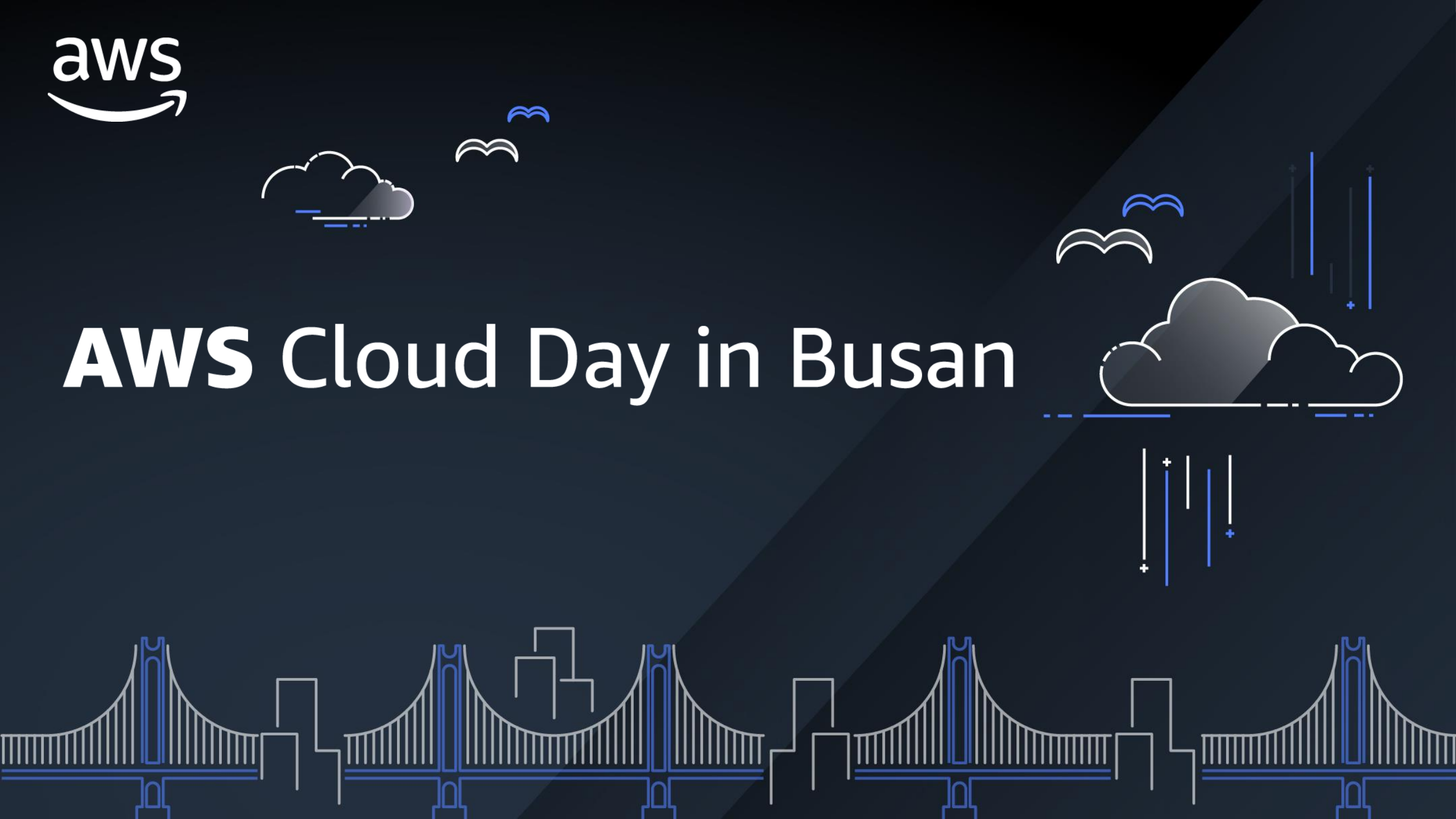
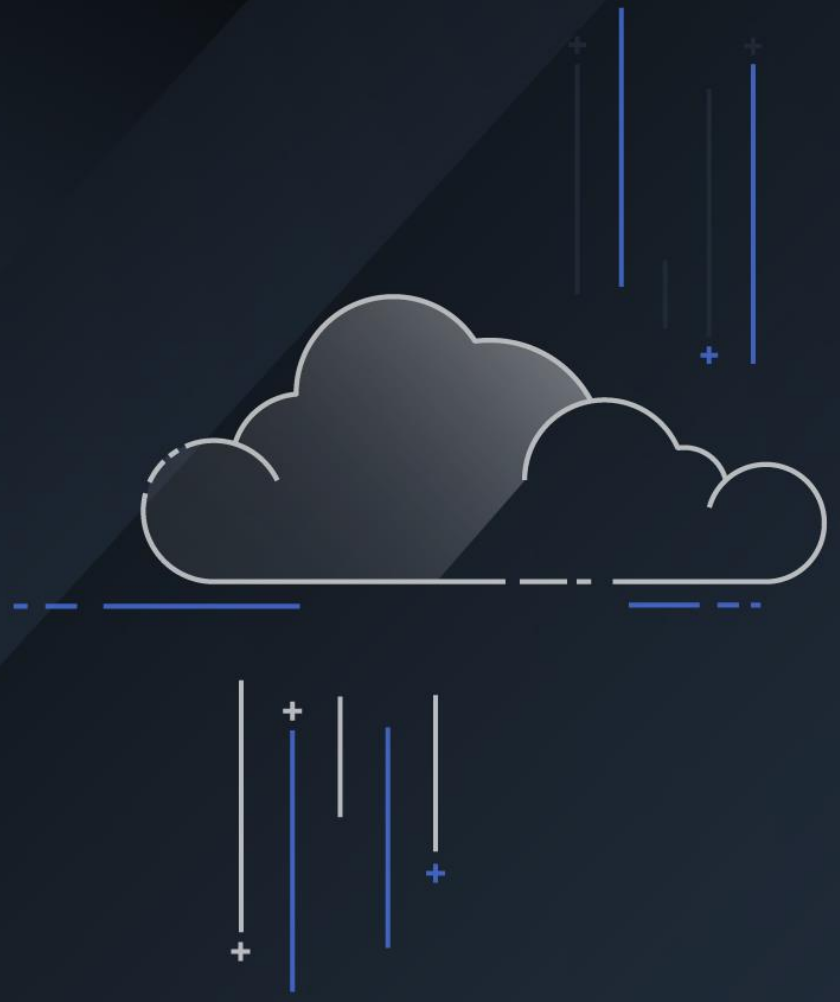

aws
AWS Cloud Day in Busan

AWS Cloud Day in Busan

Cloud 기반의 Database (데이터베이스의 자유를)



강민석

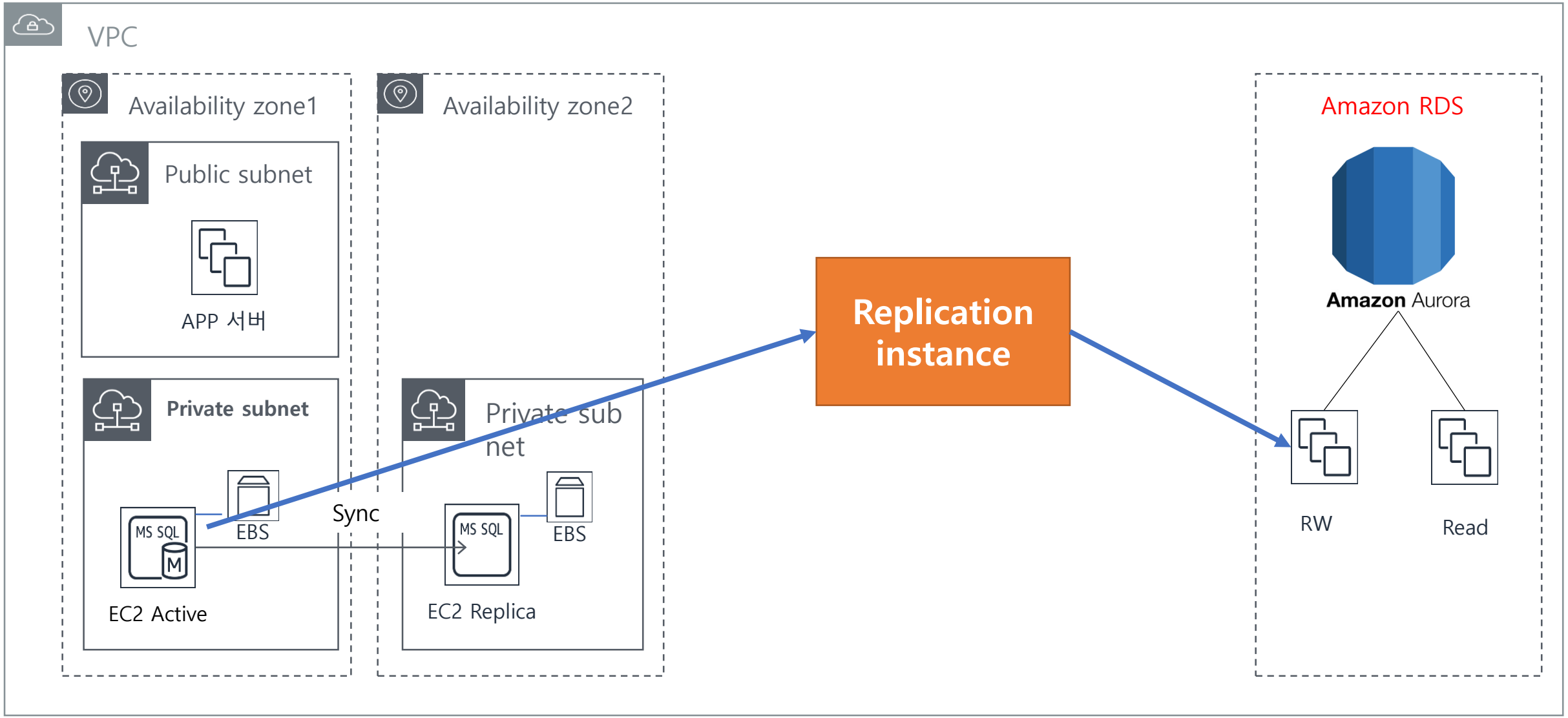
Database Specialist SA

AWS

본 강연에서 다룰 내용

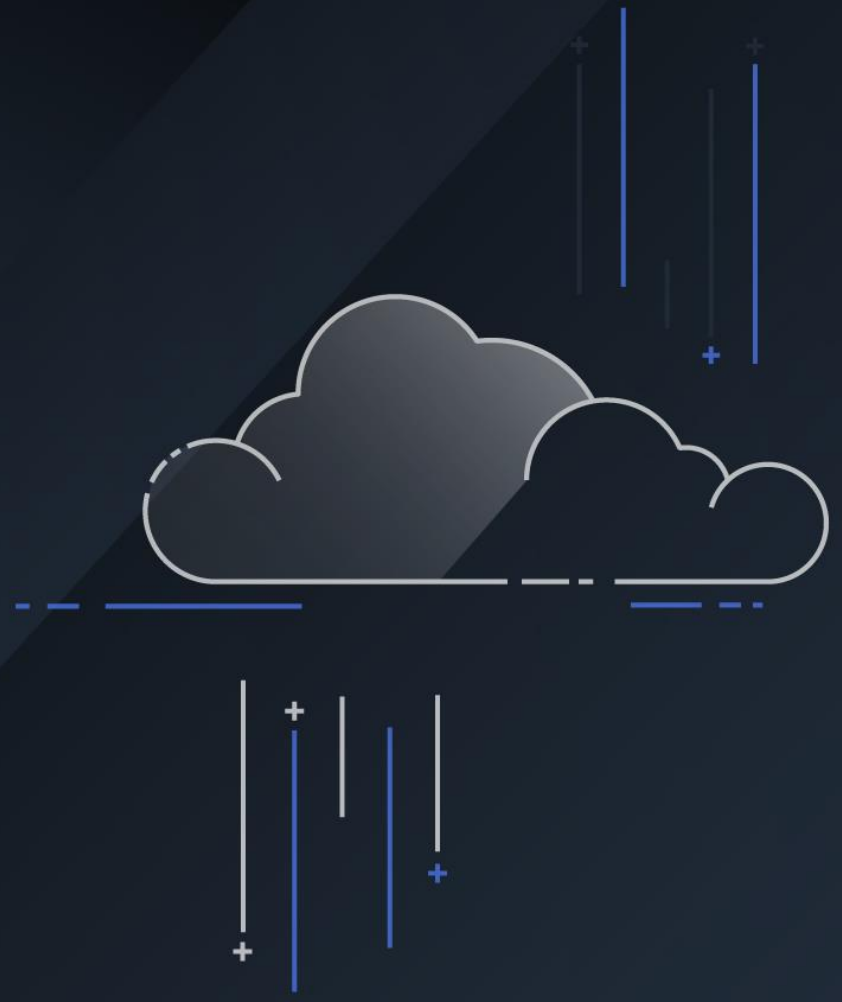
- Data Platform
- AWS가 제공하는 관계형 데이터베이스
- AWS로의 데이터베이스 이전
- Database Migration Service(DMS)

Demo – SCT & DMS



AWS Cloud Day in Busan

Data Platform



고객이 DBMS에게 바라는 것?

죽지마~
✓Zero data loss
✓다운타임 최소화

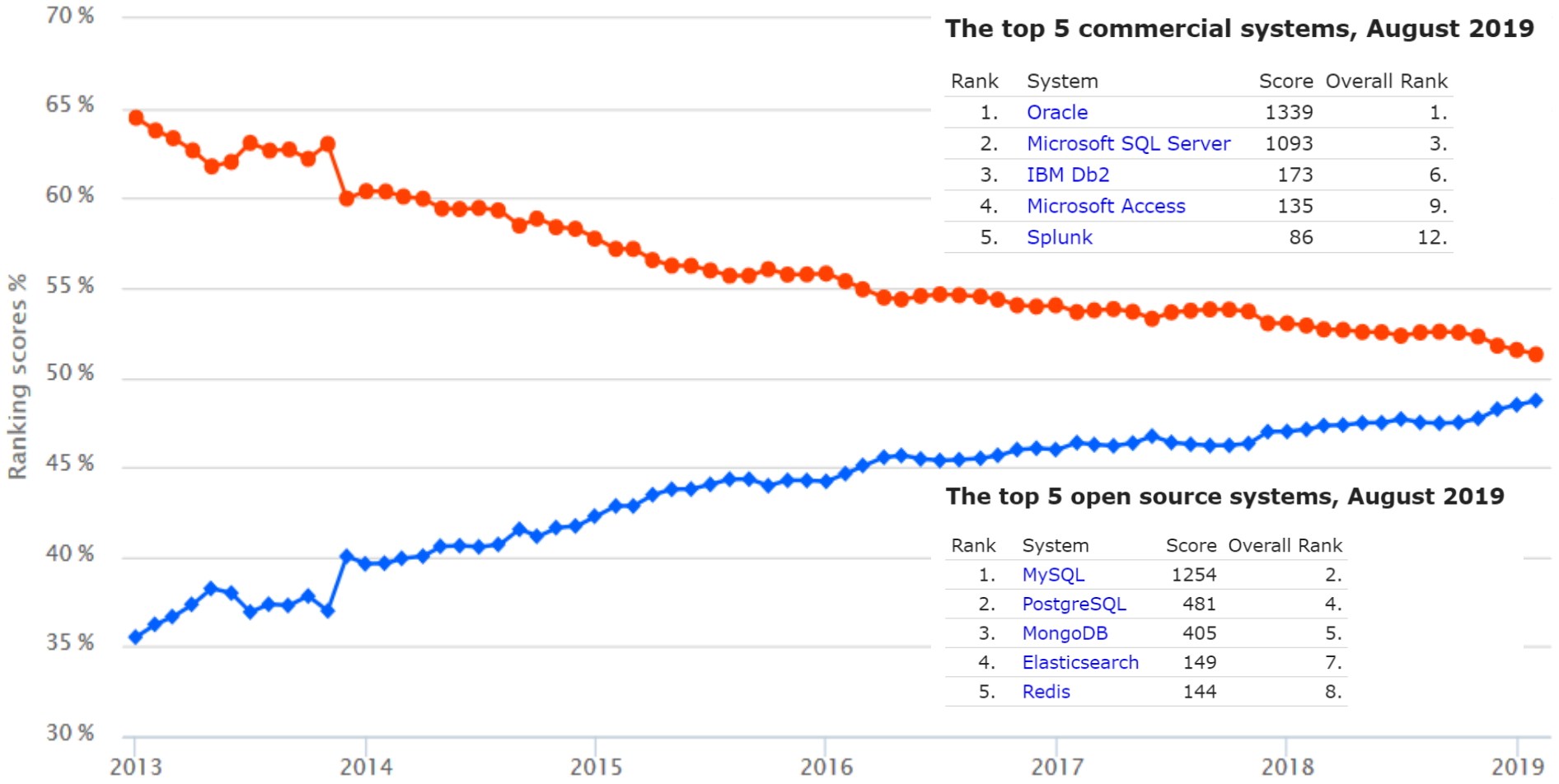
빨리빨리~
✓서비스 지연 없어야....
✓배치가 빨리 끝나서 집에 가고 싶은데..

