29일차

DP 900

구조화

반구조화: Document Data? – 문서 내부의 항목들이 문서들마다 약간씩 다른 경우

비구조화: 이미지, 텍스트 메시지, 음성 등

OLTP: 한번에 한 트랜잭션으로 싹 처리

OLAP: 주기적으로 분석해서 큐브에 저장해두고 처리

원자성 – 각 트랜잭션은 전부 성공하거나 전부 실패하거나. 모든 작업이 단일 단위로 처리

일관성 -

격리 – 여러 트랜잭션을 동시에 실행해도 데이터베이스는 동일하게 유지된다

영속성

일괄 처리 : 모아서 한번에 데이터 처리

스트림 처리 : 실시간으로 바로바로 처리

데이터의 역할

데이터베이스 (서버) 관리자

DB 관리, 보안 구현, 백업, 계정 관리, 성능 모니터링 등

데이터 엔지니어

데이터 수집하고 모아서 분석자가 분석할 수 있도록 데이터를 정리해둠

데이터 분석자

모인 데이터들을 분석하고 도움되는 정보를 만들어냄

데이터가 흐름에 따라서 그림으로 그려진 것: Entity Relationship Diagram

비관계형 데이터베이스 사용 사례

쭉~ 쌓아뒀다가 한번 열어서 살펴보는 용도로 주로 사용됨

게임에서도 주로 모바일 게임, 그 중에서도 턴제 게임에서 많이 사용

그래프 데이터베이스? -> 관계 중심으로 엔터티 저장 -> 계층적으로 표시하기 좋은 것?

-> 조직도… 네트워크 구성 등

Azure Data Services에서 지원하는 DB서버

구름모양: Microsoft SQL Server – 기업 운영

My: MySQL -> 가장 많이 사용됨

물개: MariaDB

코끼리: Postgre SQL

* 오픈소스이지만, MS가 책임지고 서비스 중

가장 호환성 높은건 당연히 IaaS인 Azure Virtual machines의 SQL Server

그러나 PaaS에서는 Azure SQL Managed Instance

아예 새로 구성할 때는 그냥 Azure SQL Database를 사용하면 된다!

인스턴스와 데이터베이스이 차이…?

단일 인스턴스: 한 대를 나 혼자 점유하여 사용

인스턴스 풀: 이것저것 쌓아 뒀다가 필요할 때 꺼내 쓰는 방식을 생각하면 됨

* 인력 풀처럼 그냥 일 하다가 뭔가 해야할 때 사람들을 모아서 프로젝트 진행하는 것

탄력적 풀: 우리 회사 데이터 베이스랑 다른 회사들의 데이터 베이스를 묶어서 서로 필요할 때에 리소스를 공유해서 사용 -> 다같이 바쁠 때에는 서로 다 리소스가 모자란 상황이 올 수도 있음…

가격은 가장 쌈

프로비전?

기술적 한계로 엄청 좋은 컴퓨터 1대보다 값싼 컴퓨터 여러 대가 훨씬 싸게 먹힘

그래서 저렴한 컴퓨터 여러 대를 그리드 방식을 통하여 뭉친 뒤, 가상화시켜서 그 위에 서버를 올리면 어떨까? -> 이게 현재 MS의 클라우드 서비스