



Amazon Redshift

학습 내용

강의의 핵심

배울 내용은 다음과 같습니다.

- 데이터 웨어하우스 서비스인 Amazon Redshift에 대해 자세히 설명하기
- Amazon Redshift의 주요 기능 살펴보기

주요 용어:

Amazon Redshift



Amazon Redshift는 클라우드 내에서 사용되는 페타바이트 규모의 완전관리형 데이터 웨어하우스 서비스입니다.

페타바이트(PB) 및 데이터 웨어하우스 이해

페타바이트 및 데이터 웨어하우스

페타바이트란 무엇입니까?

- 페타바이트는 10^{15} 바이트의 정보입니다.
- 1페타바이트는 대략 1,000테라바이트(TB)와 같습니다.
- 1테라바이트는 대략 1,000기가바이트(GB)와 같습니다.

데이터 웨어하우스란 무엇입니까?

- 데이터 웨어하우스는 보다 정보에 입각한 의사 결정을 내릴 수 있도록 분석 가능한 정보의 중앙 리포지토리입니다.
- 데이터 웨어하우스는 보고서, 대시보드 및 분석 도구를 강화합니다. 데이터 웨어하우스를 통해 사용자는 동시에 수백, 수천 명의 사용자에게 신속하게 결과를 쿼리할 수 있습니다.

데이터 웨어하우스의 작동 방식 및 이점

데이터 웨어하우스의 작동 방식

- 데이터 웨어하우스에는 여러 개의 데이터베이스가 포함될 수 있습니다.
- 각 데이터베이스 내에서 데이터는 테이블 및 열로 구성됩니다.
- 각 열 내에서 사용자는 정수, 데이터 필드 또는 문자열 등 데이터의 설명을 정의할 수 있습니다.
- 테이블은 스키마 내부에 구성될 수 있으며, 이는 폴더와 같은 역할을 합니다.

데이터 웨어하우스의 이점

- 정보에 기반한 의사 결정
- 여러 소스의 데이터 통합
- 과거 데이터 분석
- 데이터 품질, 일관성 및 정확성
- 트랜잭션 데이터베이스와 분석 처리를 분리하여 두 시스템 모두의 성능을 향상

Amazon Redshift 아키텍처

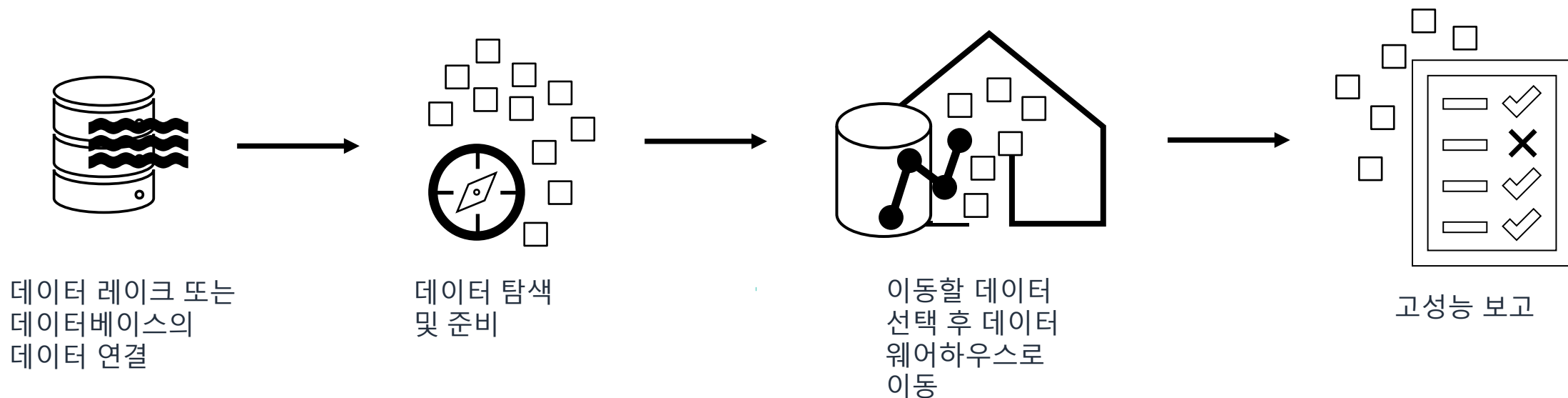
데이터 웨어하우스는 어떻게 설계되니까?

- 데이터 웨어하우스 아키텍처는 티어로 구성됩니다.
- 상단 티어는 보고, 분석 및 데이터 마이닝 도구를 통해 결과를 제시하는 프런트 엔드 클라이언트입니다.
- 중간 티어는 데이터를 액세스하고 분석하는데 사용되는 분석 엔진으로 구성됩니다.
- 아키텍처의 하단 티어는 데이터가 로드되고 저장되는 데이터베이스 서버입니다.



