

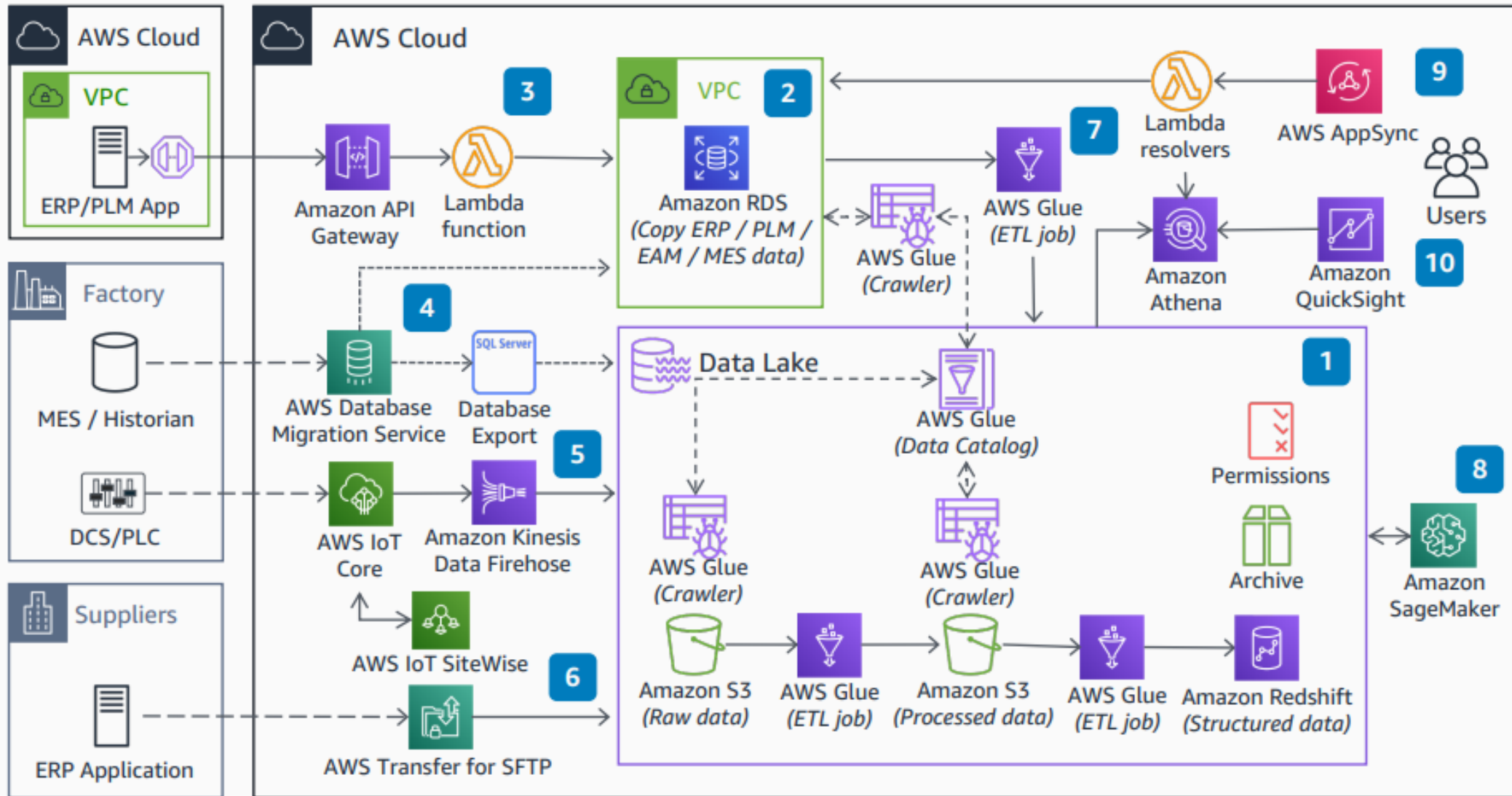
AWS Storage

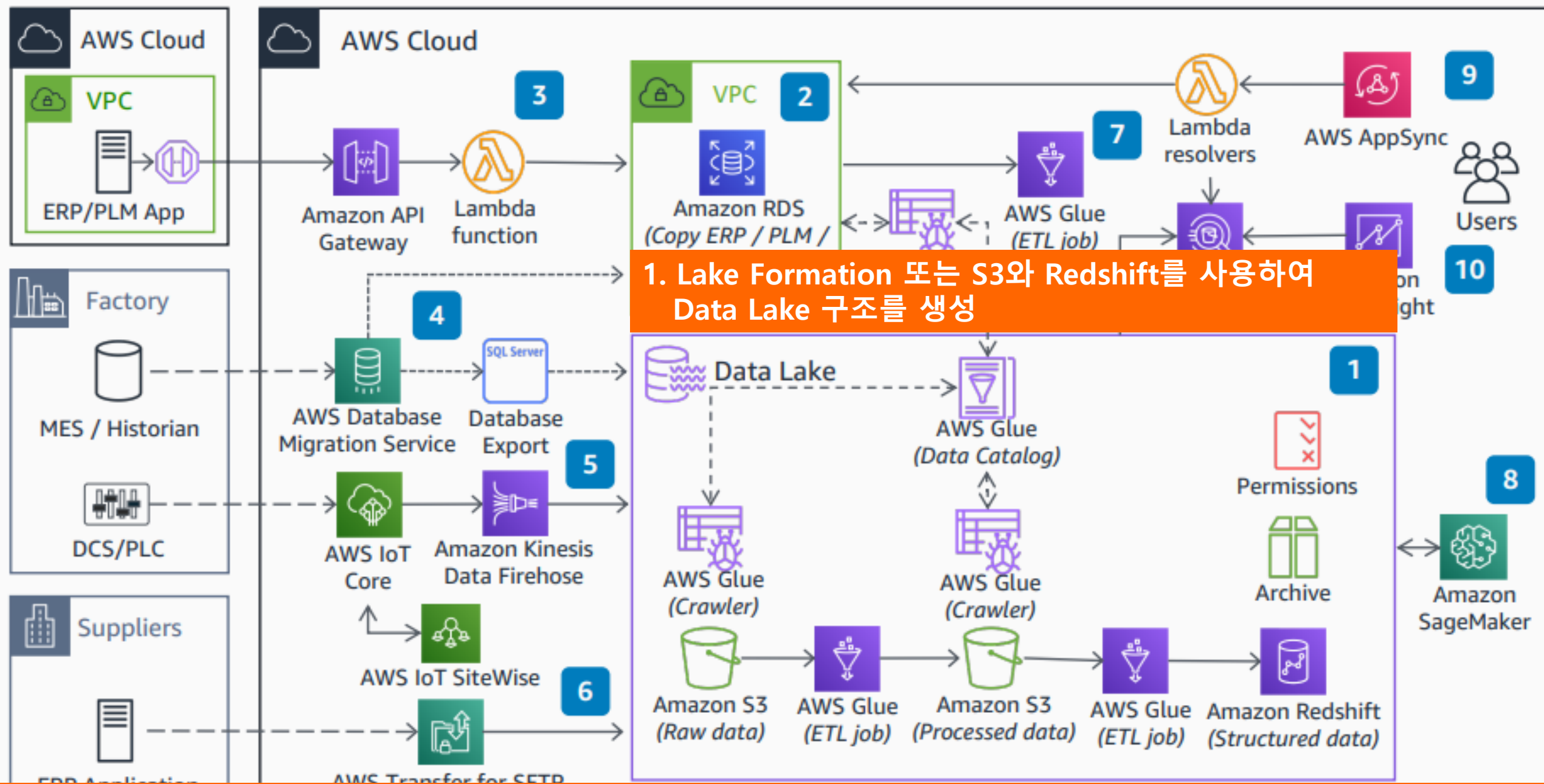
박경규



1. Reference Architecture Review

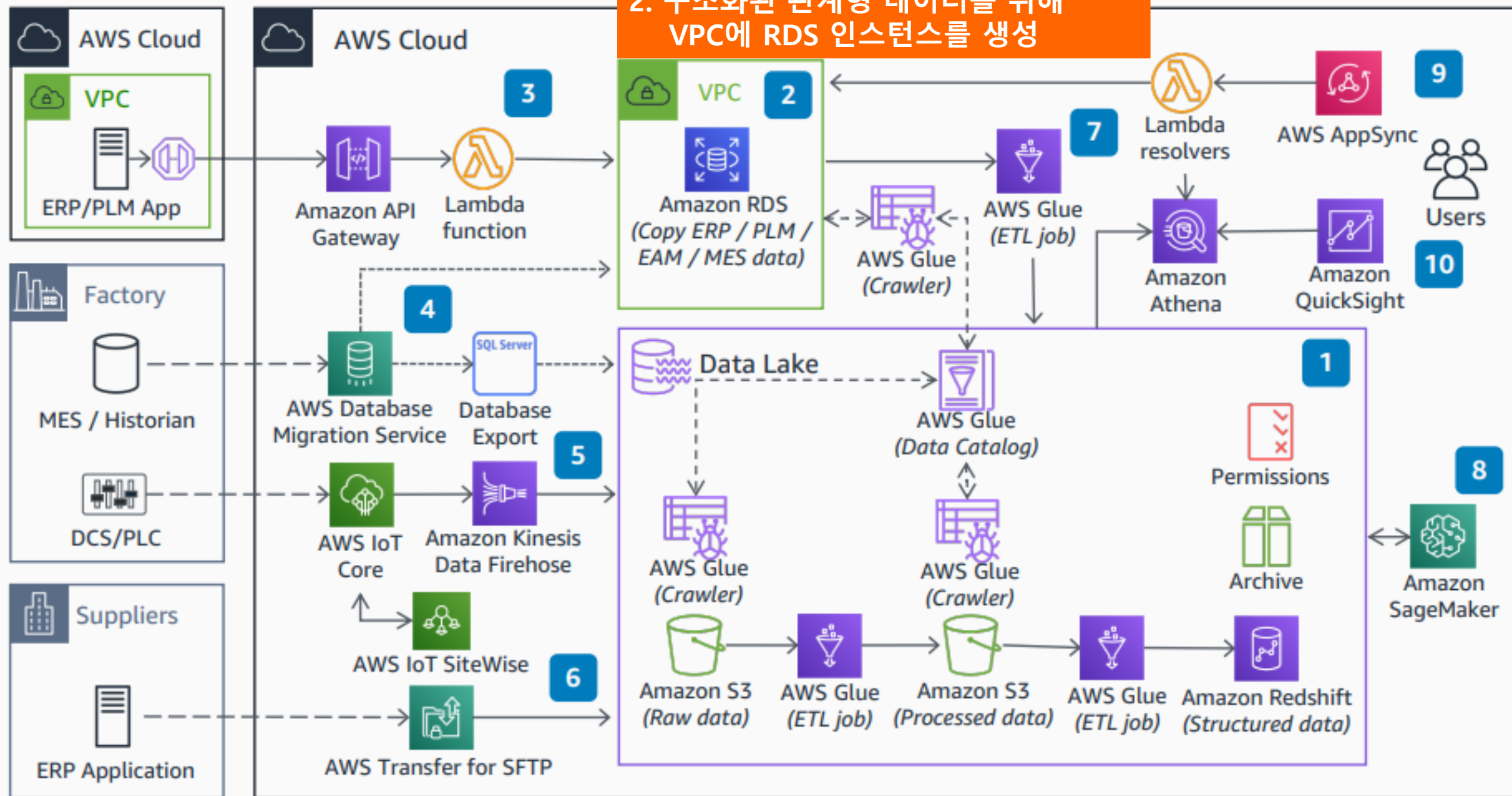
산업 분야 Data Lake 구축 모델 : IoT데이터/생산관리데이터 프로세싱 → 빅데이터 분석 → 머신러닝 모델 학습/배포