DB가 너무 커~
✓저장할 스토리지가 부족한데...
✓Data를 효율적으로 관리하고 싶은데...
✓리포트를 자유자재로 보고 싶은데...

데이터 유출 안돼~
✓암호화 해야 하는데...
✓누가 무슨 자료를 조회하는지 기록하고 싶은데...

쉽게 DB 관리할 방법 없나?
✓느려졌는데 이유를 모르겠네
✓Resource를 효율적으로 쓰고 싶은데....

상용 vs 오픈소스



The top 5 commercial systems, August 2019

Rank	System	Score	Overall Rank
1.	Oracle	1339	1.
2.	Microsoft SQL Server	1093	3.
3.	IBM Db2	173	6.
4.	Microsoft Access	135	9.
5.	Splunk	86	12.

The top 5 open source systems, August 2019

Rank	System	Score	Overall Rank
1.	MySQL	1254	2.
2.	PostgreSQL	481	4.
3.	MongoDB	405	5.
4.	Elasticsearch	149	7.
5.	Redis	144	8.

Microsoft SQL Server

ORACLE

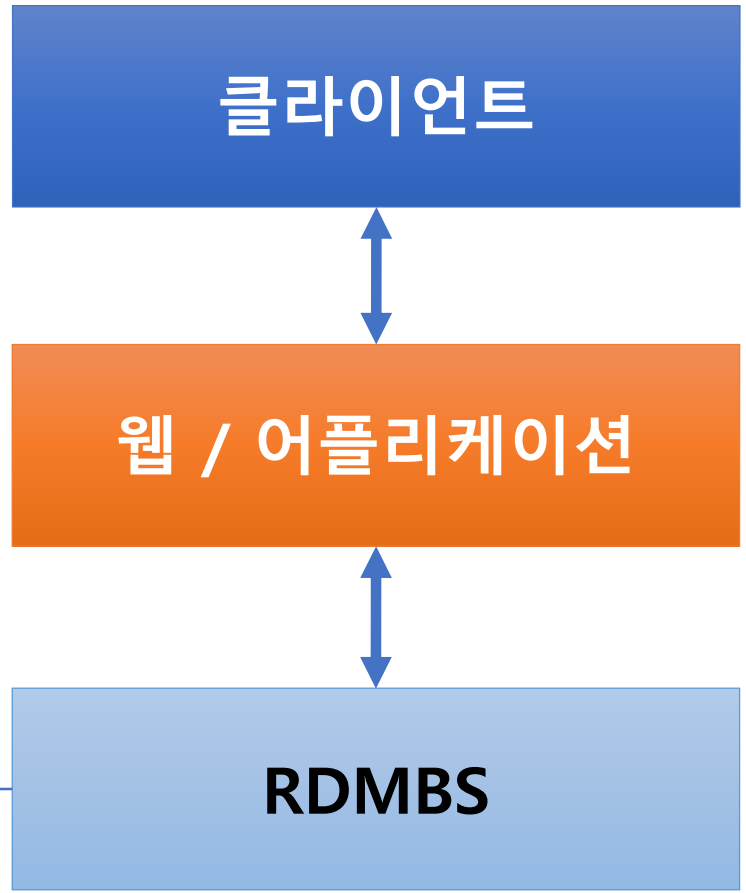
Amazon Aurora

MySQL

PostgreSQL

https://db-engines.com/en/ranking_osvsc

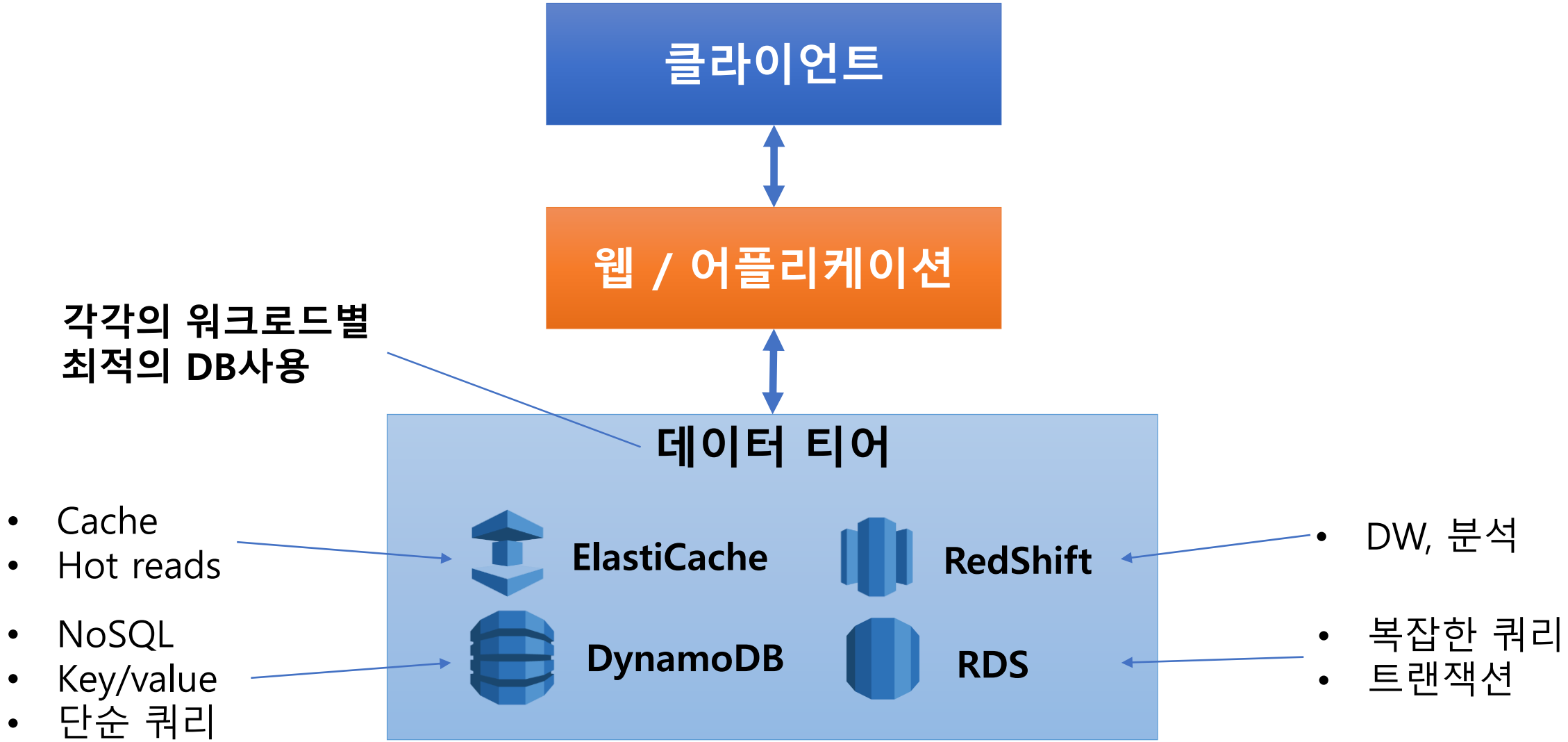
전통적인 RDBMS 아키텍쳐



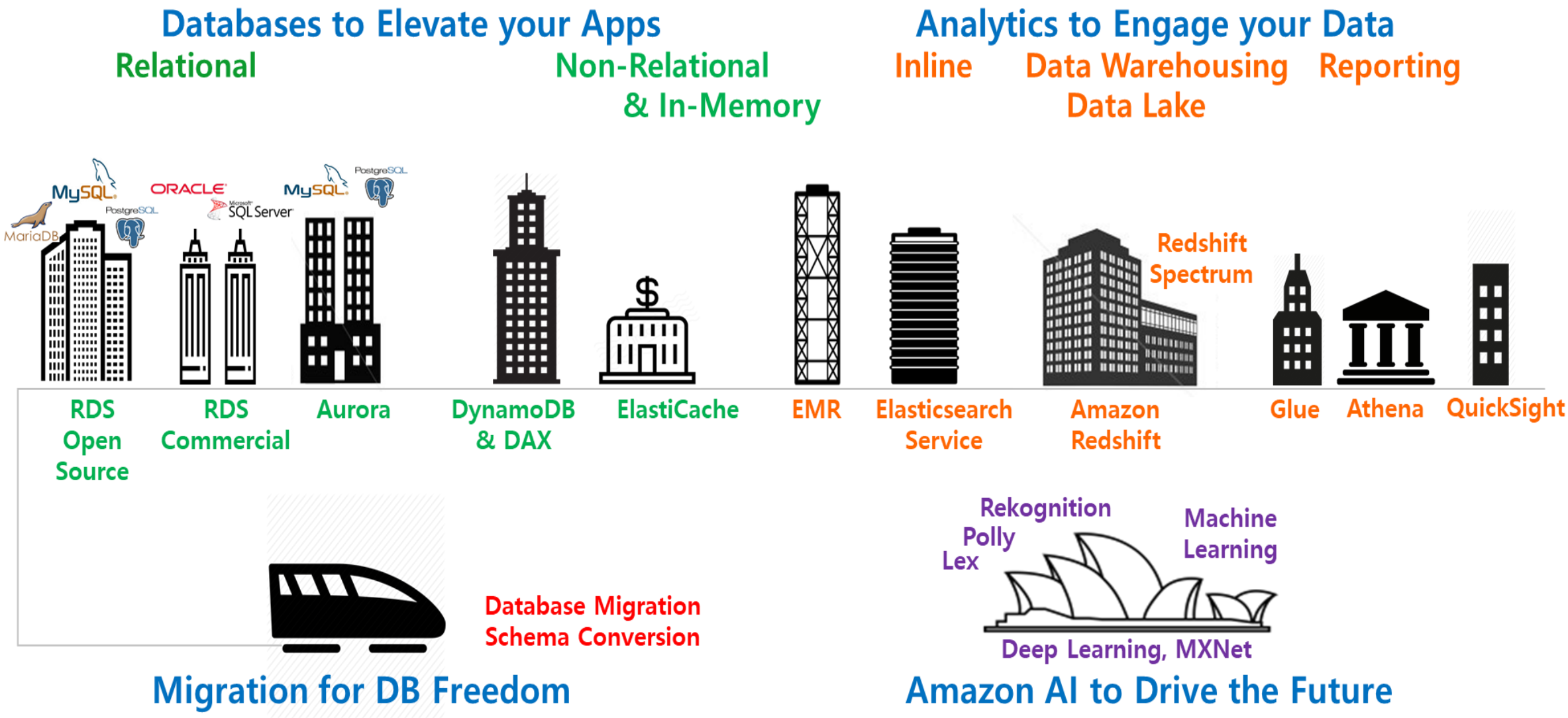
모든 워크로드를
하나의 DB에서 처리

- Key-Value 접근
- 복잡한 쿼리
- 일상적인 트랜잭션
- 데이터 분석

클라우드 환경에서 데이터 티어 아키텍쳐



AWS의 다양한 데이터 플랫폼 선택

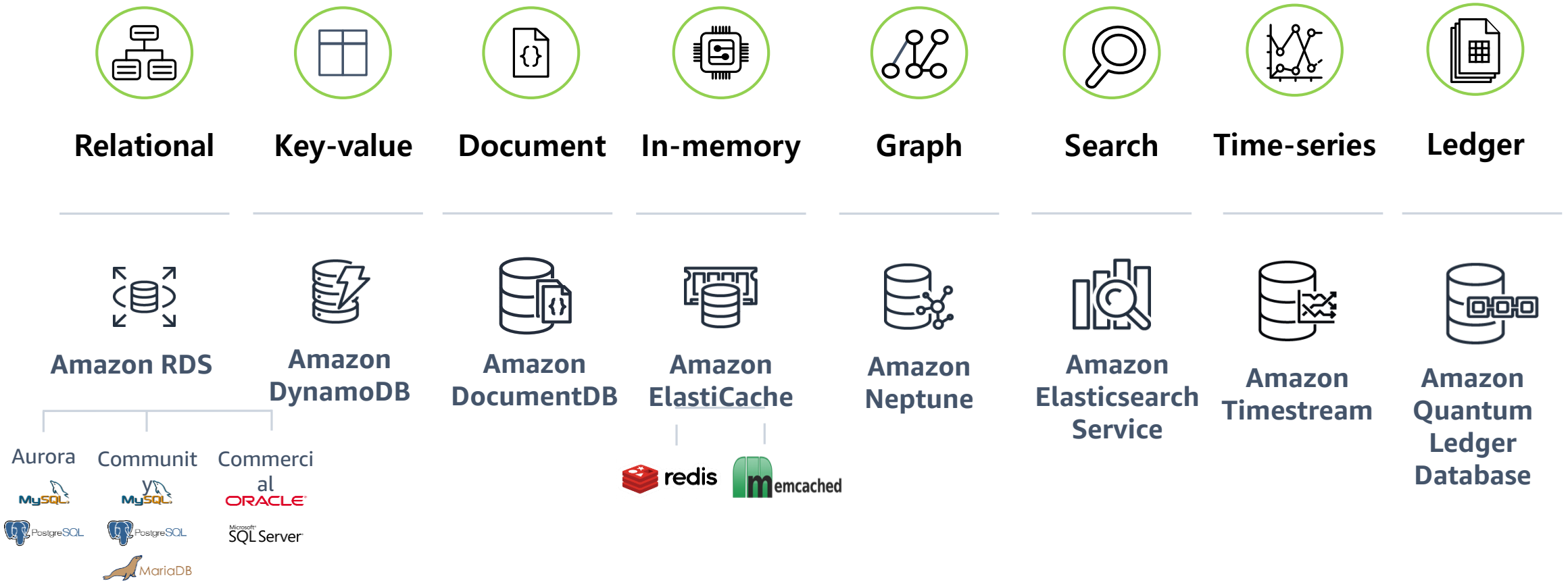


AWS Cloud Day in Busan

AWS가 제공하는 관계형 데이터베이스

AWS 데이터베이스 서비스

- 모든 애플리케이션 요구에 맞춰 특별히 구축된 데이터베이스



Amazon RDS

Amazon RDS

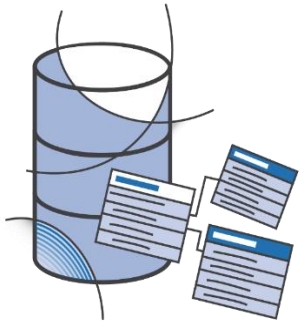
- Amazon RDS는 관리형 DB서비스
- Amazon RDS의 특징
  - 손쉬운 구성
  - 손쉬운 확장
  - 높은 안정성
  - 비용 효율성 제공
- Amazon RDS 종류

Amazon RDS 주요 기능

- 다양한 DBMS엔진 지원
  - Aurora(Mysql, PostgreSQL), MySQL, PostgreSQL, Oracle, MS-SQL, MariaDB
- 빠른 데이터베이스 시스템 구축
- 쉬운 DB 용량 변경 및 DB 패치
- 쉬운 복제본 추가
- 백업 및 로그를 99.999999999% 내구성으로 보관
- 자동화된 백업을 통한 PITR(Point in Time Recovery) 가능
- 모니터링을 위한 1초 단위의 시스템 메트릭 제공
- 데이터 저장소 암호화

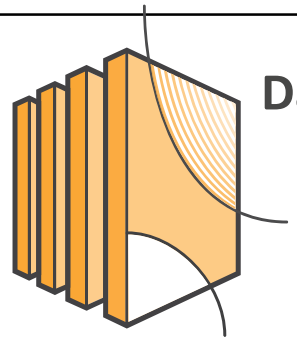
Amazon RDS

DB on EC2 vs. Amazon RDS 선택 기준



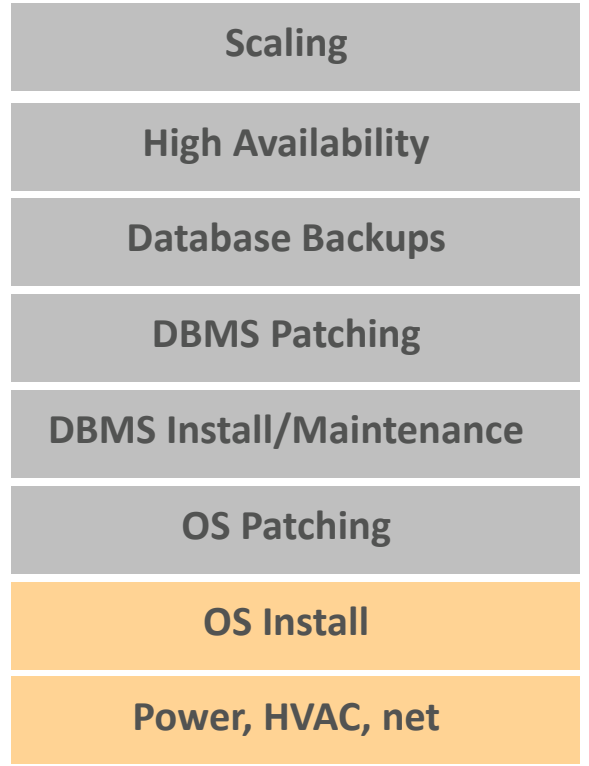
Amazon RDS

- Consider RDS first
- Focus on business value tasks
- High-level tuning asks
- Schema optimization
- No in-house database expertise



