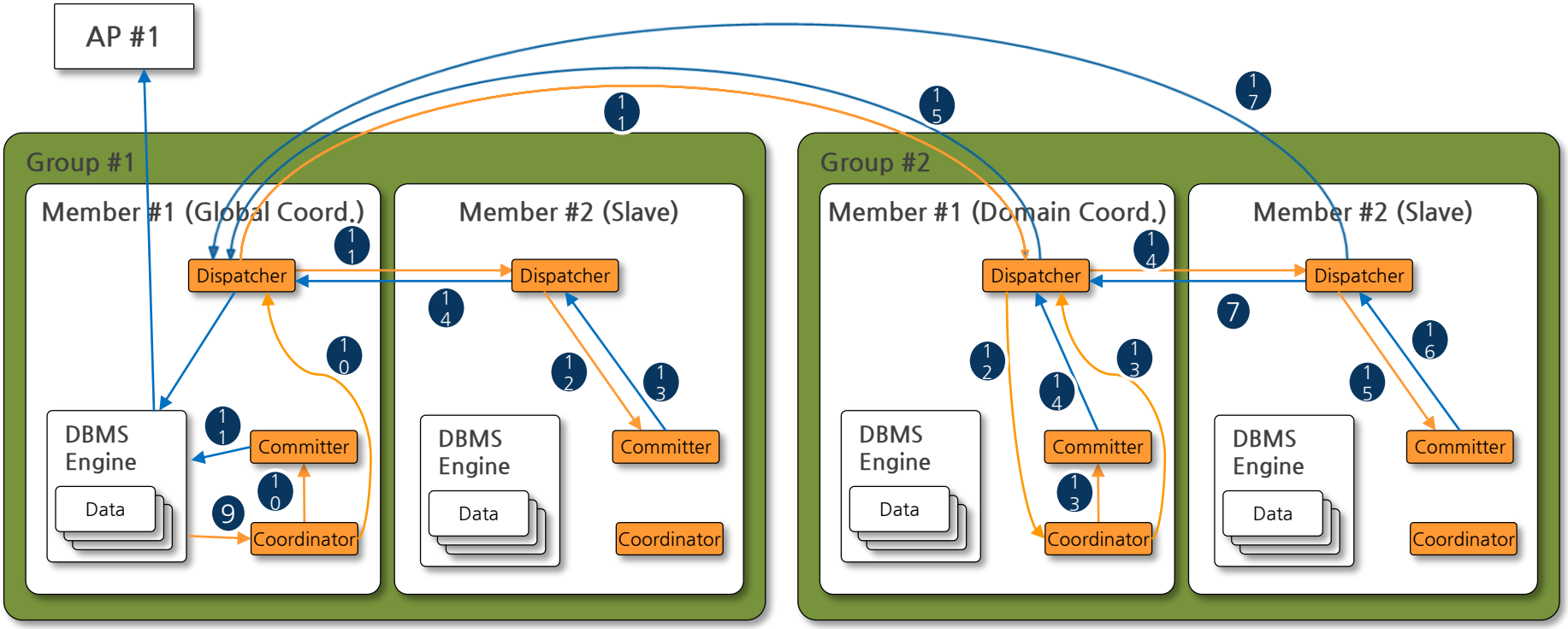
- **Prepare Phase**
  - Transaction Commit을 준비하는 단계
  - Local Transaction이 존재하는 모든 노드에 Prepare Protocol을 전파
  - Global Transaction ID를 redo log에 기록
  - 향후 Fail-back시 기록된 Global Transaction ID 를 이용해서 원격에서의 Commit 여부를 얻음



© 2017 SUNJESOFT CO., LTD. All rights reserved.

Goldilocks Cluster 핵심기술 - Global Transaction

- **Commit Phase**
  - 전체 시스템이 존재하는 모든 Local Transaction을 Commit하는 단계
  - Global Coordinator에 의해서 전파됨
  - Commit시 전체 System 내에서 유일한 번호(SCN)를 부여 받음
  - 전체 시스템 내 모든 멤버간 SCN 동기화



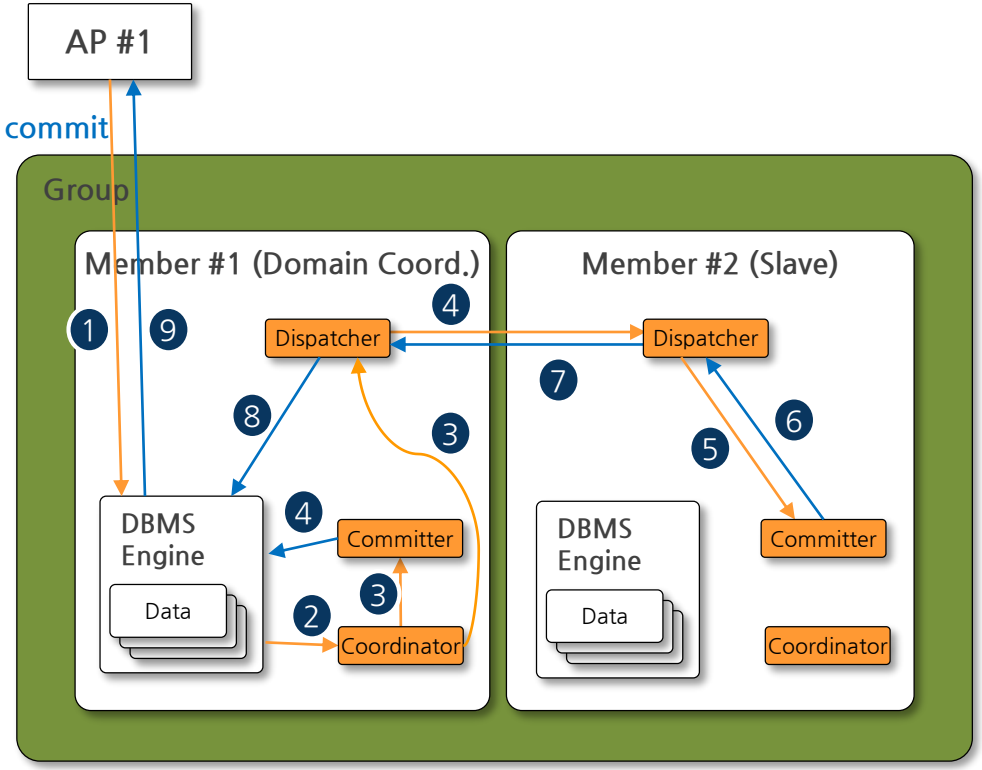
Commit Phase

© 2017 SUNJESOFT CO., LTD. All rights reserved.

Goldilocks Cluster 핵심기술 - Domain Transaction

Transaction이 변경한 데이터들이 한 개의 Group 내에서만 이루어진 경우

- 1 phase commit 사용
- Transaction Commit 절차가 한 개의 Group 으로 한정됨
- Group 별 완벽한 Scalability 보장
- Group이 증가할수록 무한한 Transaction Throughput 증가
- $Domain\ Tranaction = \sum_{k=0}^{n}(Local\ transaction\ in\ Group)_k$



- **Domain Coordinator**
  - .Domain Transaction을 처리하기 위한 조정자 프로세스
  - Group 내에 한 개의 멤버에서만 전담
  - Commit시 트랜잭션에 Cluster System 내에 유일한 번호 (SCN)를 부여함
  - 동일 그룹 내 멤버간 SCN 동기화

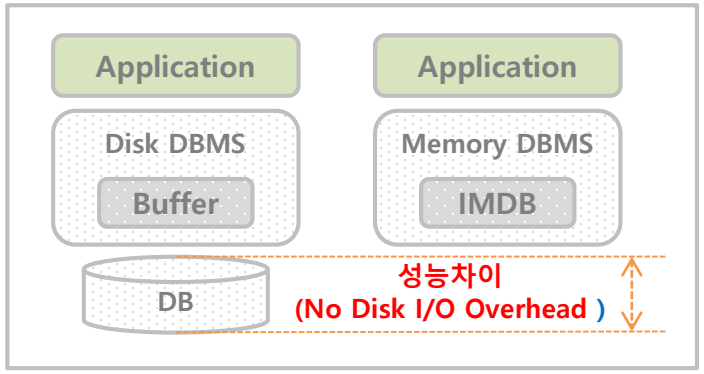
- **Committer**
  - Local Transaction을 Commit 하는 Commit 전용 프로세스
  - Coordinator가 보내준 SCN으로 Commit
  - Group/Async Commit 기법으로 Throughput 및 latency 극대화

