

# AI 기술 대응과 데이터 분산 및 통합 전략

**Quest**  
Where Next Meets Now.

**Quest Software Korea**  
유광일 차장

# Agenda

## 1. 국내 AI 시장 상황과 미래

- 1) AI와 빅데이터 기술 개발 시장
- 2) 데이터 통합 및 분산

## 2. SharePlex를 이용한 데이터 통합 및 분산 전략

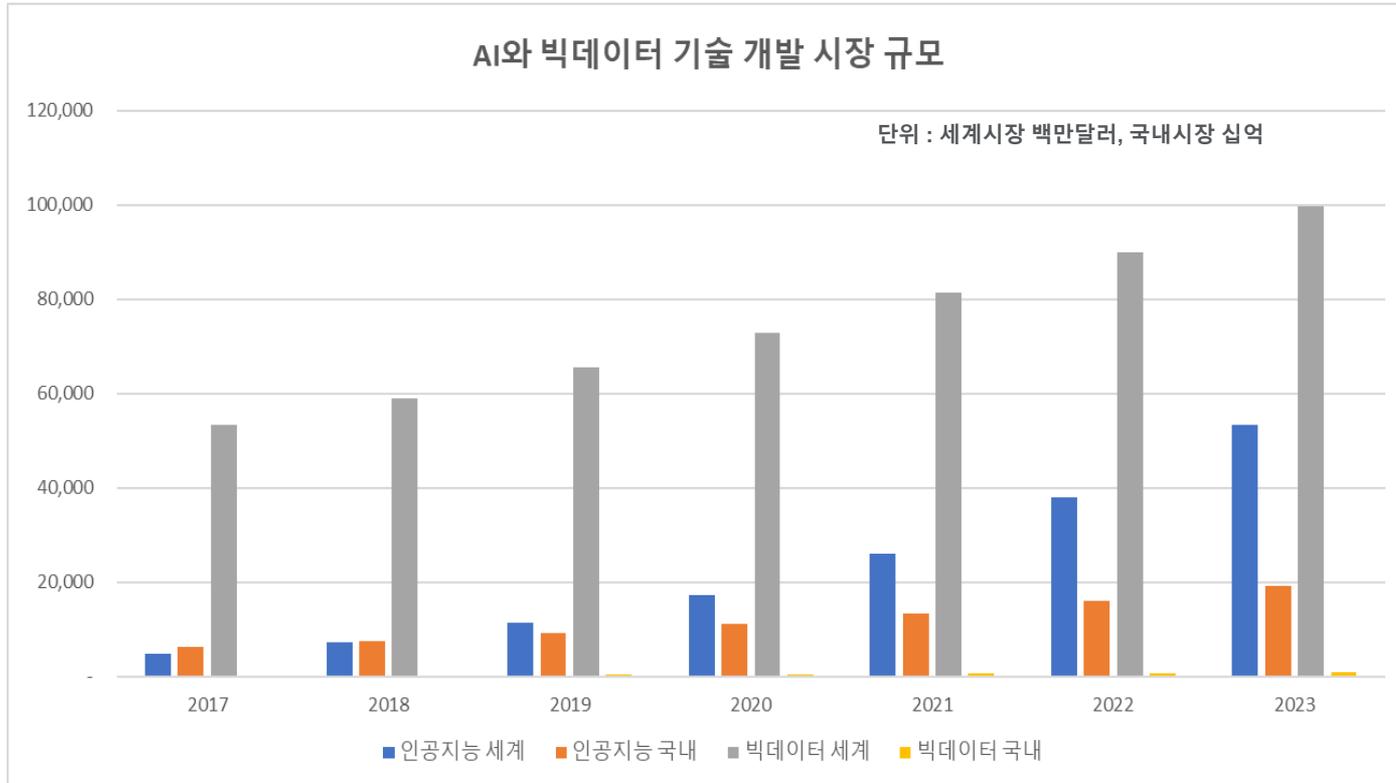
- 1) SharePlex 개요
- 2) SharePlex 특징
- 3) SharePlex 주요 활용 방안

