

# 데이터관리론 Homework#13

201823869 조성우



Hadoop과 NoSQL은 빠르게 증가하는 데이터셋을 관리하기 위한 빅데이터 관점의 데이터기술이라는 점에서 매우 유사한 포지션을 가지고 있습니다.

하지만 NoSQL은 빅데이터의 거대한요구를 다룰 수 있게 해주는 분산형 '데이터베이스 인프라'인 반면 하둡은 거대한 병렬형 컴퓨팅을 가능하게 해주는 '파일시스템'이라는 점에서 차이가 존재합니다.

다시 말해 NoSQL은 특별히 데이터베이스 프레임워크로 고안된 개념인 반면, Hadoop은 데이터베이스 시스템이 아니고 방대한 병렬컴퓨팅을 가능하게 해주는 하나의 소프트웨어 생태계 시스템인 것입니다.

기능면에서도 이 둘의 대표적인 차이점 두가지가 존재하는데 아래와 같습니다.

## 1. Analytics vs Operational

Hadoop은 기본적으로 실시간 상호작용이 발생하지 않는 컴퓨터작업 또는 데이터 분석에 용이한 시스템인 반면

NoSQL은 조직 외부의 사용자들이 실시간으로 상호작용할 수 있는 real-time application을 가능하게 하도록 고안되었습니다.

이러한 NoSQL의 특징은 User의 data query를 용이하게 합니다.

## 2. Volume vs Velocity

Hadoop과 NoSQL은 둘 다 방대한 양의 데이터를 다룰 수 있습니다.

다만 데이터로부터 추가적인 분석을 목표로 방대한 데이터를 받아오려 할 때는 Hadoop이 최고의 솔루션이 될 수 있습니다.

이러한 방대한양의 데이터 분석에 용이한 Hadoop의 transmission input에 관한 장점을 Volume이라고 한다면

NoSQL은 system이 데이터로부터 output을 내놓기까지의 속도에 집중합니다. NoSQL은 데이터의 작은 부분집합으로 작업하며 아주 빠른 처리속도를 가지고 이는 Hadoop과 비교하여 output의 Velocity면에서 상당한 장점을 가집니다.