전체 작동 원리

일반적으로 비즈니스는 데이터베이스, 데이터 레이크 및 데이터 웨어하우스의 조합을 활용하여 데이터를 저장 및 분석합니다.



데이터베이스 또는 데이터 레이크에 데이터를 저장하고 데이터를 준비하며 선택한 데이터를 데이터 웨어하우스로 이동한 후 보고를 수행합니다.

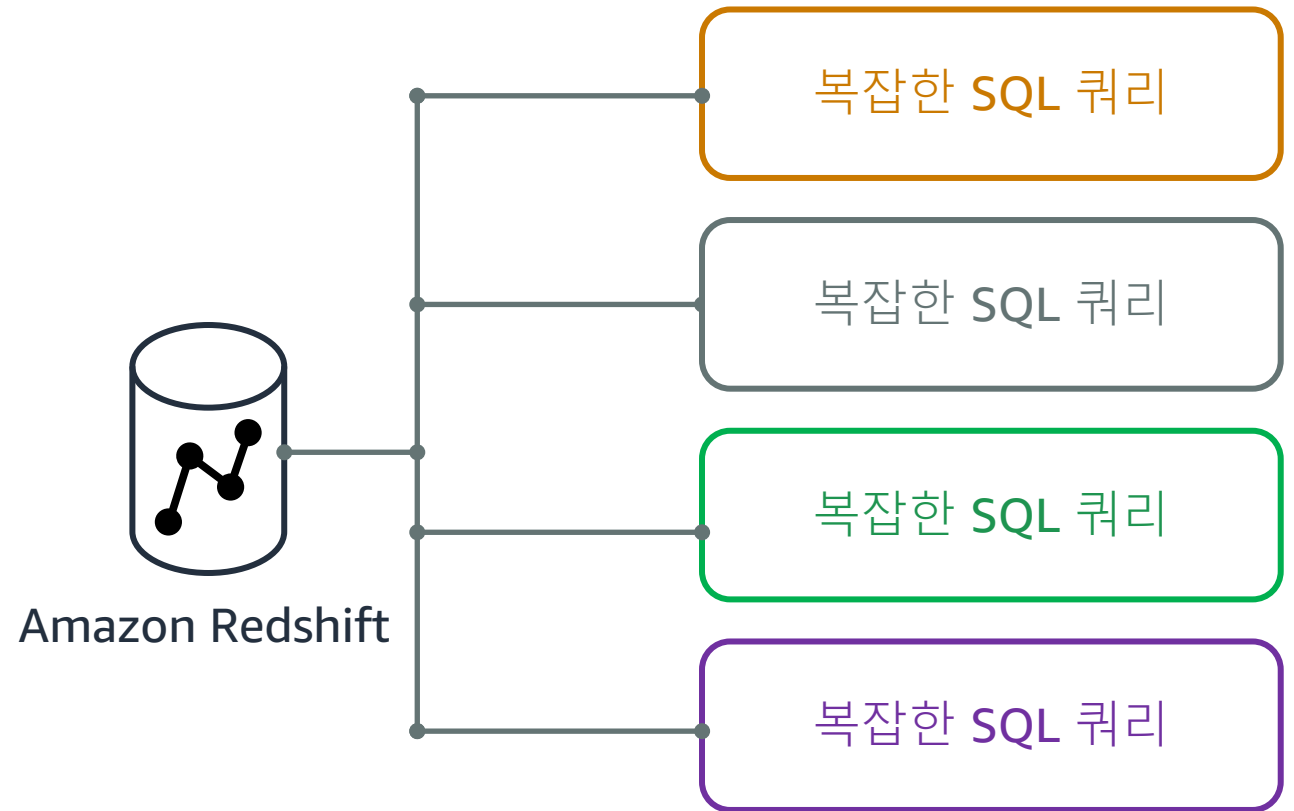


Amazon Redshift란 무엇입니까?

Amazon Redshift

주요 기능

- Amazon Redshift는 빠르고 강력합니다.
- Amazon Redshift는 완전관리형 데이터 웨어하우스입니다.
- Amazon Redshift를 사용하면 페타바이트 규모의 정형 데이터에 대해 복잡한 분석 쿼리를 실행할 수 있습니다.
- Amazon Redshift는 정교한 쿼리 최적화, 고성능 로컬 디스크의 열 기반 스토리지 및 병렬 쿼리 실행을 사용합니다.