## 3. SharePlex 사례



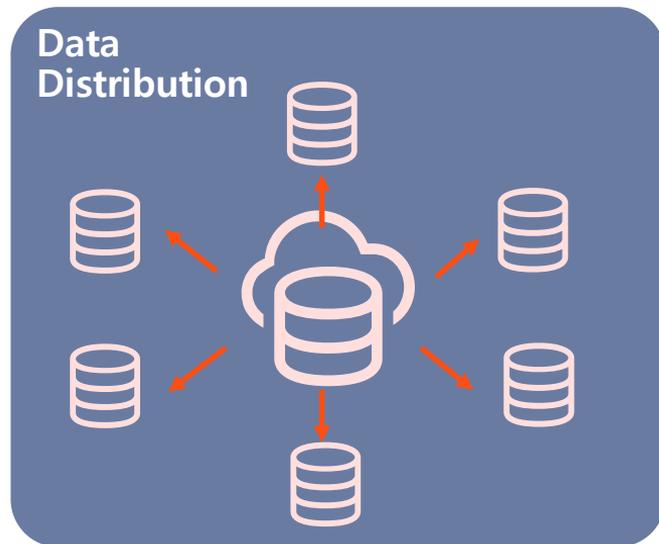
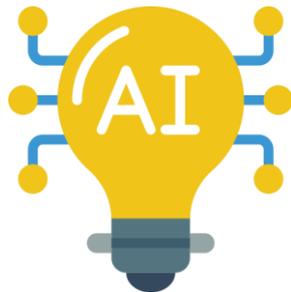
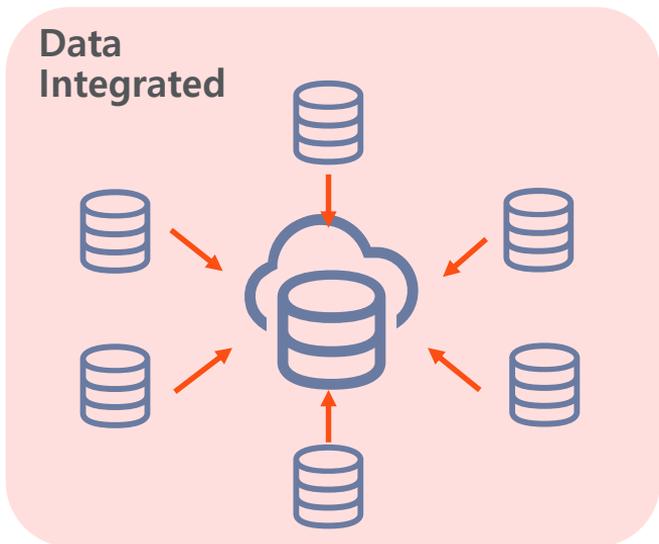
1

# 국내 AI 시장 상황과 미래



<IDS's Semianual Big Data and Analytics Software Tracker>  
 <IDC Korea Big data Technology and Services 2015-2019 Forecast>

- AI 서비스가 공공, 제조, 금융, 통신 등 여러 산업분야에서 활발하게 이용되면서 기반이 되는 데이터가 폭발적으로 증가하고, 이 모든 데이터가 저장되고 관리되는 곳이 데이터베이스
- 이 막대한 양의 데이터는 적절하게 실시간으로 통합하고 분산 시켜야 될 필요성이 있음





