30일차

SQL 문의 유형

DML: 데이터 조작 언어 (Manipulate)

* SELECT, INSERT, UPDATE, DELETE

DDL: 데이터 정의 언어 (Define)

* CREATE, ALTER, DROP, RENAME

DCL: 데이터 컨트롤 언어 (Control)

* GRANT, REVOKE, DENY

쿼리도구

Azure Portal

* Web browser

SQL management Studio

* Windows

SQL Server Data Tools

* Windows

Azure Data Studio

* Windows/Linux/Mac

SQLCMD

* Terminal

Azure CLI/Cloud Shell

* Terminal + Web browser

Azure Table Storage

키는 일정하게 지정되지만, 값(Column)은 일정하지 않음! -> 비정형, 비관계형 데이터

Azure Storage Account

* Blob - Container – 어쩌구저쩌구
* Files
* Table
* Queue

블록 Blob – 최대 크기 4.7T, 고객센터에 요청시 늘려주기도 함, 100MB 블록 5만개로 구성

페이지 Blob – 최대 8T, 512 바이트의 페이지 모음으로 구성

어펜드 Blob – 최대 195T 정도, 추가 작업에 최적화, 쓰기 작업에 유용, 주로 로그 저장

Azure File Storage

가상의 스토리지를 만들어, 가상머신에 저장 영역으로 사용하거나, 다른 사람들과 같은 디스크를 공유하여 사용할 수도 있음. 네트워크 디스크의 역할, SMB(서버 메시지 블록) 3.0 사용

부하 분산기를 사용하여 다수의 서버를 구성할 경우, 서로 다른 서버마다 스토리지를 연결하면 서버 간의 저장공간이 상이하여 데이터 공유가 어려우므로, 가상 스토리지를 이용하여 모든 서버에 같은 저장공간을 연결!!

Azure Cosmos DB?

이름은 DB이지만, 실제로는 DB가 아니라 관리 프로그램! API의 데이터가 많아질수록 원하는 데이터를 얻는 데는 시간이 길어질 수밖에 없음-> 더 많은 자원을 투자해야 응답시간을 줄일 수 있다.

Cosmos DB는 API의 응답속도를 보장, 데이터의 복제를 통해 타지역에서도 접근 가능하며 복제되어있는 지역에서도 읽기와 쓰기가 모두 가능!(메인과 서브의 구분이 없다) 동시에 수정해도 순차적으로 싱크가 다 맞춰진다. 또한 쓰기의 비용이 읽기이 약 3배이므로 종종 수정하고 읽기만 자주 하는 서비스에 적합하다. 주로 로그인 등의 서비스. 로그 자료를 남기는 경우 하루에 수천 번씩 입력해야 하므로 부적절하다.

NoSQL은 MongoDB API, JSON형식은 SQL API, 계층적인 자료는 Gremlin API를 사용!

Azure Databricks?

Apache Spark 기반의 서비스, 아파치니까 오픈소스와 연동 잘 됨

다양한 소스의 데이터를 가져와서 머신러닝, 데이터 엔지니어링 등 애저에서 제공하는 다양한 작업들을 할 수 있고, 그를 위한 작업 공간과 런타임, 보안 등이 제공됨

작업한 내용들을 다시 다양한 형태로 출력시킬 수 있다.

Azure HDInsight?

HD는 고화질이 아니라 Hadoop… 하둡을 애저에 올려놓은 프로그램이라고 생각하면 편하다

사실 하둡이 캐릭터는 귀엽게 생겼지만, 하나 돌리려면 서버가 최소 4대는 필요한 진짜 코끼리 같은 녀석…