

Data Engineering Day

Database Migration and Replication

With AWS Database Migration Service

신석주 Solutions Architect Oct 2022

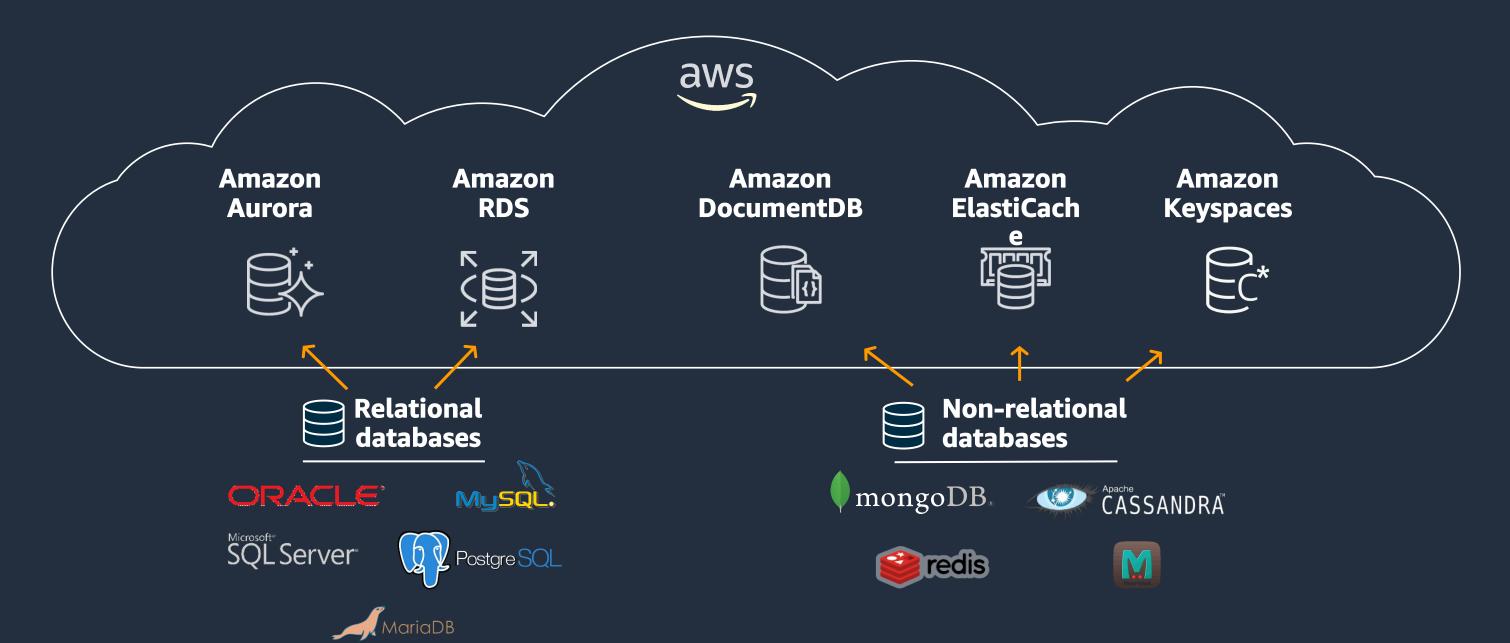


## Agenda

- Database Migration 개요
- AWS Schema Conversion Tool
- AWS Database Migration Service