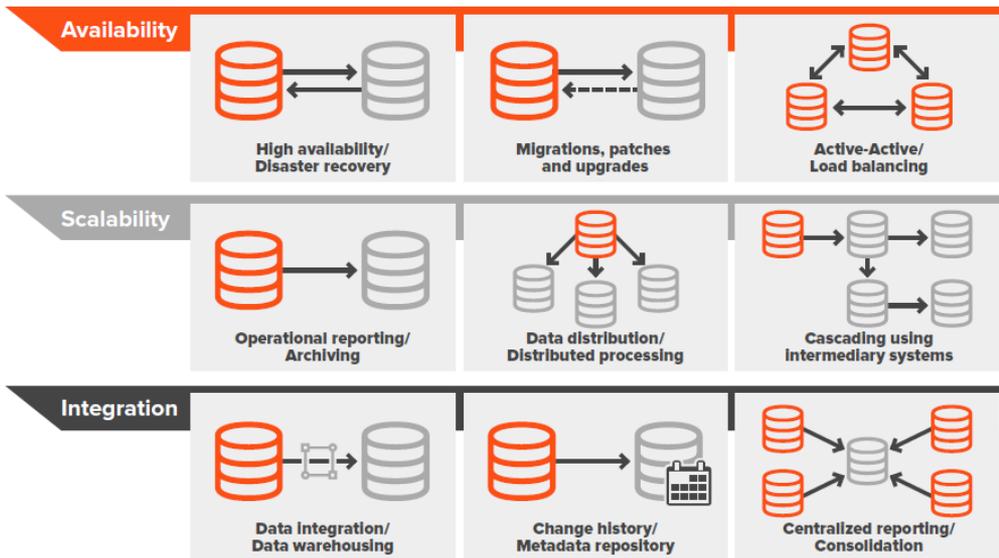
# SharePlex를 이용한 데이터 통합 및 분산 전략



## SharePlex를 이용한 데이터 통합 및 분산 전략

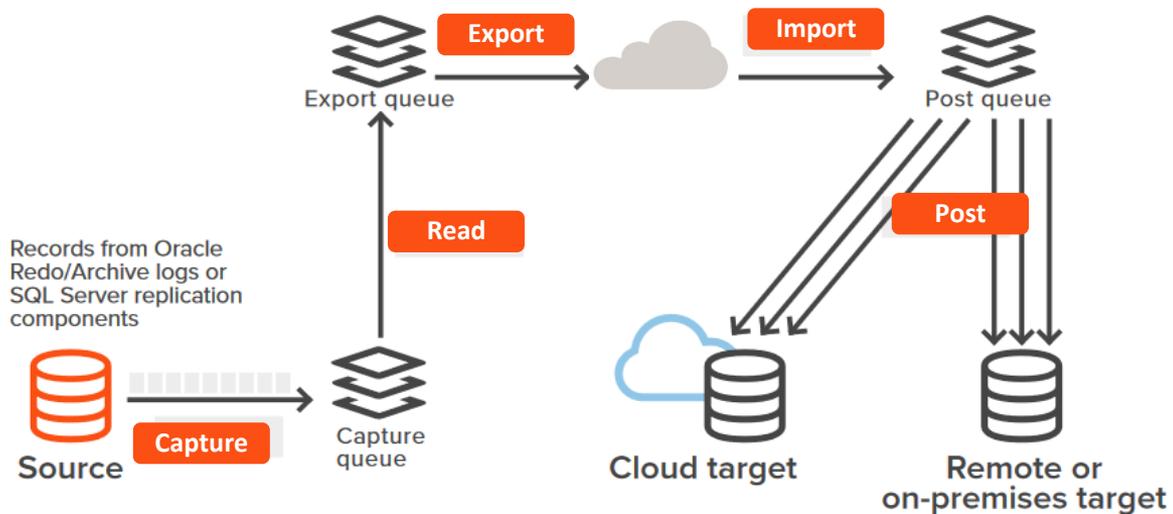
- 1) SharePlex 개요
- 2) SharePlex 특징
- 3) SharePlex 주요 활용 방안

- ✓ SharePlex는 원본 데이터베이스로 부터 트랜잭션을 추출하여 대상 데이터베이스로 복제하는 CDC 솔루션
- ✓ 다양한 고객의 니즈에 맞춰 HA, 마이그레이션, 로드밸런싱 등 유연한 구조로 동기화 구성이 가능
- ✓ 빠르고, 정확하고, 최소한의 비용으로 데이터를 복제하여 비즈니스 연속성과 생산성을 향상 시킴



5개의 Process 와 3개의 Queue 로 구성됩니다.

- **(Capture/ Read) Process** : Redo Log로 부터 변경 데이터 추출 및 라우팅 정보 매핑
- **(Export / Import) Process** : TCP/IP로 지정된 IP를 통해서 메시지를 전송
- **(Post) Process** : 최종 타겟 DBMS에 변경 데이터를 SQL로 변환하여 적용
- **Queue** : 데이터 처리/전송을 위한 임시 저장 영역



### 시스템 자원 사용을 최소화

- DBMS에 부하를 주는 Log Miner를 유틸리티를 사용하지 않고 Redo log에 Direct Access
- Queue mechanism을 사용하여 Disk IO 최소화
- Commit 전 데이터를 복제, 변경 즉시 Target으로 전송하여 Target System에서의 자원 사용을 최소화

구분	SharePlex	타사
복제 데이터 보관/전송	<ul style="list-style-type: none"><li>▪ Queue - 메모리(성능) + 디스크(안정성)</li></ul>	<ul style="list-style-type: none"><li>▪ 디스크 보관/전송</li></ul>
데이터 복제 처리 방식	<ul style="list-style-type: none"><li>▪ Commit 전 데이터 복제<ul style="list-style-type: none"><li>✓ 변경 즉시 전송/반영하여 소스/타겟 시스템 부하 최소화</li></ul></li></ul>	<ul style="list-style-type: none"><li>▪ Commit 후 데이터 복제<ul style="list-style-type: none"><li>✓ Commit 완료까지 소스시스템에 적재하여 Disk I/O, CPU 자원 사용 증가</li><li>✓ 타겟 시스템에도 동일 데이터 복제량을 보유</li><li>✓ 한꺼번에 전송함으로써 Network 부하 유발</li></ul></li></ul>

### 데이터 무결성 보장하는 기본 아키텍처

- 실시간으로 데이터 무결성을 검증하기 위해 Target에 반영하기 전, 복제 데이터의 정확성을 판단
- Before/After 이미지 관리를 통한 In-Flight(실시간) 데이터 무결성 검증 (parameter 통한 on/off 가능)