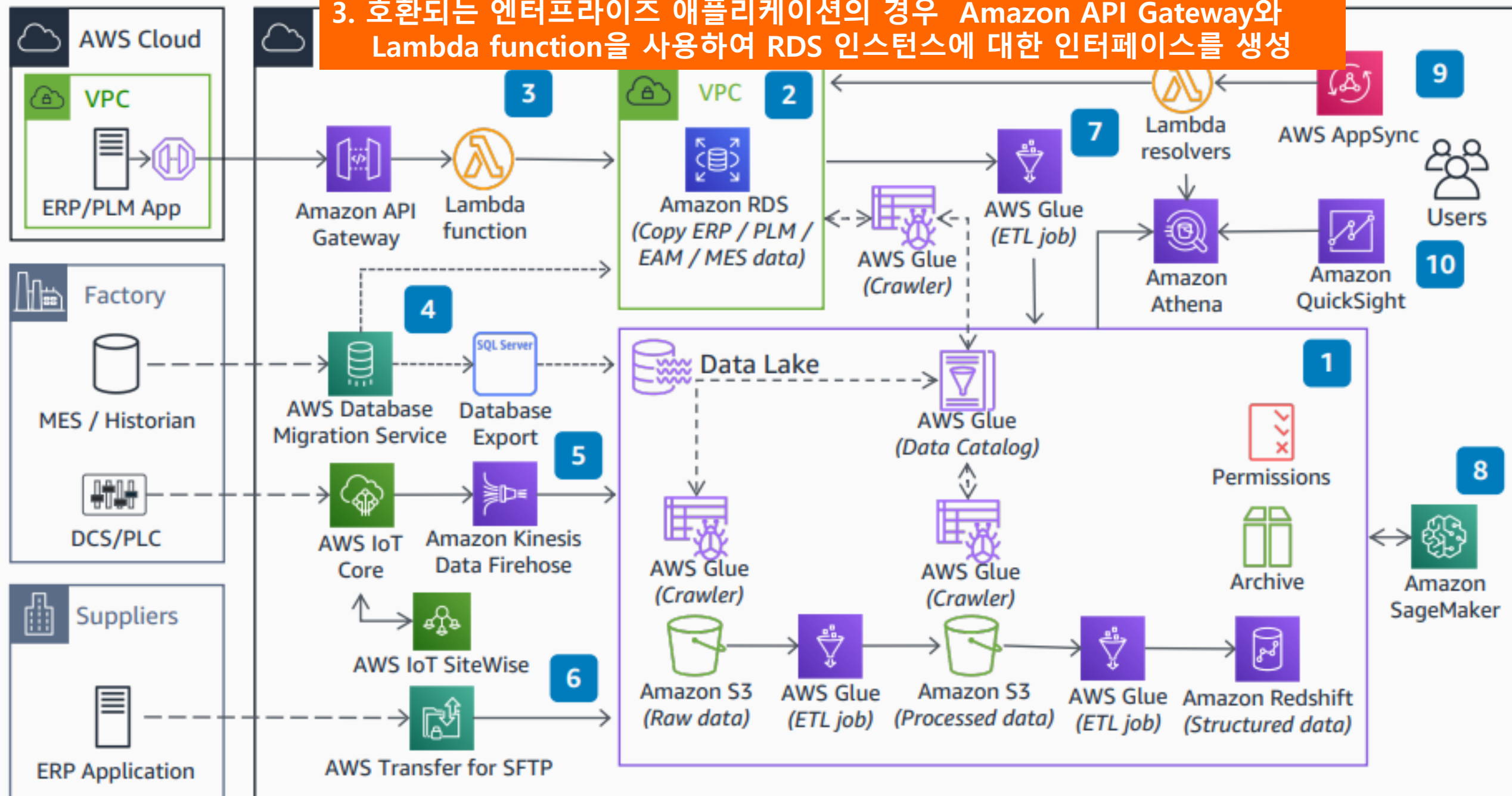


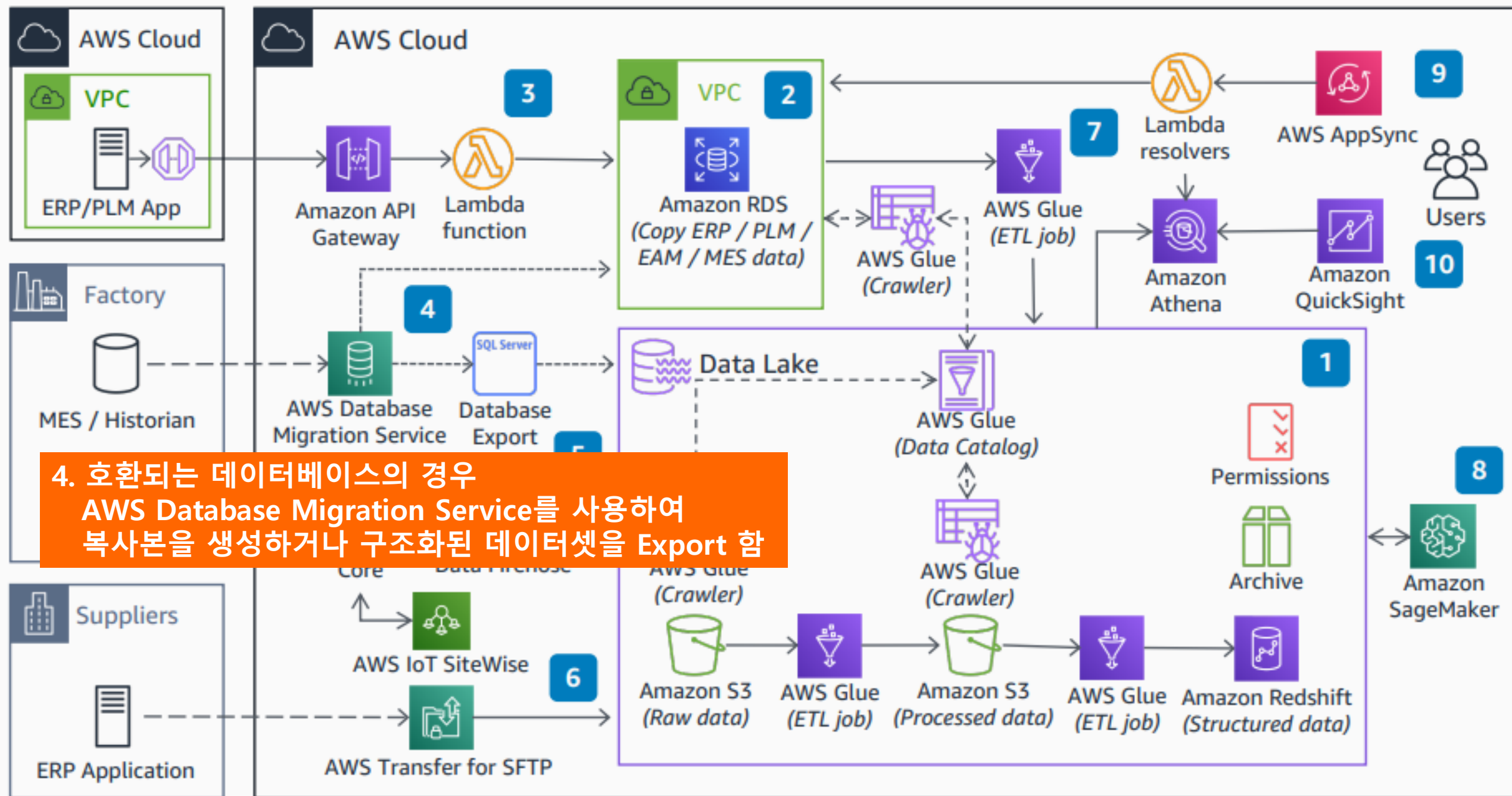
Data Lake : 조직에서 수집한 정형·반정형·비정형 데이터를 원시 형태(raw data)로 저장하는 단일한 데이터 저장소.

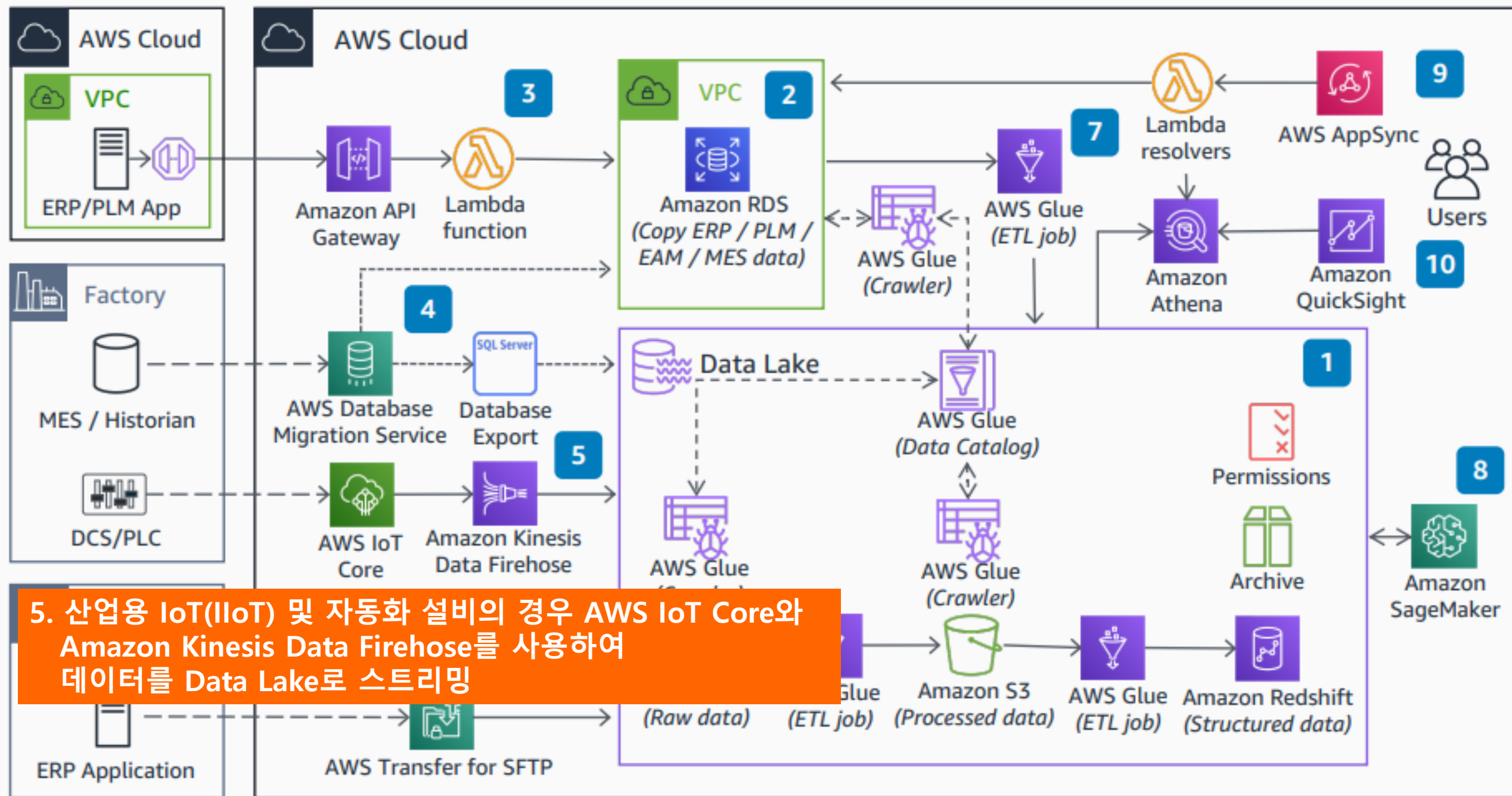