Database on Amazon EC2

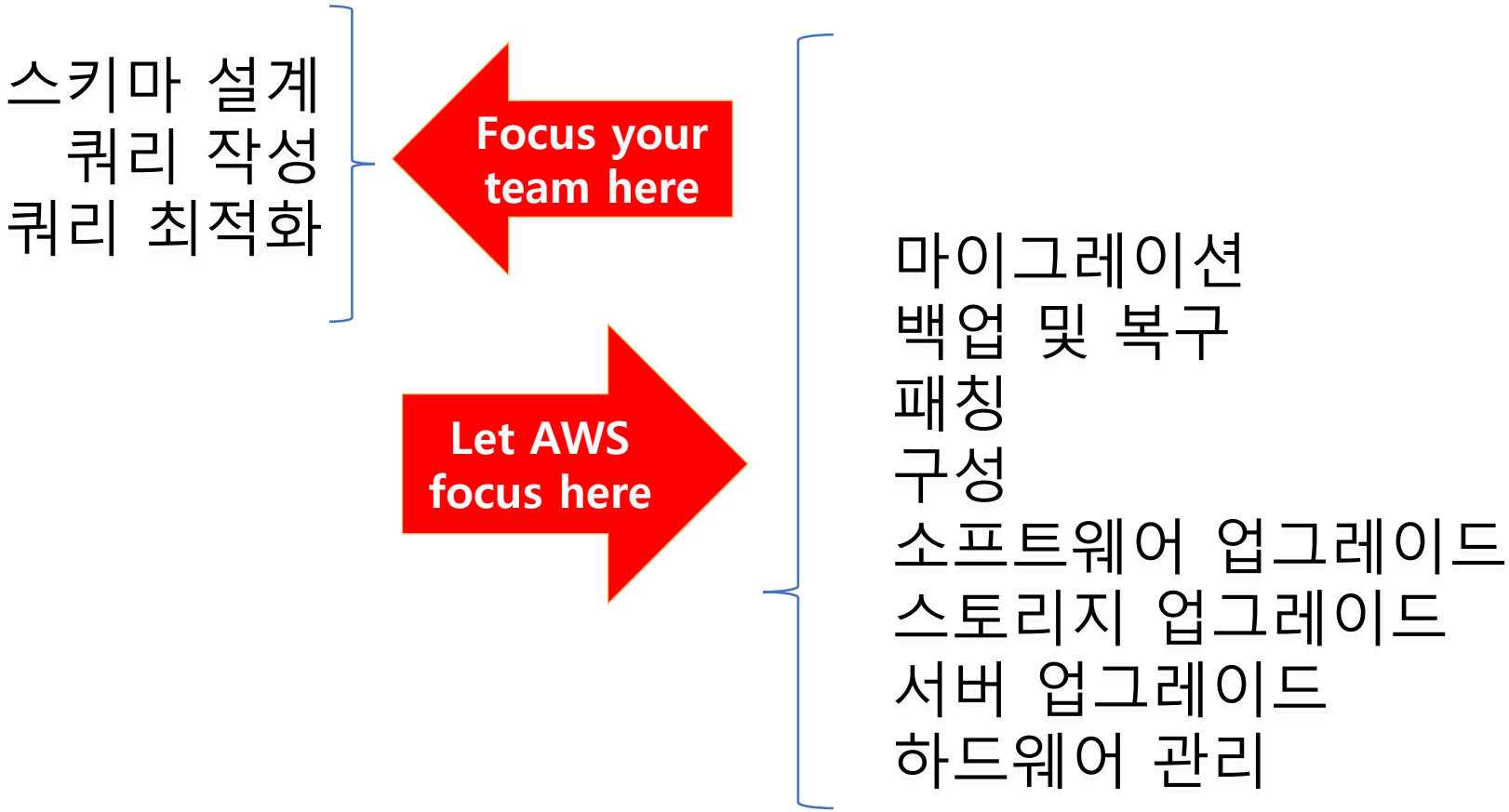
- Need full control over DB instance
- Backups
- Replication
- Clustering
- Options that are not available in RDS



AWS managed

Customer managed

Amazon RDS 서비스 사용 목적



Amazon Aurora

Amazon Aurora?

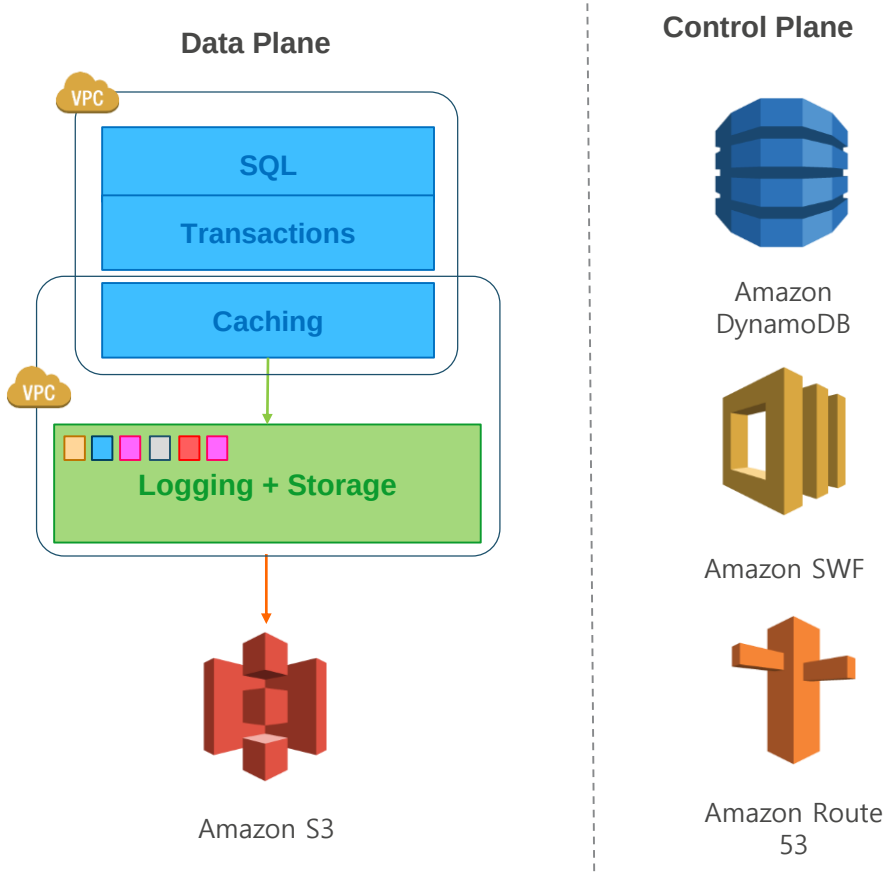
- 오픈 소스 호환 관계형 데이터베이스
  - MySQL, PostgreSQL
- 상용 데이터베이스의 성능과 가용성 제공
- 오픈소스 데이터베이스의 비용 효율성과 간단함

기존의 관계형 DBMS를 재구성

1. SQL+TRANX를 스토리지와 분리:
   로깅 및 스토리지를 스케일-아웃 가능한 스토리지 서비스로 전환

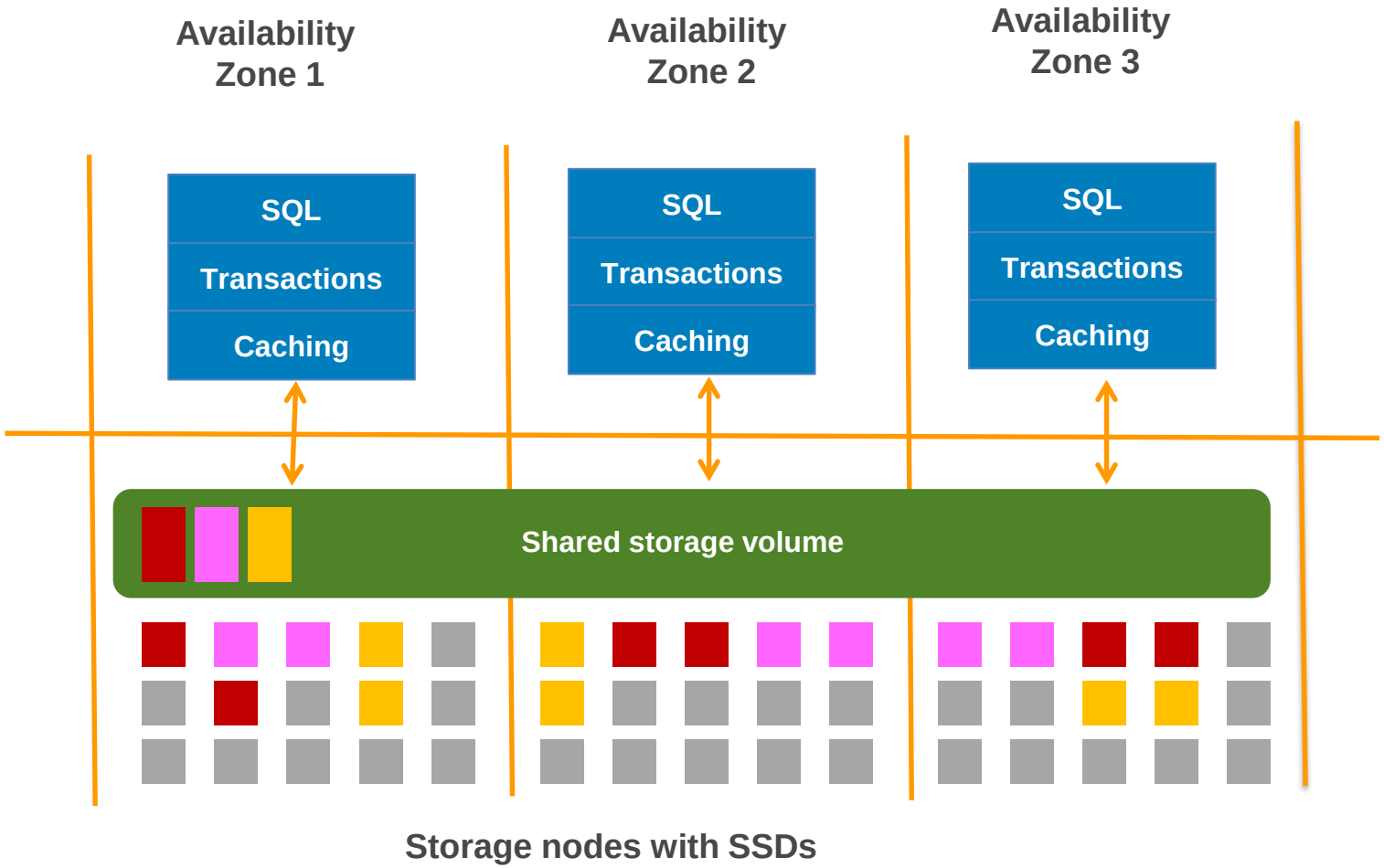
2. 서비스 내부에 Amazon EC2, VPC, DynamoDB, SWF 및 Route 53 등 다른 AWS 서비스들 사용

3. 연속적인 백업을 위한 Amazon S3 와 통합으로 99.999999999% 내구성 제공



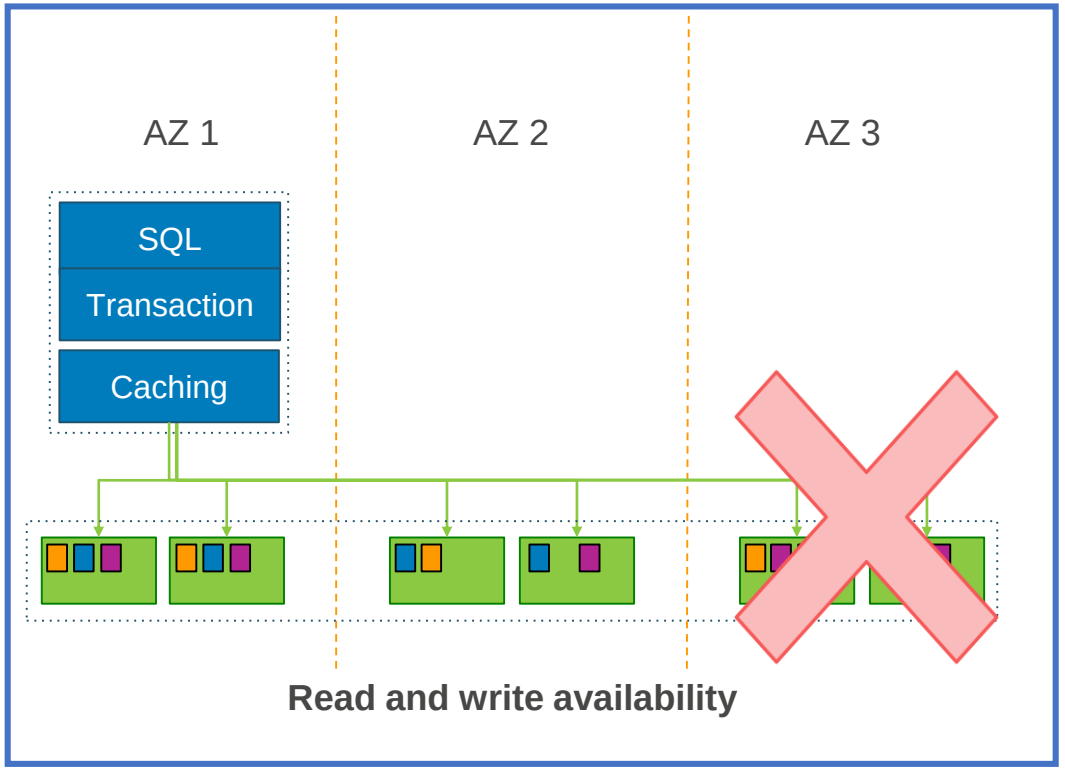
스케일 아웃 가능한 분산 스토리지

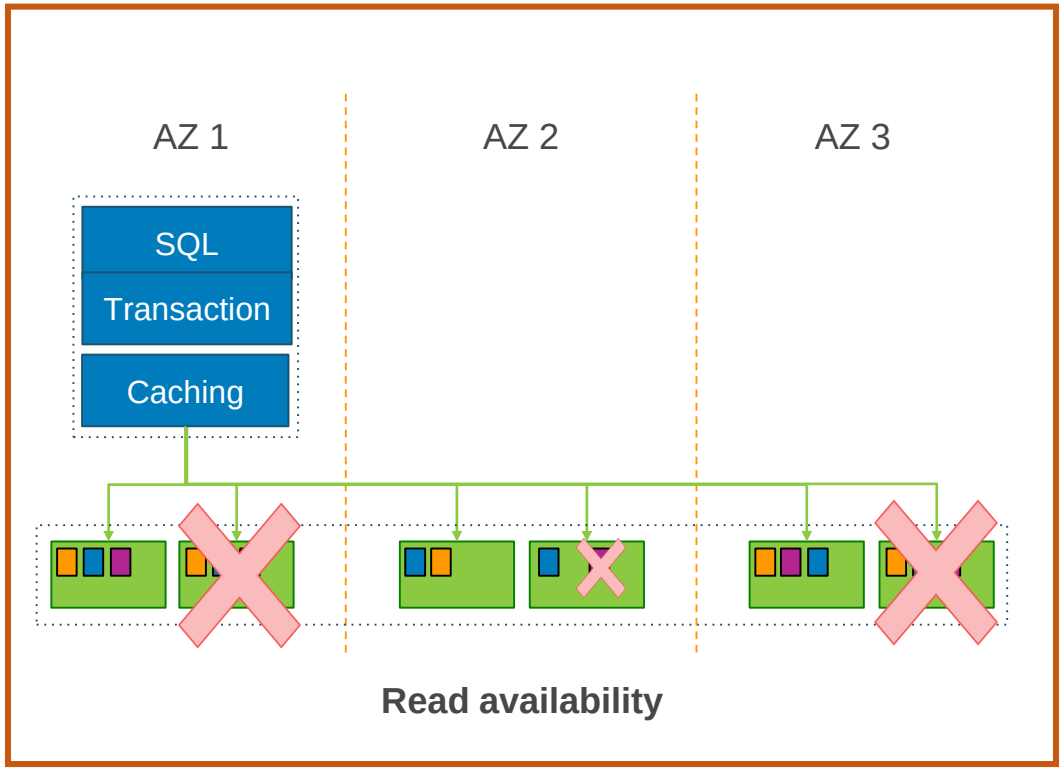
- 데이터베이스용으로 설계된 로그 구조 기반의 분산형 스토리지 시스템
- 3개의 가용영역에 걸친 수백개 이상의 스토리지 노드로 스트라이핑
- 총 6개의 복제본 유지 (각각의 가용 영역에 2개의 복제)