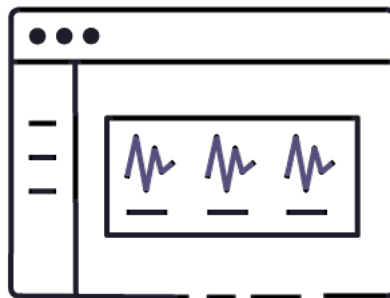


Amazon Redshift 주요 기능



관리

완전관리형 데이터
웨어하우스



암호화

저장 데이터와 전송
데이터 모두 암호화



호환성

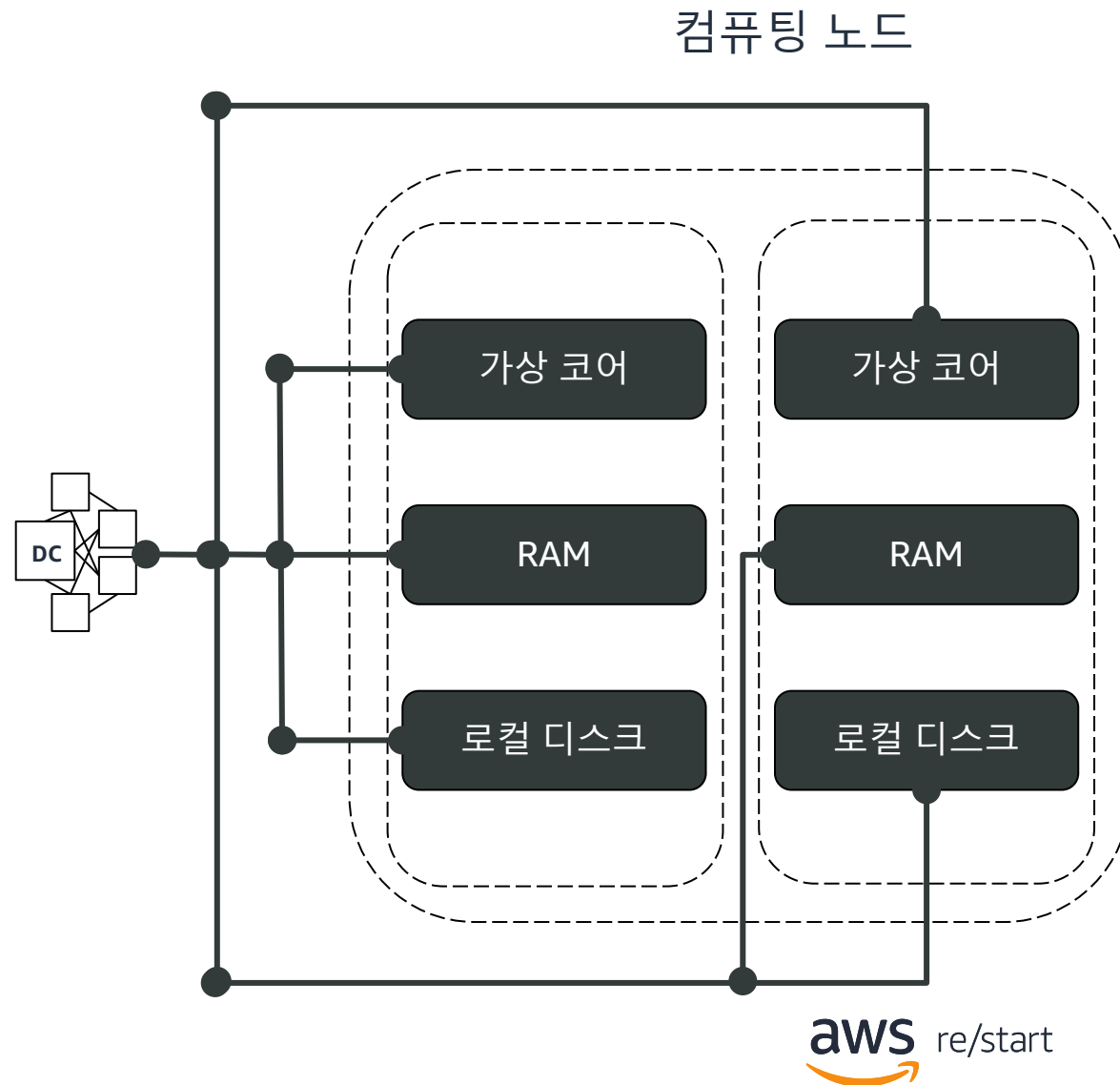
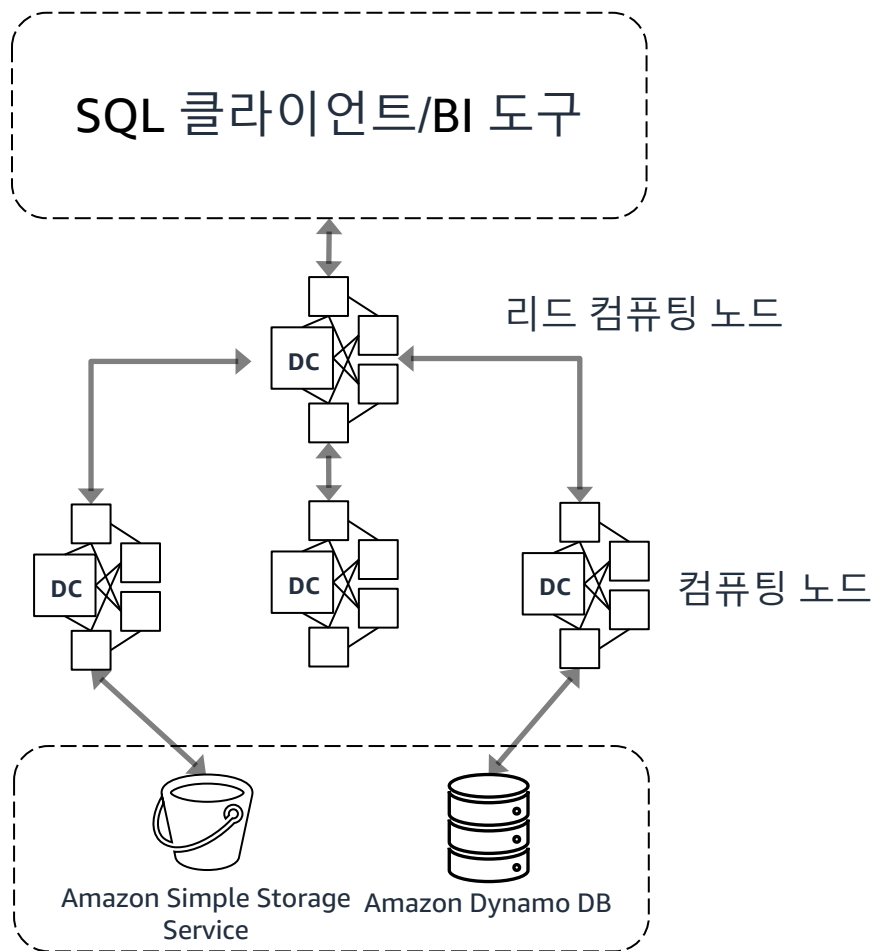
표준 정형 쿼리 언어(SQL) 지원
및 Java Database Connectivity(JDBC)
및 Open Database Connectivity(ODBC)용
고성능 커넥터 제공