2. 구조화된 관계형 데이터를 위해
VPC에 RDS 인스턴스를 생성

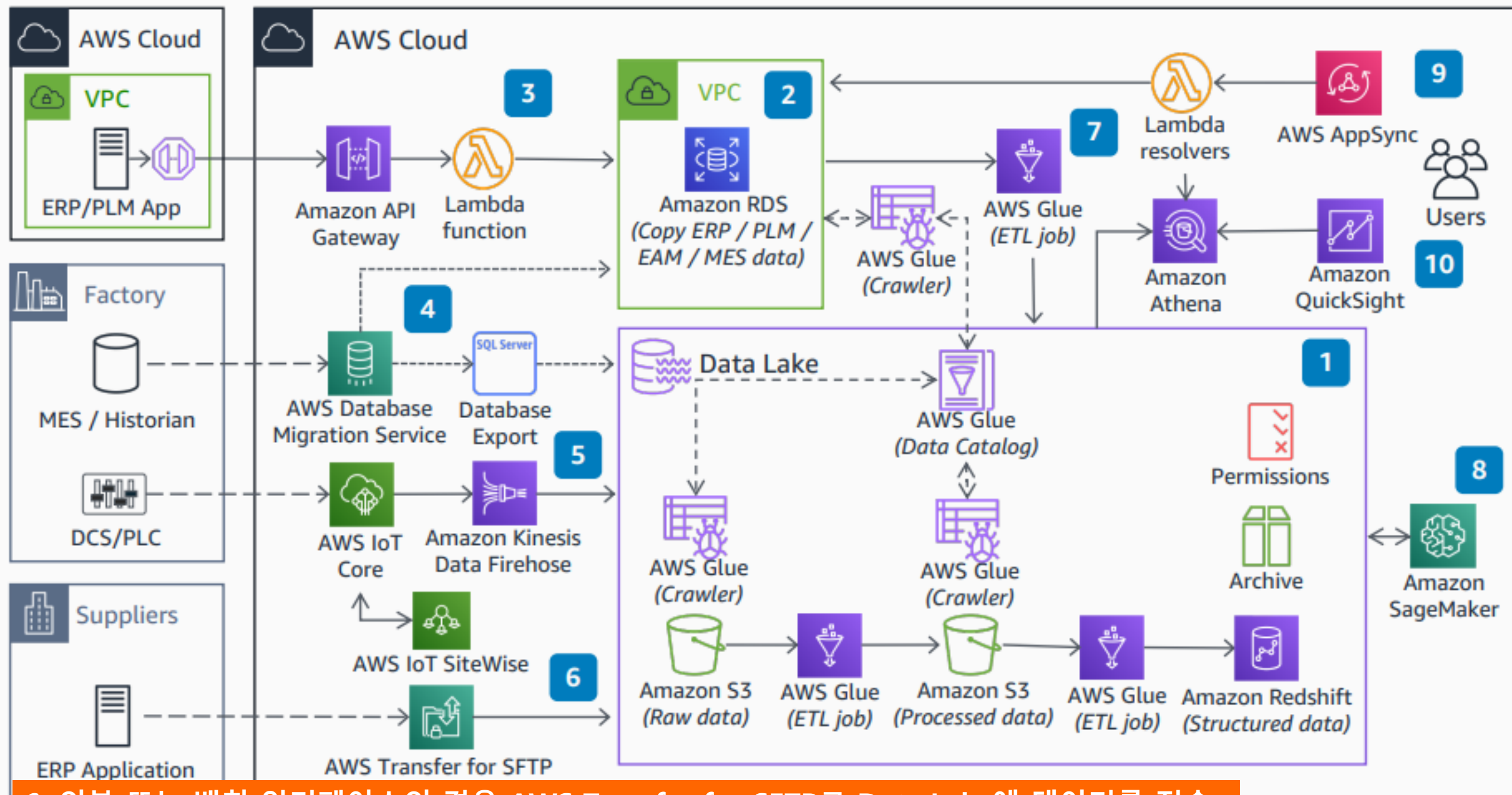


3. 호환되는 엔터프라이즈 애플리케이션의 경우 Amazon API Gateway와 Lambda function을 사용하여 RDS 인스턴스에 대한 인터페이스를 생성