스토리지 자가 치유 및 장애 내구성

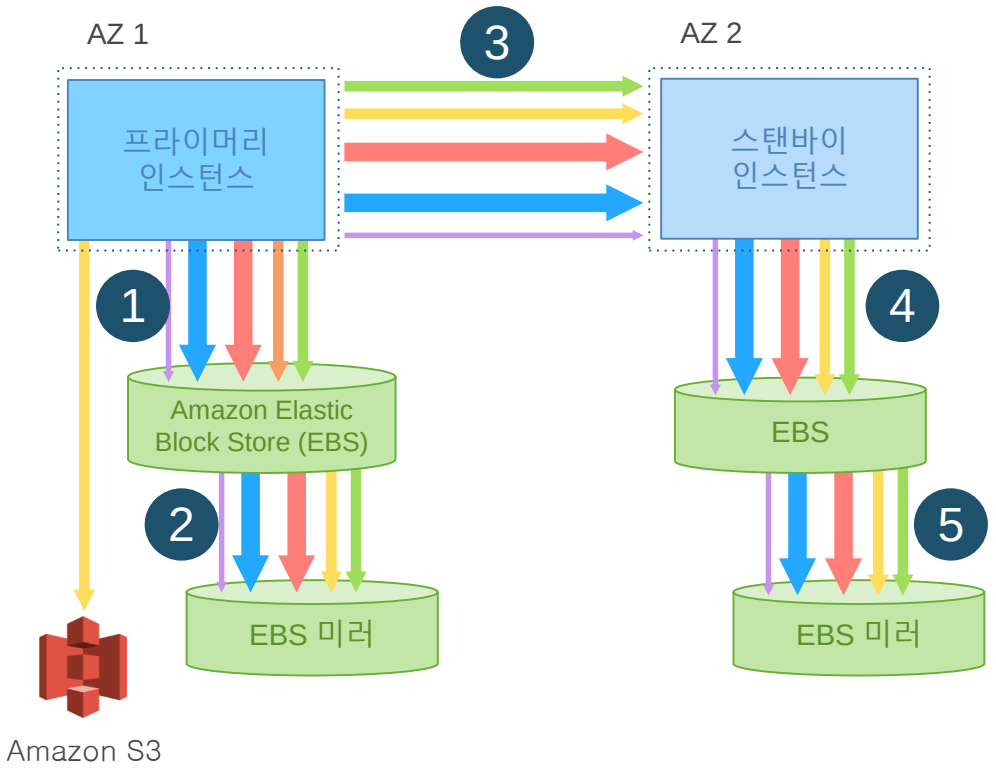
- 자동 장애 감지, 복제, 복구
- 2개의 복제 및 1개 가용 영역 장애는 읽기 및 쓰기 가용성에 영향 없음
- 3개의 복제 장애에도 읽기 가용성에 영향 없음





RDS - MySQL I/O 트래픽

MYSQL WITH STANDBY



IO FLOW

EBS 볼륨에 쓰기 수행 - EBS 는 미러에 쓰기 수행
, 둘 다 종료 시 ACK
스탠바이 인스턴스에 쓰기 복제

OBSERVATIONS

1, 3, 4 단계는 순차 및 동기
응답 속도 및 지터(Jitter) 가 증대
각 사용자 오퍼레이션을 위한 다양한 쓰기 작업들은 두 번 쓰기

PERFORMANCE

780K 트랜잭션
1백만 TX 당 7,388K I/Os (미러 및 스탠바이 제외)
TX 당 평균 7.4 I/Os

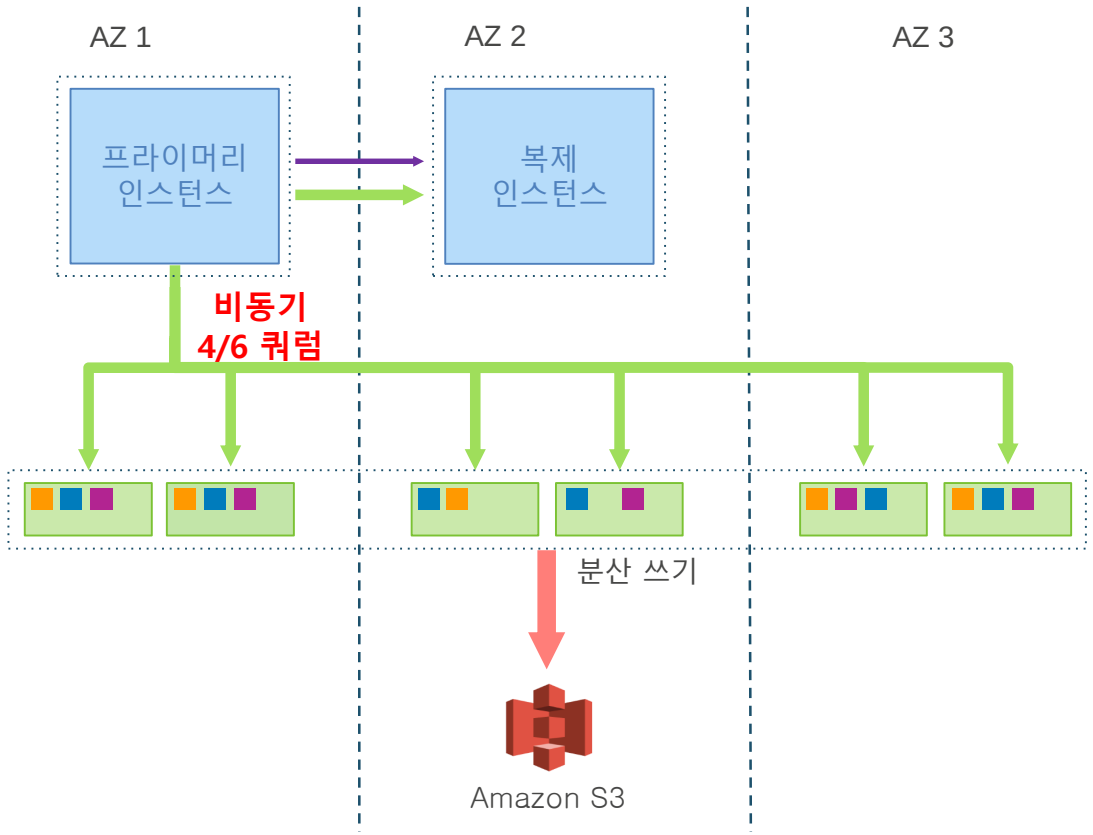
30분 SysBench 쓰기-전용 워크로드, 100 GB 데이터 셋,
RDS SingleAZ, 30K PIOPS

TYPE OF WRITE

LOG BINLOG DATA DOUBLE-WRITE FRM FILES

Aurora I/O 트래픽

AMAZON AURORA



IO FLOW

Redo 로그 레코드 전송 – LSN(Log Sequence Number)에 의해 전체 순서화
독립적인 캐시 프로세스에 업데이트
스토리지 노드에 전송 후 쓰기 수행

OBSERVATIONS

오직 Redo 로그 레코드만 쓰기, 모든 단계는 비동기화
데이터 블록 쓰기 없음 (체크포인트, 캐시 대체 등)
6X 로그 쓰기 향상, **9X** 네트워크 트래픽 감소
네트워크 및 스토리지 응답속도 증가에 내구성

PERFORMANCE

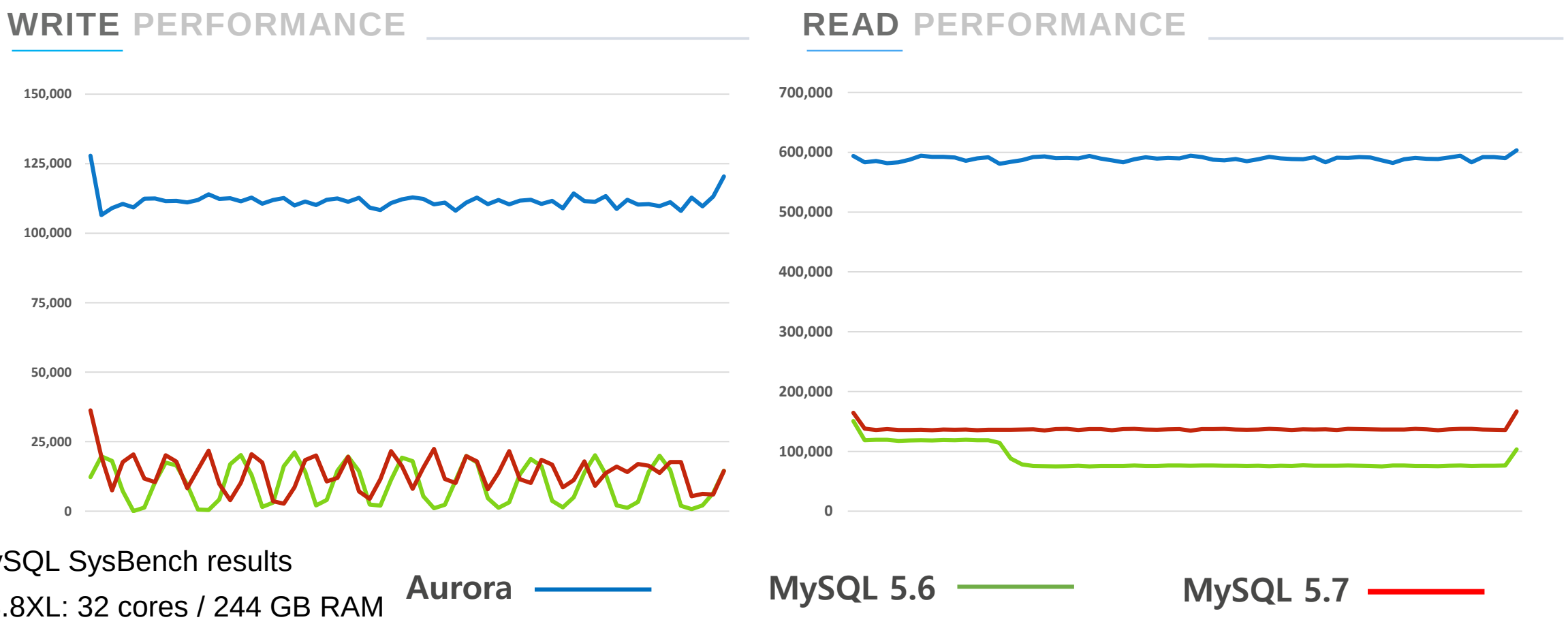
27,378K 트랜잭션	**35X** 향상
1백만 TX 당 950K I/Os (6X amplification)	**7.7X** 감소

30분 SysBench 쓰기-전용 워크로드, 100 GB 데이터 셋

TYPE OF WRITE

LOG BINLOG DATA DOUBLE-WRITE FRM FILES

RDS MySQL에 비해 5배의 성능



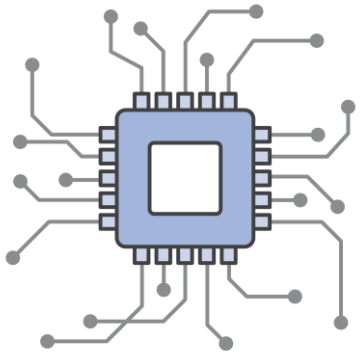
Based on industry standard benchmarks

AWS Cloud Day in Busan

AWS로의 데이터베이스 이전

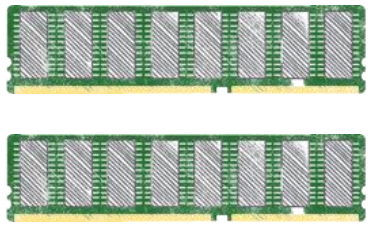
클라우드 마이그레이션

인스턴스 선택시 고려 사항



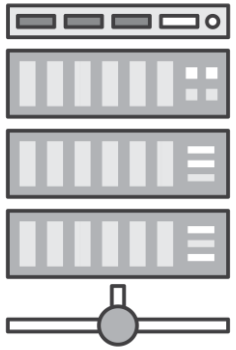
Compute Capabilities

vCPUs



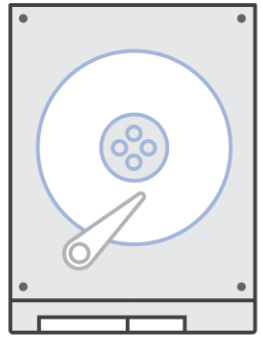
Memory Capabilities

GB of RAM



Network Performance

MB/s (Throughput)



Storage Performance

I/O Throughput

Instance generation

i3.8xlarge

Instance family

- 범용(General Purpose)
- 컴퓨팅 최적화(Compute optimized)
- 메모리 최적화(Memory Optimized)
- 스토리지 최적화(Storage and IO optimized)
- GPU enabled

Instance Size

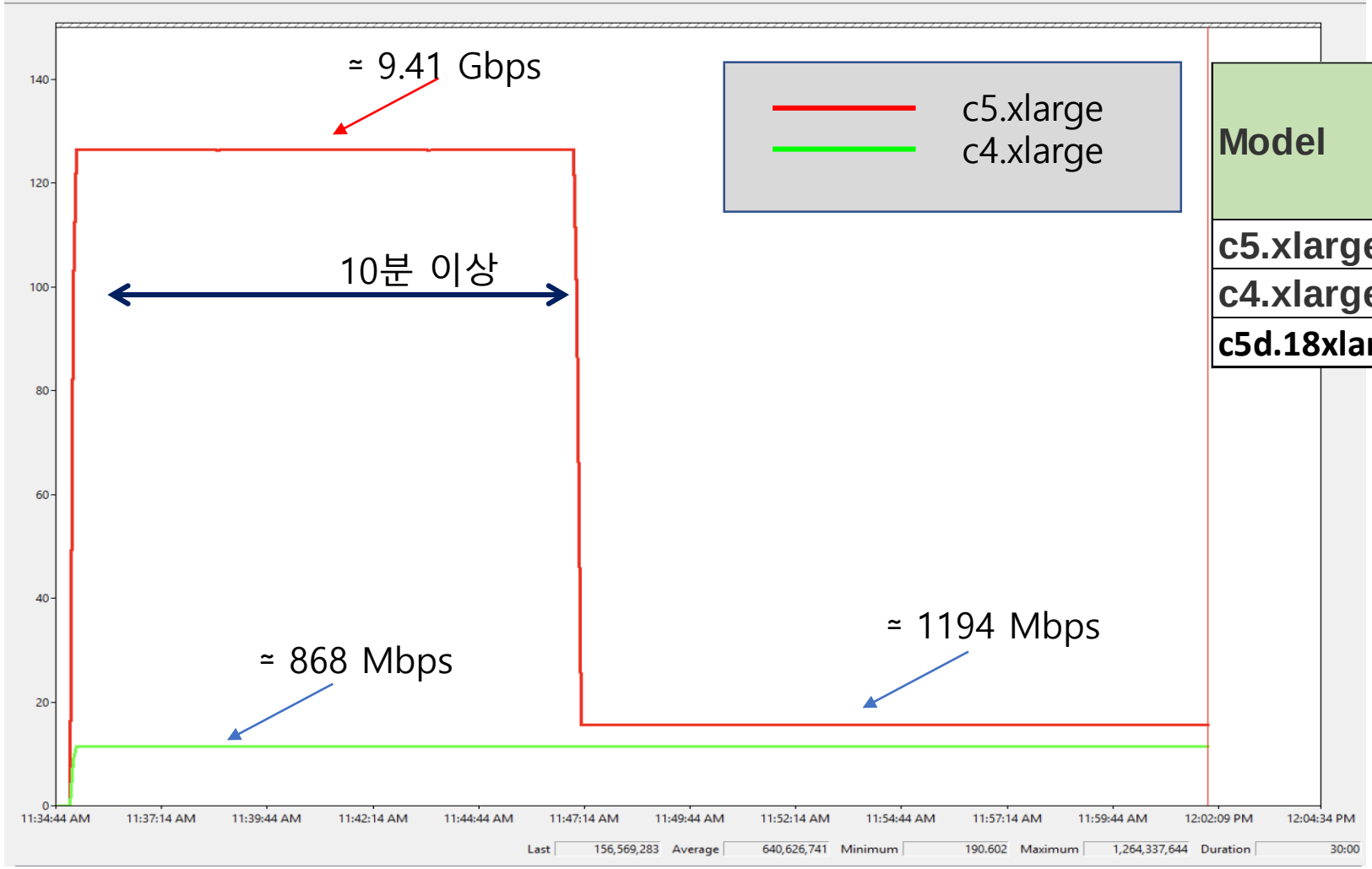
- CPU, Memory 크기 비례
- 숫자는 vCPU * 4
  (8 * 4 = 32vCPU)
- CPU 최적화 옵션

https://docs.aws.amazon.com/ko_kr/AWSEC2/latest/UserGuide/instance-optimize-cpu.html

Enhanced Networking Interface: Intel VF 과 ENA

- **[Intel VF] Intel ixgbevf 드라이버 사용**
  - EC2에서는 Intel ixgbevf 드라이버를 사용하는 Intel 82599 VF 인터페이스를 통해 향상된 네트워킹 기능을 제공
  - https://docs.aws.amazon.com/ko_kr/AWSEC2/latest/WindowsGuide/sriov-networking.html
- **[ENA] ENA(Elastic Network Adapter)**
  - RSS (Receive Side Scaling)
  - https://docs.aws.amazon.com/ko_kr/AWSEC2/latest/WindowsGuide/enhanced-networking-ena.html
- **Throughput**
  - Intel VF: 최대 10Gbps
  - ENA: 최대 100 Gbps

