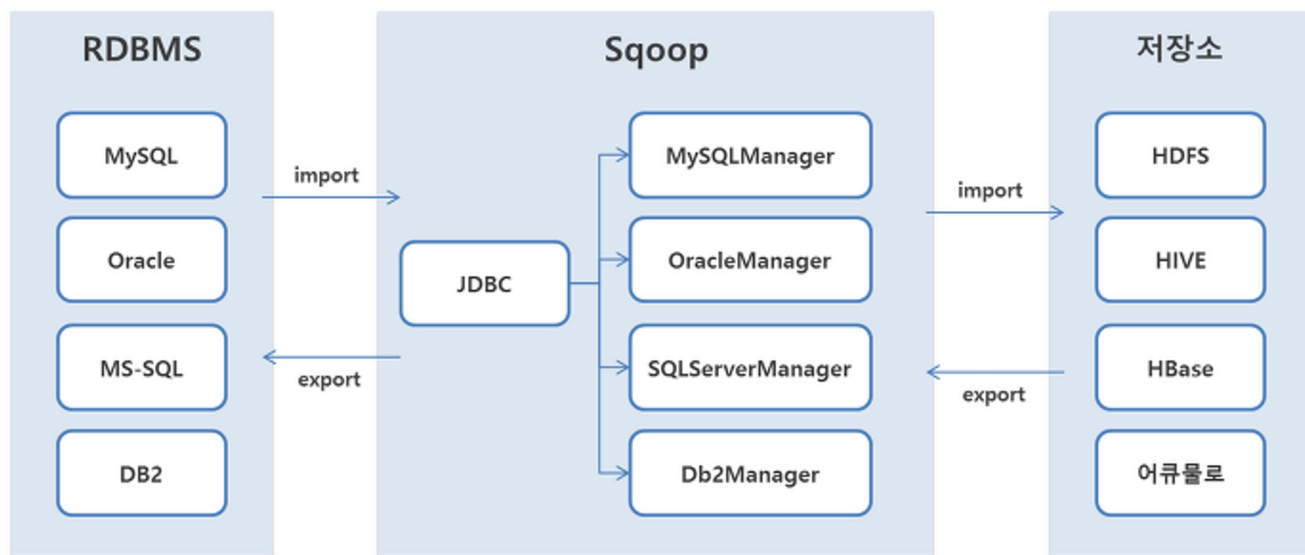


14. Sqoop

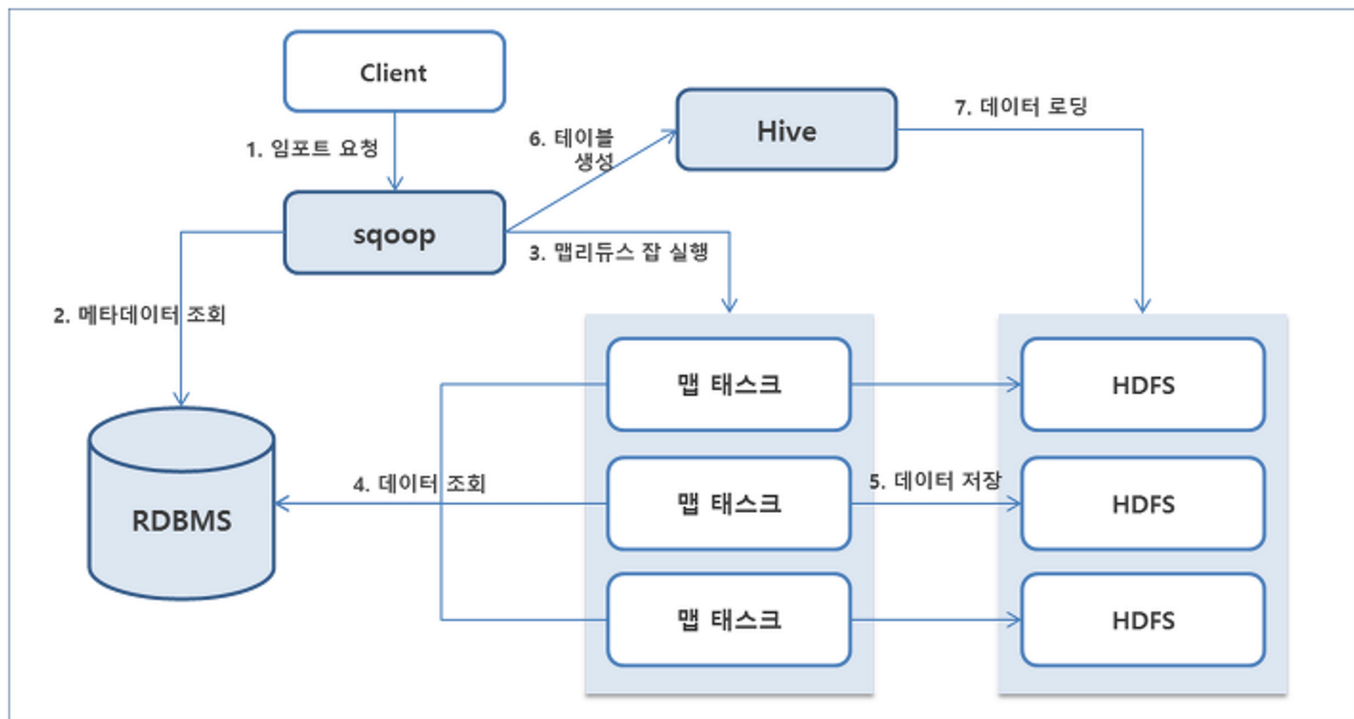
2020년 7월 26일 일요일 오후 1:59

- 구조화된 관계형 데이터 베이스와 아파치 하둡간의 대용량 데이터들을 효율적으로 변환하여 주는 명령 줄 인터페이스(Command-Line Interface) 애플리케이션
- 오라클 또는 MySQL같은 관계형 데이터 베이스에서 하둡 분산 파일 시스템으로 데이터들을 가져와서 그 데이터들을 하둡 맵리듀스로 변환을 하고, 그 변환된 데이터들을 다시 관계형 데이터 베이스로 내보낼 수 있음
- 스쿱은 데이터의 가져오기와 내보내기를 맵리듀스를 통해 처리하여 장애 허용 능력뿐만 아니라 병렬 처리가 가능하게 함

1. Sqoop 구조



- 데이터 임포트 동작



● 데이터 익스포트 동작