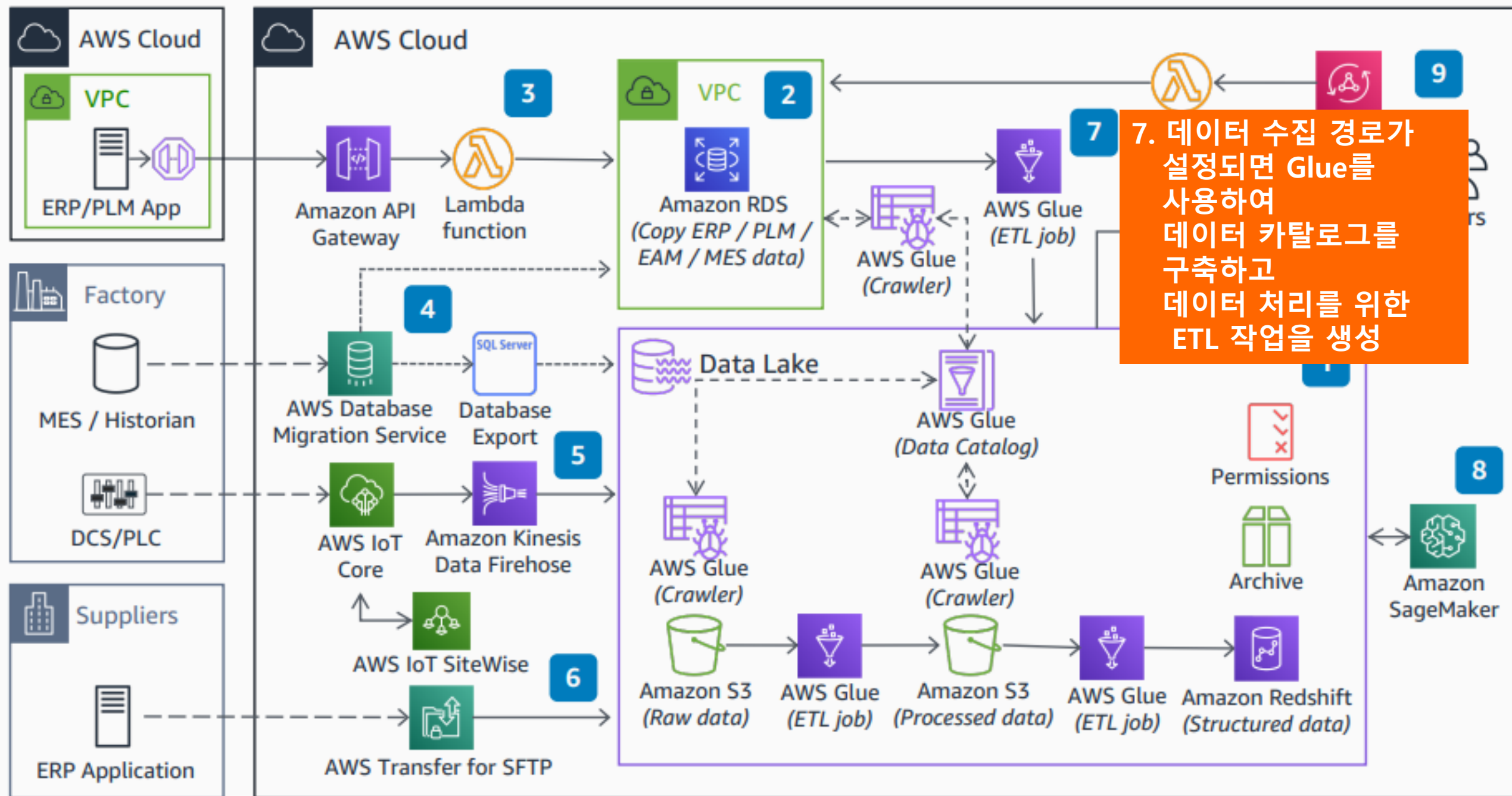


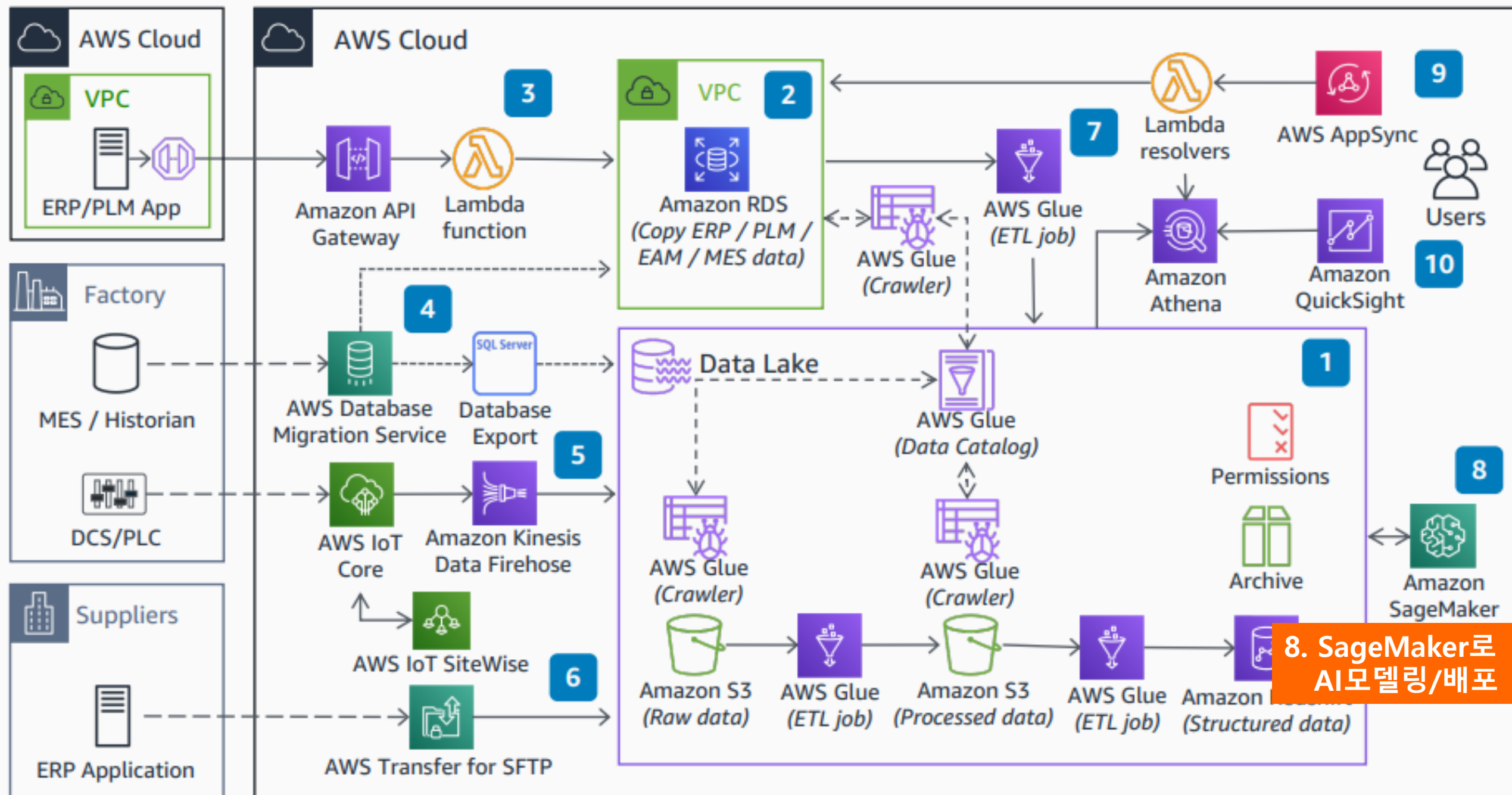


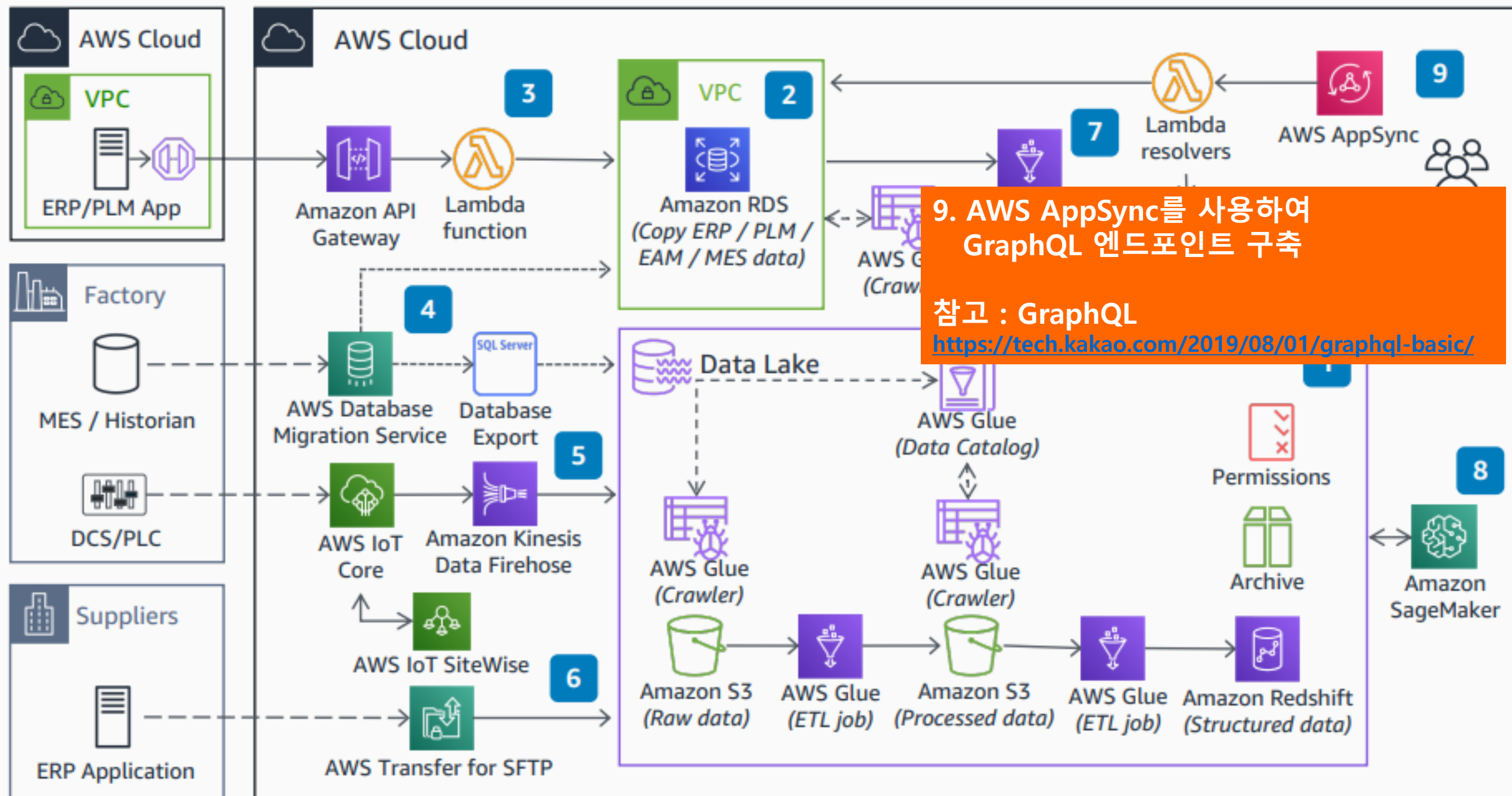




6. 외부 또는 배치 인터페이스의 경우 AWS Transfer for SFTP로 Data Lake에 데이터를 전송

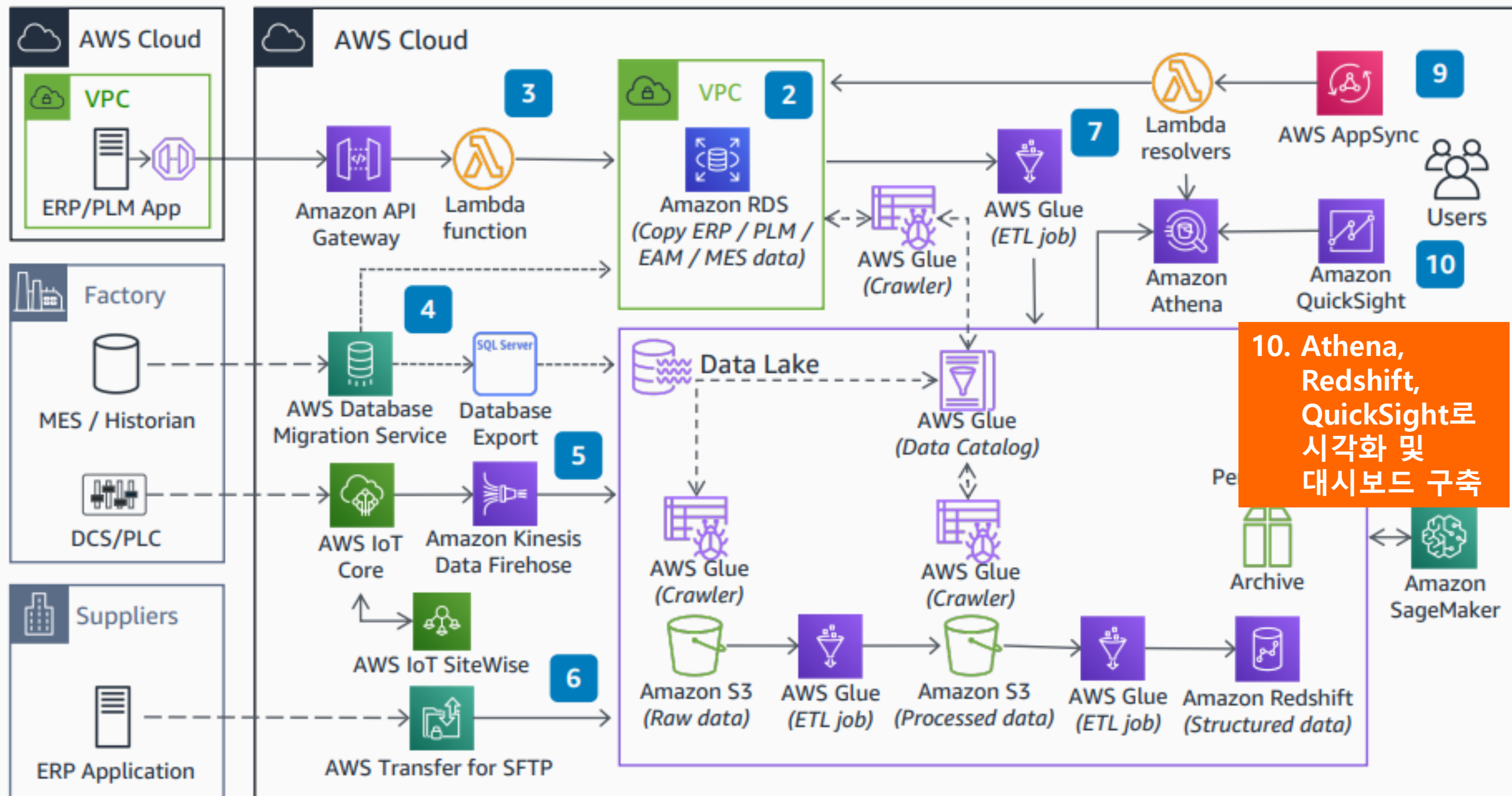






9. AWS AppSync를 사용하여 GraphQL 엔드포인트 구축

참고 : GraphQL
<https://tech.kakao.com/2019/08/01/graphql-basic/>



1. Reference Architecture Review : Storage 활용 예

■ AWS Glue 활용한 이벤트 주도 ETL 파이프라인 구성

- Lambda 함수를 통해 ETL 작업을 트리거하여 S3에 새로운 데이터가 저장되는 대로 ETL 작업이 실행되도록 할 수 있음
- 추출 : AWS Glue 데이터 카탈로그에 등록하여 모든 AWS 데이터 세트에서 검색
- 변환 : 끌어서 놓기 작업 에디터에서 ETL 프로세스를 정의하면 AWS Glue가 자동으로 코드를 생성하므로 데이터를 추출, 변환 및 로드. 이동 중인 데이터를 정리 및 변환하고, 대상 데이터 스토어에서 몇 초 안에 분석에 해당 데이터를 사용할 수 있도록 지원
- 로드 : AWS Glue 스트리밍 ETL 작업은 데이터를 보강하고 집계하며, 배치 및 스트리밍 소스를 조인하고, 다양하고 복잡한 분석 및 기계 학습 작업을 실행. AWS Glue Interactive Sessions는 작업 개발을 위한 서버리스 기능으로, 데이터 통합 작업 개발을 간소화

■ AWS Glue 데이터 카탈로그 : 모든 AWS 데이터 세트에서 검색

- 새로운 데이터소스를 AWS Glue 데이터 카탈로그에 등록 시 데이터를 이동하지 않고도 여러 AWS 데이터 세트 전체에서 데이터를 검색

■ AWS Glue Studio

- Glue ETL 작업을 시각적으로 쉽게 생성, 실행 및 모니터링.