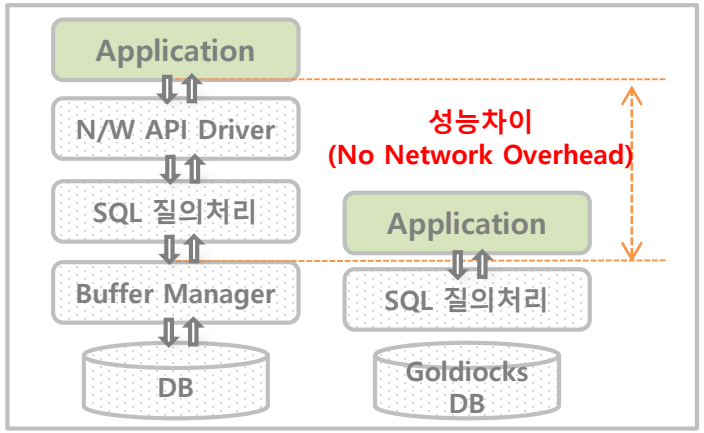
© 2017 SUNJESOFT CO., LTD. All rights reserved.

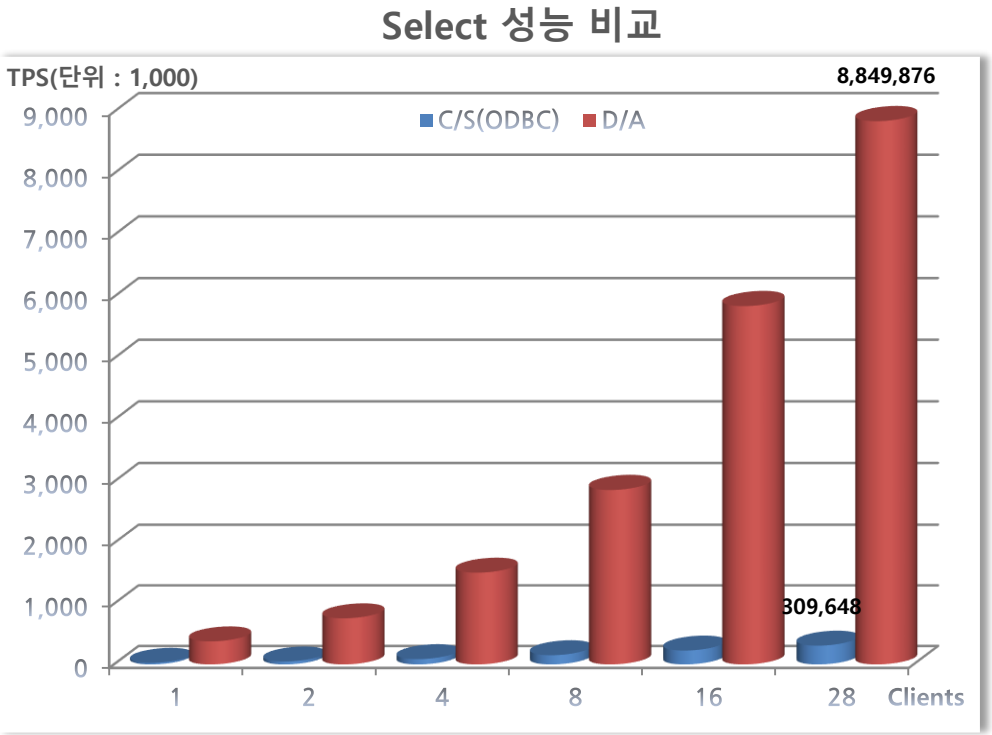
Goldilocks™ 특장점 - DA(Direct Access)모드

Network I/O 아키텍처 혁신

- DA (Direct Access)mode ： No network overhead + Optimized Query Processing
- DA mode Performance： Over 10~30Times Throughput Than Standard ODBC Interface
- 사용자는 표준 ODBC Interface 사용법과 동일하게 사용







- 테스트 장비 : Intel Xeon E5 2697v3 2.6GHz 14C * 2EA / 128GB Memory
- 데이터 건수 : 3,200,000건 , 레코드 크기 : 200 Byte/건

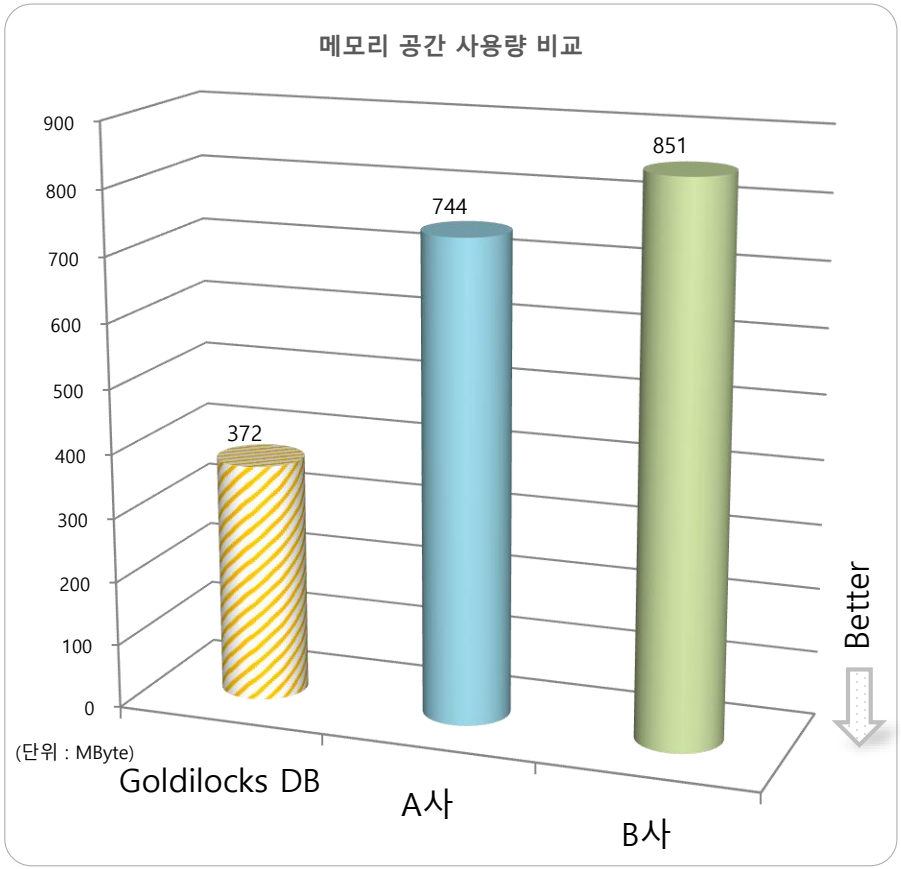
© 2017 SUNJESOFT CO., LTD. All rights reserved.