# 완전 관리형 DB 로의 전환 온프레미스/자체 관리 데이터베이스를 완전 관리형 서비스로 마이그레이션





### 전환을 위한 필요성

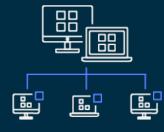


### 인프라스트럭처

- 하드웨어 EOL (End of Life)
- 이중화 또는 자동화 부족
- 다른 레이어에서의 장애



- 필요에 따라 인스턴스 크기 변경
- 필요에 따라 크로스리전 확장





### 비용 최적화

- 사용한 만큼만 지불
- 오픈소스 옵션, 라이선스 절감
- 완전 관리형 데이터베이스

### **Innovation**

- 데이터베이스 현대화
- 워크로드에 적합한 데이터베이스 선택





## AWS 데이터베이스 마이그레이션 툴



### **AWS Schema Conversion Tool (AWS SCT)**

상용 DB, DW 스키마를 오픈소스DB나 AWS 네이티브 서비스(ex. Amazon Aurora, Redshift 등) 로 변환하는 툴

### **AWS Database Migration Service (AWS DMS)**

기존 DB나 DW 를 AWS로 쉽고 안전하게 이관/복제하는 툴





## AWS DMS and AWS SCT 유즈케이스



**Migrate** 

비즈니스 크리티컬 애플리케이션 및 DW 마이그레이션

NoSQL에서 SQL로, SQL에서 NoSQL로 또는 NoSQL에서 NoSQL로 마이그레이션

마이너/메이저 버전 업그레이드

데이터 통합 및 아카이브



Replicate

리전 간 읽기 전용 복제본 생성 데이터레이크/DW 데이터 주입 스트리밍 플랫폼으로 데이터 복제



**Modernize** 

더 빠른 혁신을 위해 AWS 관리형 데이터베이스로 전환 Ex. Aurora, DynamoDB, Redshift

> 레거시 DB에서 벗어나 비용절감, 성장 및 혁신



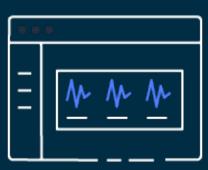
## AWS DMS 서비스 특징



**Secure** 



**Assess** 



**Validate** 



**AWS Snowball** integration



Monitor

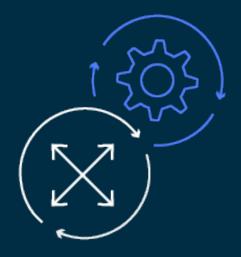








## AWS SCT 서비스 특징



**Assess** 





Plan



Convert schema and code



Migrate data warehouses



## SCT 에서 생성한 Database migration assessment 리포트

#### **Database migration assessment report**

ource database:

Oracle Database 12c Enterprise Edition 12.1.0.2.0 (64bit Production). Enterprise edition



#### **Executive summary**

We completed the analysis of your Oracle source database and estimate that 98% of the database storage objects and 96% of database code objects can be converted automatically or with minimal changes if you select Amazon RDS for PostgreSQL as your migration target. Database storage objects include schemas, tables, table constraints, indexes, types, collection types, sequences, synonyms, view-constraints, clusters and database links. Database code objects include triggers, views, materialized views, materialized view logs, procedures, functions, package, package constants, package cursors, package exceptions, package variables, package functions, package procedures, package types, package collection types, scheduler-jobs, scheduler-programs and scheduler-schedules. Based on the source code syntax analysis, we estimate 99.8% (based on # lines of code) of your code can be converted to Amazon RDS for PostgreSQL automatically. To complete the migration, we recommend 6 conversion action(s) ranging from simple tasks to medium-complexity actions to significant conversion actions.

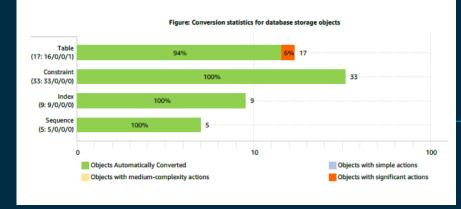
Migration guidance for database objects that could not be converted automatically can be found here

### Database objects with conversion actions for Amazon RDS for PostgreSQL

Of the total 64 database storage object(s) and 23 database code object(s) in the source database, we identified 63 (98%) database storage object(s) and 22 (96%) database code object(s) that can be converted to Amazon RDS for PostgreSQL automatically or with minimal changes.

The target database version is less than PostgreSQL 11.1 (11.1). The converted code might not work properly.

- 1 (2%) database storage object(s) require 1 significant user action(s) to complete the conversion.
- 1 (4%) database code object(s) require 1 medium user action(s) to complete the conversion.