■ AWS Glue DataBrew : 시각적 인터페이스를 사용하여 코드 없이 데이터 정규화

- 코드를 작성하지 않고도 데이터를 정리하고 정규화할 수 있도록 포인트 & 클릭 방식의 대화식 UI 제공
- Amazon S3, Amazon Redshift, Amazon Aurora 및 Amazon RDS를 비롯한 데이터 레이크, 데이터 웨어하우스 및 데이터베이스에서 직접 데이터를 시각화, 정리 및 정규화
 - Spoke VPC RT 분석



2. Hands-on Lab(HoL) : 웹애플리케이션

▼ 심화 모듈 - 웹 애플리케이션

▶ 네트워크 - Amazon VPC

▼ 컴퓨트 - Amazon EC2

웹 서버 인스턴스의 시작

오토 스케일링 웹 서비스 배포

웹 서비스 확인 및 테스트

▼ 데이터베이스 - Amazon Aurora

VPC 보안 그룹 생성

RDS 인스턴스 생성

RDS 크레덴셜 저장

웹앱 서버와 RDS 연결

(옵션) RDS 관리 기능

(옵션) RDS Aurora 연결

▼ 스토리지 - Amazon S3

S3에 Bucket 생성

버킷에 오브젝트 추가하기

오브젝트 보기

정적 웹 사이트 호스팅 사용

오브젝트 이동

버킷 버저닝 활성화

오브젝트 및 버킷 삭제

<http://web-alb-606631940.ap-northeast-2.elb.amazonaws.com/>

powered by aws

LOAD TEST	RDS
Meta-Data	Value
InstanceId	i-0f7f05f68015
Availability Zone	ap-northeast-2

Current CPU Load: 0%

powered by aws

LOAD TEST RDS

Address Book

Name	Phone	Email	Admin
Alice	571-555-4875	alice@address2.us	Edit Remove
Bob	630-555-1254	bob@fakeaddress.com	Edit Remove

[Add Contact](#)

[Load Initial Data](#)