Amazon Redshift의 병렬 처리 역학

병렬 처리는 2개 이상의 마이크로프로세서*가 동시에 프로그램을 분리하고 동시에 프로그램 태스크를 실행하는 컴퓨팅 방법입니다.

Amazon Redshift 작동 방식

아키텍처 다이어그램: 병렬 처리



Amazon Redshift 용례

Amazon Redshift 용례

엔터프라이즈 데이터 웨어하우스(EDW)

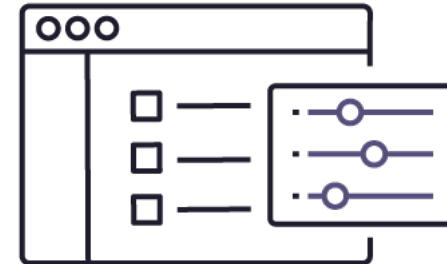
- 고객이 편한 속도로 마이그레이션
- 대규모 초기 비용이나 약정 없이 실험
- 비즈니스 요구에 더욱 신속하게 대응



Amazon Redshift 용례

빅 데이터

- 소규모 고객에게도 저렴한 비용
- 쉽게 배포 및 유지 관리할 수 있는 관리형 서비스
- 데이터베이스 관리보다 데이터에 더 집중



Amazon Redshift 용례

서비스형 소프트웨어(SaaS)

- 수요 증가에 따라 데이터 웨어하우스 용량 확장
- 애플리케이션에 분석 기능 추가
- 하드웨어 및 소프트웨어 비용 대폭 절감



학습 내용 확인



데이터 웨어하우스란 무엇입니까?

데이터 웨어하우스에는 몇 개의 티어가 있습니까?

Amazon Redshift를 사용하면 어떤 이점이 있습니까?

핵심 사항



- 데이터 웨어하우스는 보다 정보에 입각한 의사 결정을 내릴 수 있도록 분석 가능한 정보의 중앙 리포지토리입니다.
- 데이터 웨어하우스에는 여러 개의 데이터베이스가 포함될 수 있습니다.
- Amazon Redshift는 빠르고 완벽하게 관리되는 데이터 웨어하우스 서비스이며 가동 중단 시간 없이 쉽게 확장할 수 있습니다.
- Amazon Redshift는 클러스터를 거의 지속적으로 자동으로 모니터링하며 암호화가 내장되어 있습니다.

필수 항목:



이해도 점검