### • AWS SCT 를 소스/타겟 데이터베이스에 연결

- Assessment report 생성
- Executive Summary 일기
- 상세 지침에 따라 수행

#### Database migration assessment report

aws

ource database:

Oracle Database 12c Enterprise Edition 12.1.0.2.0 (64bit Production), Enterprise edition

#### Package Procedure Changes

Not all package procedures can be converted automatically. Youll need to address these issues manually.

Issue 5584: Converted functions depends on the time zone settings

Recommended action: Review the transformed code, and set time zone manually if necessary.

Issue code: 5584 | Number of occurrences: 2 | Estimated complexity: Simple

Documentation references: http://www.postgresql.org/docs/9.6/static/functions-datetime.html

Schemas.DMS\_SAMPLE.Packages.TICKETMANAGEMENT.Private procedures.TRANSFERTICKET: 1463:1469 Schemas.DMS\_SAMPLE.Packages.TICKETMANAGEMENT.Public procedures.SELLTICKETS: 1772:1778

#### Package Function Changes

Not all package functions can be converted automatically. Youll need to address these issues manually.

Issue 5644: Unable automatically convert assign operation of array or global nested table,

Recommended action: Perform a manual conversion.

Issue code: 5644 | Number of occurrences: 1 | Estimated complexity: Medium

Schemas.DMS\_SAMPLE.Packages.TICKETMANAGEMENT.Private functions.GET\_OPEN\_EVENTS: 270:297

#### **Procedure Changes**

Not all procedures can be converted automatically. Youll need to address these issues manually.

Issue 5103: Unable to convert hints

Recommended action: Use PostgreSQL methods for performance tuning.

Issue code: 5103 | Number of occurrences: 2 | Estimated complexity: Simple

Documentation references: http://www.postgresgl.org/docs/9.6/static/gego.html

Schemas.DMS\_SAMPLE.Procedures.GENERATESEATS: 2041:2089 Schemas.DMS\_SAMPLE.Procedures.GENERATE\_TICKETS: 290:1184



## **AWS Database Migration Service**



### **AWS DMS**



- 10분 이내에 첫 마이그레이션을 시작하세요
- 마이그레이션 중에도 앱은 계속 실행하세요
- AWS 내에서, AWS로 또는
   AWS에서 데이터 복제
- 동종 또는 이종 데이터베이스 엔진으로 데이터 이관

소 스	타 겟
Oracle	Oracle
SQL Server	SQL Server
Azure SQL	PostgreSQL
PostgreSQL	MySQL
MySQL	Amazon Redshift
SAP ASE	SAP ASE
MongoDB	Amazon S3
Amazon S3	Amazon DynamoDB
IBM DB2	Amazon Kinesis
	Amazon ElasticSearch



### DMS 구성요소

- Replication instance
- Endpoint
- Task

### **Dedicated EC2 instance**

• T3, C5/6, or R5/6 instance types

### **Public or private IP address**

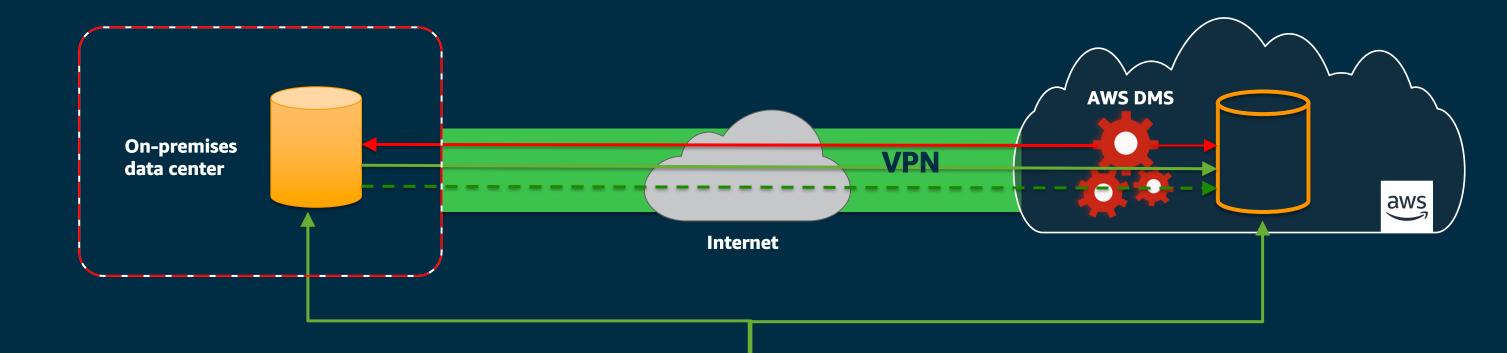
- Private: Source or Target inside VPC / VPN / DC
- Public: Source or Target outside VPC