Goldilocks™ 특장점 - 메모리 사용량 최적화

메모리 사용량 최적화

- 자체 메모리 관리자를 통한 메모리 사용 최적화
- 가변길이 Numeric, Char 타입 제공
- In-Place DDL 제공

❖ 증권(KOSPI) 시세 데이터 100만건 저장 시 메모리 사용량 비교



테스트 환경

- 목적
  - In-Memory DBMS에 데이터 저장 시 물리적인 메모리 사용량 비교
- Test suite
  - 증권(KOSPI) 시세 데이터 100만 건 저장

❖ 스키마 구조 (컬럼개수 77개)

```
CREATE TABLE STOCK_JONG_CHE (
        F_16013     VARCHAR(12)   NOT NULL
        F_99990     VARCHAR(3)    NOT NULL
        F_99991     NUMBER(20)    NOT NULL
        F_12506     VARCHAR(8)    NOT NULL
        F_14501     NUMBER(11)    NULL
        F_14511     NUMBER(11)    NULL
        F_14531     NUMBER(11)    NULL
        F_14541     NUMBER(11)    NULL
        F_14561     NUMBER(11)    NULL
        F_14563     NUMBER(11)    NULL
        F_14565     NUMBER(11)    NULL
        F_14566     NUMBER(10)    NULL
        F_14567     NUMBER(11)    NULL
        F_14568     NUMBER(10)    NULL
        F_14569     NUMBER(10)    NULL
        F_14853     VARCHAR(8)    NULL
        F_14854     NUMBER(12,2)  NULL
        F_14855     NUMBER(12,2)  NULL
        F_15001     NUMBER(11)    NULL
        F_15004     NUMBER(12,2)  NULL
        F_15006     NUMBER(4)     NULL
        F_15007     NUMBER(11)    NULL
        ...
        F_30615     NUMBER(10)    NULL
        F_30616     NUMBER(12,2)  NULL
        F_30617     NUMBER(11)    NULL
        F_30618     NUMBER(10)    NULL
        F_30626     NUMBER(4)     NULL
        F_30628     NUMBER(20)    NULL
        F_30629     NUMBER(20)    NULL
        F_30691     NUMBER(13,2)  NULL
CONSTRAINT STOCK_JONG_CHE_PK PRIMARY KEY( F_16013,F_99990,F_99991,F_12506 DESC ) );
```

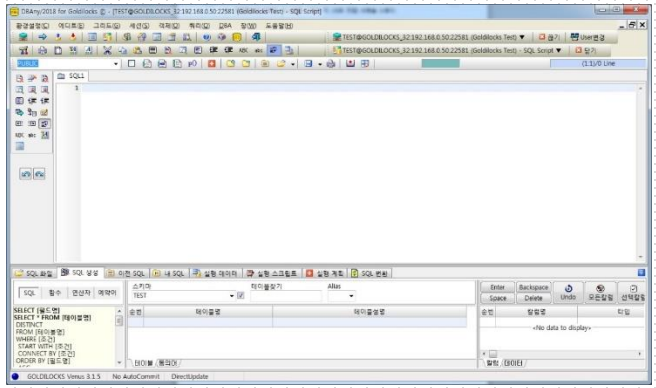

© 2017 SUNJESOFT CO., LTD. All rights reserved.

Goldilocks™ 특장점 - 산업표준 Interface 지원

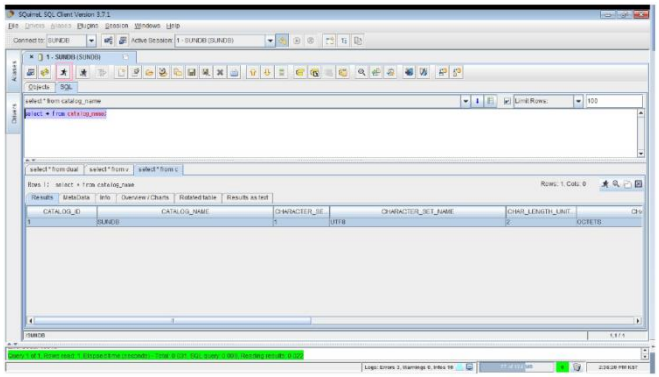
산업 표준 ODBC, JDBC 를 직접 개발 및 지원

- 다양한 3rd Party 제품들과의 손쉬운 연동 지원
- 표준화 Connection API를 통해 3rd Party 제품 확장 가능

DBany for Goldilocks

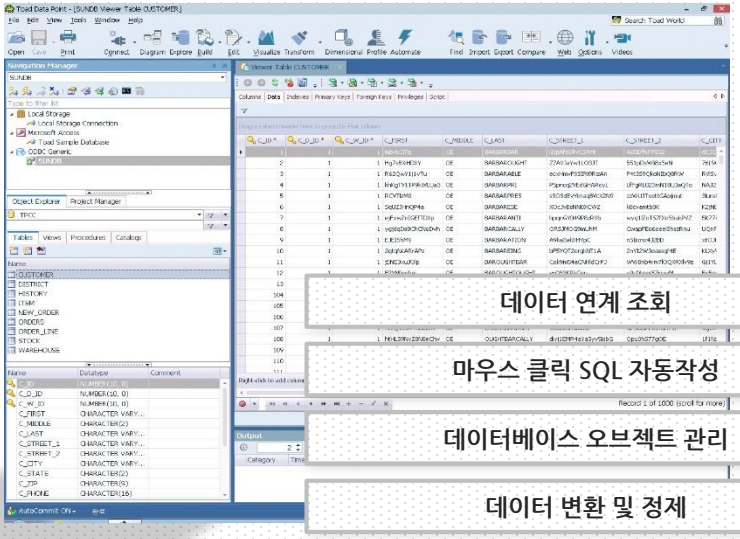


squirrel-sql (무료/오픈소스)



Toad Data Point(상용)

- Self-service data access, integration and Preparation tool
- ER Diagrammer and data browsers, SQL editor



데이터 연계 조회

마우스 클릭 SQL 자동작성

데이터베이스 오브젝트 관리

데이터 변환 및 정제

테이블간 데이터 비교 및 동기화

대량 데이터 분석 및 시각화

© 2017 SUNJESOFT CO., LTD. All rights reserved.