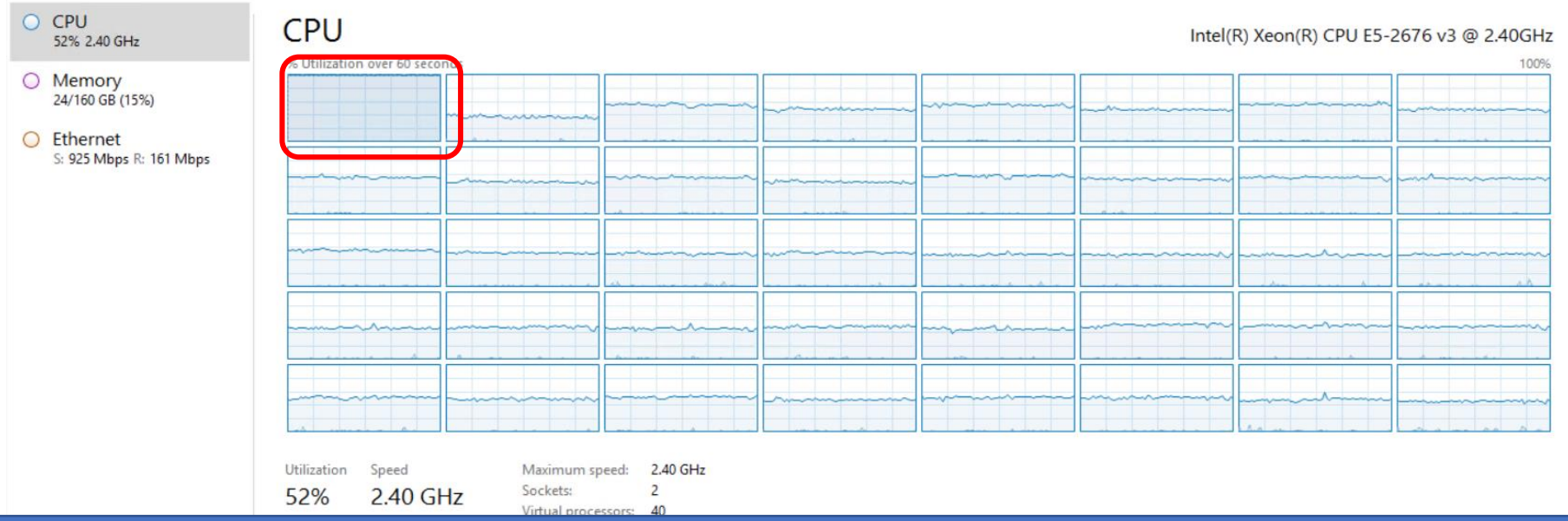
Network Throughput 테스트 결과



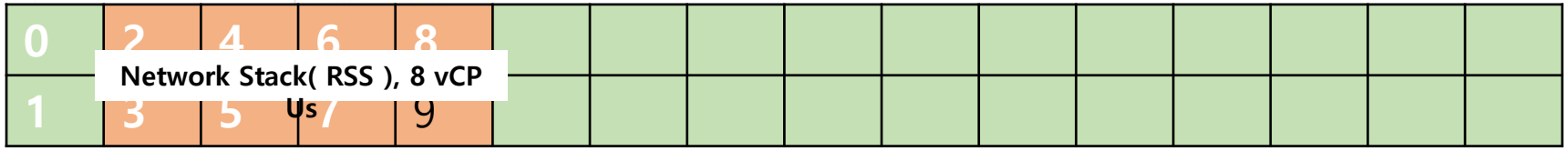
Model	vCPU*	Network Performance (Gbps)
c5.xlarge	4	Up to 10
c4.xlarge	4	High
c5d.18xlarge	72	25

ENA 설정 튜닝- RSS Group 변경

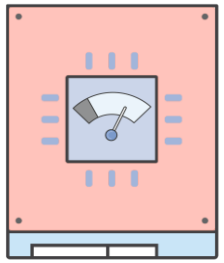
ENA Network 병목: vCPU(0)가 과도하게 사용됨




```
Set-NetAdapterRss -name (Get-NetAdapter | Where-Object {$_.InterfaceDescription -like '*Elastic*'}).Name -Baseprocessorgroup 0 -BaseProcessorNumber 2
```



EBS Volume Types: io1



Provisioned IOPS SSD

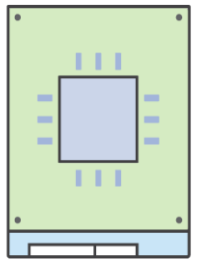
Volume Size: 4 GB to 16 TB

Max Throughput per Volume: 1000 MB/s

Max IOPS per Volume: 64,000

Ideal for critical applications and databases with sustaine d IOPS

EBS Volume Types: gp2



General Purpose SSD

Volume Size: 1 GB to 16 TB

Max Throughput per Volume: 250 MB/s

Max IOPS per Volume: 16,000

Burst: 3,000 IOPS (for volumes up to 1 TB)

Great for boot volumes, low latency application s and bursty databases

	SSD(Solid-State Drive)		하드 디스크 드라이브(HDD)	
볼륨 유형	범용 SSD(gp2)*	프로비저닝된 IOPS SSD (io1)	처리량에 최적화된 HDD (st1)	Cold HDD (sc1)
API 이름	gp2	io1	st1	sc1
볼륨 크기	1GiB - 16TiB	4GiB - 16TiB	500GiB - 16TiB	500GiB - 16TiB
최대 IOPS**/볼륨	16,000***	64,000****	500	250
최대 처리량/볼륨	250MiB/s***	1,000MiB/s†	500MiB/s	250MiB/s
최대 IOPS/인스턴스	80,000	80,000	80,000	80,000
최대 처리량/인스턴스	1,750MiB/s	1,750MiB/s	1,750MiB/s	1,750MiB/s

응답 시간에 민감한 DB를 위한 - NVMe

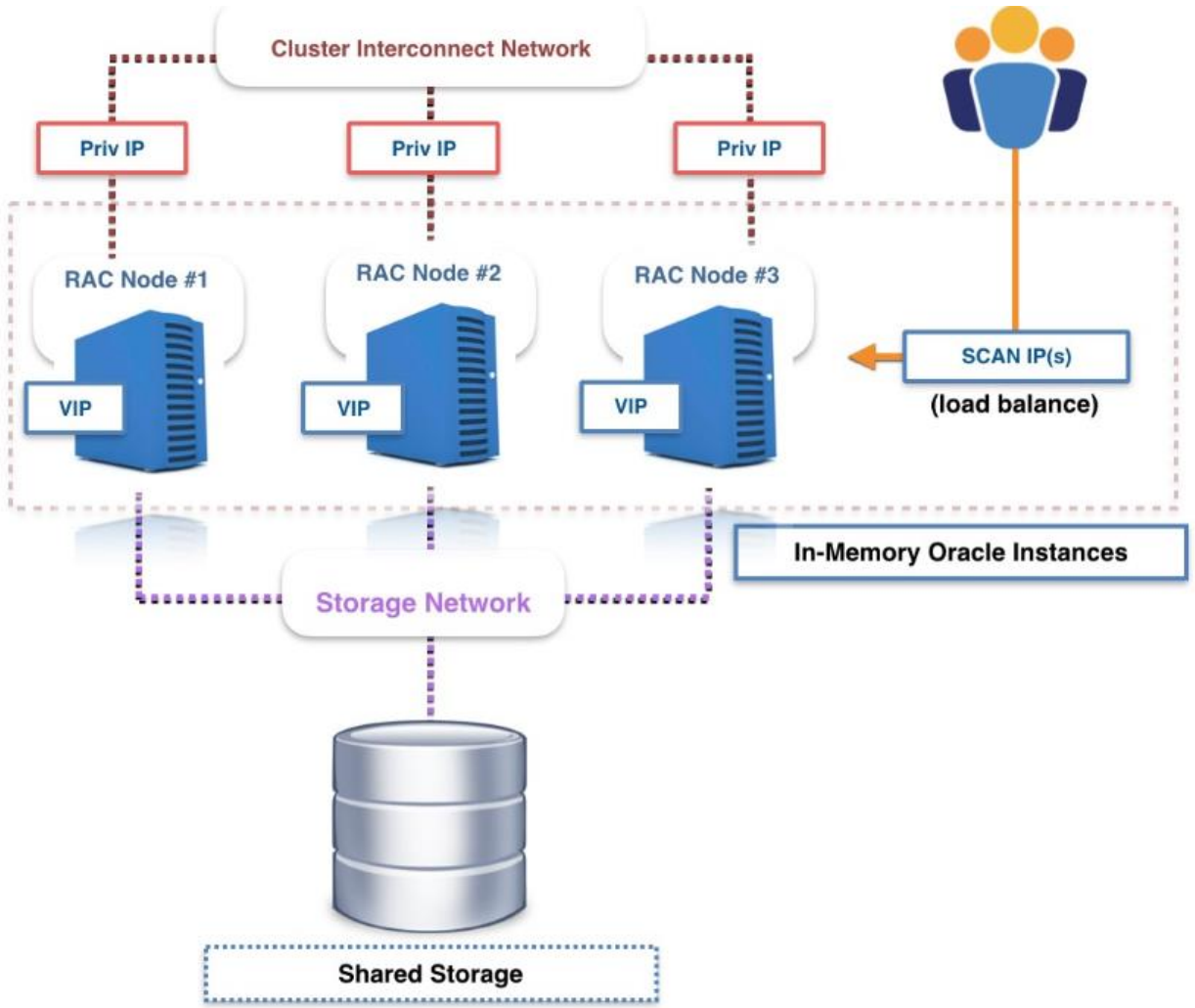
- Amazon EC2 인스턴스 스토어
  - NVMe (Non-Volatile Memory express) 타입의 SSD를 장착하여 낮은 지연속도와 높은 IOPS
  - 데이터 유지: 물리적 호스트 재시작, EC2 재부팅
  - 데이터 삭제: 인스턴스가 중지/종료, 또는 호스트 하드웨어 이슈로 호스트가 이동되는 경우
  - 지원 인스턴스: C5d, I3, F1, M5ad, M5d, p3dn.24xlarge, R5ad, R5d 및 z1d

인스턴스 타입	vCPU (CPU clock)	Mem(GiB)	인스턴스 스토어 크기
c5d	2 ~ 72 (3.0 ~ 3.5 GHz)	4 ~ 144	50 ~ 900 GiB
m5d	2 ~ 96 (2.5 GHz)	8 ~ 384	75 GiB ~ 3.6 TiB
r5d	2 ~ 96 (~3.1 GHz)	16 ~ 768	75 GiB ~ 3.6 TiB
z1d	2 ~ 48 (~4.0 GHz)	16 ~ 384	75 GiB ~ 1.8 TiB
i3	2 ~ 72 (2.3 GHz)	15.25 ~ 512	475 GiB ~ 15.2 TiB
i3en	2 ~ 96 (3.1 GHz)	16 ~ 768	1.25 TiB ~ 60 TiB

https://docs.aws.amazon.com/AWSEC2/latest/WindowsGuide/instance-types.html#ec2-nitro-instances

Oracle 마이그레이션

Oracle Migration



Oracle RDS

Amazon EC2

Oracle on EC2

Amazon Aurora

Aurora

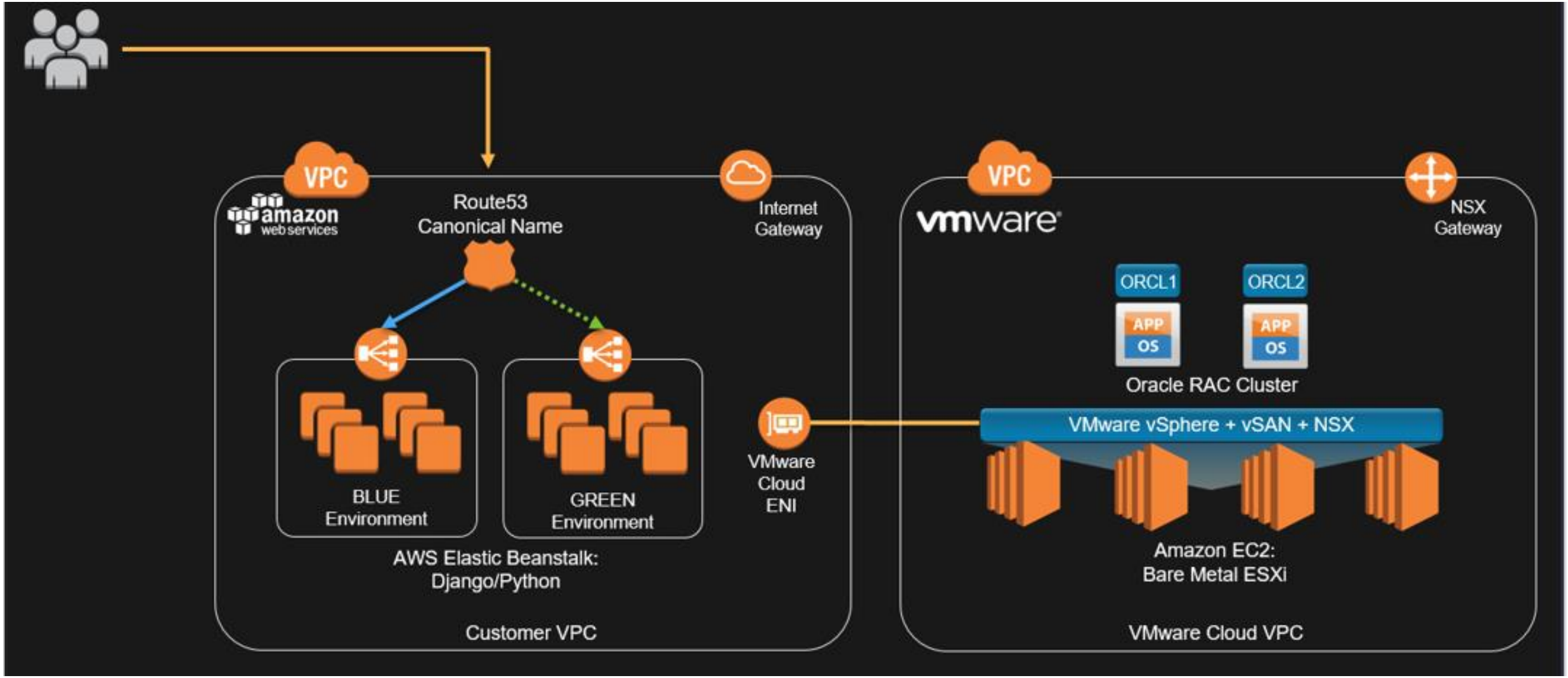
AWS migration strategies for Oracle

Solution	Oracle Databases		
	Oracle Databases on AWS EC2	Oracle Databases on AWS RDS	Oracle database migration to AWS
Migration path	Rehost	Replatform	Refactor
Post migration	Customer runs Oracle Database EE, SE, NoSQL, TimesTen, MySQL, Golden Gate on AWS	Customer shifts Oracle EE, SE to AWS RDS for Oracle	Customer migrates from Oracle EE, SE, NoSQL to AWS RDS OSS, Aurora or Redshift
AWS Services	EC2/EBS, VPC	Oracle RDS	RDS OSS, Aurora, Redshift, Schema Conversion Tool, Database Migration Service

Oracle to AWS license and support

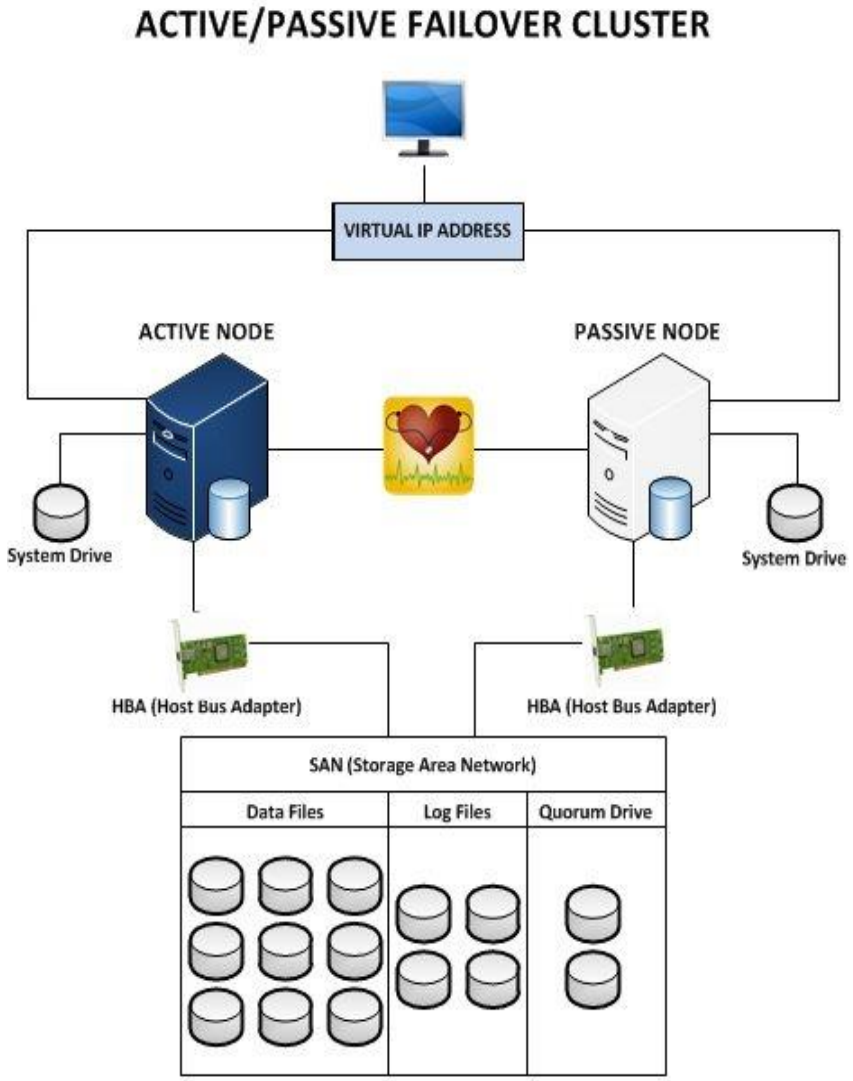
Solution	Oracle Databases		
	Oracle Databases on AWS EC2	**Oracle Databases on AWS RDS**	**Oracle database migration to AWS**
Migration path	Rehost	Replatform	Refactor
License consideration	BYOL. Review Oracle Cloud Licensing Policy. **2 vCPU= 1 Oracle Proc** with Hyper threading enabled	License included or BYOL	MySQL and PostgreSQL open source
Support consideration	Standard Oracle support. Oracle Database >=11.2.0.4 & >= 12.1.0.2. AWS EC2 DB optimized instances.	Supported for Oracle Database >=11.2.0.4 & >= 12.1.0.2. AWS EC2 DB optimized instances.	AWS RDS and Redshift are managed services.