### **Task execution**

- Moves the data from the source to the target
- Support for multiple tasks



## 데이터 마이그레이션 과정



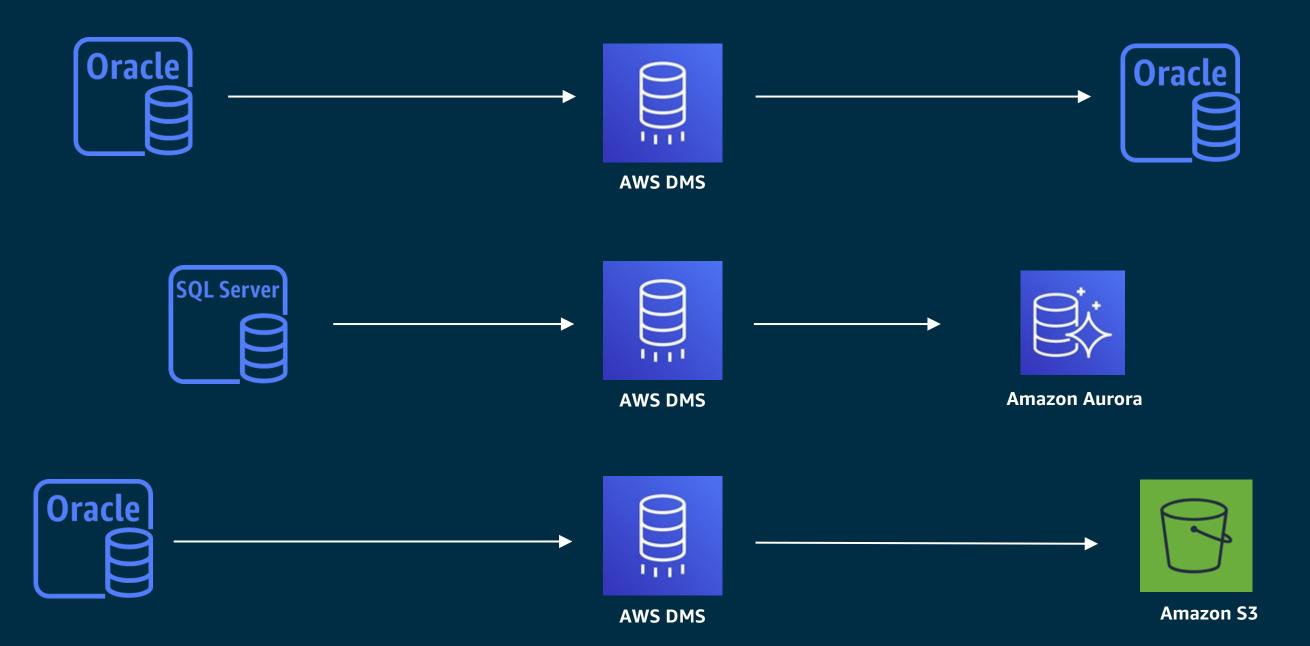
- 복제 인스턴스 시작
- 소스 및 타겟 데이터베이스에 연결
- 대상 테이블, 스키마 또는 데이터베이스 선택