2. Hands-on Lab(HoL) : 정적 웹 사이트 호스팅

▼ 심화 모듈 - 웹 애플리케이션

▶ 네트워크 - Amazon VPC

▼ 컴퓨트 - Amazon EC2

웹 서버 인스턴스의 시작

오토 스케일링 웹 서비스 배포

웹 서비스 확인 및 테스트

▼ 데이터베이스 - Amazon Aurora

VPC 보안 그룹 생성

RDS 인스턴스 생성

RDS 크레덴셜 저장

웹앱 서버와 RDS 연결

(옵션) RDS 관리 기능

(옵션) RDS Aurora 연결

▼ 스토리지 - Amazon S3

S3에 Bucket 생성

버킷에 오브젝트 추가하기

오브젝트 보기

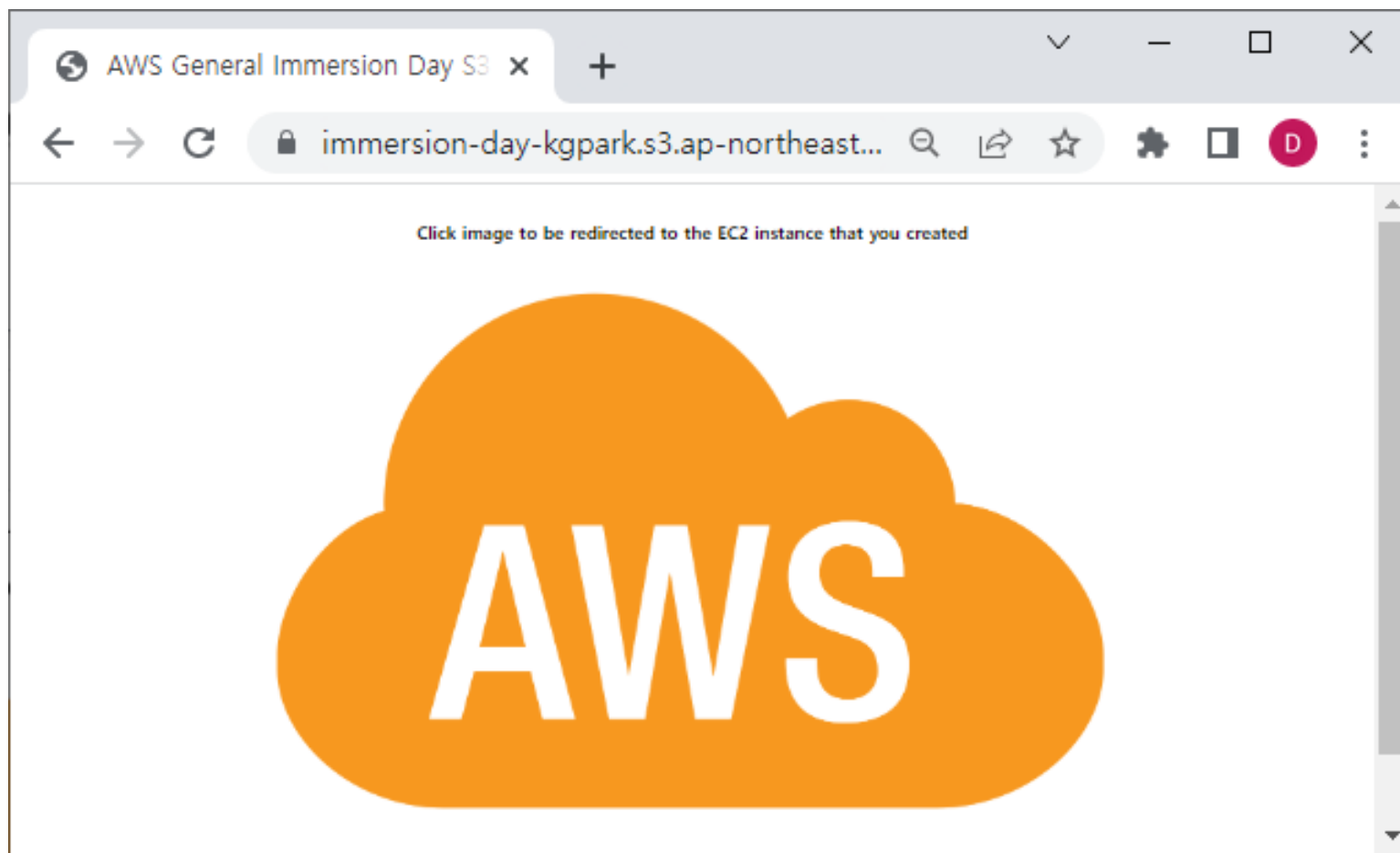
정적 웹 사이트 호스팅 사용

오브젝트 이동

버킷 버저닝 활성화

오브젝트 및 버킷 삭제

<http://immersion-day-kgpark.s3-website.ap-northeast-2.amazonaws.com>

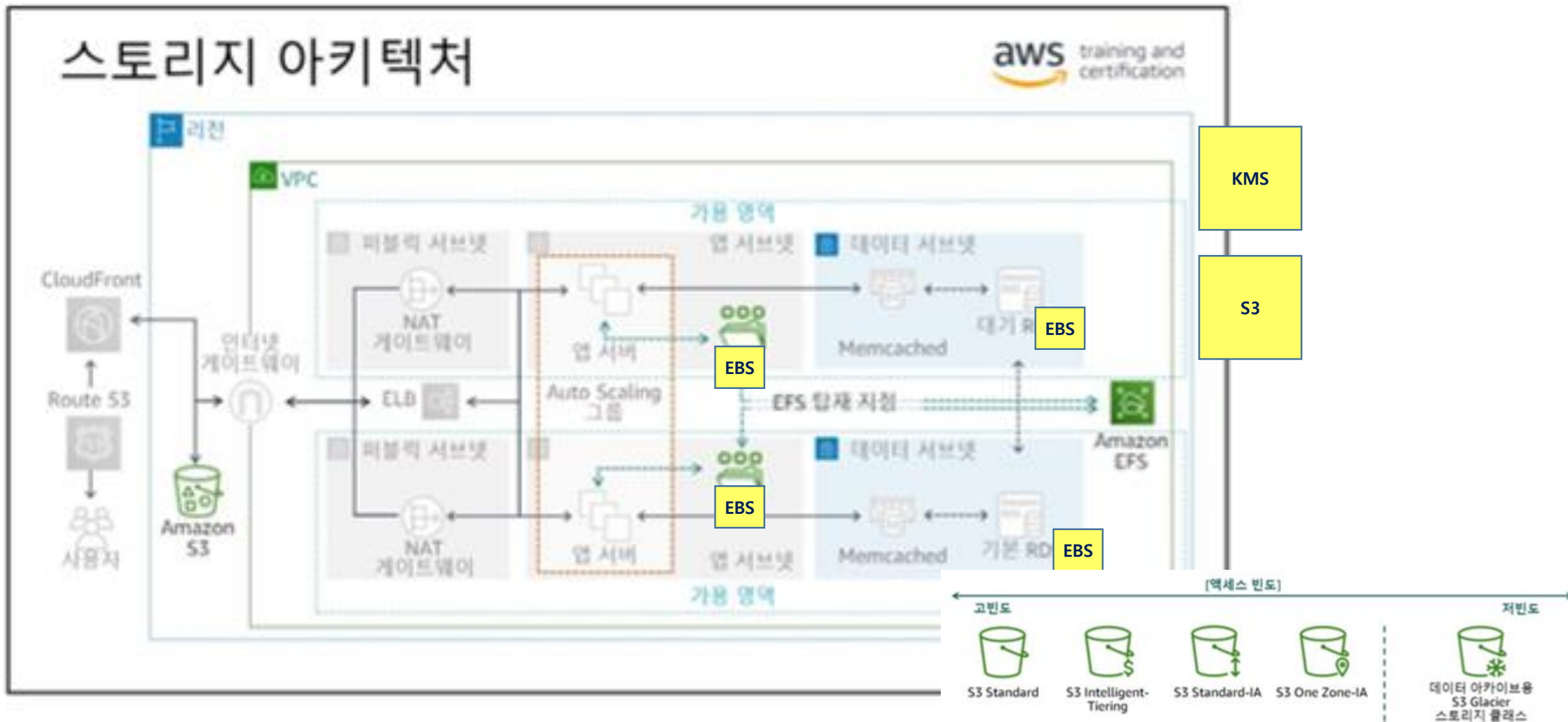


3. Design Architecture

■ Condition

- 두 가용 영역(Availability Zone)에 이중화된 3-tier Web Application 구성
[Public Subnet]Web → [Private Subnet]Application → [Private Subnet]DB
- 각 Tier는 Subnet으로 구분되며 포함하는 Instance들의 Storage로 EBS(Elastic Block Store)를 배치하고
각 EBS에 대한 Snapshot 을 S3로 Back-up하는 흐름을 표현
+ Database의 Storage:EBS에 대한 Snapshot은 AWS KMS(Key Management Service)를 사용하여 암호화, Back-up
- S3의 EBS 증분 Back-up은 일정 기간 후 Archiving to S3 Glacier

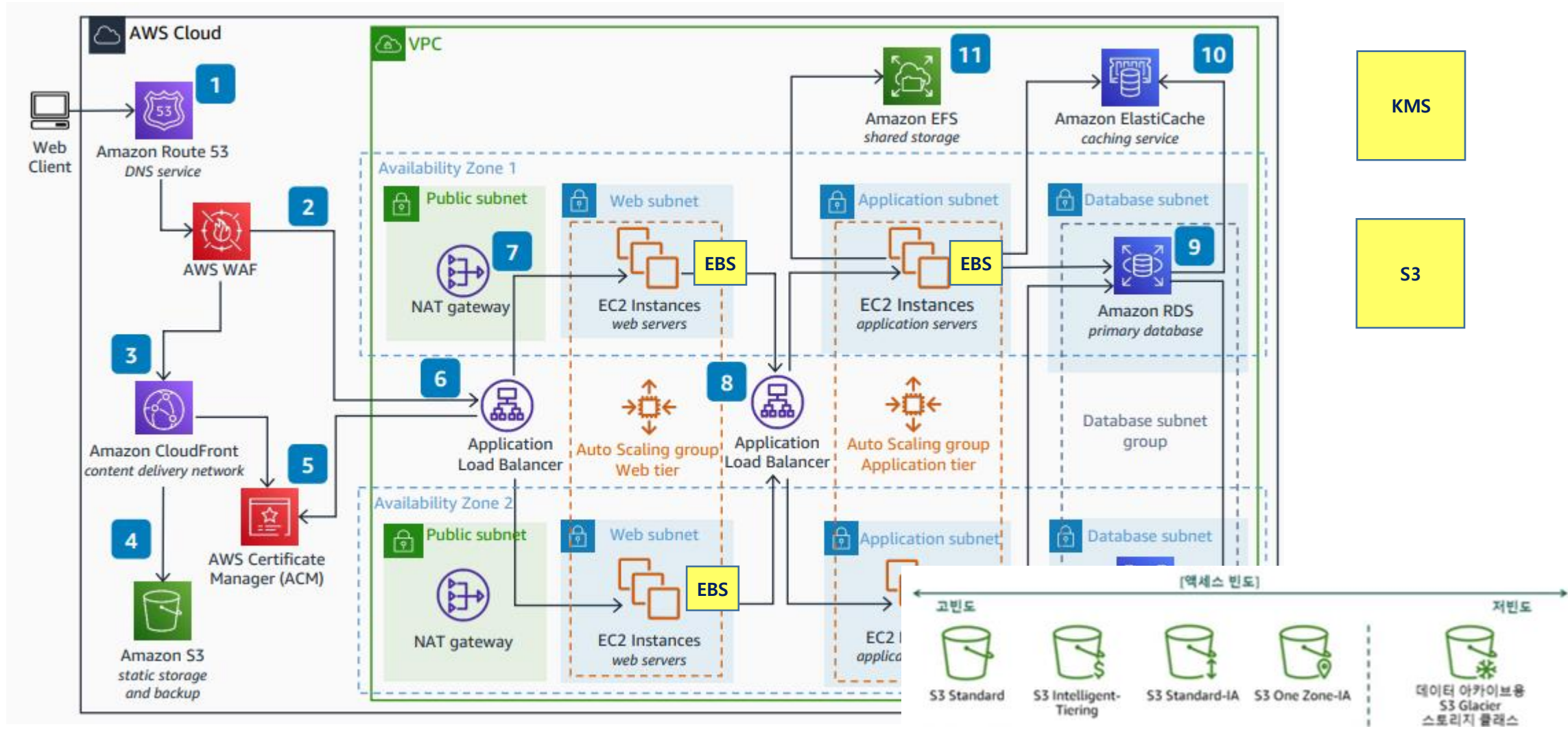
3. Design Architecture



3. Design Architecture

https://d1.awsstatic.com/architecture-diagrams/ArchitectureDiagrams/web-application-architecture-on-aws-ra.pdf?did=wp_card&trk=wp_card