Oracle RAC – VMWare cloud on AWS



SQL Server 마이그레이션

SQL Server Migration



Microsoft SQL Server
Amazon RDS

SQL Server RDS

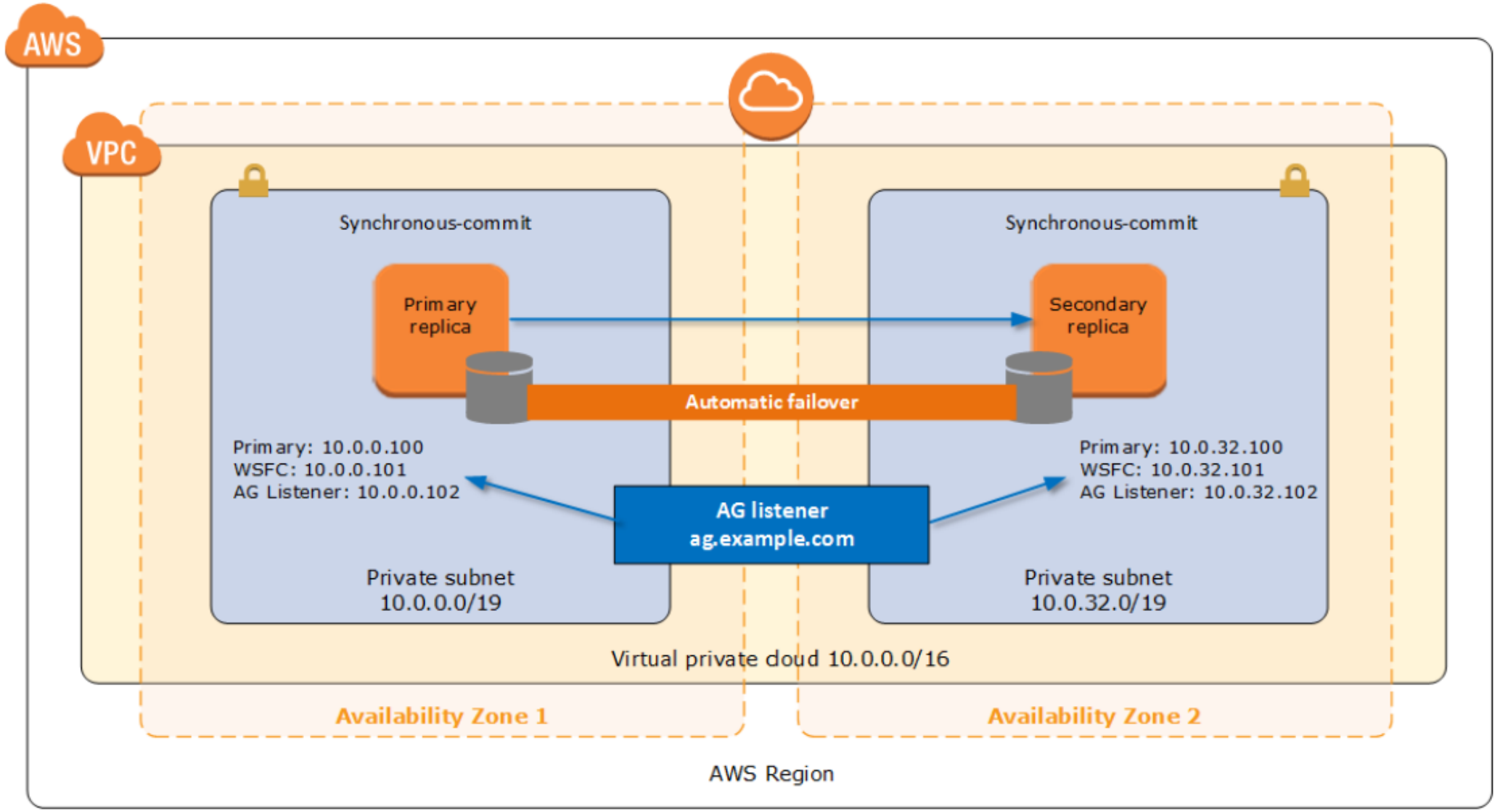
Amazon EC2

SQL Server on EC2

Amazon Aurora

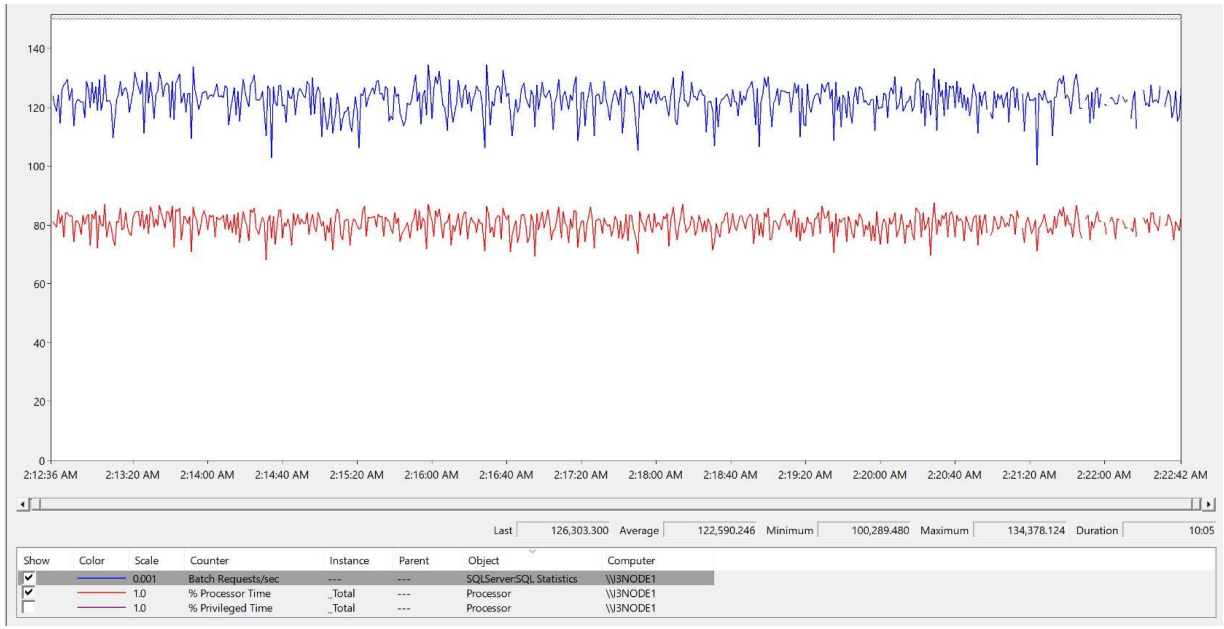
Aurora

아키텍쳐 - SQL Server AlwaysOn on AWS

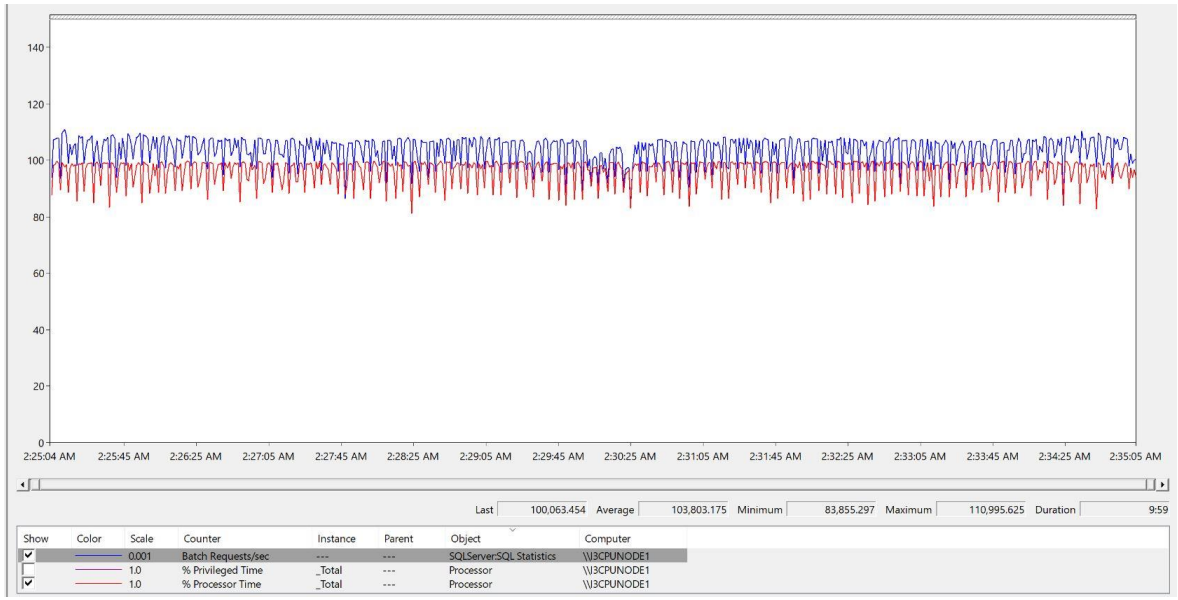


Optimize vCPU 성능 비교

I3.8xlarge with NVME- enabled HyperThreads



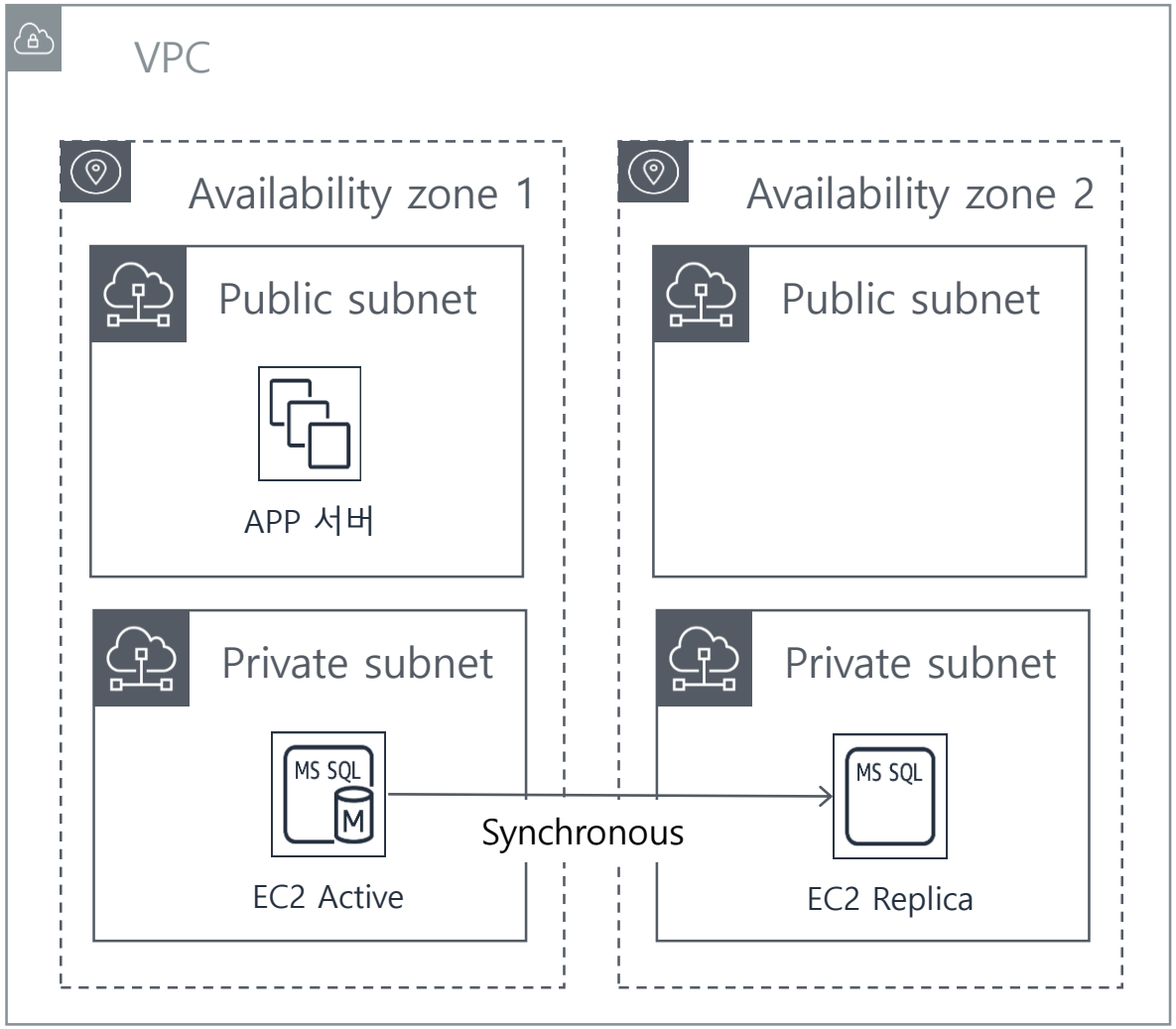
I3.8xlarge with NVME- disabed HyperThreads