- AWS DMS 가 데이터를 로드하고 동기화된 상태로 유지
- 동기화된 후 원할 때에 타겟으로 애플리케이션을 스위칭

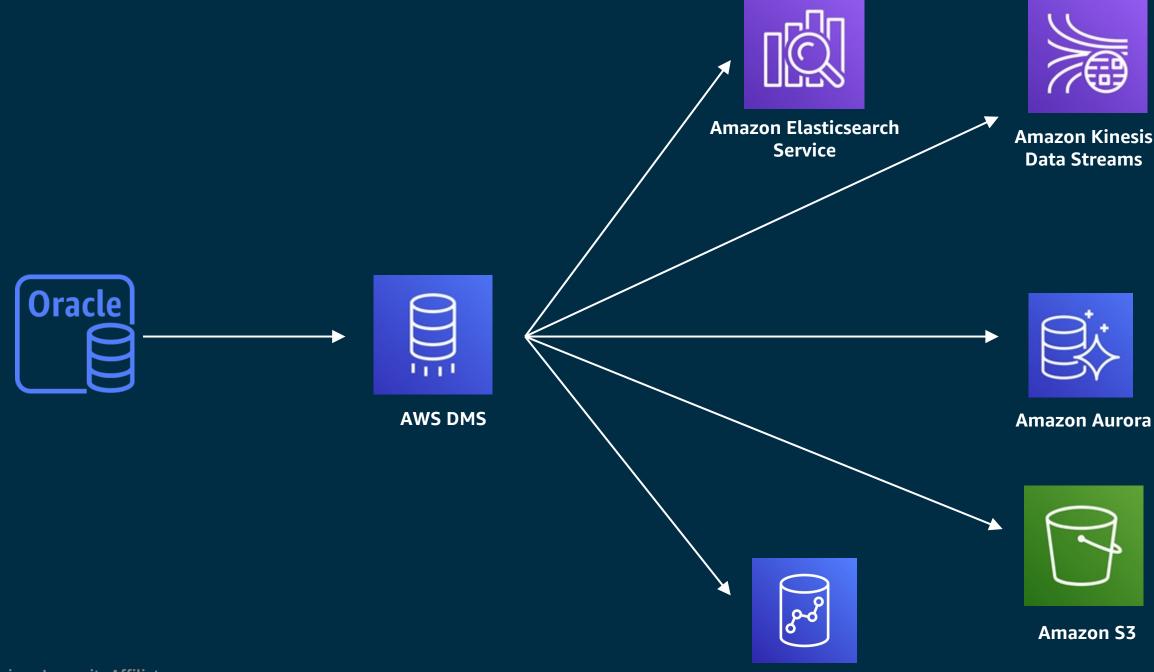


## 동종DB간 또는 이종DB간 데이터 이관





## Fan out: Microservices architecture 목적별 DB로 분화



**Amazon Redshift** 



## 지원하는 소스, 타겟

#### **NOSQL** RELATIONAL **DATA WAREHOUSE** Oracle Oracle SQL Server SQL Server Oracle Oracle SQL Server SQL Server PostgreSQL MongoDB Amazon S3 SOURCES MySQL MySQL Cassandra MariaDB MariaDB Teradata **AWS Snowball** IBM Db2 LUW IBM Db2 z/OS **Azure Synapse** Vertica **DocumentDB** GCP MySQL **SQL Azure** Amazon SQL Server SQL Server PostgreSQL Oracle **OpenSearch Amazon Redshift DynamoDB Service TARGETS Amazon Kinesis** MySQL MySQL Amazon MariaDB MariaDB Amazon Aurora **DocumentDB Data Streams** SAP ASE **Amazon Neptune Amazon S3 Amazon ElastiCache** Managed **Streaming for** Kafka

### AWS DMS 를 사용한 65 만건의 데이터베이스 이관사례



SAMSUNG

















































**Expedia** 























































































# Thank you!