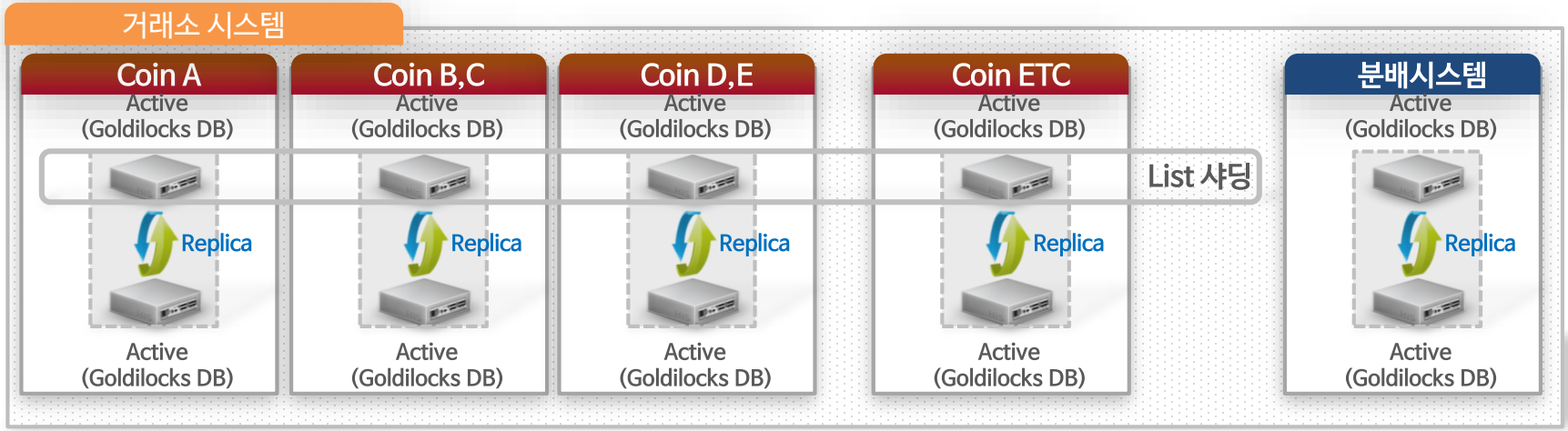
3. 구축 사례

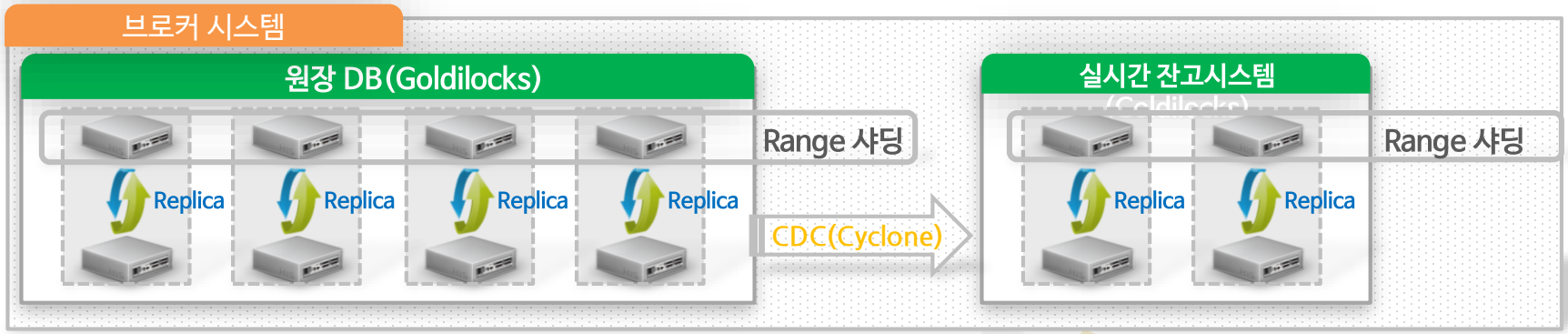
© 2017 SUNJESOFT CO., LTD. All rights reserved.

Goldilocks Use Case - 빗썸거래소 매매체결시스템

➢ 거래소 DB, 원장DB 모두 확장 가능한 Cluster 구조로 구성

- 거래소DB : List 샤딩 기술을 이용하여 추가 상장되는 코인에 대한 즉시 대응 가능, 시세 분배를 위해 분배시스템 구축
- 원장DB : Range 샤딩 기술을 이용하여 고객 증가에 따른 성능 향상 유연한 대응 및 실시간 잔고시스템을 별도로 구축하여 조회 속도 보장 및 지불 서비스에 활용
- 향후, 코인/고객 추가에 대해 유연한 시스템 확대 가능 구조





© 2017 SUNJESOFT CO., LTD. All rights reserved.

Goldilocks Use Case - 경찰청 안전 Dream

사회적 약자를 관리하고자 하는 시스템으로 실시간 지문인식 및 대용량 데이터 검색 등을 위하여 Goldilocks적용

- 최적의 데이터 분산 환경과, DB 구조를 통하여 AS-IS 모델에 대한 효율적인 고도화 사업에 적용
- 고비용 구조의 외산 데이터베이스를 대체
- 전자정부 솔루션 해외 수출 시 인프라 기반 기술도 동반 진출 효과

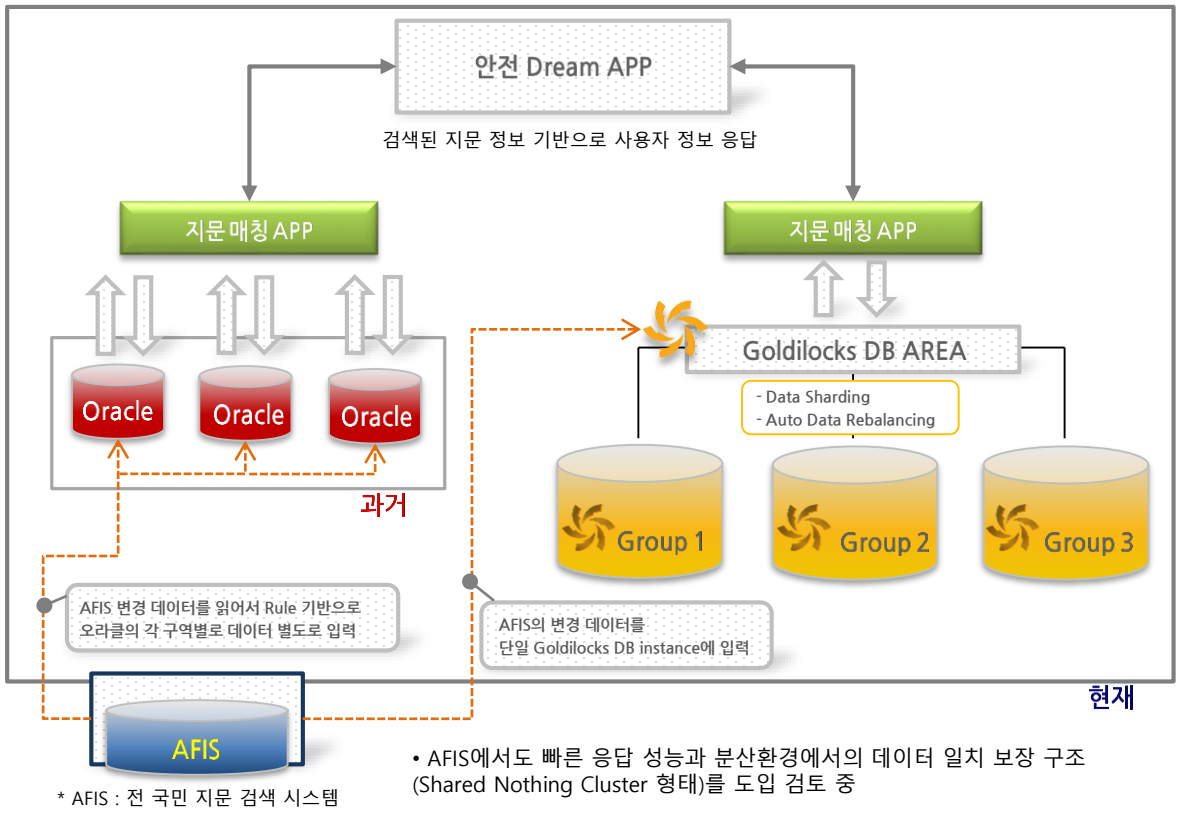
❖ 경찰청 안전 Dream APP 고도화

• 요구 사항
- 데이터 분산 저장
- 빠른 응답 속도
- 균일한 성능 보장
- 운영/관리 용이성 필요
- 데이터 불일치 문제 회피 필요

• 도입 결과
- **매칭 TX 건수 :약 3억 건/day**
- **3개의 노드에 데이터 분산(3X1)**
- **샤딩 정책(Hash Sharding)**
- PeakTime Test시 안정성 확보
- 단일 인스턴스 접근으로 데이터 정합성 보장 및 효율적인 APP 구조 개선

• 비용 결과
- 시스템 성능 저하에 대한 Cost 감소
- 빠른 검색 및 응답으로 인한 업무 지연 Cost 감소
- 데이터 정합성 보장으로 안정적 시스템 구현으로 장애 복구를 위한 Cost 감소