Primary Update OK			Secondary Update OK		
Id	Name	Address	Id	Name	Address
1	John	22 Second St	1	John	11 First St

```
UPDATE customer
  SET address = '22 Second St'
  WHERE id = 1;
COMMIT;
```

Write to Redo

```
UPDATE customer
  SET address = '22 Second St'
  WHERE id = 1
  AND address = '11 First St'
  AND rownum = 1;
COMMIT;
```

SQL to Post

<ROWID>:<OBJ\_ID>:<COL\_ID>:22 Second St:11 First St:1

### UPDATE OOS Example

Primary Update OK			Secondary Out-Of-Sync		
Id	Name	Address	Id	Name	Address
1	John	22 Second St	1	John	33 Third St

```
UPDATE customer
  SET address = '22 Second St'
  WHERE id = 1;
COMMIT;
```

Write to Redo

```
UPDATE customer
  SET address = '22 Second St'
  WHERE id = 1
  AND address = '11 First St'
  AND rownum = 1;
COMMIT;
```

SQL to Post

Record Not Found

<ROWID>:<OBJ\_ID>:<COL\_ID>:22 Second St:11 First St:1

### Compare/Repair

All-in-One (1)

- 다양한 이유로 정합성이 맞지 않게 될 경우 'event log' 파일에 해당 내용이 logging
- Online으로 SharePlex에 내장 된 Compare/Repair Utility를 통하여 정합성 검증 및 보정 기능 제공

#### Insert OOS Example in prod\_errlog.sql

```
-- [3] Tue Oct 9 14:26:38 2011
-- redolog seq#/offset 81319/159081124
-- redolog timestamp 635305391 (10/07/11)
-- original rowid AAaKivAGQAAAEJABo
-- ORA-00001: unique constraint
(OSI_SIEBEL2.S_ORDER_ITEM_XA_F1) violate
-- CONSTRAINT VIOLATION
insert into "OSI_SIEBEL2"."S_ORDER_ITEM_XA_F1"
("ROW_ID", "CREATED", "CREATED_BY", "LAST_UPDATE_DATE",
"MODIFICATION_NUM", "CONFLICT_ID", "ATTR_NUM", "ORDER_ITEM_ID",
"READ_ONLY_FLG", "REQUIRE", "ATTR_ID", "CHAR_VAL", "DATA_TYPE_CD", "DISPATCH_CD",
"VLDTN_LOV_TYPE_CD")
values ('1-43RHNT8', '10072011014311', '1-10072007014311', '1-T8E-750', '0', '0', 'La', 'N', '1-43RHNT7', 'N', 'Y', 'Update', '1-08DV', 'Thai', 'Text', 'Language', 'AIS_VMS_LANGUA
```

#### ✓ Compare 결과 확인

```
Host: SPSRC
ID      Tablename      Since      Repair
-----
6208   "SPXUSER"."TB_CUST"  03-Mar-11 11:28:30 Off

Source table : "SPXUSER"."TB_CUST"
Source node  : SPSRC
Target table : "SPXUSER"."TB_CUST"
Target route : SPTGT@ORCL9
Where clause :
Select hint  :
Status       : Out Sync   데이터 비교 결과 18,000 Row 가 다름
              ( Target 에 INSERT 18,000 row 가 필요함 )

* Compare Results
* 18000 source and 0 target rows compared successfully (L387).
* 18000 out-of-sync row(s) found in this table.
* Inserts : 18000
* Updates : 0
* Deletes : 0
*/
```