- Blue: Batch Request/sec (Scale-0.001)
- Red: % Processor time

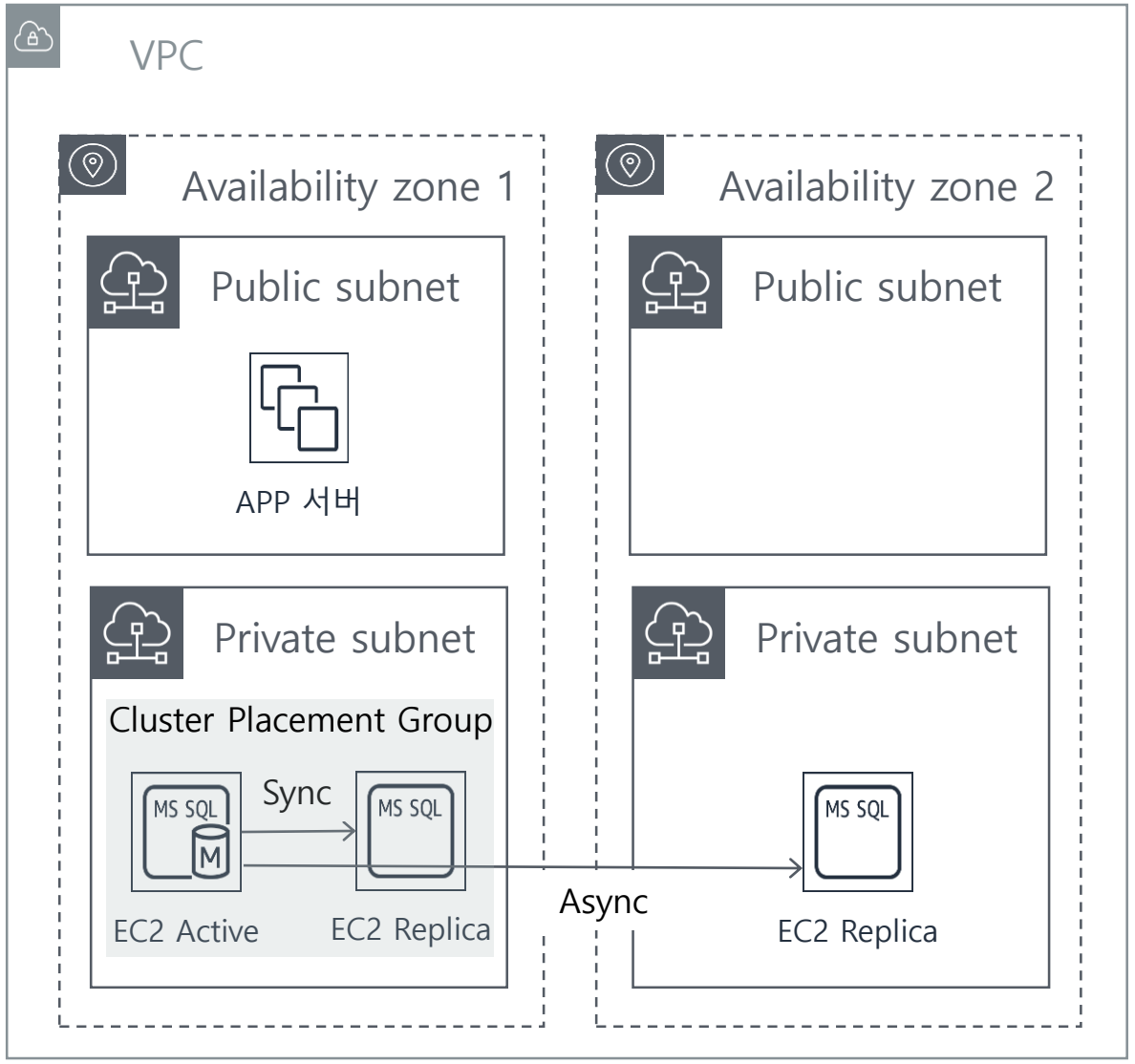
Architecture Sample – 비용 효율성 위주

- **SQL 서버의 스탠다드 라이선스 사용 (SE)**
  - 2대 인스턴스 구성
  - 복수 AZ 구성 권고
  - 인스턴스간 Synchronous 복제 (EBS의 간헐적인 Latency Spike 영향 최소화)
  - 인스턴스 스토어에 데이터 파일 및 로그 파일 구성



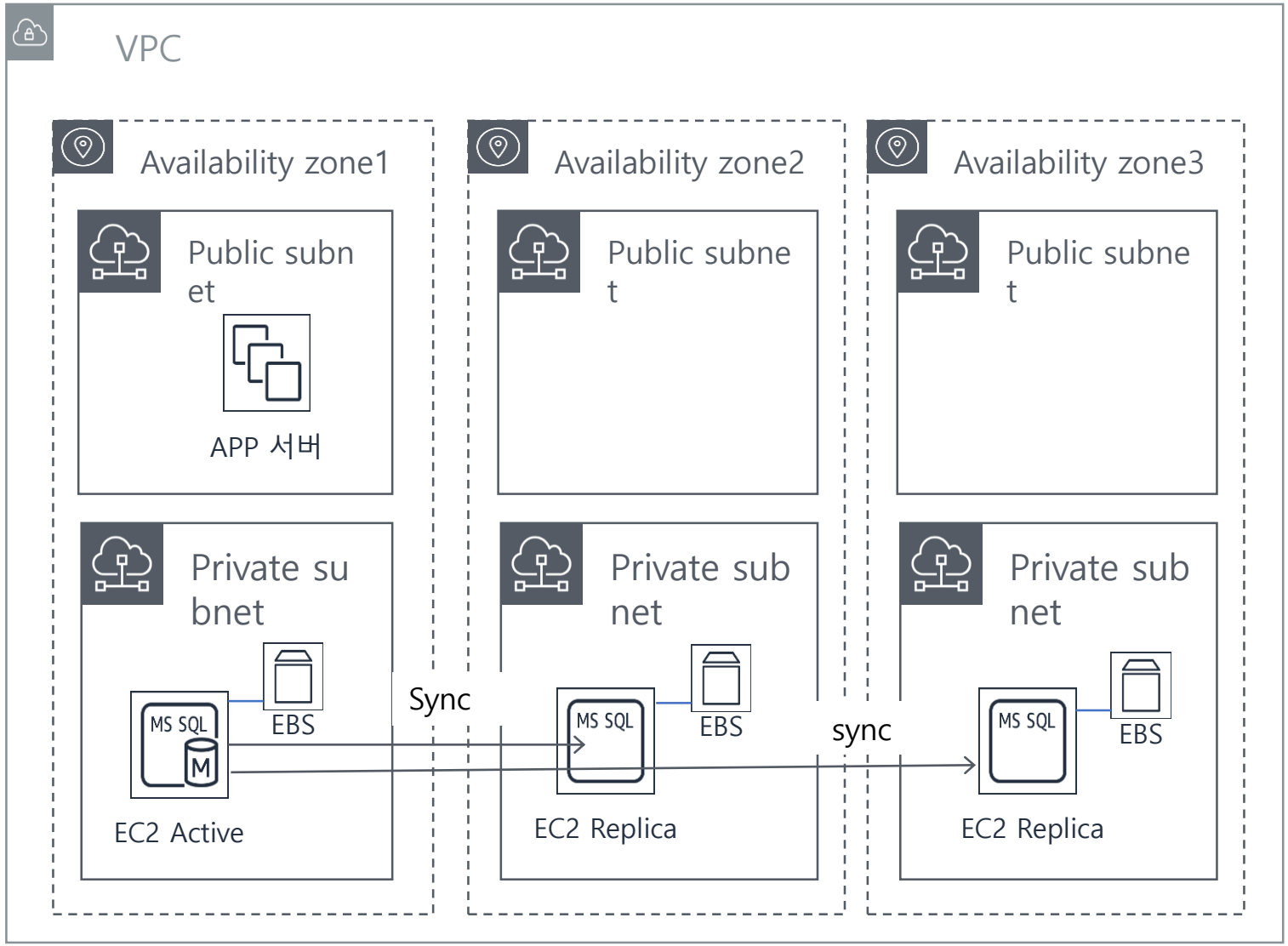
Architecture Sample – 고성능 위주

- **SQL 서버 엔터프라이즈 라이선스 사용(EE)**
  - 3대 인스턴스 사용
  - 동일 AZ에 Cluster Placement Group으로 2대 인스턴스 배치
  - 동일 AZ 인스턴스간 Synchronous 복제
  - 다른 AZ에 추가 1대 인스턴스 배치
  - 다른 AZ 인스턴스간 Asynchronous 복제
  - 복제 인스턴스에서 데이터 및 로그 백업



Architecture Sample – 데이터 안정성 위주

- **SQL 서버 엔터프라이즈 라이선스 사용(EE)**
  - 3대 이상 인스턴스 사용
  - 데이터는 EBS 볼륨, 로그는 인스턴스 스토어에 배치
  - 다수의 AZ에 인스턴스 배치
  - 모든 인스턴스간 Synchronous 복제
  - 복제 인스턴스에서 데이터 및 로그 백업
  - 동일 AZ에 인스턴스 배치할 경우, Spread Placement Group 사용



AWS Cloud Day in Busan

Database Migration Service

DMS: AWS Database Migration Service
SCT: AWS Schema Conversion Tool

AWS Database Migration Service(DMS) 개요

AWS
Database Migration
Service

- 데이터베이스 마이그레이션을 손쉽게 시작
- 마이그레이션 중 어플리케이션 서비스 유지
- Amazon EC2 또는 RDS를 원본 및 대상으로 복제
- 동일 또는 이기종 데이터베이스간 복제

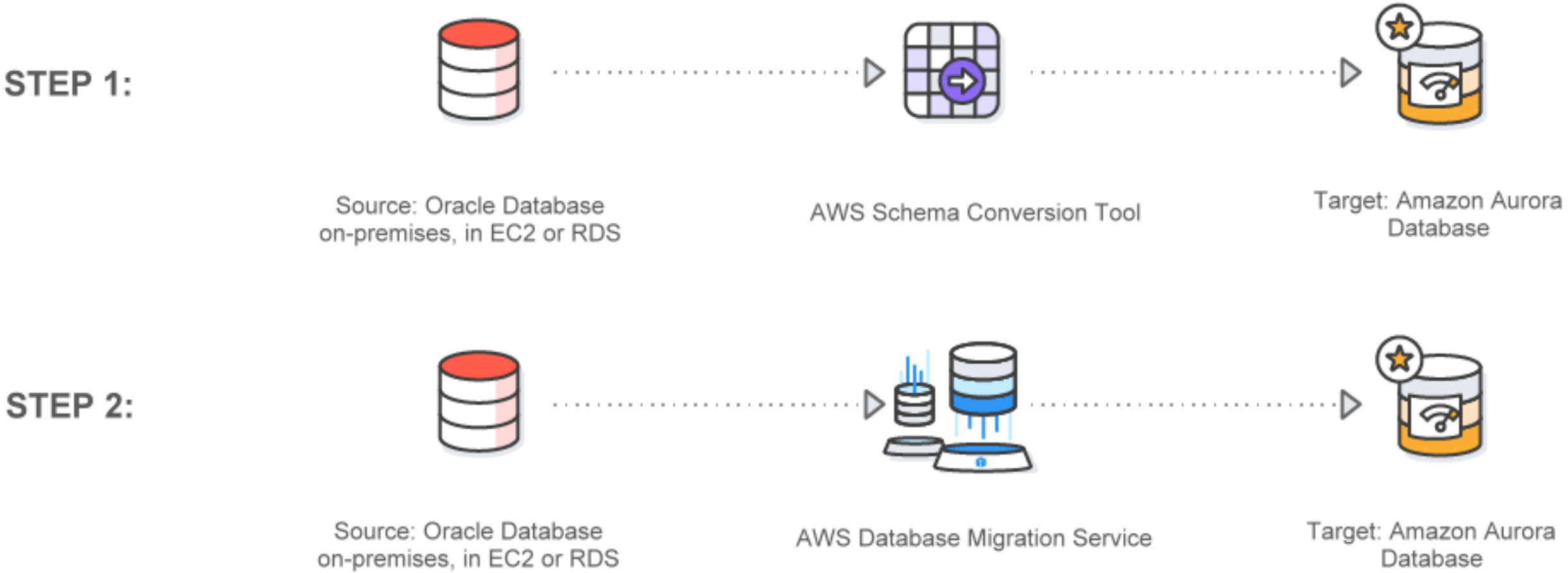
What are DMS and SCT?

AWS Database Migration Service (DMS) easily and securely migrate and/or replicate your databases *an d* data warehouses to AWS

AWS Schema Conversion Tool (SCT) convert your comm ercial database and data warehouse schemas to open-sou rce engines or AWS-native services, such as Amazon Auro ra and Redshift

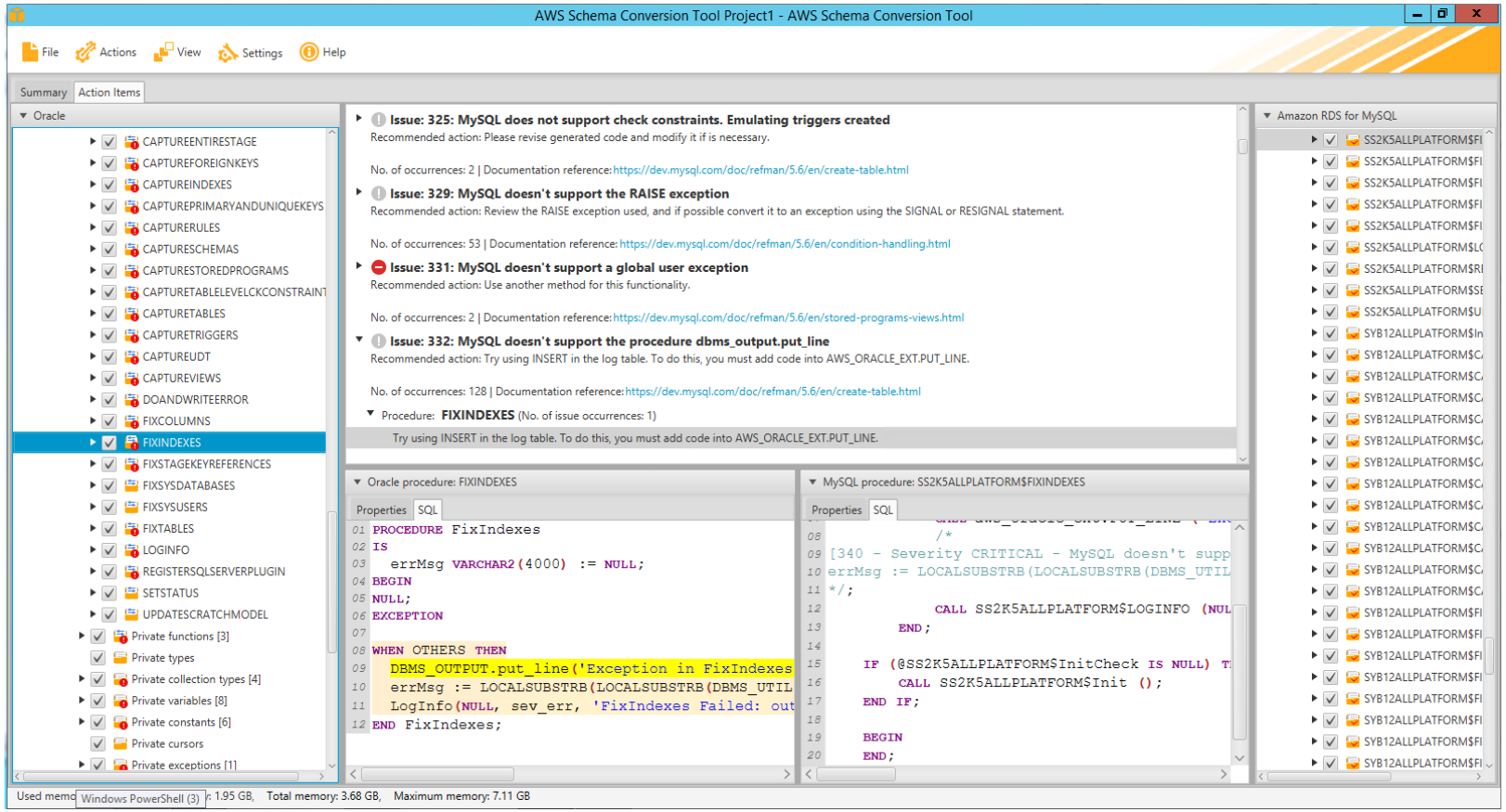
We migrated over 160,000 databases using AWS Database Migration Service

Database migration process



SCT를 통한 테이블, 뷰 및 코드 변환

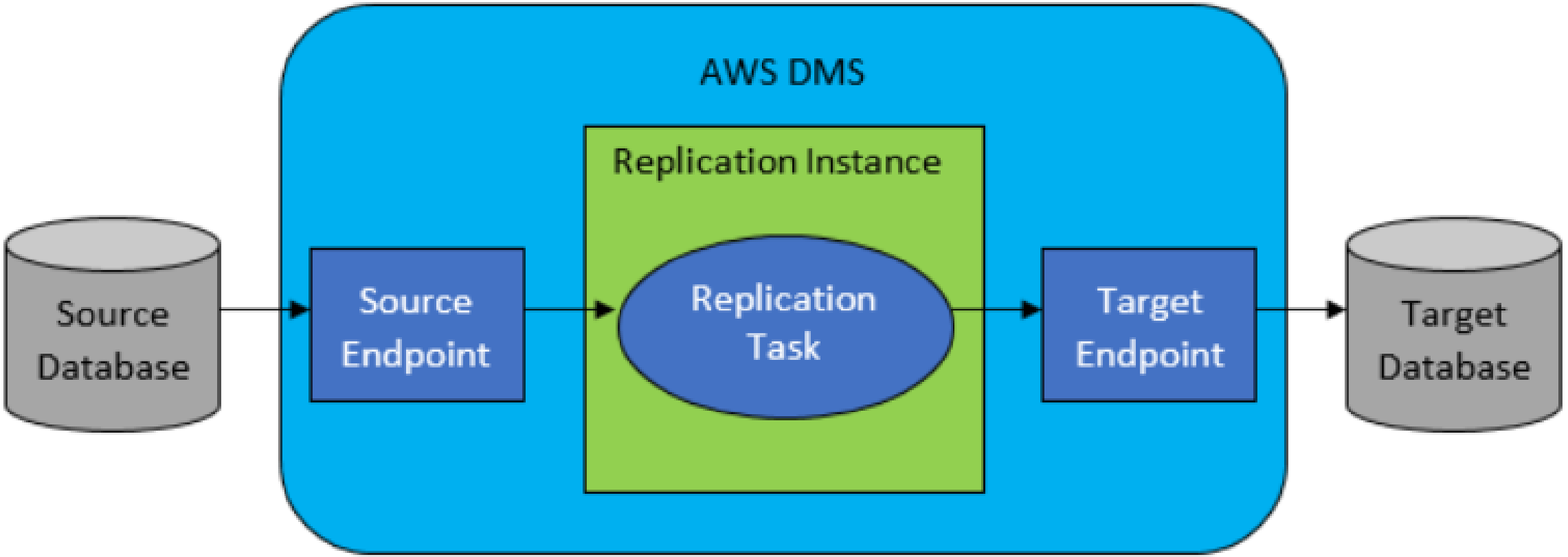
ORACLE Microsoft SQL Server ⇨ PostgreSQL MariaDB Amazon Aurora MySQL