• 기 도입 S/W의 이식성
- 타 시스템 연동성에 문제 없음



* AFIS : 전 국민 지문 검색 시스템

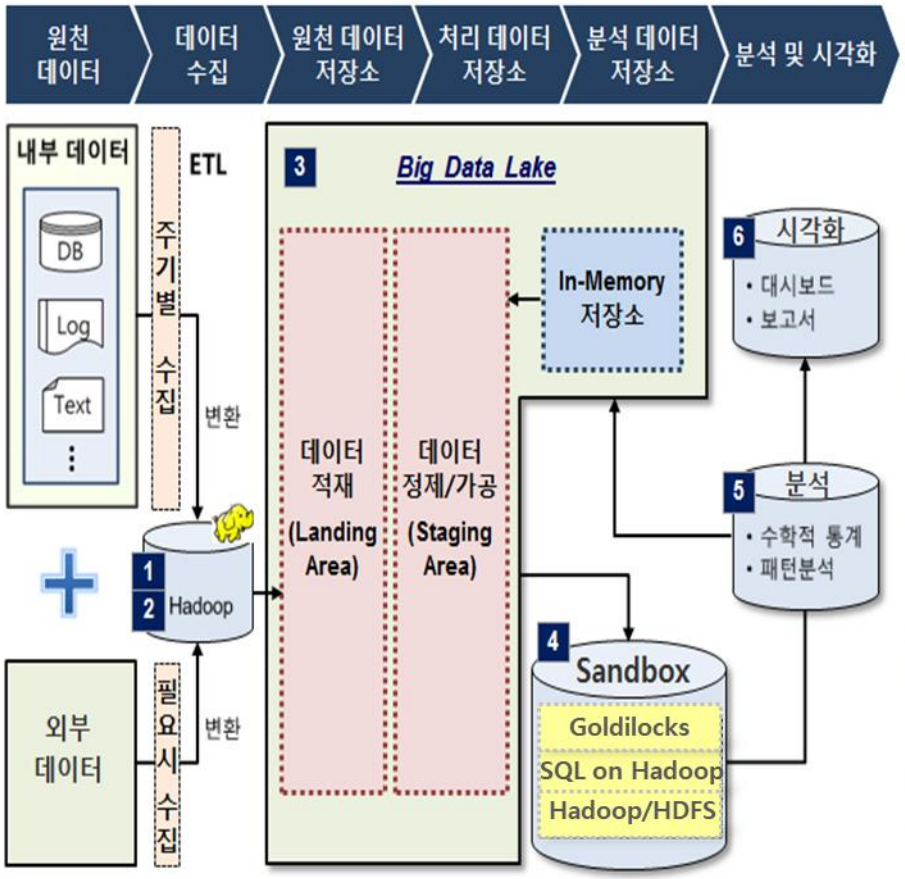
• AFIS에서도 빠른 응답 성능과 분산환경에서의 데이터 일치 보장 구조 (Shared Nothing Cluster 형태)를 도입 검토 중

© 2017 SUNJESOFT CO., LTD. All rights reserved.

Goldilocks Use Case - 금융결제원 빅데이터

CD/ATM기 거래시, 금융결제원 CD공동망을 통해 저장되는 데이터를 추출하여 데이터 마트를 구성하고, 이를 Goldilocks에서 가공/분석하는 시스템

- CD/ATM기 원천 데이터 추출 ---> HIVE, 분석 Mart ---> 시각화 ---> ATM재배치 등 효율화



영역별 기능 구성

1 내부 데이터 수집/저장
- 내부 정형/비정형 데이터를 주기별로 수집하여 하둡 시스템으로 변환·저장

2 외부 데이터 수집/저장
- 수집도구를 통해 적재된 외부 데이터를 하둡 시스템으로 변환·저장

3 빅데이터 레이크
- 원천 데이터를 적재(landing), 정제/가공하여 staging 영역에 저장

4 샌드박스(보호영역)
- 빅데이터 고급 분석을 위한 샌드박스의 구성

5 정형/비정형 데이터 분석
- 수학적 통계기법 등을 활용하여 Insight 도출

6 시각화
- 빅데이터 분석결과를 활용한 시각화, 보고서 생성 영역

© 2017 SUNJESOFT CO., LTD. All rights reserved.

한국거래소 차세대 매매체결시스템(1/2)

선진금융 시스템을 벤치마킹 하여 금융기술과 국내 최고의 S/W기술을 융합, 글로벌 거래소 수준대비 최상위 시스템

- 고효율, 저비용, 초고성능, 고품질 트레이딩 시스템 구축을 목표
- 최신 IT인프라, 초고속 H/W, S/W기반 기술 적용, 선진기술에 기반한 아키텍처 수립

선재소프트 개발 구축 범위

- In-memory 기술 적용
- 데이터 동기화 방식 설계 및 적용
- 업무 프로세스 방식 설계 및 적용

기술구분	EXTURE	EXTURE+
거래처리	동기식	비동기식
프레임워크	멀티프로세싱	멀티쓰레딩
데이터저장장치	Oracle	메모리DBMS
메시지미들웨어	Disk Queue방식	RDMA,LDMA 방식
개발언어	C	C, C++
시스템 이중화	디스크 공유방식	메시지 공유방식
OS/하드웨어	UNIX/UNIX서버	LINUX/x86서버
네트워크	1 GbE	10 GbE, Infiniband

KRX KOREA EXCHANGE

非 동기 기술구조 기반의 거래처리

- 거래소↔회원사, 가동↔예비시스템, 가동↔DR 센터 등 비동기 기술구조 기반의 거래처리로 Throughput 개선

㎲(1/100만) 성능의 기반기술 개발

- 메모리 호가집계장,RDMA 기반 주문/체결 메시지 통신, 멀티쓰레딩 기반 프레임워크 등 초고속 기반 기술 개발로 Latency Time 획기적 개선
- Millisecond → Microsecond로 처리속도 진입

최신 인프라 도입 및 적용

- x86 기반 리눅스 서버, 초고속 네트워크(10GbE, Infiniband 등) 기반의 최신 기술의 인프라 구축
- EXTURE 보다 비용 대비 성능이 획기적으로 개선된 저비용/고효율의 IT인프라로 전환

<출처: 오픈플랫폼 기반의 차세대 증권거래시스템. 한국거래소, 2015, 이창진>

© 2017 SUNJESOFT CO., LTD. All rights reserved.

한국거래소 차세대 매매체결시스템(2/2)

차세대 프로젝트를 통하여 글로벌 수준을 뛰어 넘는 초고성능 시스템을 구축 할 수 있었으며 증권. 금융업계의 변화. 혁신을 주도한 선도적 사례가 되었습니다.

- 글로벌 거래소 가운데 최상위 시스템 구축
- 고품질. 고성능 시스템으로 대외 수출 경쟁력 확보

구 분	EXTURE	EXTURE+	비 고
Throughput (단일종목 처리)	250 TPS	20,000 TPS (글로벌수준: 10,000TPS)	기존대비 80 배 향상
Latency Time	20,000 ㎲	70 ㎲ (글로벌수준: 100-150㎲)	기존대비 285 배 향상
전체 처리용량	4,000 만건	최초 20,000만건 (시장 상황에 따라 확장)	기존대비 5 배 향상

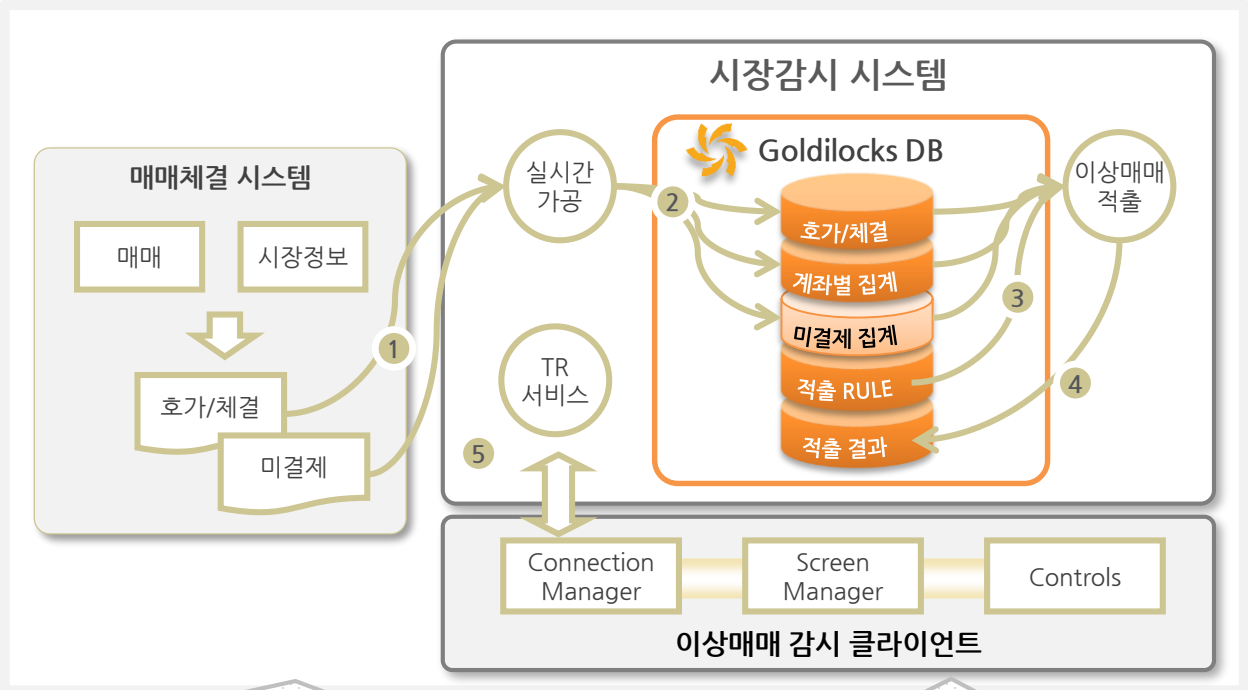
<출처: 오픈플랫폼 기반의 차세대 증권거래시스템. 한국거래소, 2015, 이창진>

© 2017 SUNJESOFT CO., LTD. All rights reserved.