#### ✓ Compare 결과 Repair 를 위한 SQL 문 확인

```
/*
* Compare Report
* Session ID : 6208
* Schema : SPXUSER
* Table : TB_CUST
* Repair : Off
* Target Route: SPTGT@ORCL9
* Key Compare : Off
* Select Hint :
* Log File : /sp1ex75/var/log/ORCL9_SPXUSER-TB_CUST-6208-20353.sql
* Date : Thu Mar 3 11:28:35 2011
*/

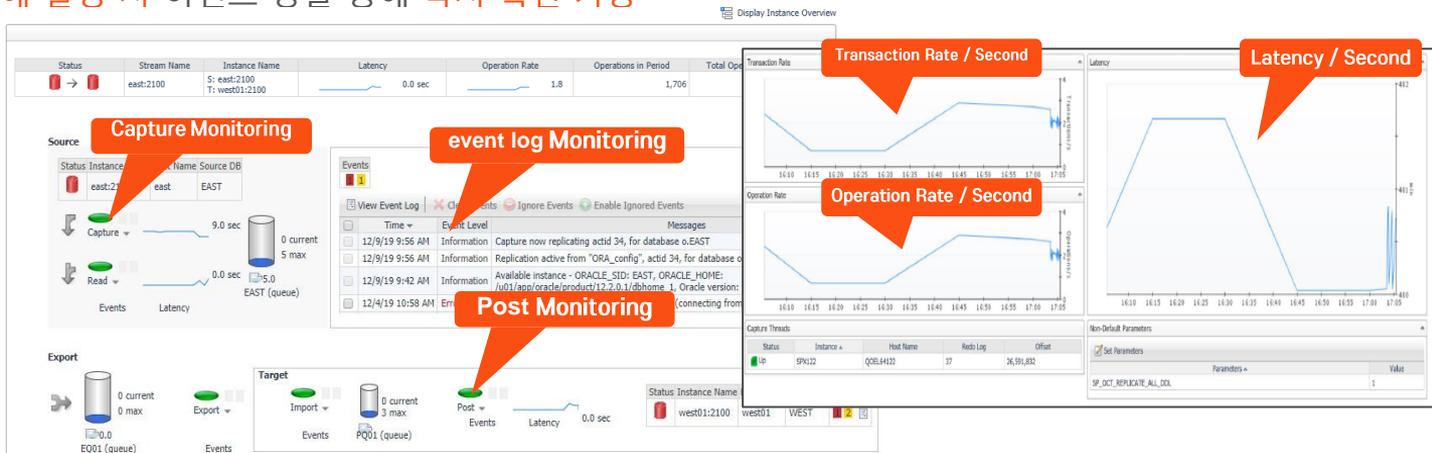
데이터 비교 결과 INSERT 18,000 Row SQL문

/* source rowid=AAAH0VAALAAAAPnAAA */
insert into "SPXUSER"."TB_CUST" ("CUST_NO","NAME","ADDRESS","IN_DATE","IN_TIME"
) values ('3721','Cust Name_1','Cust Address_1', '20110215', '135616 ');
/* source rowid=AAAH0VAALAAAAPnAAB */
insert into "SPXUSER"."TB_CUST" ("CUST_NO","NAME","ADDRESS","IN_DATE","IN_TIME"
) values ('3723','Cust Name_1','Cust Address_1', '20110215', '135616 ');
```

### Monitoring

All-in-One (2)

- 대시보드 화면을 통해 프로세스 상태, 처리 건수, 지연 시간 등을 쉽게 확인
- 장애 발생 시 이벤트 창을 통해 즉시 확인 가능



### 데이터 암호화 지원

- 보안에 대한 이슈가 커지고 있는 만큼 필요한 경우 네트워크 전송구간의 데이터를 암호화 처리 가능
- sp\_security 유틸리티를 사용하여 SharePlex가 사용하는 네트워크에 대한 SSL/TLS 설정을 활성화



- ✓ 다양한 복제 단위 - **Schema, Table, Column, Row** 단위의 복제 대상을 지정합니다.
- ✓ **Schema** 단위 복제 - 기존 테이블 단위의 복제뿐만 아니라, Schema단위의 복제가 가능하며 새로운 Table 을 생성하거나 변경, 삭제 할 경우 자동으로 복제 대상을 업데이트합니다.
- ✓ **DDL(Data Definition Language)**복제의 경우, **Parameter 옵션**으로 복제 대상을 지정합니다.

Object	DDL Operation	SharePlex Parameter
Table and IOT	Create, Alter, Drop, Truncate, Comment, CTAS	SP_OCT_REPLICATE_DDL=3 SP_OCT_AUTOADD_ENABLE=1
Sequence	Create, Alter, Drop	SP_OCT_AUTOADD_SEQ=1 SP_OCT_AUTOADD_ENABLE=1 SP_SYS_TARGET_COMPATIBILITY=8.6.2 or higher
Partition/Sub-partition	Add, Split, Merge, Drop, Modify, Coalesce, Exchange, Move, Truncate, Rename, Set	SP_OCT_REPLICATE_DDL=3 SP_OCT_AUTOADD_ENABLE=1
Index	Create, Alter, Drop	SP_OCT_REPLICATE_DDL=3 SP_OCT_AUTOADD_ENABLE=1
View	Create, Alter, Drop, Comment	SP_OCT_REPLICATE_ALL_DDL=1
Stored procedure	Create, Alter, Drop	SP_OCT_REPLICATE_ALL_DDL=1

### ✓SharePlex Process fail

- 사용자 실수에 의한 Process kill 이나 강제 종료와 같은 장애 발생시,  
자동으로 **SharePlex 프로세스가 재 시작**되어 정상적인 복제 시작

### ✓DB fail

- **Source DB fail**: Redo Log로부터 추출한 데이터까지의 Offset 정보 관리를 통한 **데이터 연속성 자동 유지**
- **Target DB fail**: DB에 적용된 데이터에 대한 Meta 정보를 확인하여 **장애시점부터의 연속적 복제 Data 반영**