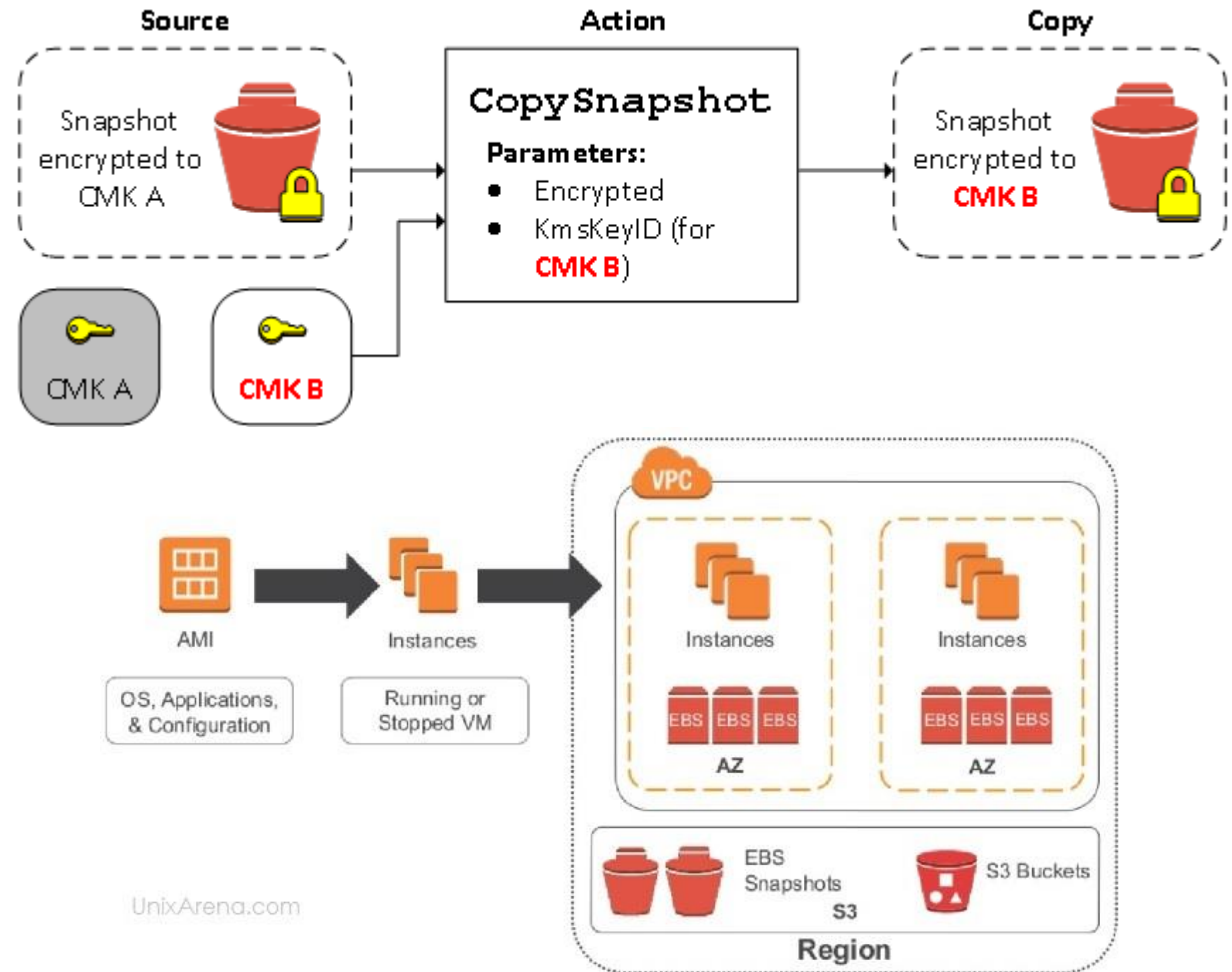
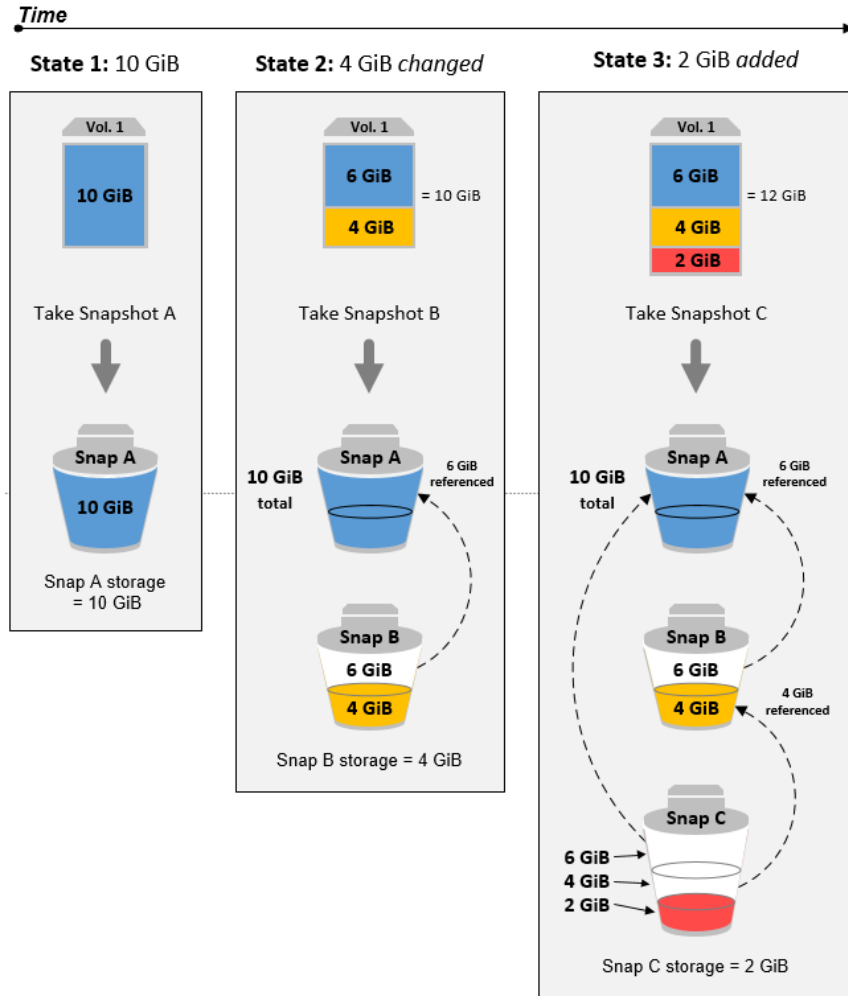
<https://aws.amazon.com/ko/architecture/reference-architecture-diagrams/>



3. Design Architecture : 증분백업 관련 자료

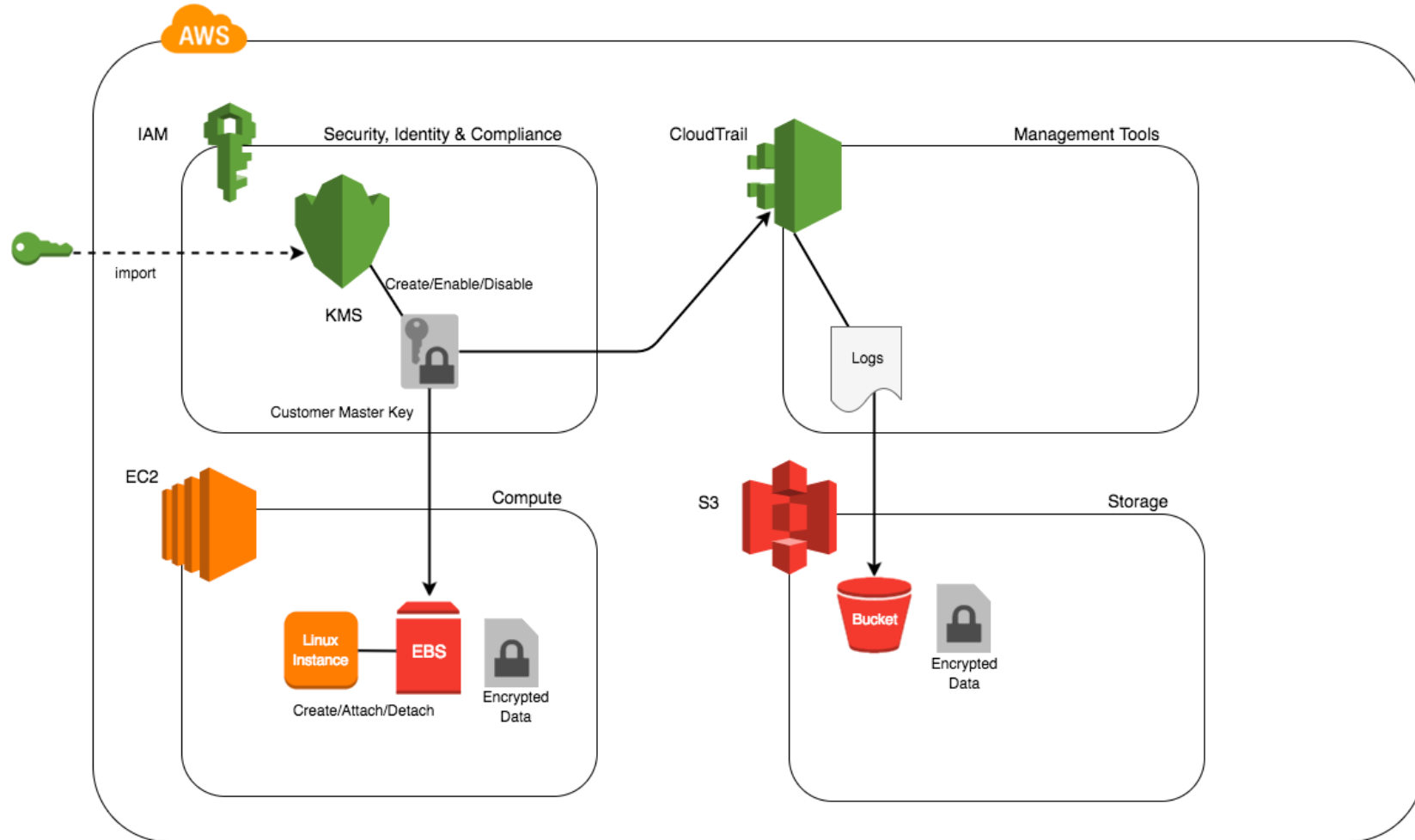
https://docs.aws.amazon.com/ko_kr/AWSEC2/latest/UserGuide/EBSSnapshots.html

<https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/EBSEncryption.html>



3. Design Architecture : 증분백업 관련 자료

<https://cloudacademy.com/lab/using-amazon-key-management-service-encrypt-s3-and-ebs-data/>



Thank you