한국거래소 시장감시 시스템

신속한 시장감시 업무를 위해 대용량 메모리를 이용하여 거래소 매매체결 시스템에서 체결 및 호가 정보를 실시간으로 받아, 실시간 가공 및 적출을 통해 시장감시 시스템을 구축 하였습니다.

- 국내 최초 3.5Tbyte 대용량 IMDB적용



1 실시간 가공

✓호가/체결 미결제 파일 처리

2 가공 후 저장

✓호가/체결, 계좌별 집계, 미결제 집계 가공 및 결과 저장

3 이상매매 적출

✓가공 데이터 및 적출Rule을 이용하여 적출처리

4 적출 결과 저장

✓이상매매 징후로 적출된 결과를 저장

5 이상매매 감시

✓이상매매 시장감시 및 서비스 수행

- 성능 개선 사항

업 무	Oracle 사용	Goldilocks DB 사용
실시간 가공 처리	처리 프로세스 : 60개 프로세스 당 18건/초 처리 전체 처리량 : 1,080건/초	처리 프로세스 : 60개 프로세스 당 1,060건/초 처리 전체 처리량 : 63,600건/초
적 출	지연	지연 없음
시장감시 결과	수 분 ~ 수 일	수 초 ~ 수 분

- 운영환경

- H/W : HP Server (10core/1cpu) * - 8CPU, Mem TOTAL 3.2TB
- OS : Linux (RHEL 6.4)
- DBMS : Goldilocks DB
- DISK : PCI-e SSD (Fusion-IO)

© 2017 SUNJESOFT CO., LTD. All rights reserved.

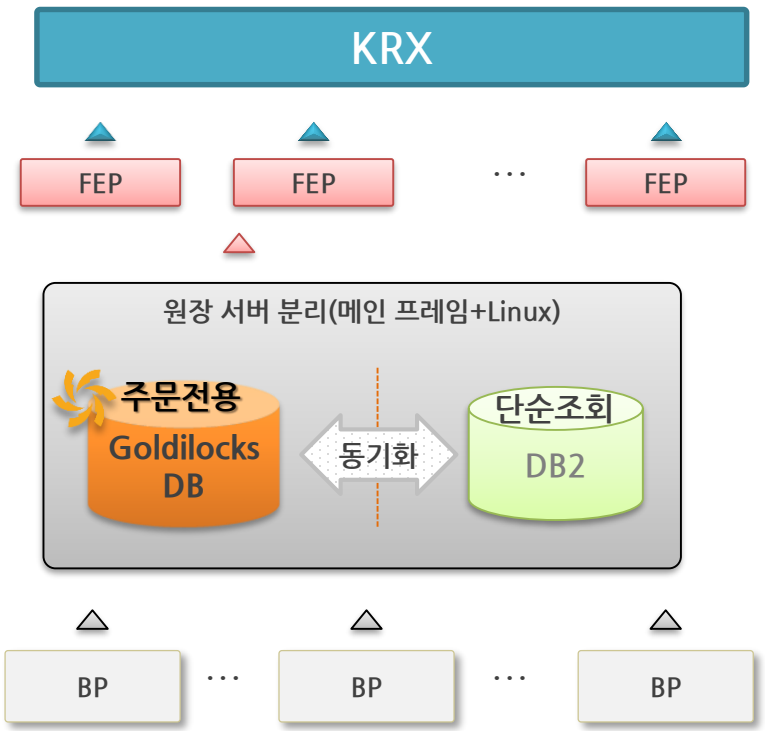
S증권 주문시스템 A-A프로젝트(1/2)

기존 메인 프레임의 원장 시스템에서 주문과 체결관련 업무를 분리하여 Trade Book과 Position정보를 국내 최초로 IMDB로 구성하여 고속 주문시스템을 구축 하였습니다.

- '원장데이터에 국내최초 IMDB적용' 사례 - 초고속 주문시스템 구축

❖ As-Is 시스템

❖ To-Be 시스템

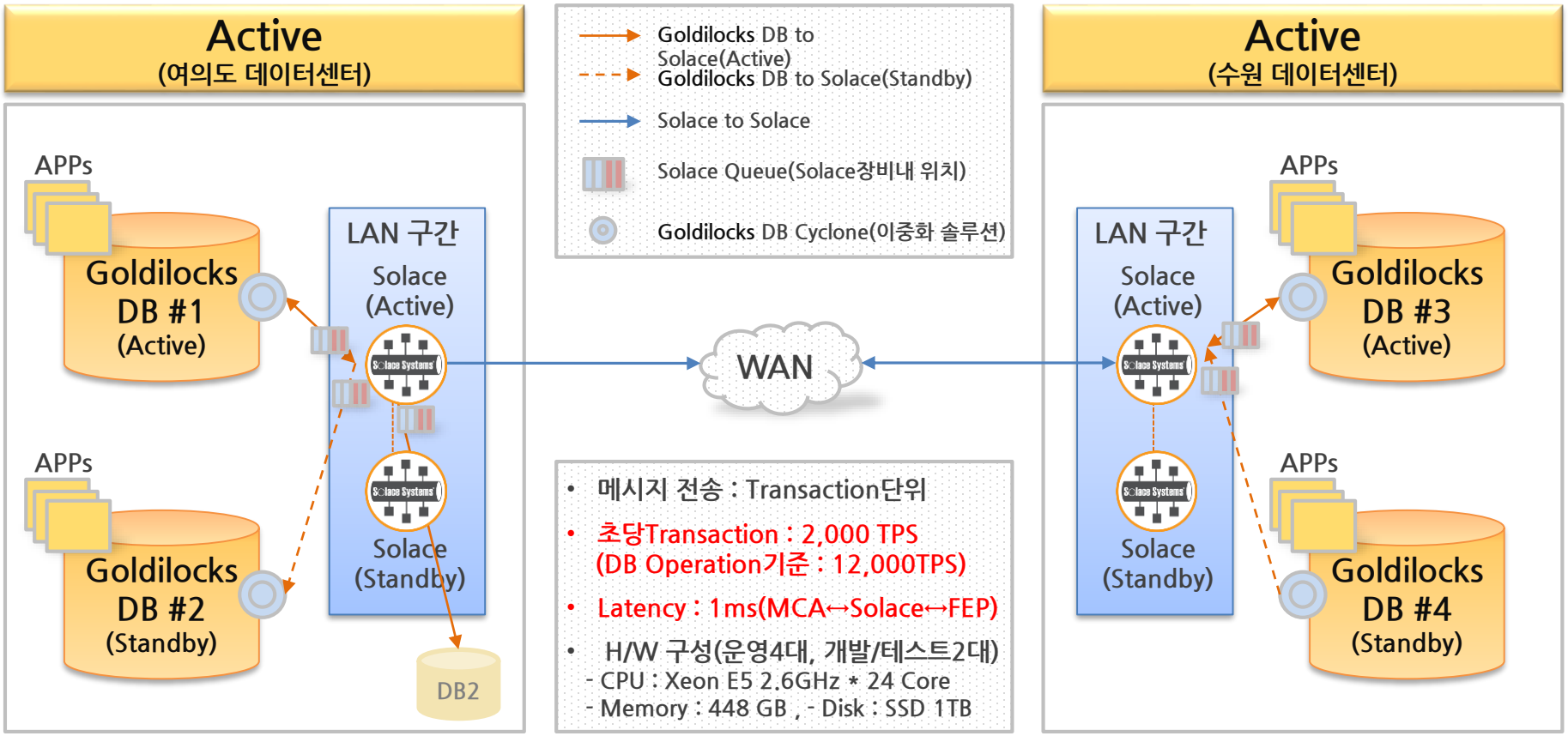


© 2017 SUNJESOFT CO., LTD. All rights reserved.