### ✓Network fail

- Network 정상 여부 자동 체크 및 복구를 통한 **복제 연속성 유지**
- Network 장애에 대비하여, Queue가 보관되는 **Disk 공간에 대한 적절한 산정 필요**

DBMS	Version	Platform (only 64-bit)
Oracle (Source)	11gR2, 12cR1, 12cR2, 18c, 19c(for Linux 7 later, and not supported HP-UX) EE & SE2	AIX 7.1, 7.2 CentOS Linux 5.x, 6.x, 7.x, 8.x Oracle Linux 5.x, 6.x, 7.x, 8.x RHEL 5.x, 6.x, 7.x, 8.x SuSE SLES 11.x, 12.x, 15.x Oracle Solaris(SPARC and x86) 10, 11, 11.4 Windows Server 2008, 2008 R2, 2012, 2012 R2, 2016, 2019 HP-UX 11.31 Itanium IaaS Azure and AWS
MS SQL Server (Target)	SQL Server 2008, 2008 R2, 2012, 2014, 2016, 2017, and 2019	Windows Server 2008, 2008 R2, 2012, 2012 R2, 2016, and 2019
Kafka (Target)	0.8.x, 0.9.x, 0.10.x, 0.11.x, 1.0.x, 1.1.x, 2.0.x, 2.1.0, and 2.5.0	-
MySQL (Target)	5.5, 5.6, 5.7, and 8	RHEL 7.x IaaS Azure and AWS
PostgreSQL (Target)	PostgreSQL 9.4, 9.5, 9.6, 10, 13.4, and 14.0	RHEL 7.x, 8.x IaaS Azure and AWS

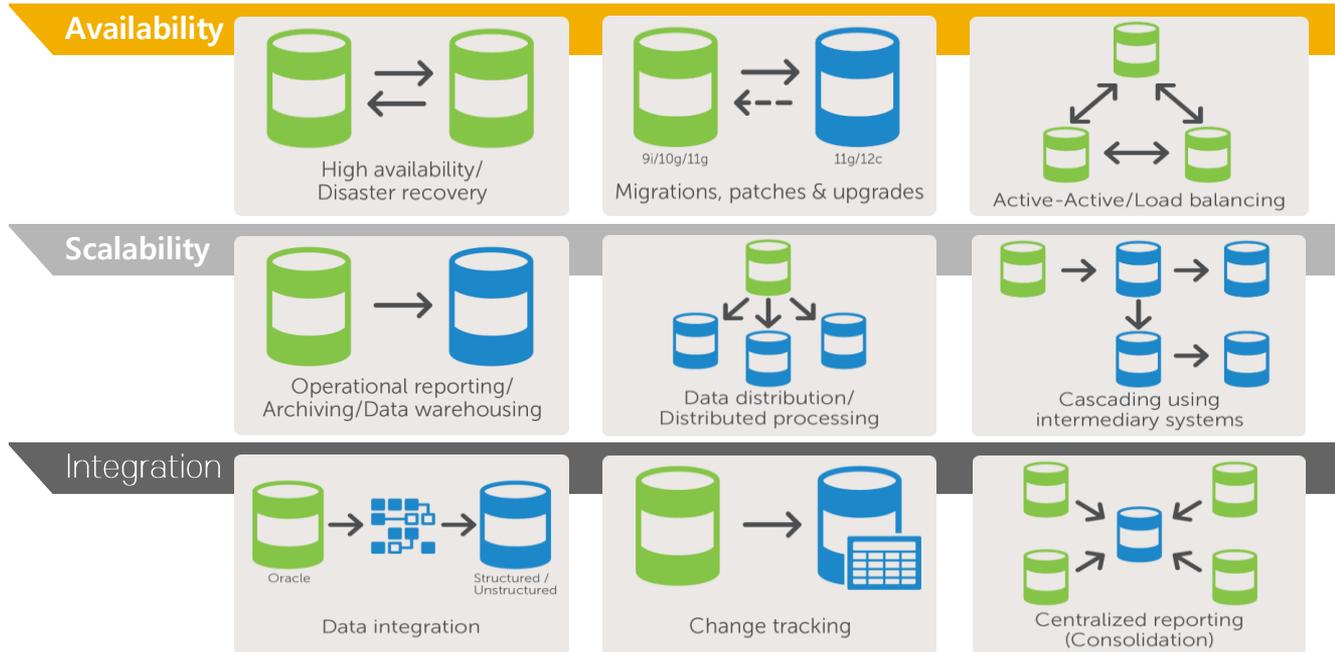
The background features a complex network of interconnected nodes and lines. The nodes are represented by small spheres, some in a vibrant orange and others in a muted blue. The lines connecting them are thin and light-colored, creating a web-like structure against a dark, almost black background. The overall aesthetic is modern and technological.

