

### \*필요 개념

1. Data Warehouse (DW) : 하나의 통합된 데이터 저장공간으로서, 다양한 운영 환경의 시스템들로부터 데이터를 추출, 변환, 통합해서 요약한 데이터베이스

- 데이터베이스가 관련 있는 업무 데이터는 잘 저장하지만, 저장된 데이터들을 제대로 활용하지 못 하는 것에서 착안하여 시작

- 기본적으로 관계형 데이터베이스가 있는 상태를 가정하여 DW 를 구성하며, 동영상이나 음악처럼 DB 에 저장할 수 없는 파일도 필요한 부분을 추출하여 보여주어야 함

2. ETL(Extract, Transform, Load) : 데이터를 추출하고, 변형하여, (Data Warehouse 에) 적재하는 과정을 일컫는 말

3. BI(Business Intelligence) : 데이터 추출/통합/리포팅을 위한 기본도구 집합, DW 에서 분석된 데이터를 통해 숨겨진 패턴을 찾아냄

== > ETL 을 통해 뽑아낸 데이터를 DW 에 적재하고, BI 를 이용하여 분석하는 기본 과정을 거침

### \* Redshift 란?(모든 개념 중요!!!)

- PostgreSQL 을 기반으로 하는 AWS 의 Data Warehouse Service

- 모든 데이터를 표준 SQL 혹은 BI 도구를 사용하여 효율적으로 분석할 수 있도록 지원

- 대량 병렬처리(MPP)를 통해 복잡한 쿼리라도 빠른 속도로 실행하여 대용량 처리 가능

- 열(Column) 단위 데이터 저장방식

- COPY 명령어를 통해 Amazon EMR, Amazon DynamoDB, S3로부터 데이터를 병렬 로드 가능

- Enhanced VPC Routing 을 통해 클러스터와 VPC 외부의 COPY, UNLOAD 트래픽을 모니터링할 수 있음

- WLM(Workload Management)를 통해 사용자가 작업 부하 내 우선 순위를 유연하게 관리하도록 지원

- 보존기간이 1 일인 자동 백업을 지원하며, 최대 35 일까지 설정 가능

- 단일 AZ 배포만을 지원함

### \* Redshift 의 구성

- 클러스터 : Redshift 의 핵심 요소로, 하나의 리더 노드와 다수의 컴퓨팅 노드를 가지고 있는 구성 요소

- 리더 노드 : 클라이언트 프로그램과 일어나는 통신을 비롯해 컴퓨팅 노드간의 모든 통신/작업 관리

- 컴퓨팅 노드 : 실제 작업을 수행하는 노드로, 각 노드마다 전용 CPU 와 메모리 내장 디스크 스토리지를 따로 보유함

\* Redshift vs RDS

- Redshift 는 보고 및 분석에 사용되지만, RDS 는 OLTP (온라인 트랜잭션) 워크로드에 사용

- Redshift 는 대용량 데이터 세트를 대상을 복합적인 분석 쿼리를 빠르게 실행하는 것에 목표를, RDS 는 단일 행 트랜잭션에 목표를 둠

\*\*참고링크

[https://www.youtube.com/watch?v=BEML\\_jLP8xQ&list=PLpSOEOJwTzsp8JKrUkMMTV3naJfm\\_0XyO&index=8&t=2050s](https://www.youtube.com/watch?v=BEML_jLP8xQ&list=PLpSOEOJwTzsp8JKrUkMMTV3naJfm_0XyO&index=8&t=2050s) (현대백화점의 Redshift 활용 사례)

[https://d1.awsstatic.com/whitepapers/ko\\_KR/enterprise-data-warehousing-on-aws.pdf?trk=gs\\_card](https://d1.awsstatic.com/whitepapers/ko_KR/enterprise-data-warehousing-on-aws.pdf?trk=gs_card)

<https://digitalbourgeois.tistory.com/71>

데이터웨어하우스의 개념과 성능.pdf (최용락, 이정일)

[https://ko.wikipedia.org/wiki/%EB%8D%B0%EC%9D%B4%ED%84%B0\\_%EC%9B%A8%EC%96%B4%ED%95%98%EC%9A%B0%EC%8A%A4](https://ko.wikipedia.org/wiki/%EB%8D%B0%EC%9D%B4%ED%84%B0_%EC%9B%A8%EC%96%B4%ED%95%98%EC%9A%B0%EC%8A%A4)

<https://pizzaplanet.tistory.com/entry/Data-Warehouse-ETL-%EA%B0%84%EB%9E%B5-%EA%B0%9C%EB%85%90-%EC%A0%95%EB%A6%AC>

[https://ko.wikipedia.org/wiki/%EC%B6%94%EC%B6%9C,%\\_EB%B3%80%ED%99%98,%\\_EC%A0%81%EC%9E%AC](https://ko.wikipedia.org/wiki/%EC%B6%94%EC%B6%9C,%_EB%B3%80%ED%99%98,%_EC%A0%81%EC%9E%AC)

<https://namu.wiki/w/%EB%8D%B0%EC%9D%B4%ED%84%B0%20%EC%9B%A8%EC%96%B4%ED%95%98%EC%9A%B0%EC%8A%A4>