S증권 주문시스템 A-A프로젝트(2/2)

주문업무의 비즈니스 연속성을 보장하고 재해복구 시스템에 상시 가용성을 구현 하였습니다.

- 증권업계 최초로 원거리 Active-to-Active 주문 시스템 체계 구축
- 지역 및 원격간 통신은 LLM(Low Latency Messaging Middleware)을 통한 Guarantied Messaging 구현

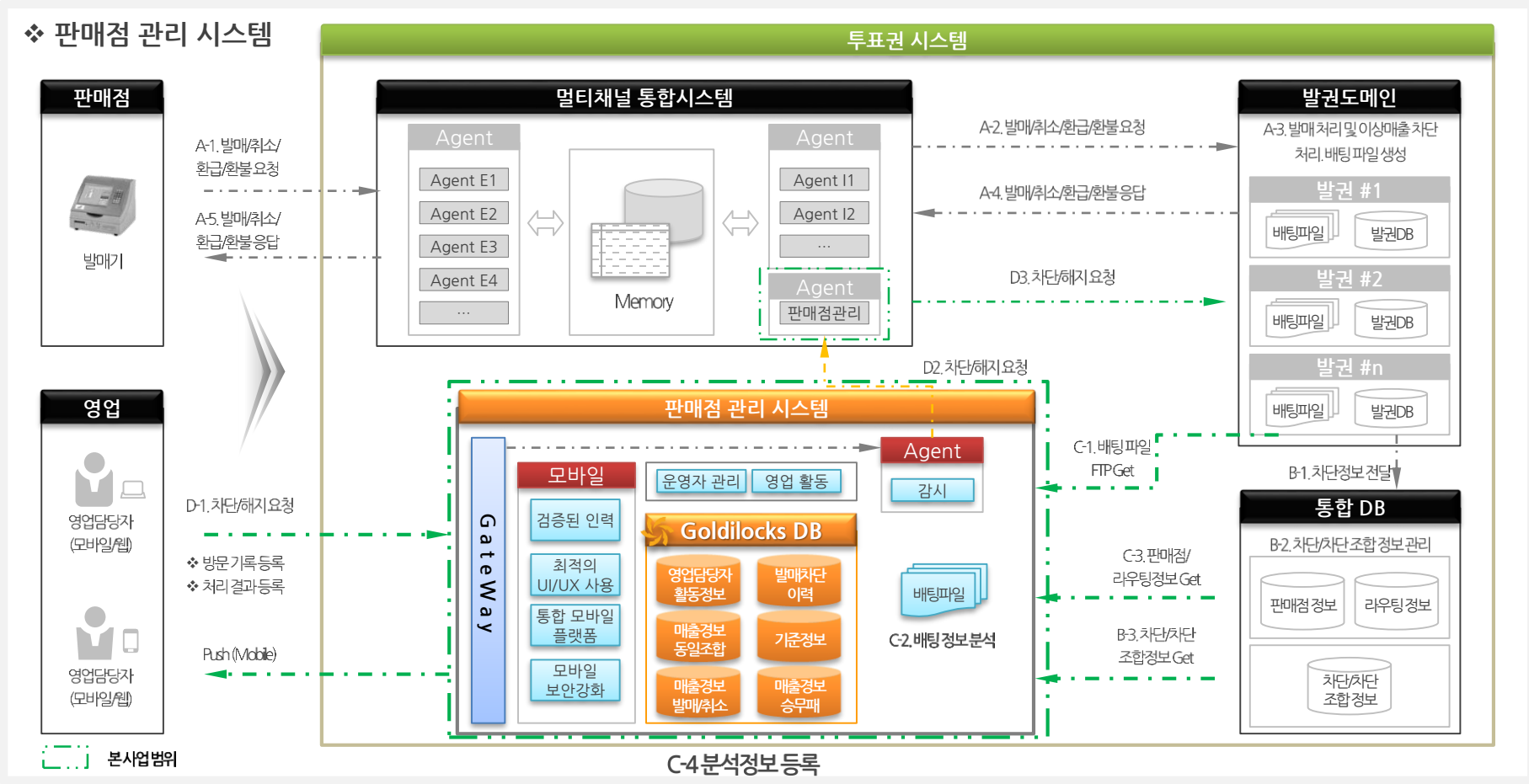


© 2017 SUNJESOFT CO., LTD. All rights reserved.

스포츠토토 판매점 관리시스템

체육진흥투표권 사업의 불법적인 행위를 규제하기 위하여 판매점관리시스템을 구축 하였습니다.

- 판매점의 불법행위를 감시하고 통제와 판매점의 부정행위를 사전에 차단
- 투표권사업의 발매건전성 강화



© 2017 SUNJESOFT CO., LTD. All rights reserved.

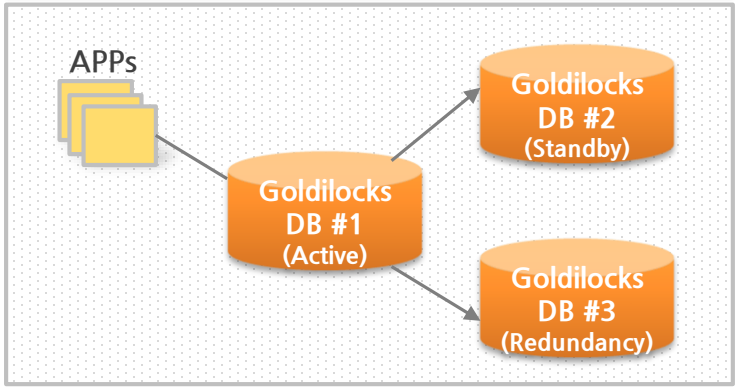
China Unicom 차세대 cBSS 사업

全 중국 31개 행정구역(22개성, 4대 직할시,5개 자치구)대상으로 실시간 정산,과금을 구현 하였습니다.

- 총 8개 권역(Area1~Area8)으로 권역당 2-5개 도시 정산, 과금 (권역별 가입자 수 3천만~5천 만명)
- 권역별 Active-Standby-Redundancy 구성으로 안정성 및 고가용성 확보(2016.06 최종 오픈 완료)

권역별 구성

권역당 시스템 구성 및 환경



- H/W 사양 (3대 구성)
  - CPU : IBM Power Linux P8 3.4GHz * 10 CPU
  - Memory : 512GB
  - Disk : SSD 2.2TB
- N/W
  - 10GbE

© 2017 SUNJESOFT CO., LTD. All rights reserved.

China Telco 사업

시스템명	적용업무
차이나유니콤 OCC	[장비] 총 8대 운영(master 4 / slave 4) [업무] 선불 Billing 처리 시스템
허난모바일 CRM	[장비] 64 core / 250G / 10G 총 12대 -> 메인 DB(master 1 / slave 1) 조회 전용 DB (master 1 / slave 9) [업무] CRM 업무 시스템
심천모바일 국제로밍빌링	[장비] 총 2대 (master 1 /slave 1) [업무] 국제 빌링 업무시스템

© 2017 SUNJESOFT CO., LTD. All rights reserved.

Appendix
Goldilocks DB 구성의 특장점

© 2017 SUNJESOFT CO., LTD. All rights reserved.