# 쉐어플렉스 주요 활용방안

# SharePlex - 다양한 복제 구성

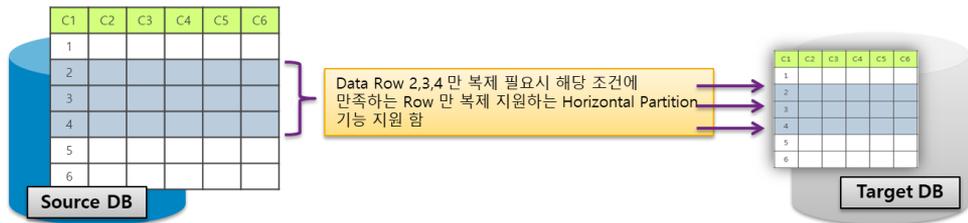
## 2. SharePlex 주요 활용방안

1:N N:1 N:M 등의 유연한 복제 구조를 가지고 있어 ODS, 마이그레이션, 리포팅, HA/DR 등 고객의 모든 요구사항을 만족할 수 있는 다양한 복제 구성을 제공합니다.



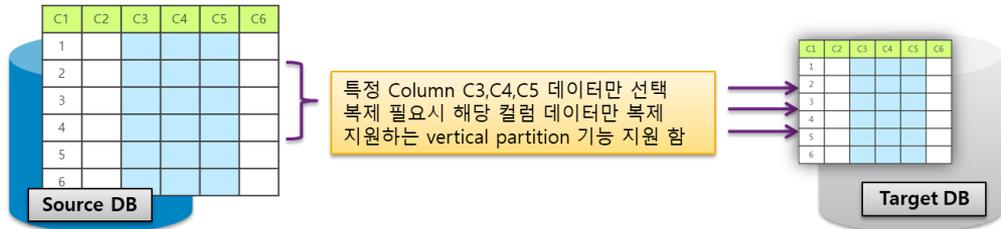
### ✓ Row 단위의 복제 (Horizontally partitioned replication)

- 테이블의 복제 대상이 되는 조건(where)을 추가하여 조건에 해당하는 Row (C1=2,3,4)만 복제



### ✓ Column 단위의 복제 (Vertically Partitioned replication)

- 특정 Column (C3,C4,C5) 의 데이터만 복제하는 vertical partition 기능 지원

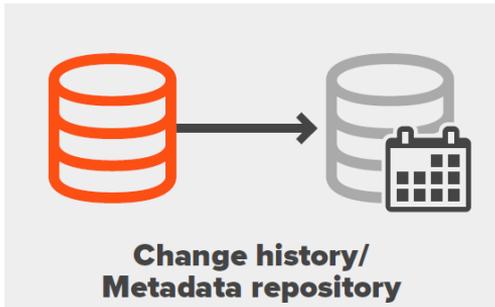


✓ SharePlex 를 통해 Migration 작업을 수행 시 Downtime을 최소화 할 수 있습니다.

### ✓ Migration Steps

1. 기존 운영 DB에서 변경 분 Capture 시작
2. 신규 운영 DB에 데이터를 초기 적재 (RMAN, Exp/Imp, Etc)
3. 초기 적재 이후 변경 데이터를 동기화 (SharePlex)
4. 신규 운영 DB로 서비스 전환 (Network, Application)
5. 신규 운영 DB 에서 서비스 모니터링





- **Change History**
  - 변경이력 추적 및 감사 목적의 데이터 수집 (For auditing and discovery purposes)
  - 공.금융기관에서 활용 중 (국내/해외 포함)
- **Source DBMS 의 Metadata**
  - user id, time, operation type, scn, row id, transaction id 등
- **Operations supported**
  - INSERT/UPDATE/DELETE/TRUNCATE/DROP COLUMN