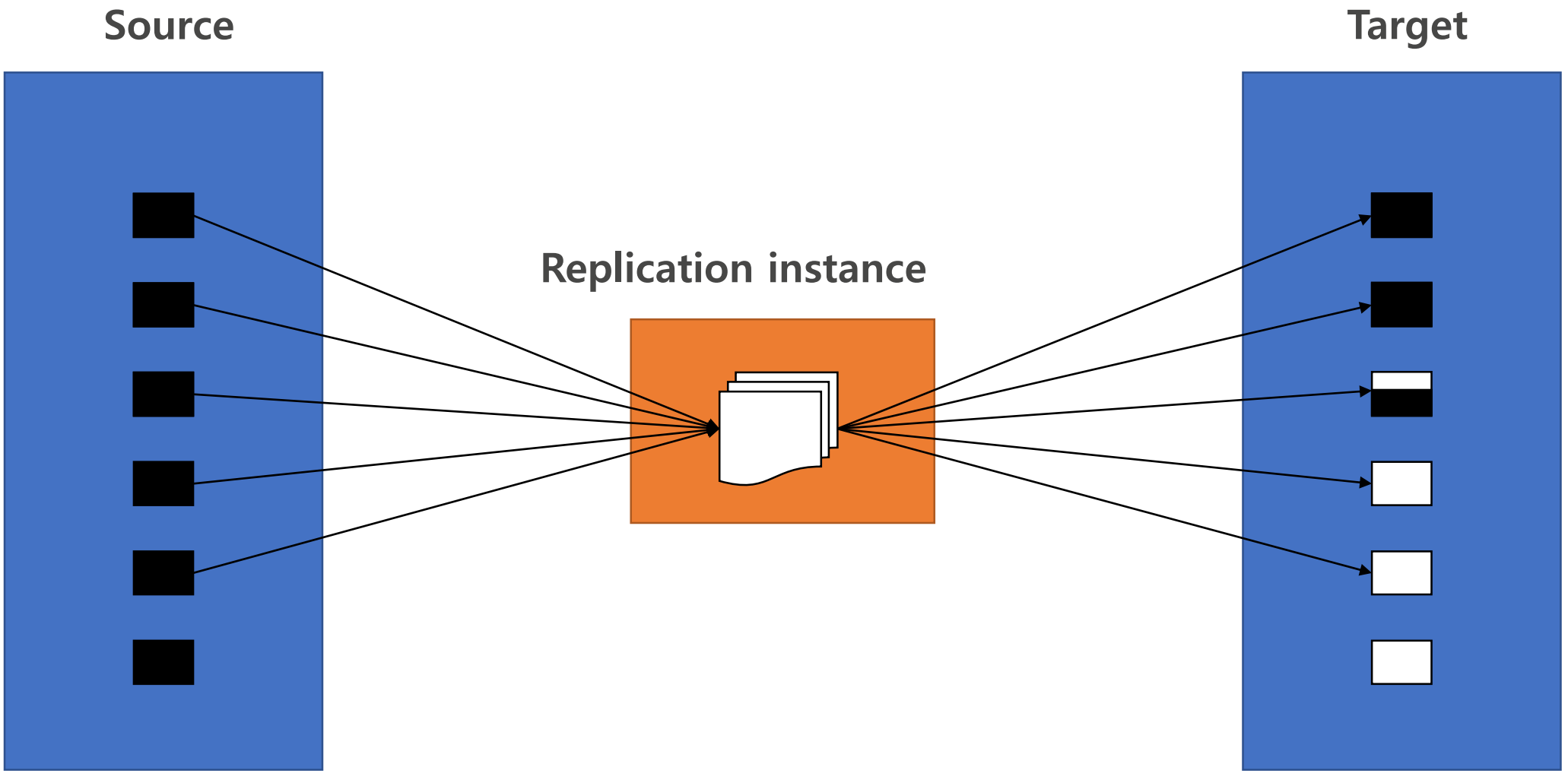
- 시퀀스
- 사용자 정의 타입
- 패키지
- 스토어드 프로시저
- 함수
- 트리거
- 스키마
- 테이블
- 인덱스
- 뷰

https://aws.amazon.com/blogs/database/the-database-migration-playbook-has-landed/

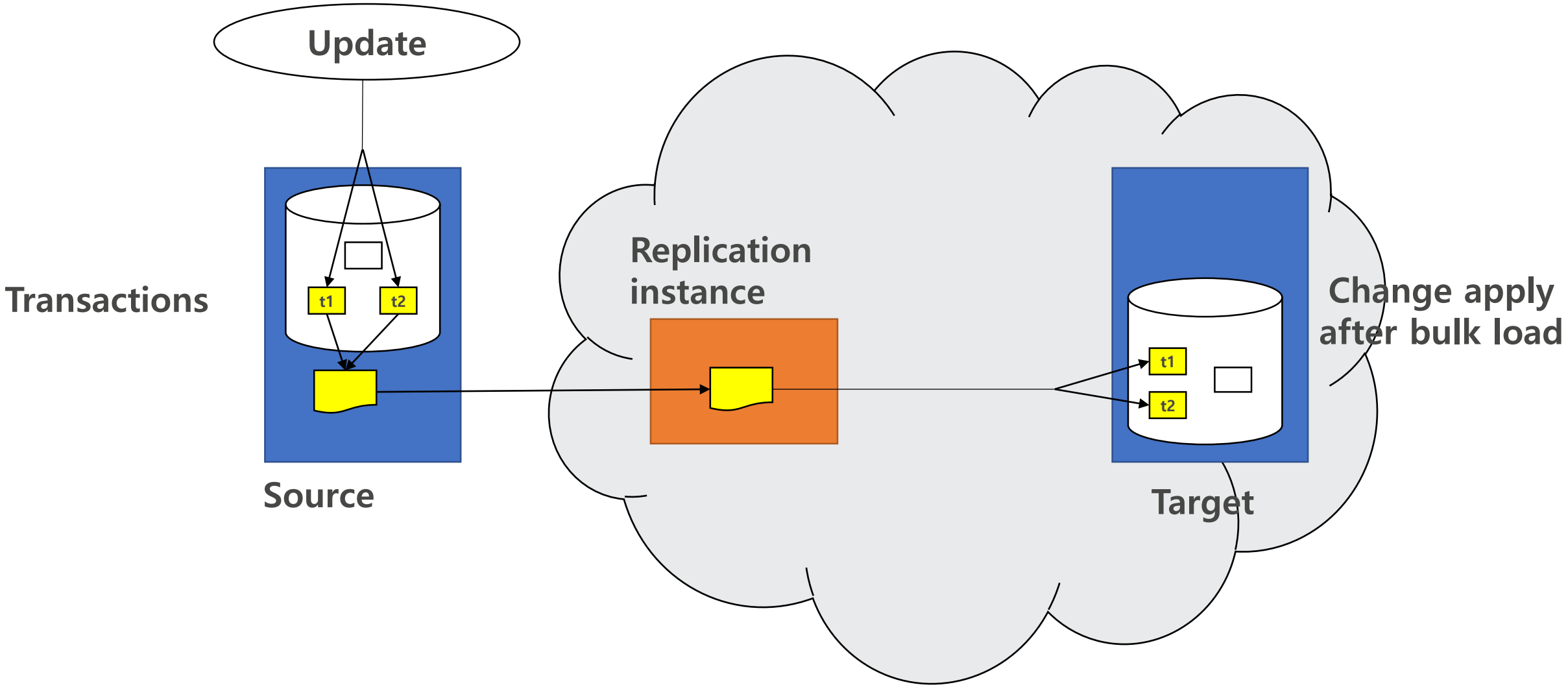
AWS DMS를 이용한 데이터 전송



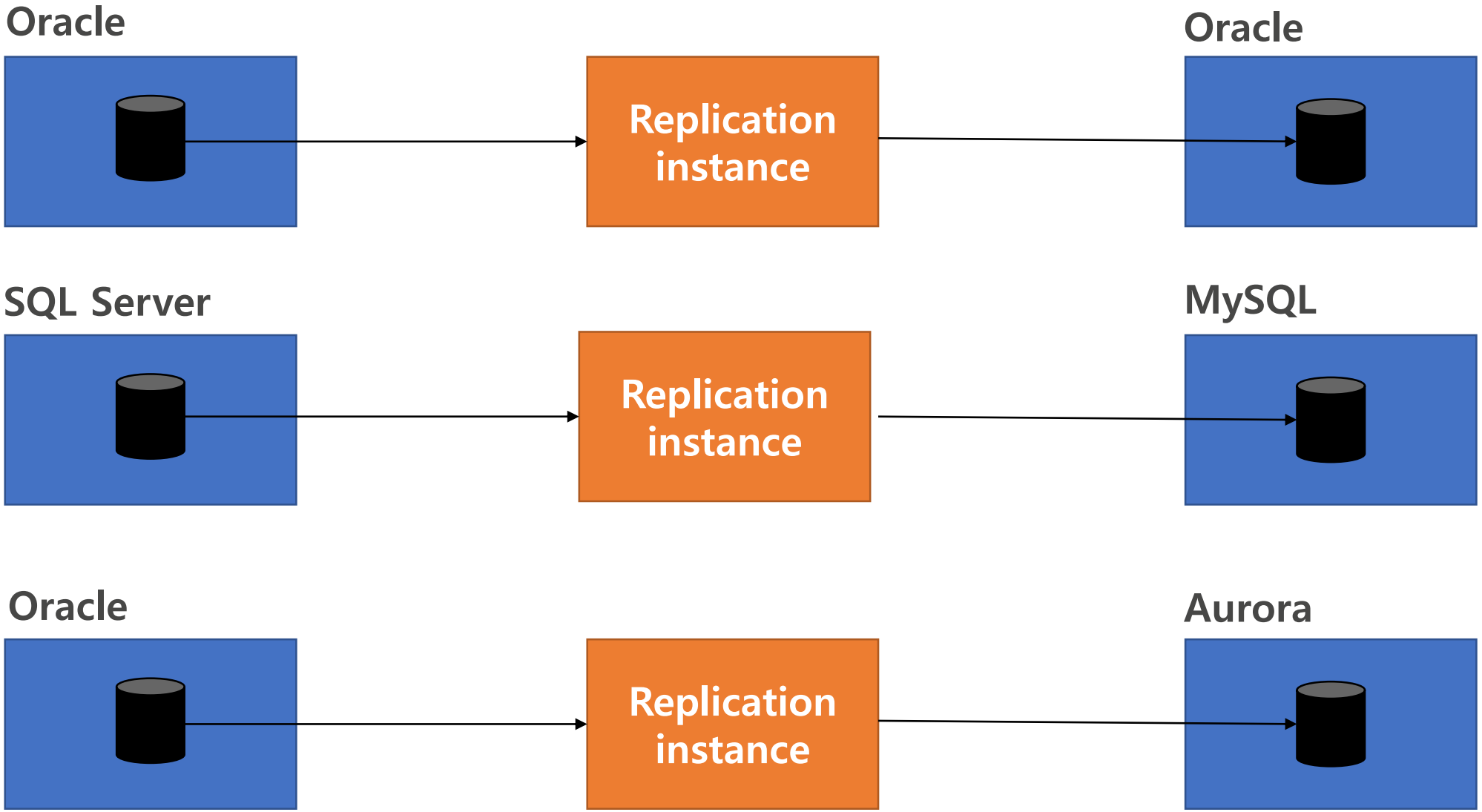
최초 Load is table by table



실시간 동기화 - Change data capture (CDC)



Homogenous or heterogeneous



Sources for AWS Database Migration Service

Customers use the following databases as a source for data migration using AWS DMS:

On-premises and Amazon EC2 instance databases:

- Oracle Database 10g–12c
- Microsoft SQL Server 2005–2014
- MySQL 5.5–5.7
- MariaDB (MySQL-compatible data source)
- PostgreSQL 9.4–9.5
- SAP ASE 15.7+

RDS instance databases:

- Oracle Database 11g–12c
- Microsoft SQL Server 2008R2–2014. CDC operations are not supported yet.
- MySQL versions 5.5–5.7
- MariaDB (MySQL-compatible data source)
- PostgreSQL 9.4–9.5. CDC operations are not supported yet.
- Amazon Aurora (MySQL-compatible data source)

Targets for AWS Database Migration Service

Customers can use the following databases as a target for data replication using AWS DMS:

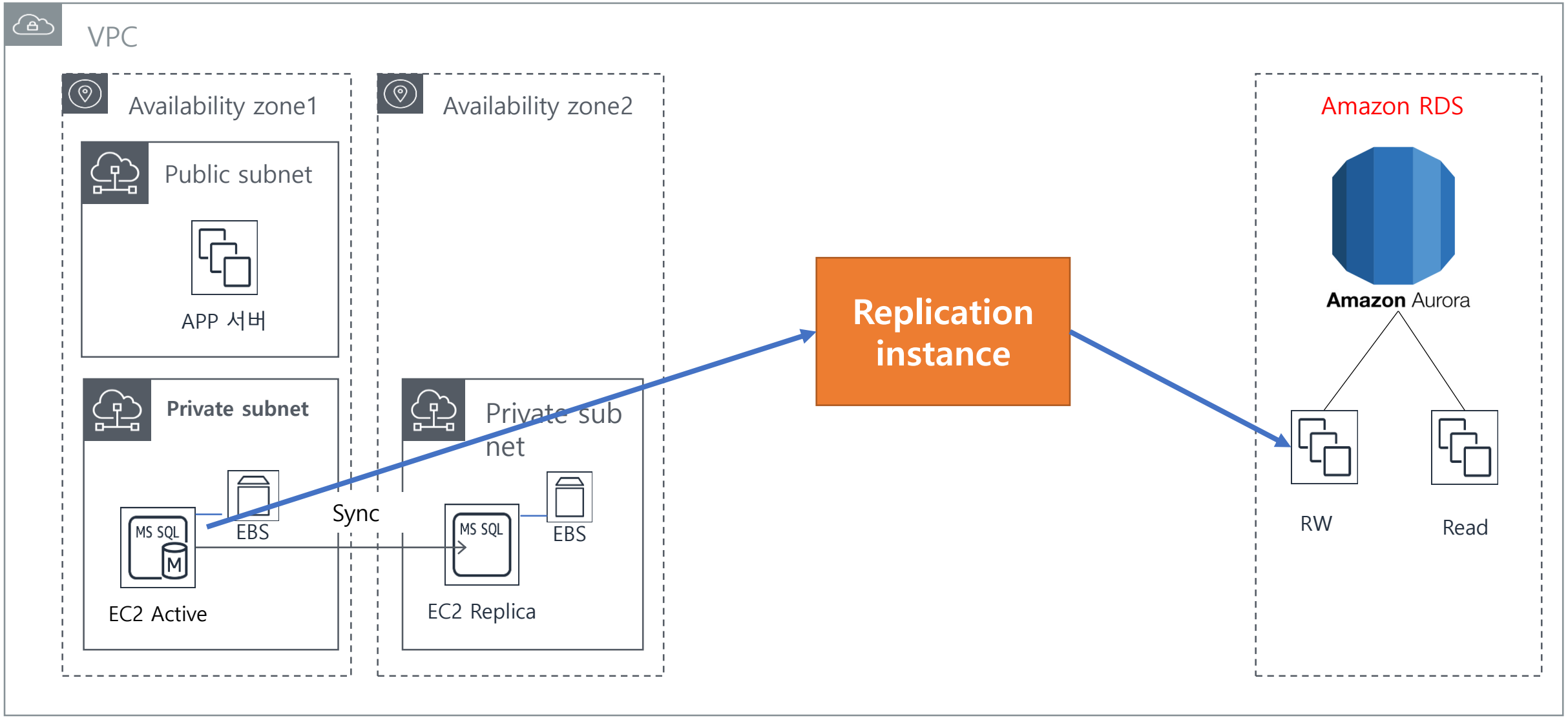
On-premises and EC2 instance databases:

- Oracle Database 10g–12c
- Microsoft SQL Server 2005–2014
- MySQL 5.5–5.7
- MariaDB (MySQL-compatible data source)
- PostgreSQL 9.3–9.5
- SAP ASE 15.7+

RDS instance databases:

- Oracle Database 11g–12c
- Microsoft SQL Server 2008 R2 - 2014
- MySQL 5.5–5.7
- MariaDB (MySQL-compatible data source)
- PostgreSQL 9.3–9.5
- Amazon Aurora (MySQL-compatible data source)
- Amazon Redshift

Demo – SCT & DMS



AWS database migration partners

AWS Cloud Day in Busan

감사합니다.