Goldilocks DB 구성의 특장점(1/4)

① *확장성(Scale-Out)*

- ✓ High Throughput
  - Sharding을 통한 데이터 분산
  - 서버 추가에 따른 total throughput 향상
- ✓ Online Scale Out
  - Online Add/Remove Nodes(Server추가 및 제거 시 서비스 중단 없는 Scale Out)
  - Auto Data Re-balancing

② *정합성(ACID Transaction)*

- ✓ Global Transaction
  - 여러 노드에 데이터가 분산되어 있어도 하나의 DB처럼 동작 (Multi-Node Transaction Support)
  - ACID Compliant With MVCC
  - DBMS엔진에서 2Phase Commits Processing

© 2017 SUNJESOFT CO., LTD. All rights reserved.

Goldilocks DB 구성의 특장점(2/4)

③ *편의성(SQL)*

✓ ANSI 표준 SQL 지원
- SQL 표준 ISO/IEC 9075:2011준수
- ODBC 3.x for C/C++
- Embedded SQL for C/C++
- JDBC 1.x for Java

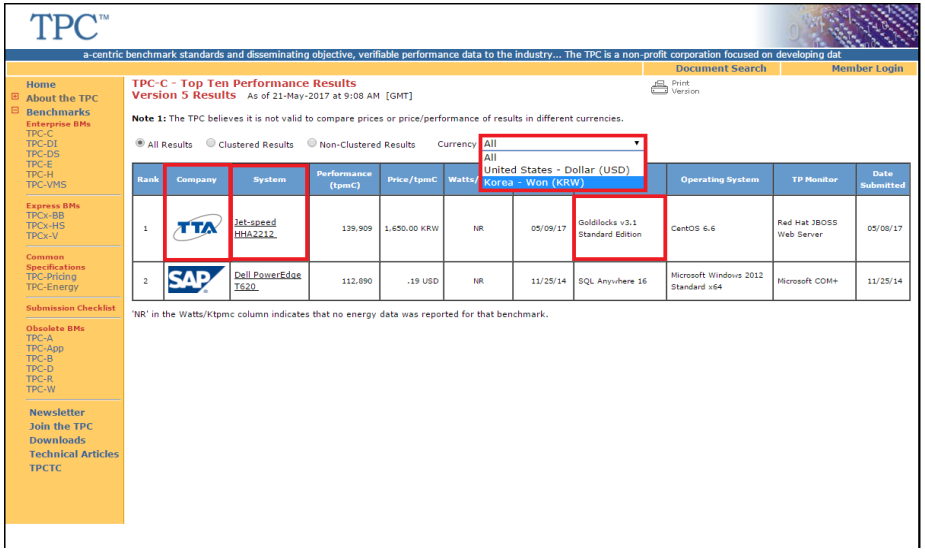
✓ 개발 및 운영의 편의성
- C,C++,java,C#등 다양한 언어 지원
- Data 저널링 시 Parallel Checkpoint 기법으로 일정한 서비스 성능 유지
- MVCC지원(Record Level Locking 지원)
- DB 개발자 TOOL 제공 (DBany for Goldilocks)

© 2017 SUNJESOFT CO., LTD. All rights reserved.

Goldilocks DB 구성의 특장점(3/4)

④ *안정성(TPC 국제 인증)*

- ✓ TPC-C 국제공인 인증 획득 : 인증비용 4억,품질검사 소요기간 : 8개월, Auditor : 프랑소와(TPC위원회 지정) => 철저한 품질테스트, TPC-C인증을 받은 제품은 OLTP환경에 적합한 DBMS라고 글로벌하게 인정함.



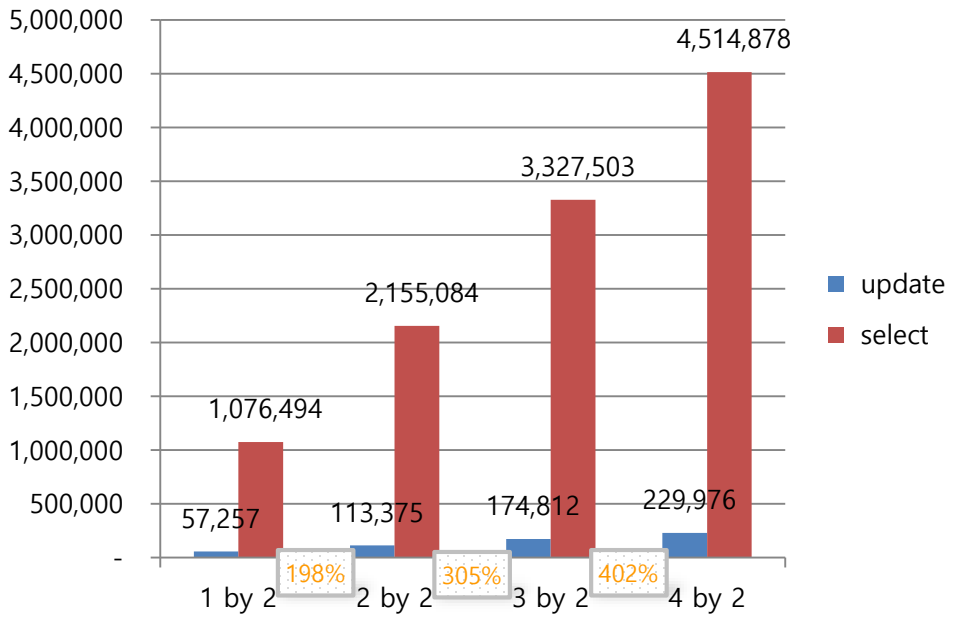
- ✓ High Availability
  - 데이터 다중화를 통한 고가용성 확보(일반적인 메모리DB가 쓰던 불완전한 Replication방식이 아니라 Replica방식으로 완벽한 Active-Active 운영 가능, 장애 시 Fail-Over를 통해 서비스 연속성 보장,이중화로 인한 Data Gap 없음)
  - No Single-Point Failure(단일 Point 장애로 전체 서비스가 정지하지 않음)

© 2017 SUNJESOFT CO., LTD. All rights reserved.

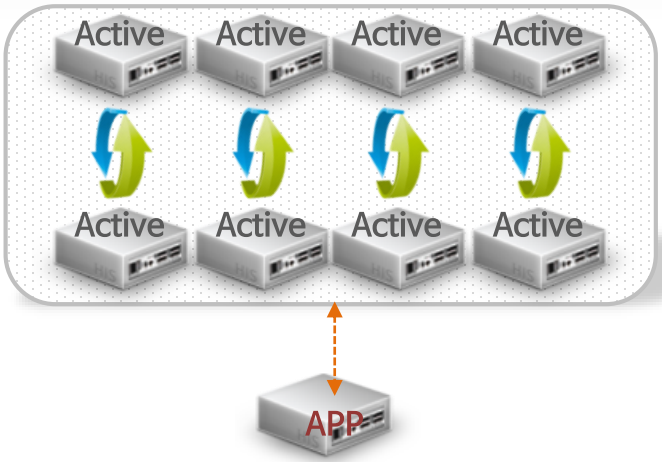
Goldilocks DB 구성의 특장점(4/4)

⑤Goldilocks DBMS는 Scale-Out 가능한 DBMS로써 노드를 증설하면 전체 시스템의 Throughput을 확장 할 수 있습니다.

Goldilocks 성능시험 (단위TPS)



성능 시험 환경



- DB서버 Spec : Intel Xeon E5 2697v3 2.6GHz 14C * 2EA / 128G mem
- Network 환경 : 10G
- 데이터 건수 : 8,000,000건
- 데이터 스키마 : int,char,date 등 11개 컬럼으로 구성
- 샤딩 정책 : RANGE 샤딩
- APP Query : 1건 당 update 후 commit
- 1by2 , 2by2, 3by2 ,4by2 테스트
- App thread 개수(per Node) : Upd :10 / Sel 40
- HA모드 : sync 모드 : (async) = 90 : 100

© 2017 SUNJESOFT CO., LTD. All rights reserved.

(주) 선재소프트

서울시 영등포구 당산로 171 금강펜테리움IT타워 604호
Tel 02-322-6288 / 070-4238-6288 Fax 02-322-6788

영업 문의 : sales@sunjesoft.com / 02-322-6288

SUNJESOFT
SUNJESOFT Corp.

© 2017 SUNJESOFT CO., LTD. All rights reserved.