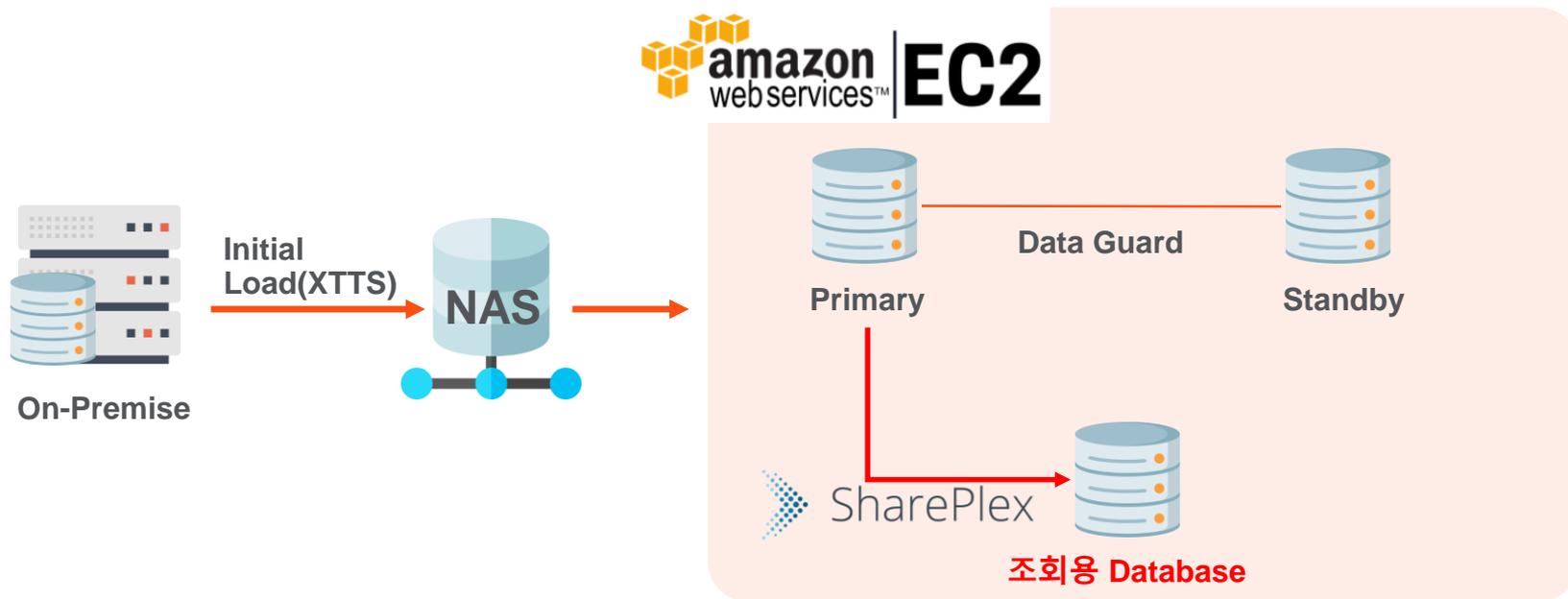
```
SHAREPLEX_SOURCE_TIME  
SHAREPLEX_SOURCE_USERID  
SHAREPLEX_SOURCE_OPERATION  
SHAREPLEX_SOURCE_SCN  
SHAREPLEX_SOURCE_ROWID  
SHAREPLEX_SOURCE_TRANS  
SHAREPLEX_OPERATION_SEQ  
SHAREPLEX_SOURCE_HOST  
SHAREPLEX_QUEUE_NAME  
SHAREPLEX_SOURCE_ID
```



## SharePlex 사례 소개

- 도입 배경 및 효과

운영 서버의 조회성 업무를 조회 DB로 분산함으로써 본 운영 서버의 안정성과 성능 확보



### • 도입 배경 및 효과

- 실시간 DW 구성, 기간계 데이터를 실시간으로 복제하여 정보계 EDW 고도화
- 정보계의 실시간 마케팅 업무 용도로 활용 → 실시간 정보 제공으로 인한 분석 능력 강화



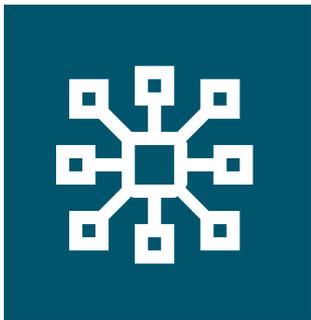
### • 도입 배경 및 효과

- 조회성 업무를 리포팅 및 아카이빙 서버로 분산 → 본 ERP 서버의 안정적 운영
- 월 단위 매출/결산 정보의 실시간 처리 및 분석 보고서를 리포팅 서버에서 제공
- Source Database 11 : Target Database 1



### 데이터 분산(Distribution)

- ✓ Data Distribution
- ✓ Data cascading
- ✓ Data Reporting & Archiving



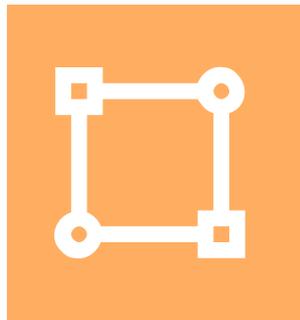
### 데이터 고가용성 (High Availability)

- ✓ Directional
- ✓ Bi-Directional



### 데이터 통합(Integrated)

- ✓ 정보계
- ✓ Data Warehouse 구축
- ✓ Data lake



### 데이터 마이그레이션(Migration)

- ✓ On-Premise to On-Premise
- ✓ On-Premise to Cloud
- ✓ Cloud to Cloud
- ✓ Unix to Linux ...





Thank You